Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Nash Strategies with Distance Discount Factor in Target Selection Problems

Yong Liu

Department of Electrical Engineering
The Ohio State University
Columbus, OH 43210
liuy@ee.eng.ohio-state.edu

David Galati, and Marwan A. Simaan

Department of Electrical Engineering
University of Pittsburgh
Pittsburgh, PA 15261
dgalati,simaan@engr.pitt.edu

ABSTRACT — In this paper, we consider a game theoretic approach for the task assignment problem associated with a team of semiautonomous unmanned aerial vehicles (UAVs) targeting as set of enemy ground units. A typical task would be to attack a unit or on the other side given a probability of kill of the UAV weapon against that unit. We introduce the concept of Distance Discount Factor (DDF) to address the fact that targeting close but less significant units could be more rewarding than targeting far but more significant units. To illustrate the effect of the DDF on the target assignment process, we consider a simulation example, involving a team of UAVs engaged in battle with a group of surface to air missiles (SAMs). We illustrate and compare the simulation results with and without DDF using feedback control.

I. Introduction

The planning and management of a military operation conducted by teams of semi-autonomous entities, such as unmanned aerial vehicles (UAVs) must take into account the uncertainty associated with the environment in which the operation will take place. One important type of uncertainty that is always present in a military operation is the presence and impact of an intelligent adversary. This uncertainty is accounted for by including the adversary in the mathematical model used to describe the battle space. Clearly, all the control decisions for the friendly units will be influenced by the presence of the adversary. The battle space model [1] simply indicates where the friendly controls enter in the mathematical model. The calculations are performed after each team's objectives are modeled and then the team optimal tactics are derived. This neutrality of the battle dynamics is also preserved with respect to the adversarial controls. The framework of non-cooperative nonzero-sum games is used to calculate the optimal tactics for each friendly team. In each team, the cooperative strategies among team members can be best analyzed using team theory [2]. Necessarily, the forecasted tactics of the adversary are calculated simultaneously. In this framework, the objective function of the adversary is considered as known or estimated. In [3], an efficient search algorithm was developed to calculate the Nash strategies for target assignment for both the friendly team (which will be referred to as the Blue force) and the adversarial team (which will be referred to as the Red force). Two implementations of this algorithm as open-loop and feedback controllers are described in [4]. The performance of these controllers and the advantages and disadvantages of each are evaluated in simulations performed on various scenarios using a simulator developed by Boeing [4].

The objective functions used in finding the Nash solutions are linear combination of the strength of the remaining units after the battle. The relative importance, or worth, of the corresponding units is often taken as the weight in the objective functions [4]. Using this formulation, a problem will arise when there are critical Red units with very high worth that are far from the attacking Blue units, and at the same time there are less important Red units which have strong killing capabilities but are closer to the attacking Blue force. In such a situation, the target selection algorithm, which assigns Red targets to the Blue units, would very likely ignore the less important Red units and assigns all Blue units to target the high worth Red units. This issue will be even more critical if the Blue force has limited resources as compared to the Red force. It is clear that, by ignoring the close Red units, the Blue force could itself be targeted by these units and may suffer considerable losses before it can reach the important Red targets. In this paper, we introduce a new parameter in the objective functions, which we call the distance discount factor (DDF). For each Red target, this factor is selected based on its specific geographic distribution with respect to the other Red targets. The paper is organized as follows. In section 2, we formulate the target assignment problem as a varying moving horizon game and define its corresponding Nash strategies. In section 3, we discuss the introduction of the DDF and its effect on the Nash strategy when implemented in feedback form. In section 4, we give a simulation example performed on a Boeing simulator to illustrate and compare the results with and without the DDF. Finally, in section 5 we present some concluding remarks.

