43rd IEEE Conference on Decision and Control
December 14-17, 2004
Atlantis, Paradise Island, Bahamas

A Game Theoretic Approach to Team Dynamics and Tactics in Mixed Initiative Control of Automa-Teams

Yong Liu, *Member, IEEE*, David G. Galati, *Member, IEEE*, and Marwan A. Simaan, *Fellow, IEEE*

***Abstract*—The planning and operational hierarchy in future combat systems that involve unmanned aerial vehicles will very likely consist of three levels: a Team Composition and Tasking (TCT) level, a Team Dynamics and Tactics (TDT) level, and a Cooperative Path Planning (CPP) level. In this paper, we discuss a game theoretic approach for the target assignment problem at the TDT level. This problem considers the issue of assigning tasks to individual UAVs which could involve attacking one or more fighting units on the other side. The approach considered in this paper involves estimating the reaction of the other side for every possible UAV target assignment that could be made and calculating the resulting Nash equilibrium solution. We discuss an algorithm for determining the Nash solution which overcomes issues related to scalability and is capable of handling target assignments with a large number of non-homogeneous units on each side. In this paper, we discuss both the open-loop and feedback implementations of this algorithm and present simulation results that can be used to assess their performance. Our simulation results show that the availability of sensor information on target damage, as the battle progresses, will allow the feedback implementation at the TDT level to optimally allocate the available UAV resources by avoiding the assignment of tasks that have already been satisfactorily accomplished, either fully or partially. We also introduce the concept of distance discount factor (DDF) to address the fact that targeting close but less significant targets could be more rewarding than targeting far but more significant units. We discuss and compare the results of the feedback implementation with and without DDF.**

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.

Yong Liu is with the Department of Electrical and Computer Engineering, Ohio State University, Columbus, OH 43202 USA. (e-mail: liu.609@ osu.edu).

David G. Galati. is with the Department of Electrical Engineering, University of Pittsburgh, PA 15261(e-mail: dgalati@engr.pitt.edu).

Marwan A. Simaan is with the Department of Electrical Engineering, University of Pittsburgh, PA 15261 (Correspondence author, phone: 412-624-8099; fax: 412-624-8099; e-mail: simaan@engr.pitt.edu).

I. Introduction

THE challenge of using hierarchical decision-making and control for planning and battle management in future combat systems that involve mixed control of automa-teams is addressed in a recently developed system by researchers at the Ohio State University, the University of Pittsburgh, the University of Cincinnati, Iterativity Inc, and the Air Force Institute of Technology, called SHARED [1]. The architecture of SHARED is shown in Fig. 1. An important level in the operational hierarchy of this system is the Team Dynamics and Tactics (TDT) level. At this level, individual autonomous entities, such as unmanned aerial vehicles (UAVs) within a team are assigned to a given set of subtasks in order to accomplish an overall team task. These assignments could target specific units on the enemy side so as to support a collective team objective.

The planning and management of a military operation conducted by several teams of semi-autonomous entities (such as 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 (the enemy). In the SHARED system, this uncertainty is accounted for by including the adversary in the mathematical model used to describe the battle space [2]. The battle space model does not pre-judge how the friendly team strategies are obtained. It simply indicates where the friendly controls enter. The calculations are performed after the team objectives are modeled and then the optimal tactics are derived. The framework of non-cooperative nonzero-sum games is then used to calculate the optimal strategies of the team. Necessarily, the forecasted strategies of the adversary are calculated simultaneously. In this framework, the objective function of the adversary is considered as known or estimated.

In this paper, a search algorithm, called Unit Level Team Resource Allocation (ULTRA) [3], used for calculating the Nash strategies for target assignment for both the friendly units and the adversarial units is briefly reviewed. Two control implementations of this algorithm in open-loop and

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.

