The protection of defended assets such as military bases and population centers against hostile targets (e.g., aircrafts, missiles, and rockets) is a highly relevant problem in the military conflicts of today and tomorrow. In order to neutralize threats of this kind, they have to be detected and engaged before causing any damage to the defended assets. We review algorithms for solving the resource allocation problem in real-time, and empirically investigate their performance using the open source testbed SWARD. The results show that many of the tested algorithms produce high quality solutions for small-scale problems. A novel variant of particle swarm optimization seeded with an enhanced greedy algorithm is described and is shown to perform best for large instances of the real-time allocation problem.

1. INTRODUCTION

A severe threat encountered in many international peacekeeping and peace forcing operations is that of rockets, artillery, and mortars (RAM) fired by insurgents towards military bases, troops, and other assets. Attacks like these have cost many human lives in places like Iraq and Afghanistan during recent years. Similar attacks are faced by civilians in some parts of Israel on a regular basis, where so called Katyusha and Qassam rockets are fired against Israeli population centers such as Sderot and Ashkelon. Asymmetrical threats like these have caused an increased interest in systems for detecting and tracking incoming RAM before they hit their intended targets. The detection and tracking of RAM makes it possible to estimate the point of impact, so that any troops or civilians in the impact area can be alerted. However, such a warning is not always enough, due to very quick course of events, and that buildings and infrastructure will be destroyed no matter how early warnings come, given that active countermeasures are not taken. Hence, one would like to destroy incoming RAM before they hit their intended targets (and before they risk causing collateral damage upon destruction). Systems for detecting, tracking, and engaging RAM are often referred to as Counter Rocket, Artillery, and Mortar (C-RAM) systems. An example of such a system is the recently deployed Israeli Iron Dome system. Another kind of air defense situation is that in which we would like to protect defended assets against maneuvering targets such as fighter aircrafts, attack helicopters, and non-ballistic missiles. For such kind of threats, we can in general not easily predict which defended asset (if any) is the intended target of an attack, making it necessary to estimate the level of threat posed by detected targets to the defended assets in a so-called threat evaluation process.

When faced with many simultaneous threats, it is unlikely that the defenders can take action against all incoming threats, since there often are fewer firing units available than there are threats. Even when this is not the case, a problem is to know which firing unit to use against which threat in order to maximize the survivability of the defended assets or minimize the total expected target value of surviving hostile targets. This can be described as a resource allocation problem, known as the weapon allocation problem [7] within the field of operations research. Unfortunately, the allocation of defensive firing units to targets has been shown to be NP-complete [23].

The time available for weapon allocation depends on many factors such as the type of RAM used, the range from which it is fired, type of detection radar and type of defensive weapons (rapid-fire guns, lasers, radar-guided missiles, etc.). However, taking into account that the incoming threats often have high speed and are fired from a range of only a few kilometers, very short time is available for detection, weapon allocation, and
interception. Hence, empirical results for how weapon allocation algorithms perform on problem instances of various size are needed.

We have reviewed the available literature in order to identify suitable weapon allocation algorithms, and we have implemented and systematically evaluated the real-time performance of a selection of the identified algorithms on static asset-based weapon allocation problems. The results show that especially particle swarm optimization algorithms produce high quality solutions for small-scale problems. In this article, we describe a novel variant of particle swarm optimization seeded with an enhanced greedy algorithm and show that the seeded version performs very well relative to previously tested algorithms also for large-scale instances of the real-time allocation problem.

The rest of this article is structured as follows. In Section 2, we present the static asset-based weapon allocation problem, which is a suitable optimization model when the impact area of a threat can be assumed to be known. We also present its target-based counterpart which is more suitable for air defense situations involving maneuvering targets. In Section 3, we present a literature survey of algorithms that have been suggested for static weapon allocation (both target-based and static-based). Based on this survey, we have implemented algorithms for static asset-based weapon allocation which are presented in Section 4. Experiments in which we compare the real-time performance of the implemented algorithms are presented in Section 5, and we conclude the article in Section 6.