0-7803-8335-4/04/$17.00

II. Varying Moving Horizon Nash Strategies

We consider a military operation, in which there are two opposing forces, referred to as Blue and Red, respectively. The Blue force consists of air power such as unmanned aerial vehicles and its objective is to attack some fixed targets that are defended by a ground based Red force. The Red force consists of Red ground troops, such as tanks and mobile vehicles, Red air defenses such as surface to air missiles (SAM's), and extremely important (or valuable) fixed Red targets. The objective of the Blue force is to destroy the Red fixed targets. Let I_B and I_R represent the number of Blue units and Red units, respectively. Let

$u^B(t) = \left(u_1^B(t) \quad \cdots \quad u_{I_B}^B(t)\right)^T$ and

$u^R(t) = \left(u_1^R(t) \quad \cdots \quad u_{I_R}^R(t)\right)^T$ be the control vectors,

and $s^B(t) = \left(s_1^B(t) \quad \cdots \quad s_{I_B}^B(t)\right)^T$ and

$s^R(t) = \left(s_1^R(t) \quad \cdots \quad s_{I_R}^R(t)\right)^T$ be state vectors for the Blue and Red forces respectively. The general objective function for each force is assumed to be of the form [5]:

$$\max_{u^B(t)} \left\{ \phi^B(s^B(T), s^R(T), T) + \int_{t_0}^{T} L^B(s^B(t), s^R(t), u^B(t))dt \right\}$$

for Blue force;

$$\max_{u^R(t)} \left\{ \phi^R(s^B(T), s^R(T), T) + \int_{t_0}^{T} L^R(s^B(t), s^R(t), u^R(t))dt \right\}$$

for Red force.

where T denotes the duration of the battle. The Nash strategies [6] are determined based on these objective functions. We note that ϕ^B and ϕ^R in the above expressions depend on the final state and final time. Also L^B and L^R depend on the states and control inputs at intermediate times in [t_0, T]. The general system attrition model can be described by the following dynamical equations:

$$\begin{cases} \dot{s}^B(t) = f^B(s^B(t), s^R(t), u^B(t), u^R(t), t) \\ \dot{s}^R(t) = f^R(s^B(t), s^R(t), u^B(t), u^R(t), t) \end{cases}$$

For the purpose of analyzing problem in the level considered in this paper, we will not consider the pointwise objective function L^B and L^R in the objective functions and we can do this by maintaining desirable vehicle control during those intermediate times. In other words, we consider only the terminal terms ϕ^B and ϕ^R in our objective functions. However, the Nash strategies over the entire time horizon T are extremely difficult to determine and even when determined may not be very useful for practical situations because of the uncertainty in the battle environment such as unknown or pop-up targets. To deal with this issue, the objective functions are optimized over a shorter time horizon $H < T$ instead of the entire the duration of the battle. The horizon H is called an optimization or planning horizon [7]. Moreover, in most applications, even though the Nash strategies are determined over the optimization horizon H , they are implemented over a much shorter time horizon $h \le H$. This horizon is called the implementation horizon [7]. An illustration of both types of horizons is given in Fig. 1. Note that as illustrated in Fig. 1, both the optimization and implementation horizons can vary with time. Let $\{t_k\}, k = 0,1,\cdots,$ denote the time instants when the optimization horizon $H(t_k)$ is updated and a new Nash strategy is solved over $H(t_k)$. It is clear that

$$t_{k+1} = \sum_{l=0}^{k} h(t_l)\,. \tag{1}$$

Let the control vectors $u_i^B(t_k)$ (or $u_j^R(t_k)$) denotes the target selection of the i[th] Blue unit (or the j[th] Red unit) at time t_k . We will assume that the objective of each force is to assign for each unit in that force a target in the adversarial force so that:

(1) The total strength of units and the total number of weapons of its own forces remaining at the end of each optimization horizon are maximized, and

(2) The total strength of units in the adversarial force remaining at the end of each optimization horizon is minimized.

Thus, the objective functions for the Blue and Red forces can be written as:

$$\max_{\substack{u^B(t)\in U^B(t)\\ t\in[t_k, t_k+H(t_k)]}} J_B(u^B, u^R) = \sum_{i=1}^{I_B}\left(w_i^B p_i^B(t_k+H(t_k)) + l_i^B q_i^B(t_k+H(t_k))\right) - \sum_{j=1}^{I_R} w_j^{BR} p_j^R(t_k+H(t_k)) \tag{2a}$$

for the Blue force, and

$$\max_{\substack{u^R(t)\in U^R(t)\\ t\in[t_k, t_k+H(t_k)]}} J_R(u^B, u^R) = \sum_{i=1}^{I_R}\left(w_i^R p_i^R(t_k+H(t_k)) + l_i^R q_i^R(t_k+H(t_k))\right) - \sum_{j=1}^{I_B} w_j^{RB} p_j^B(t_k+H(t_k)) \tag{2b}$$

for the Red force.