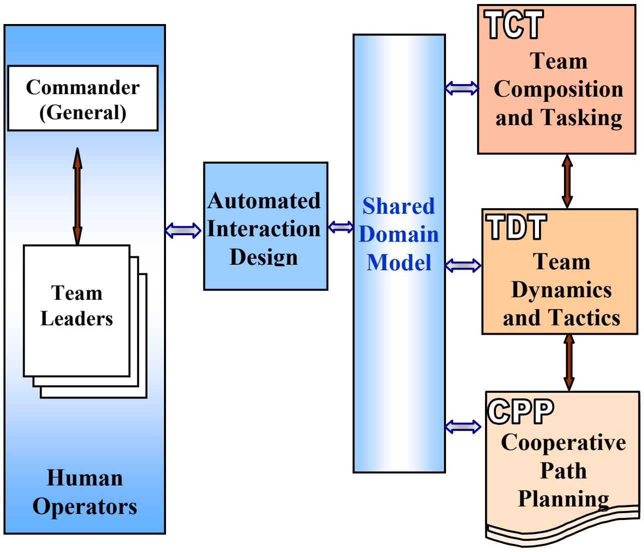

Fig.1. Hierarchical architecture of SHARED

feedback forms in the TDT level are also described. 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. The concept of distance discount factor (DDF) is then introduced to account for the fact that two identical targets have different values if they are at different distances from the UAV. We illustrate and compare its effect on the final outcome when ULTRA is implemented in feedback form.

II. THE ULTRA ALGORITHM

An important consideration in using a game theoretic approach at the TDT level is the dimensionality of the search space. Consider, for example, a team of N friendly units (referred to as Blue units from now on) engaged in combat with a team of M enemy units (referred to as Red units from now on). Assume that the units on each side are non-homogeneous, so that it would not be feasible to group them into a smaller number of sub-teams. A game theoretic approach will therefore need to examine all the possible combinations of target assignments (or control) options for units on one side against all units on the other side. On the Blue side, each unit has $M+1$ choices of targets consisting of the M Red units and the option of no target. Thus, the Blue side will have $(M+1)^N$ options. Similarly, the Red side will have $(N+1)^M$ options. Normally, these options are arranged in a matrix where the Blue options are represented as rows and the Red options as columns. If the effectiveness of the control options is assessed using objective functions $J_B(u,v)$ and $J_R(u,v)$ for the Blue and Red sides respectively, then each entry in this matrix will be a pair of real numbers $\{J_B(u,v),\ J_R(u,v)\}$ that correspond to the pair of options $\{u,v\}$ where u and v represent the Blue and Red control variables, respectively. Table I illustrates the dimensionality of this matrix for several values of N and M.

A pair of control options $\{u^N, v^N\}$ will represent a Nash equilibrium solution [4] if the following two inequalities hold:

$$J_B(u^N,v^N) \geq J_B(u,v^N) \quad \text{for all possible Blue control options } u;$$
$$J_R(u^N,v^N) \geq J_R(u^N,v) \quad \text{for all possible Red control options } v.$$

As is clear from Table I, an important issue that needs to be addressed in determining the Nash solution is scalability. An exhaustive search for the Nash solution over the entire space of control options is feasible only if the number of units on each side is small (less than 10). When the number of units on each side is larger than 10, the search space becomes too large and an exhaustive search becomes computationally not feasible. An efficient search algorithm called Unit Level Team Resource Allocation (ULTRA) that overcomes this scalability issue is described in [3] and is implemented in the TDT level. Essentially, ULTRA takes advantage of the structure of the control options available to each side. For a given control option on one side, it first optimizes the objective function for each member of the team on the other side, and then iterates among the remaining members of that team by changing once and then twice their target assignments while keeping the remaining assignments fixed. These iterations will continue until an optimum team response is reached. In this sense, this algorithm shares some of the properties of the Hooke-Jeeves and the Rosenbrock search algorithms for function minimization [5]. Once the team optimum response is determined, the roles of the two sides are interchanged, and the process repeated to determine the corresponding optimum team response for the other side.

TABLE I
DIMENSIONALITY OF THE TARGET ASSIGNMENT GAME MATRIX

Number of Blue Units N	Number of Red Units M	Size of Game Matrix: $(M+1)^N \times (N+1)^M$	Size of Search Space
4	3	256×125	32×10^3
8	6	$(5.76\times10^6)\times(0.53\times10^6)$	3.06×10^{12}
16	12	$(6.65\times10^{17})\times(5.82\times10^{14})$	38.71×10^{31}

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.