2. WEAPON ALLOCATION

Informally, weapon allocation (often also referred to as weapon assignment or weapon-target allocation) can be defined as the reactive assignment of defensive weapon resources (firing units) to engage or counter identified threats (e.g., aircrafts, air-to-surface missiles, and rockets) [29]. More formally, the weapon allocation problem can be stated as a non-linear optimization problem in which we aim to allocate firing units so as to minimize the expected total value of the targets, or, alternatively, to maximize the expected survivability of the defended assets. These alternative views are referred to as target-based (weighted subtractive) defense and asset-based (preferential) defense, respectively. The asset-based formulation demands knowledge of which targets that are headed for which defended assets and thereby assumes a high level of situation awareness [24]. Therefore, the static asset-based weapon allocation problem formulation is suitable for problems involving defense against ballistic weapons, while the target-based formulation is more appropriate when the intended aims of the targets are not known [28]. In Section 2.1 we describe the static asset-based weapon allocation problem, and in Section 2.2 we give a similar description of the static target-based weapon allocation problem.

2.1. The Static Asset-Based Weapon Allocation Problem

When presenting the static asset-based weapon allocation problem, the following notation will be used:

- \(|A|\) denotes the number of defended assets.
- \(|W|\) denotes the number of firing units.
- \(|T|\) denotes the number of targets.
- \(\omega_j\) denotes the protection value of defended asset \(A_j\).
- \(P_{ik}\) denotes the probability that firing unit \(W_k\) destroys target \(T_i\) if assigned to it.
- \(\tau_i\) denotes the probability that target \(T_i\) destroys the asset it is aimed for.
- \(x_{ik}\) is defined as:
  - 1 if firing unit \(W_k\) is assigned to target \(T_i\),
  - 0 otherwise.

In the static asset-based weapon allocation problem, each offensive target is assumed to be aimed at a defended asset, where each defended asset is associated with a protection value \(\omega_j\). Each target has an associated lethality probability \(\tau_i\), indicating the probability that \(T_i\) destroys the defended asset it is aimed for, given that it is not successfully engaged. This probability depends on the accuracy of the targets as well as the nature of the defended assets [8]. As can be seen, we are assuming that such probabilities are target dependent only, i.e., we do not take the type of the defended asset into consideration. The defenders are equipped with firing units, where each pair of firing unit and target is assigned a kill probability \(P_{ik}\). Now, the objective of the defense is to allocate the available firing units so as to maximize the total expected protection value of surviving defended assets [7]:

\[
\max J = \sum_{j=1}^{\mid A \mid} \omega_j \prod_{i \in G_j} \left( 1 - \tau_i \prod_{k=1}^{\mid W \mid} (1 - P_{ik})^{\omega_a} \right) \tag{1}
\]

subject to:

\[
\sum_{i=1}^{\mid T \mid} x_{ik} = 1, \quad \forall k \tag{2}
\]

\[
x_{ik} \in \{0, 1\}, \quad \forall i \forall k.
\]

In (1), the inner product \(\prod_{k=1}^{\mid W \mid} (1 - P_{ik})^{\omega_a}\) should be interpreted as the probability that target \(T_i\) survives the countermeasures taken against it. Hence, the product \(\prod_{i \in G_j} (1 - \tau_i \prod_{k=1}^{\mid W \mid} (1 - P_{ik})^{\omega_a})\) is the probability that the defended asset \(A_j\) survives the attack of all targets aimed for it.
A solution to a static weapon allocation problem can be represented as a matrix of decision variables

$$
\mathbf{x} = \begin{bmatrix}
   x_{11} & x_{12} & \ldots & x_{1|W|} \\
   x_{21} & x_{22} & \ldots & x_{2|W|} \\
   \vdots & \vdots & \ddots & \vdots \\
   x_{|T|1} & x_{|T|2} & \ldots & x_{|T||W|}
\end{bmatrix}.
$$

(3)

Such a solution is feasible if it fulfills the constraints given in (2), i.e., that the entries of each column in (3) sum to one. For a problem instance consisting of $|T|$ targets and $|W|$ firing units, there are $|T|^{|W|}$ feasible solutions.

From the solution of the static asset-based weapon allocation problem, we can discover which of the defended assets that should be protected, and in which way each of the defended assets should be protected (preferential defense).