In the above expressions, $p_i^X(t)$ is the strength of the i[th] unit in force X (where X=B or R), $q_i^X(t)$ is the number of weapons carried by the i[th] unit in force X , and $U^X(\cdot)$ is the set of admissible controls for force X. In each of the previous expressions, we will assume that the w 's and l 's are all non-negative coefficients that represent weights to assign relative importance to the terms in the objective function. For example, w_i^B represents the value (or worth) of the i^{th} Blue unit, and l_j^R represents the value (or worth) of each weapon carried by the j^{th} Red unit.

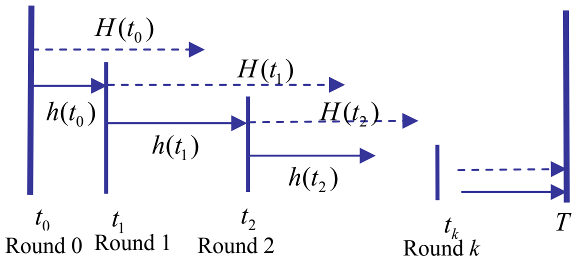

Fig. 1: Diagram of horizons H and h

The expressions of each weapon carried by the j^{th} Red unit. The expressions in (2) are linear combinations of the strength of units and weapons at the end of the optimization horizon and express the objective of each force to maximize the remaining strength of its own units and weapons while minimizing the remaining strength of units in the opposing force. The attrition model for $p_i^X(t)$ and $q_i^X(t)$ has been developed in [4]. Suppose that the target selections $u_i^B(t_k)$ and $u_j^R(t_k)$ ($i=1,\cdots,I_B; j=1,\cdots,I_R$) are fixed during the optimization horizon $[t_k, t_k+H(t_k)]$ resulting in one-step look-ahead Nash strategies. As was mentioned earlier, both optimization horizon $H(t_k)$ and implementation horizon $h(t_k)$ can be varied at different time instants t_k's. Therefore, the corresponding solutions are called varying moving horizon (VMH) one-step look-ahead Nash strategies.

Considering straight-line flight path of Blue air vehicles from the current location to the target location, the optimization horizon can be calculated as:

$$H(t_k) = \max_{\{B_i \in O_B(t_k), R_j \in O_R(t_k)\}} \left(d(b_i^B(t_k), b_j^R(t_k)) / v_i^B(t_k) \right) \quad (3)$$

where $b_i^B(t_k) = \begin{pmatrix} x_i^B(t_k) & y_i^B(t_k) & z_i^B(t_k) \end{pmatrix}^T$ (or $b_j^R(t_k) = \begin{pmatrix} x_j^R(t_k) & y_j^R(t_k) & z_j^R(t_k) \end{pmatrix}^T$) is the location coordinate of the i^{th} Blue unit (or the j^{th} Red unit) in three dimensional space at time t_k, and $O_B(t_k)$ (or $O_R(t_k)$) is the set of observed Blue units (or Red units) at time t_k, and $v_i^B(\cdot)$ is the flight speed of the i[th] Blue unit. Note that all Red units are fixed during the progress of battle. In expression (3), the Euclidean distance between the location of the i^{th} Blue unit and the location of the j^{th} Red unit is given by

$$d(b_i^B(t_k), b_j^R(t_k)) = \sqrt{(b_i^B(t_k) - b_j^R(t_k))^T (b_i^B(t_k) - b_j^R(t_k))} \quad (4)$$

The optimization horizon $H(t_k)$ represents the maximal time required by the slowest Blue UAV to reach the farthest Red target in one step. After determining the Nash strategy $\{u^{B*}, u^{R*}\}$ based on the objective functions given in (2), the implementation horizon can be calculated as:

$$h(t_k) = \min_{\{B_i \in O_B(t_k), R_j \in u_i^{B*}(t_k)\}} \left(d(b_i^B(t_k), b_j^R(t_k)) / v_i^B(t_k) \right). \quad (5)$$

In the above expression, $R_j \in u_i^{B*}(t_k)$ means that the i^{th} Blue unit selects the j^{th} Red unit as its target at time t_k. The implementation horizon $h(t_k)$ denotes the duration when the first Blue unit arrives at its selected target, and $t_k + h(t_k)$, i.e., t_{k+1}, is the next updated time instant.