III. Open-Loop And Feedback Implementations Of The ULTRA TDT Controller

Fig. 2 shows a block diagram illustrating an open-loop implementation of the ULTRA controller at the TDT level. In this implementation, sensor information from the battlefield about damage assessment is either not available or cannot be obtained. This could be due to several reasons, including the breakdown of communication between the sensor UAVs and the TDT, or possibly the destruction of the sensor UAVs. In this case, and without such information, ULTRA will use an attrition model [2,6,7] to predict the battle damages, and uses these predictions to calculate the target assignments (controls) at the next step. Clearly, because of the probabilistic nature of any attrition model used in this context, ULTRA's prediction of the damages on each side could be considerably different from the actual damages in the battlefield. Consequently, the resulting target assignments at subsequent steps may not be the most effective. On the other hand, when real time information from the battlefield about unit damage assessment are available and can be transmitted to the algorithm, ULTRA can be implemented as a feedback controller as illustrated in Fig 3. In this implementation, the unit damage information from the battlefield, which are fed back to the algorithm at the end of every step, are used to calculate the target assignments at the next step.

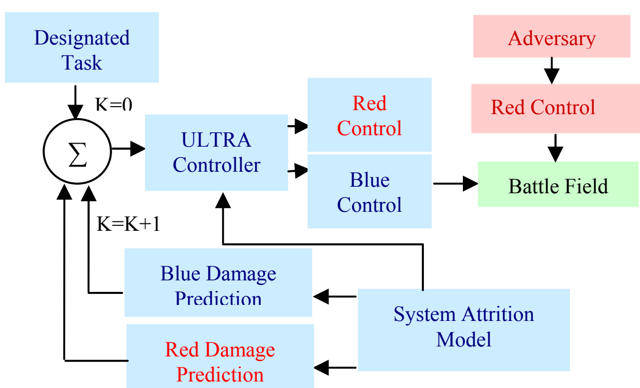

Fig.2. Block diagram of ULTRA open-loop controller

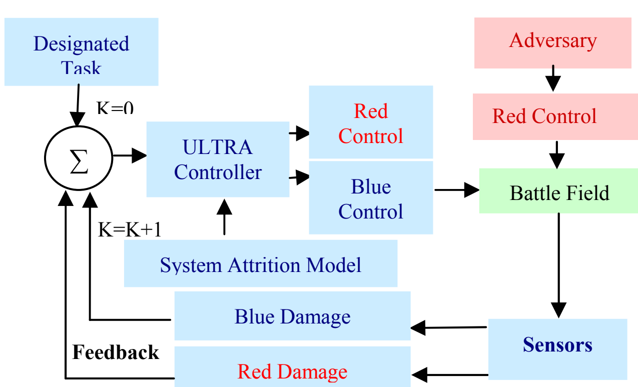

Fig. 3. Block diagram of ULTRA feedback controller

IV. Feedback Implementation With Distance Discount Factor

An important characteristic of the way the Blue objective functions are constructed is that the more important a Red unit is, the higher the weight it has in the objective function, and the more likely that it will be selected as a target. A problem will arise if the geometric distribution of the Red targets is such that some extremely important targets are farther than other less important targets. In this case, the Blue targeting strategy could select the more important units and ignore the less important ones. This will be even more crucial if the number of Blue UAVs is less than the number of Red units. By ignoring the less important, but geographically close, Red units it is very likely that these could destroy the Blue UAVs even 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 issue, a distance discount factor with respect to the i^{th} Red unit will be introduced in the form:

$$\xi(\bar{b}^B, b_i^R) = \exp\left(-\frac{d(\bar{b}^B, b_i^R) - \min_j\{d(\bar{b}^B, b_j^R)\}}{c}\right) \qquad (1)$$

where c is an adjustable positive constant, $\bar{b}^B$ is the center of gravity of the Blue team which is calculated as

$$\bar{b}^B = \frac{1}{N}\sum_{i=1}^{N} b_i^B \qquad (2)$$

and $b_i^X = (x_i^X, y_i^X, z_i^X)$ is the location coordinate vector of unit i in the Blue $(X = B)$ or Red $(X = R)$ force. In (1), $d(b_i^B, b_j^R)$ is the Euclidean distance between the i$^{\text{th}}$ Blue and j$^{\text{th}}$ Red units and is given by

$$d(b_i^B, b_j^R) = \sqrt{(b_i^B - b_j^R)^T (b_i^B - b_j^R)} \qquad (3)$$

Clearly, the distance discount factor is an exponentially decaying function introduced in the objective functions. As illustrated in Fig. 4, the Red force consists of an extremely important Red unit R1 which is farther from the UAV than two less important Red units R2 and R3. Because of the importance of R1, it is very likely that this Red unit will be assigned to the UAV as a target, hence leaving the UAV as a possible target for either R2 or R3. The DDF discount factor given by (1) will prevent this from happening.