2.2. The Static Target-Based Weapon Allocation Problem

Using the same notation as in Section 2.1, but with the additional definition:

- $V^T_{i}$ = target value of target $T_i$,

we can define the static target-based weapon allocation problem as:

$$
\min F = \sum_{i=1}^{|T|} V^T_i \prod_{k=1}^{|W|} (1 - P_{ik})^{x_{ik}}
$$

(4)

subject to the constraints given in (2). Since the product as before is the probability that target $T_i$ survives the countermeasures taken against it, the objective function should be interpreted as the minimization of the total expected target value of surviving targets.

The estimation of target values is far from trivial, and can be seen as a very important high-level information fusion problem. A survey of how threat values $V_j$’s can be estimated (representing the threat posed by target $T_j$ to defended asset $A_i$) is presented in [11, 12]. Once such threat values have been calculated, these can be aggregated into target values using weighted averages such as:

$$
V_i = \frac{\sum_{j=1}^{|A|} V^A_j \omega_j}{\sum_{j=1}^{|A|} \omega_j}.
$$

(5)

Nevertheless, this is only one choice of how to aggregate threat values into target values. Furthermore, the original target values rely on coarse models of what is threatening behavior or not (typically parameters such as distance between the target and the defended asset, the speed and heading of the target, threat type, etc.). To complicate matters, expert air defense operators frequently disagree about the threat of individual aircraft [31]. Consequently, it should be remembered that target values will always be associated with uncertainty, and that they to a large degree are subjective.

2.3. Properties of Weapon Allocation Problems

A few assumptions are made in the static weapon allocation formulations. Firstly, all firing units have to be assigned to targets, as indicated in the constraint given in (2). Moreover, all the firing units have to be assigned simultaneously, i.e., we can not observe the outcome of some of the engagements before a remaining subset of firing units are allocated. This is what is meant by static weapon allocation, as opposed to dynamic weapon allocation. The static formulation makes sense for the problem domain studied here, since the high speed of short-range RAM does not allow for several engagement cycles. We also assume that an engagement will not affect other engagements (e.g., that a firing unit can destroy another target than it is allocated to, or that targets can destroy other assets than they are aimed for). Without the last assumption, the geometry of the problem must be taken into account, creating an extremely complex problem. We also ignore the risk of collateral damage to the protected area when intercepting the targets.

Despite the assumptions, there is a combination of factors that make the static weapon allocation problems hard to solve. Firstly, the objective functions given in (1) and (4) are non-linear, so that well-known linear programming techniques such as the simplex algorithm can not be used to solve the problems. Secondly, the problems are discrete, since they only allow for integer valued feasible solutions due to the second constraint in (2) (i.e., fractional allocations are not possible). In general, this kind of integer programming problems are hard to solve. Thirdly, the problems are stochastic, due to kill probabilities (and lethality probabilities) not equal to zero or one. This non-determinism further complicates the problems. Fourthly, it is not unusual with large-scale problem instances, i.e., problems consisting of a large number of firing units, defended assets, and/or targets. The asset-based formulation can be shown to be a generalization of the static target-based weapon allocation formulation [7] presented in Section 2.2. The NP-completeness of the static target-based weapon allocation problem was established in [23], and hence, we can conclude that the static asset-based weapon allocation problem is NP-complete as well [7]. These properties taken together show that finding good solutions in real-time to static weapon allocation problems is indeed a very hard problem, and according to [7], rule out any hope of obtaining efficient optimal algorithms.

3. A SURVEY OF ALGORITHMS FOR WEAPON ALLOCATION

Initial research on the static target-based weapon allocation problem dates back as far as the end of the 1950s (cf. [5, 25]). Much of the initial research on the
problem seems to have been motivated by the threat from intercontinental ballistic missiles during the Cold War era [24]. Despite the end of the Cold War, research on defensive weapon allocation still remains a very active area [24]. The static target-based weapon allocation problem has been quite well studied, especially within the field of operations research. Despite the extensive research, static weapon allocation is an example of a classical operations research problem that still remains unsolved [1], in the sense that effective methods for real-time allocation are lacking. Moreover, the asset-based version of the problem is much less studied than its target-based counterpart.