III. Nash Strategies with DDF

We notice that the more important the Red unit, the higher weight it has in the Blue objective function and the more likely that it will be selected as a target. A problem will arise if the geographic distribution of the Red targets is such that some extremely important targets are farther than other less important targets. In that case, the Blue targeting strategy could select the more important Red units and ignore the less important ones. This will even be more crucial if the number of Blue UAVs is less than the number of Red targets. By ignoring the close Red units, it is very likely that these could destroy the Blue UAVs before they reach the important Red units. As a result, this may cause the Blue force to incur considerable losses before accomplishing its objectives. In order to deal with this problem, we will introduce a distance discount factor (DDF) with respect to the i^{th} Red unit in the following form:

$$\xi(\bar{b}^B(t_k), b_i^R(t_k)) = \exp\left(-\left(d(\bar{b}^B(t_k), b_i^R(t_k)) - \min_j \{ d(\bar{b}^B(t_k), b_j^R(t_k)) \} \right) \Big/ c \right) \quad (6)$$

where c is an adjustable positive constant, and $\bar{b}^B(\cdot)$ is the center location[1] of the grouped Blue units, which can be calculated as:

$$\bar{b}^B(t_k) = \frac{1}{I_B} \sum_{i=1}^{I_B} b_i^B(t_k) \quad (7)$$

The DDF concept is illustrated in Fig. 2 where the Red force consists of one extremely important Red unit R1 and two less important units R2 and R3. We notice that R2 and R3 are closer to one Blue UAV than R1. Because of the importance of R1, it is very likely that this UAV will be assigned R1 as a target leaving it to be a possible target for either R2 or R3. The corresponding DDF given by (6) is shown in Fig.3. We observe that the DDF is a decreasing function of the distance between the Blue UAV and the corresponding Red target.

[1] In reality, we could have a DDF between individual units; however for simplicity in this paper, we consider the distance between the center location of the Blue units and each Red unit.

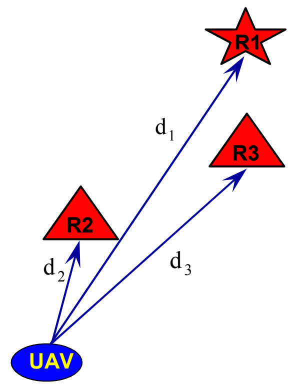

Fig. 2: Scenario of illustrating DDF

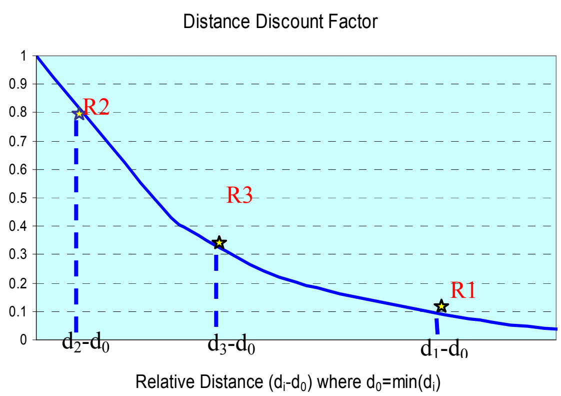

Fig. 3: DDF calculated for Scenario in Fig. 2

The objective function for the Blue force equipped with the DDF can be written as:

$$\max_{\substack{u^B(t)\in U^B(t)\\ t\in[t_k,t_k+H(t_k)]}} J_B^{(d)}(u^B,u^R) = \sum_{i=1}^{I_B}\left(w_i^B p_i^B(t_k+H(t_k)) + l_i^B q_i^B(t_k+H(t_k))\right)$$
$$-\sum_{j=1}^{I_R}\xi(\bar{b}^B(t_k),b_j^R(t_k))w_j^{BR}p_j^R(t_k+H(t_k))$$

We observe, in above expression, that the weighting coefficients corresponding to the Red units can be modified at the updated time t_k.