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.

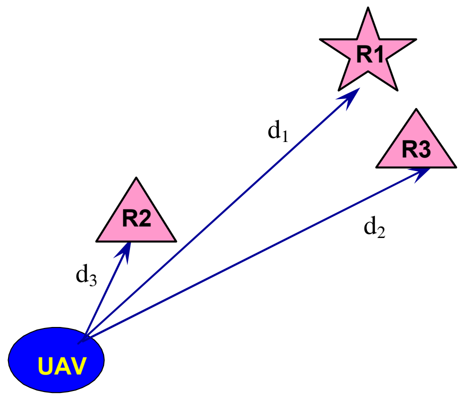

Fig. 4. Red units at different distances from a UAV

V. Illustrative Examples

In this section, we will illustrate the performance of the developed control strategies on a test bed scenario described in Fig. 5. In this scenario, it is assumed that the Red side (the enemy) is occupying Red areas 1, 2 and 3 and has integrated its air defenses into a neighboring region (Red area 0). These air defenses consist of EW radars and C2 structures and are deemed as acceptable targets. The Blue force consists of a limited number of ground forces and UAVs. The Blue base is centered in Blue area 3 as shown in Fig. 5 and other supporting Blue forces can be deployed in Blue areas 1 and 2. The Red force has a limited number of long range and medium range, surface-to-air missiles (SAMs). The strategic objective for the Blue force is to protect the Blue operating base from attack by the Red surface-to-surface missiles (SSMs) and armor; and to eliminate the Red SAM sites.

A. Experiment 1: Open-Loop vs. Feedback

In order to compare the performance of the developed game-theoretic strategies in open-loop and feedback form, we will consider a specific detailed experiment performed in Red area 2. Suppose that one Blue team of UAVs is dispatched to neutralize the ground forces and the integrated air defenses (IADs) in Red area 2.

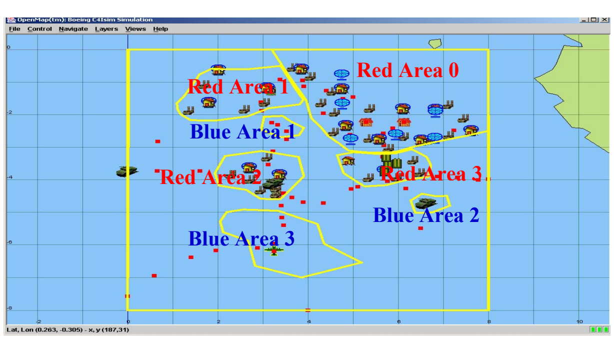

Fig. 5. Scenario battlefield

The ground forces include tanks, personal carriers, communication vans, etc., and the IADs include long range, medium range, and mobile SAM sites. The deployment of Red forces in Red area 2 is shown in Fig. 5. A complete description including the initial equipment, the number of units, the worth of units, the weapon types and quantities for each Red unit in that area is given in Table II. The Blue team, on the other hand, consists of a total of 11 UAVs: 3 UAVs equipped with large weapons (16 seeker missiles), 4 UAVs equipped with small weapons (6 seeker missiles) and 4 UAVs equipped with small combos (4 seeker missiles). A complete description of the Blue units is given in Table III (the data in tables II and III are obtained from a simulator provided by Boeing).

The objective functions $J_B(u,v)$ and $J_R(u,v)$ are given by

$$J_B(u,v) = \sum_{k=1}^{K} W_B(k) \text{ and } W_B(k) = \sum_{i=1}^{I_B} w_i^B p_i^B(k) \text{ for Blue} \quad (4a)$$

$$J_R(u,v) = \sum_{k=1}^{K} W_R(k) \text{ and } W_R(k) = \sum_{i=1}^{I_R} w_i^R p_i^R(k) \text{ for Red} \quad (4b)$$

where the worth values of units in the third column of Table II and Table III are used as weighting coefficients w_i^B and w_i^R in the above objective functions. I_X is the total number of units for the force *X*, and $p_i^X(x)$ is the number of the remaining ith unit in the force *X* at step *k* (*X*=B,R). The ULTRA algorithm calculated the open-loop target selections for the first four steps for the Blue UAVs and these are then implemented on the Boeing simulator. For comparison purposes, a partial outcome of the battle in both the open-loop and feedback cases is given in Table IV. A complete comparison of the outcome of the battle after every step of the 4 steps is given in [8]. Note that with feedback, four Long-SAM-13 launchers, one medium SAM site, and all the ground troops are destroyed. In addition, six Blue UAVs are preserved.