Much of the original work on weapon allocation focused on the allocation of missiles to defended assets, rather than the other way around. Hence, the problems were often modeled from an attacker’s side, instead of from the defending side. A brief summary and review of unclassified literature from the first years of research on the problem is given in [26]. Some years later, a monograph describing many of the developed mathematical models for weapon allocation problems was published in [6]. Unlike Matlin’s review, the monograph by Eckler and Burr takes on the weapon allocation problem from a defender’s view. The authors present a number of useful techniques for weapon allocation, such as relaxing the integer constraint and then make use of linear programming to solve the resulting continuous problem. This is a technique that is still in use (cf. [17]). It should be noted however, that fractional assignments of firing units to targets does not make sense, and rounding off the optimal solution to the relaxation of an nonlinear integer programming problem can yield solutions that are infeasible or far from the optimal solution to the original nonlinear problem [36]. Other kinds of tools such as the use of Lagrange multipliers and dynamic programming are also described in [6]. As the authors make clear, their focus is on analytical approaches, since it is argued that what they refer to as computer-oriented solutions give less insight into the weapon allocation problem than analytical approaches. A somewhat more recent survey of work within weapon allocation is presented in [3]. As in the earlier mentioned surveys, its focus is on analytical approaches to weapon allocation. However, it is mentioned that a shift towards various techniques such as implicit enumeration algorithms and nonlinear programming algorithms had been started at that time, since mathematical formulations of the weapon allocation problem are not generally amenable to solution in closed form [3, p. 66]. In later years, advanced computer-based techniques have been developed which are better suited for real-time weapon allocation [9]. In the following, we will focus on modern heuristic/approximate approaches, but will first present enumerative techniques, since such approaches can be very useful for special cases of static weapon allocation problems.

3.1. Exact Approaches

For small values of $|T|$ and $|W|$, the optimal solution to a static weapon allocation problem can easily be found by exhaustive search (also referred to as explicit enumeration), i.e., a brute-force enumeration where all feasible solutions are tested one after the other. However, as a static weapon allocation problem consists of $|T|^W$ feasible solutions, this is not a viable approach for air defense scenarios involving a large number of targets and firing units.

Exact polynomial time algorithms have been identified for the special case of the static target-based weapon allocation problem in which the kill probabilities of all firing units are assumed to be identical, i.e., $P_{ik} = P_i$. For this special case, the well known maximum marginal return (MMR) algorithm suggested in [5], and the local search algorithm suggested in [7] can be proven to be optimal. Some other special cases of the static target-based weapon allocation problem can be formulated as network flow optimization problems. If we assume the constraint that all firing units have kill probabilities $P_{ik} \in \{0, P_i\}$, i.e., that firing units either can or cannot reach a target, and in the former case, the kill probability only depend upon the target, the problem can be transformed into a minimum cost network flow problem with linear arc costs, for which several efficient algorithms exist [7]. A similar transformation can be done for the special case of the static target-based weapon allocation problem where we assume that $|T| \leq |W|$, and that at most one firing unit is to be allocated to each target. In this case, we can convert the problem into a so called transportation problem, for which efficient algorithms exist [7]. However, the general static target-based weapon allocation problem has been proved to be NP-complete [23], as have been discussed earlier. This also holds true for the asset-based version of the static weapon allocation problem, since this can be seen as a generalization of the static target-based version.

Another exact approach is to use branch-and-bound algorithms for finding the optimal solution. Branch-and-bound algorithms use tree representations of the solution space and are often able to prune away large subsets of feasible solutions through calculation of lower and upper bounds on different branches of the tree. In a recent article by [1], three branch-and-bound algorithms (using different lower-bound schemes) are investigated and are shown to give short computation times on average. The results are impressive, however, in theory the risk exists that the algorithm will require branching the full tree for some problem instances. This means that in worst-case, the performance of the branch-and-bound algorithm can be at least as bad as the performance of more naive exhaustive search algorithms. Although it in practice is unlikely that this worst-case scenario will appear, it is unfortunately not possible to in advance compute an upper bound on the computational time it will...
take to find the optimal solution to a problem instance when using a branch-and-bound algorithm. Hence, as can be seen in the results reported in [1], some problem instances of large size can be solved very quickly, while considerably smaller problem sizes can demand considerably more time for the optimal solution to be found. In other words, we have to rely on heuristic algorithms for large-scale problems when real-time guarantees are needed [1, 7].