The one-step look-ahead Nash strategy with DDF $\{u^{B*},u^{R*}\}$ should satisfy the following inequalities:

$$J_B^{(d)}(u^{B*},u^{R*}) \ge J_B^{(d)}(u^B,u^{R*}) \quad \forall u^B(t)\in U^B(t), t\in[t_k,t_k+H(t_k)] \tag{8a}$$

$$J_R(u^{B*},u^{R*}) \ge J_R(u^{B*},u^R) \quad \forall u^R(t)\in U^R(t), t\in[t_k,t_k+H(t_k)] \tag{8b}$$

After calculating the Nash strategies for both forces, i.e., $\{u^{B*}(t),u^{R*}(t)\}, t\in[t_k,t_k+H(t_k)]$, only $\{u^{B*}(t),u^{R*}(t)\}$ in the time horizon $[t_k,t_k+h(t_k)]$ (or $[t_k,t_{k+1}]$) are implemented in the battle.

The varying moving horizon Nash strategies with DDF are implemented in feedback form as illustrated in [4]. In this implementation, the unit damage information from the battlefield at the updated time sequence $\{t_k\}$, which are fed back to the algorithm at the end of every step, are used to calculate the target assignments at the next step. In [4], our simulation results show that, the availability of feedback sensor information on target damage, as the battle progresses, will allow the Blue force to optimally allocate the available resources by avoiding the assignment of tasks that have already been satisfactorily accomplished, either fully or partially. In this paper, we assume that the sensor information on target damage can be obtained as the battle progresses. Therefore, we consider only the case of feedback control implementation.

An important issue that needs to be addressed in determining the Nash solution, even for the one-step look-ahead case, is scalability. An exhaustive search over the entire space of control options is feasible only if the number of units on each side is small. When the number of units on each side is larger than 6 or 7, the search space becomes too large and computationally not feasible to search within for the Nash solution. An efficient search algorithm that overcomes this scalability issue is described in [3] and will be implemented in determining the Nash strategy for the Blue feedback controller.

IV. Illustrative Example

In this section, we will illustrate the performance of the feedback controllers with and without DDF on a test bed scenario. In this scenario, the Blue force consists of a limited number of UAVs. The Red force has a limited number of long range and medium range, surface-to-air missiles (SAMs) and is centered in the Red area also labeled 3 shown in Fig. 4. The strategic objective for the Blue force is to eliminate the SAM sites in the Red area.

In order to compare the performance of the feedback controllers with and without DDF, we will consider a specific detailed experiment. Suppose that the Blue team of UAVs is dispatched to neutralize the Transport Erector Launchers (TELs), which are carrying SSMs, and the integrated air defenses (IADs) in Red area 3. TELs are critical offensive units and bring most risk to the Blue base, and the IADs are defending units including long range and medium range SAM sites. The deployment of Red forces in Red area 3 is shown in Fig. 4. The Blue team consists of a total of 5 UAVs equipped as described in Table 1.

The objective functions without DDF, i.e., $J_B(u^B,u^R)$ and $J_R(u^B,u^R)$, are given in (2). With respect to the calculated feedback target selections without DDF for the first four rounds for the Blue UAVs, it is not surprising that