TABLE II
Red Force In Red Area 2

Red Unit (Red Area 2)	Number of Unit (total=38)	Worth of each unit	Weapon Type	Weapon Quantity Per Unit
Long Range SAM sites (2 sites)	8	10	long_sam _missile	4
Medium Range SAM sites(6 sites)	6	7.5	medium_ sam_missile	8
Tanks	10	10	tank_projectile	50
SPARTY	4	10	artillery_proj-ectile	100
Personnel Carriers	5	10	small_arms	20
Communication Vans	1	10	surface_to_su-rface_missile	4
Mobile SAM sites	4	7.5	medium_sam_ missile	8

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.

TABLE III
BLUE TEAM 1 ASSIGNED TO RED AREA 2

Blue UAVs in Team 1 (assigned to Red Area 2)	Number of Unit (total=11)	Worth of each UAV	Weapon Type	Weapon Quantity Per Unit
Large Weapon	3	20	seeker missile	16
Small Weapon	4	20	seeker missile	6
Small Combo	4	20	seeker missile	4

TABLE IV
PARTIAL OUTCOME OF THE BATTLE IN RED AREA 2

	Open-Loop Controller	Feedback Controller
8 long sam launchers	1 destroyed	4 destroyed
6 medium sam sites	0 destroyed	1 destroyed
11 UAVs	3 preserved	6 preserved

In contrast, in the open-loop case, only one Long-SAM-13 launcher is destroyed, and three Blue UAVs are preserved.

B. Experiment 2- Feedback With and Without DDF

In this section, we will illustrate the performance of the feedback controllers with and without DDF. 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 3 in Fig. 6. 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. Assume that each TEL has an extremely high value. The Blue team consists of a total of five UAVs.

The objective functions without DDF, i.e., $J_B(u^B,u^R)$ and $J_R(u^B,u^R)$, are given in (4). With respect to the calculated feedback target selections without DDF for it is not surprising that all Blue UAVs have been assigned the critical targets TELs, instead of the defending SAM sites. 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.

We now introduce the DDF given by expression (1) in

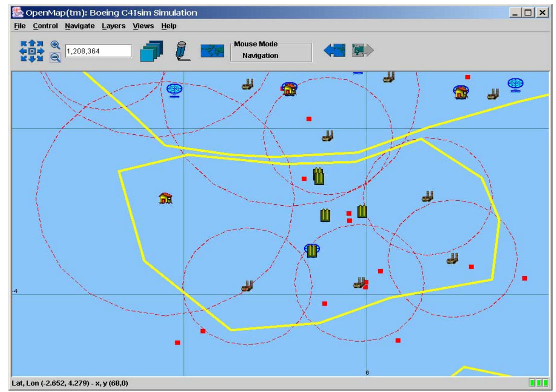

Fig. 6. Deployment of red units in red area 3

the objective function of the Blue force so as to reduce the relative importance of the critical Red targets that are far from the Blue team. The objective function of the Red force remains unchanged. We then calculate the corresponding control choices for the first four rounds for the 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.7 and Fig.8 respectively. The total worth of the Red and Blue force at step *k* is given in expressions (4). 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, which we also used in the experiment, is the net performance of the Blue controller. This is the total gain of the Blue force plus the total loss of the Red force as given by

$$Net(k) = \left(W_B(k) - W_B(0)\right) - \left(W_R(k) - W_R(0)\right) \qquad (5)$$

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.

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.9. As we expected, the net performance of the Blue force tends to improve when using feedback controller with DDF as the battle progresses.

VI. CONCLUSIONS

In this paper, we discussed a game theoretic approach for the TDT level in the SHARED system. We used an efficient search algorithm called ULTRA for calculating the Nash target assignments for all units on one side against all units on the other side. We also described two different implementations of the ULTRA controller depending on whether information about battle damage is continuously available to the TDT level or not. We also introduced the concept of distance discount factor so as account for the fact that targeting close but less significant enemy units could me more rewarding than targeting far but more significant units.