3.2. Heuristic Approaches

A well-known heuristic approach for static target-based weapon allocation is the greedy maximum marginal return algorithm, originally suggested in [5]. A similar greedy algorithm is presented in [18]. Basically, the maximum marginal return algorithm works sequentially by greedily allocating firing units to the target maximizing the reduction of the expected value. It starts with allocating the first firing unit to the target for which the reduction in value is maximal, whereupon the value of the target is reduced to the new expected value. Once the first firing unit is allocated, the same procedure is repeated for the second firing unit, and so on, until all firing units have been allocated to targets. Pseudo code for the maximum marginal return algorithm is shown in Section 4.1. Obviously, the maximum marginal return algorithm is very simple and fast. This is a general advantage of greedy algorithms, but due to their greedy nature they are also very likely to end up with suboptimal solutions. Since the algorithm uses target values for choosing which target to be allocated next, it cannot be used as is for static asset-based weapon allocation. However, in [27] a number of greedy algorithms for asset-based weapon allocation are described. These algorithms basically work by approximating the asset-based problem with its target-based counterpart, by using the protection value of the defended asset to which the target is aimed for as the target value. When the problem has been approximated by a target-based problem, it is suggested that the maximum marginal return algorithm returns a solution that can be used as an approximative solution to the asset-based problem. Another suggested approach in [27] is to use the solution returned from the maximum marginal return algorithm and to apply local search on the solution so that the target allocated by one weapon can be swapped to the target allocated by another weapon, and vice versa.

Another kind of heuristic approach to a constrained version of the target-based weapon allocation problem has been suggested in [35], in which artificial neural networks are used. It is stated that solutions close to global optima are found by the algorithm, but results are only presented for a few small-scale problem instances, from which it in the authors’ view is not possible to generalize. It is in [9] also argued that artificial neural network algorithms for weapon allocation sometimes are unsteady and non-convergent, leading to that obtained solutions may be both suboptimal and infeasible.

As an alternative, the use of genetic algorithms seems to be popular. Such an algorithm for static target-based weapon allocation is described in [16], while a genetic algorithm combined with local search is presented in [22] and [21]. The quality of the solutions returned by the greedy maximum marginal return algorithm presented in [18] is in [16] compared to the solutions returned by genetic algorithms. However, the algorithms are only evaluated on target-based weapon allocation problems. The standard genetic algorithm is outperformed on large-scale problem sizes, but only one problem instance is tested for each problem size, so the possibility to generalize the results can be questioned. Even though, the results seem to indicate that greedy search works better than standard genetic algorithms on large target-based problem instances. It is in [16] suggested that genetic algorithms can be seeded with the solution returned from a greedy algorithm, which seems to be a suitable approach to improve the quality of genetic algorithms on large problem sizes. In [2], a genetic algorithm combined with local search is suggested for a dynamic version of the asset-based weapon allocation problem. It is shown that local search improves the results, but that the computational time needed is increased. The effects of real-time requirements on the algorithms are not tested.

The use of ant colony optimization for target-based weapon allocation is suggested in [19, 20]. Reported results in [19] and [20] indicate that ant colony optimization algorithms perform better than standard genetic algorithms on large-scale problems, and that the algorithms can be improved upon by using local search. However, the algorithms were allowed to run for two hours, so it is unclear how this generalizes to settings with real-time requirements.

A simulated annealing algorithm for static asset-based weapon allocation is presented in [4]. Basically, simulated annealing is based on an analogy of thermodynamics with the way metals cool and anneal, in which a liquid that is cooled slowly is likely to form a pure crystal corresponding to a state of minimum energy for the metal, while a quick cooling phase is likely to result in states of higher energy levels [32]. By controlling an artificial “temperature” when making the optimization (corresponding to the minimization of energy levels), it becomes possible to escape from local minima in the hunt for the optimal solution (the purest crystal in the thermodynamics analogy). However, no evaluation of the quality of the solutions obtained by the algorithm is presented in [4], so it is unknown how good their implemented algorithm performs. Another implementation of a simulated annealing algorithm provides solutions of lower quality than ant colony optimization and genetic algorithms in a static target-based weapon allocation experiment described in [19]. The algorithms were, as describe above, allowed to run for two hours,
The algorithms for static asset-based weapon allocation evaluated in this article share the same kind of representation, in which a solution is represented as a vector of length $|W|$. Each element $k$ in the vector points out the target $T_i$ to which the weapon is allocated. As an example of this, the vector $[2, 3, 2, 1]$ represents a solution in which $W_1$ and $W_3$ are allocated to $T_2$, $W_2$ is allocated to $T_3$, and $W_4$ is allocated to $T_1$.