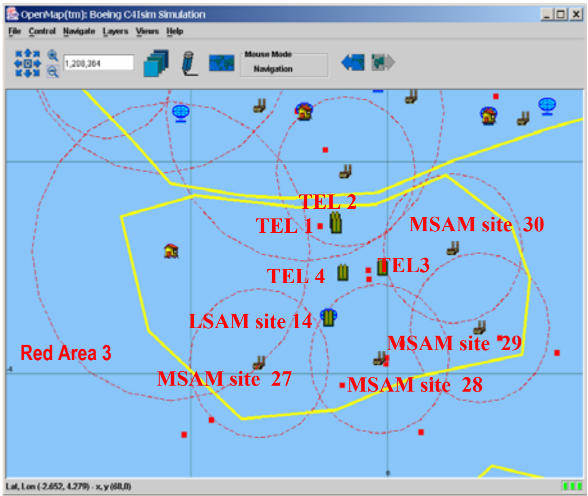

Fig. 4: Deployment of Red units in Red Area 3

Table 1. Blue Team assigned to Red Area 3

Blue UAVs in Team (assigned to Red Area 3)	Number of Unit (total=5)	Worth of each UAV	Weapon Type	Weapon Quantity Per Unit
Large Weapon	1	20	seeker missile	20
Small Weapon	2	20	seeker missile	8
Small Combo	2	20	seeker missile	4

all Blue UAVs have been assigned the critical targets TELs, instead of the defending SAM sites. This is so because, compared to the worth of other units, each TEL has an extremely high value. When these controls are implemented, after the first round of engagement, we observe that all the Blue UAVs except Small Combo 2 on their ways to the assigned targets are destroyed by the Red defending units. In the following rounds, Small Combo 2 still selects one TEL as its target and gets destroyed as expected.

In the next experiment, we use the DDF given by (6) to reduce the relative importance of the critical Red targets far from the Blue team. The corresponding objective function for the Blue force is now $J_B^{(d)}(u^B, u^R)$ and the objective function for the Red force remains $J_R(u^B, u^R)$ is given by (2b). We then calculate corresponding control choices for the first 5 steps for Blue UAVs. When these controls are implemented, the target assignment for the Blue units at the first round includes neutralizing those medium SAM site and long range SAM site, which are much closer to Blue team than the TELs. Similarly, at the second round, the Blue team continues to weaken the defending units and subsequently upon detecting that all TELs are now near the Blue team, the surviving Blue UAVs are now able to attack those important targets. Clearly, the feedback controller with DDF allows for more reasonable decisions to be made by the Blue team in the battlefield. Note that with DDF, four Long-SAM-14 launchers, one medium SAM site, and four TELs are destroyed. In addition, one Blue UAV is preserved. In contrast, in the case using the feedback controller without DDF, only two TELs are destroyed and none of Blue UAVs is preserved.

We also compared the worth of the remaining Red force and the remaining Blue team at the end of each round using the feedback controllers with and without DDF. These are shown in Fig.5 and Fig.6, respectively. The total worth of the Red and Blue force at step k is given by

$$W_X(k) = \sum_{i=1}^{I_X} w_i^X \, p_i^X(t_k) \text{ for X=B,R at round k} \qquad (9)$$

where the worth values of units w_i^B are obtained from the third column of Table 1, and w_i^R =75 for TELs, w_i^R =10 for LSAMs and w_i^R =7.5 for MSAMs , which are also used as weighting coefficients[2] in the objective functions (2). We note that the worth of the Red force when using the feedback controller with DDF is higher in the early stages of the battle, and then becomes lower than that of Red force when using the feedback controller without DDF as the battle progresses. This essentially confirms that only close and less important defending units are weakened at the early stages, and those critical targets with high values are then destroyed in the end. Similarly, more Blue UAVs are preserved when using the feedback controller with DDF.

Another measure is the total gain of the Blue force plus the total loss of the Red force. We refer to this as the net performance of the Blue controller and at round k is calculated according to:

$$Net(k) = \left(W_B(k) - W_B(0)\right) - \left(W_R(k) - W_R(0)\right) \qquad (10)$$

We compared the net performance of the Blue force when controls are implemented with and without DDF in feedback form. The results are shown in Fig.7. As we expected, the net performance of the Blue force tends to improve when using feedback controller with DDF as the battle progresses. However, this was not the case for the feedback controller without DDF.

V. Concluding Remarks

In this paper, we discussed a varying moving horizon Nash strategy with distance discount factor (DDF) for the target selection for the Blue force in a battle system. We introduced a distance discount factor to the objective

[2] These data are obtained from a simulator developed by Boeing as a part of the test bed.

functions used to calculate the optimal target selection, and presented simulation results to assess the performance of two implementations of the feedback controller with and without DDF. We used an efficient search algorithm for calculating the Nash target assignments. Our simulation results show that the DDF is important in target selection especially when the target units have different worth and the more valuable targets are farther than the less valuable ones. The DDF provides a tradeoff between the cost and reachability of the target.