We presented two experiments performed on the Boeing simulator to assess the performance of these two implementations of the ULTRA controllers and the effects of the distance discount factor. We note that in this simulator, the Red controls are provided from within and are not known in the process of generating the Blue controls. Thus, the Blue controller must guess or estimate the Red controls. In the first experiment, we compared the worth of the remaining Blue and Red units, and the net performance of the Blue controls in the open-loop and feedback cases. Our simulation results show that the performance of the Blue force is considerably improved when feedback controls are used. In the second experiment, we compared the results when the controls are implemented in feedback form with and without the distance discount factor. 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.

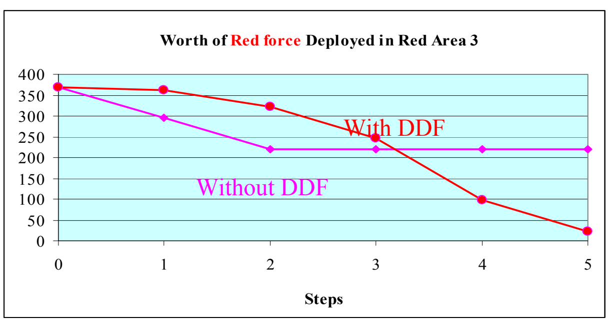

Fig.7. Worth of Red Force deployed in Red Area 3

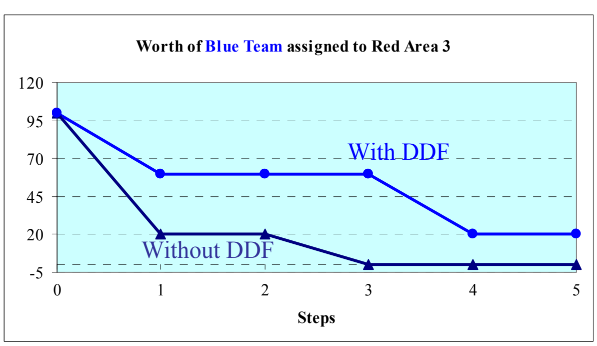

Fig.8. Worth of blue team assigned to red area 3

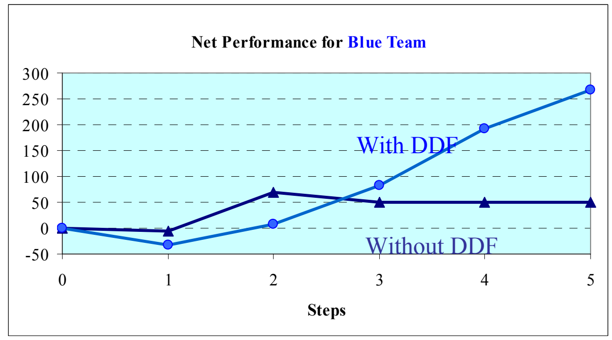

Fig. 9. Net performance for blue team

REFERENCES

[1] Strategies for Human-Automaton Resource Entity Deployment, Special Invited Session (ThAPI), *Proceedings of the 42nd IEEE CDC*, Hawaii, December 2003, pp. 3549-3590

[2] J. B. Cruz, Jr., M. A. Simaan, A. Gacic, H. Jiang, B. Letellier, M. Li, and Y. Liu "Game- Theoretic Modeling and Control of Military Operations" *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 37, No. 4, 2001, pp. 1393-1405.

[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*, Hawaii, December 2003, pp. 2872-2877.

[4] A.W. Starr, and Y.C. Ho, "Nonzero-Sum Differential Games", *Journal of Optimization Theory and Applications*, Vol.3, No.3, 1969, pp.184-206.

[5] M. Aoki, *Introduction to Optimization Techniques*, Macmillan: New York, NY, 1971.

[6] J. B. Cruz, Jr., M. A. Simaan, A. Gacic, and Y. Liu "Moving Horizon Nash Strategies for a Military Air Operation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 3, 2002, pp. 989-999.

[7] Y. Liu, M.A.Simaan, and J.B.Cruz, Jr. "An Application of Dynamic Nash Task Assignment Strategies to Multi-Team Military Air Operations," *Automatica*, Vol. 39, 2003, pp.1469-1478.

[8] 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*, Hawaii, December, 2003, pp. 3561-3566.

Authorized licensed use limited to: University of Melbourne. Downloaded on December 01,2022 at 03:58:13 UTC from IEEE Xplore. Restrictions apply.