VI. Acknowledgement

This research was sponsored in part by the Defense Advanced Research Projects Agency (DARPA) under contract number F33615-01-C-3151. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the DARPA, or the U.S. Government.

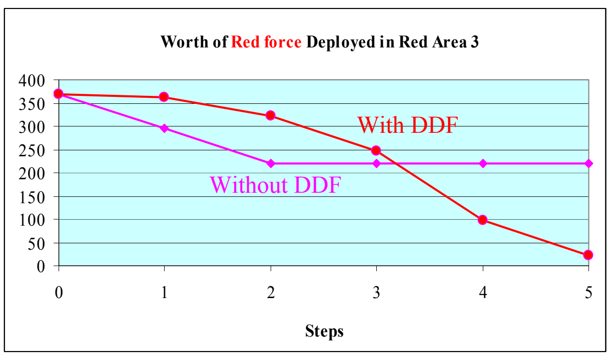

Fig.5: Worth of Red Force deployed in Red Area 3

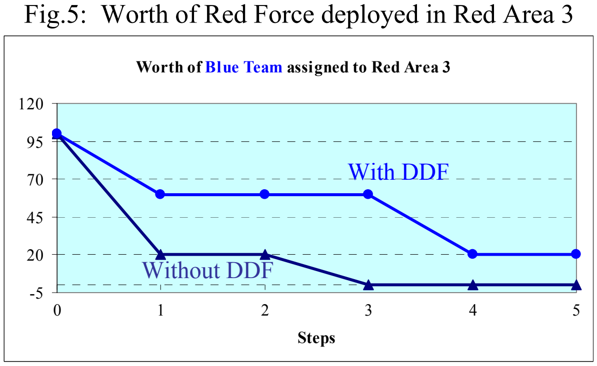

Fig.6: Worth of Blue Team assigned to Red Area 3

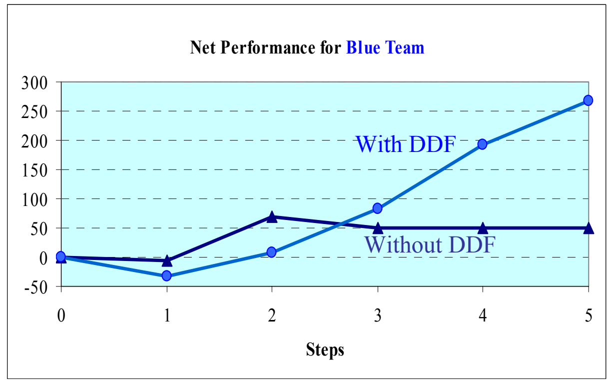

Fig. 7: Net Performance for Blue Team

VII. References

[1] J. S. Przemieniecki, Mathematical Methods in Defense Analysis, AIAA Education Series, 3rd Edition, 2000.

[2] Y. Ho and K. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems-Part I", IEEE Transactions on Automatic Control, AC-17, 1972, pp.15-22.

[3] D. Galati, Y. Liu and M. A. Simaan, "A Fast Algorithm for Unit Level Team Resource Allocation in a Game Environment", Proceedings of the 42nd IEEE CDC, Maui, Hawaii, December 2003, pp.2872-2877.

[4] Y. Liu, D. Galati and M.A. Simaan, "Team Dynamics and Tactics in Strategies for Human-Automaton Resource Entity Deployment", Proceedings of the 42nd IEEE CDC, Maui, Hawaii, December 2003, pp.3561-3566.

[5] Frank L. Lewis and Vassilis L. Syrmos, "Optimal Control," John Wiley & Son, Inc., 1995.

[6] T. Basar and G. J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, San Diego, CA, 1981.

[7] Christos G. Cassandras and Wei Li, "A Receding Horizon Approach for Solving Some Cooperative Control Problems", Proceedings of the 41st IEEE Conference on Decision and Control, Nevada USA, December 2002, pp.3760-3765.