4.1. A Maximum Marginal Return Algorithm for Static Weapon Allocation

A greedy algorithm for static target-based weapon allocation, known as the maximum marginal return (MMR) algorithm, was initially suggested in [5]. This algorithm (described with pseudo code in Algorithm 1) is very simple since it as already explained works greedily by assigning weapons sequentially to the target that maximizes the reduction of the expected target value. When the first weapon has been allocated to the target for which the reduction in value is maximal, the target value is reduced to the new expected value. After that, the same procedure is repeated for the second weapon, and so on, until all weapons have been allocated to targets, yielding a computational complexity of $O(|W| \times |T|)$.

**Algorithm 1 Maximum marginal return algorithm**

```
for all $k$ such that $1 \leq k \leq |W|$ do
    highestValue ← $-\infty$
    allocatedTarget ← 0
    for all $i$ such that $1 \leq i \leq |T|$ do
        value ← $V_i \times P_{ik}$
        if value $>$ highestValue then
            highestValue ← value
            allocatedTarget ← $i$
            assign $W_k$ to target $T_{allocatedTarget}$
            $V_{allocatedTarget} ← V_{allocatedTarget} − highestValue$
    return allocation
```

It is not obvious how to use the MMR algorithm for the static asset-based weapon allocation problem, since it in this version of the problem does not exist any target values. Instead, there are protection values associated with the defended assets, and lethality values associated with the targets. In [27], it is suggested that a defended asset’s weight (protection value) is equally distributed over the targets aimed for it, so that a target’s value is computed as $V_j = \omega_j / |G_j|$ (where $j$ is the index for the set $G_j$ of which target $T_j$ is a member), and that the asset-based problem is approximated with its target-based counterpart. Similar reasoning is presented in [7] where it is suggested that the value of a target is set to the expected destroyed protection value of the defended asset to which it is aimed, given that the target is not engaged and that all other targets aimed for the defended asset are destroyed.

We have here chosen to calculate the target value $V_j$ for a target $T_j$ as:

$$V_j = \omega_j \times \pi_j$$  \hspace{1cm} (6)
where \( j \) is the index of the defended asset to which target \( T_i \) is aimed. Hence, the target value has been calculated as the product of the lethality probability \( \pi_j \) of the target and the protection value \( \omega_j \) of the defended asset it is aimed at. In this way, we follow the approach suggested in [7] to use the protection value of the defended asset to impact on the target value, but we complement this with taking the lethality of the target into account, since this extra information otherwise is lost.

We have also included a variant of greedy search where we have taken the solution generated by the MMR algorithm and improved it with a simple local search (LS) that creates neighbor solutions by swapping two positions selected at random in the solution vector (this variant of the MMR algorithm, described with pseudo code in Algorithm 2, will in the following be referred to as MMR-LS). This algorithm is an implementation of the idea briefly discussed in [27].

**Algorithm 2** The MMR-LS algorithm

```
bestSolution ← MMR()
J_best ← CalculateFitness(bestSolution)
while termination criteria not met do
    neighborSolution ← neighbor(bestSolution)
    J_new ← CalculateFitness(neighborSolution)
    if J_new > J_best then
        bestSolution ← neighborSolution
        J_best ← J_new
return bestSolution
```

Clearly, the quality of the solutions generated by the MMR-LS algorithm will always be at least as good as the quality of the solutions returned by the MMR algorithm.

4.2. An Enhanced Maximum Marginal Return Algorithm for Static Weapon Allocation

What here will be referred to as the enhanced maximum marginal return algorithm (the authors’ terminology) is quite similar to the standard maximum marginal return algorithm. The difference is that in the enhanced maximum marginal return (EMMR) algorithm it is not predetermined which firing unit to allocate next. Instead, the choice of which firing unit to allocate next is based on which weapon-target pair that maximizes the marginal return. We have implemented this algorithm based on the description in [16], and the pseudo code for the algorithm is given in Algorithm 3. In the first iteration \( it = 1 \), \(|W| \times |T|\) combinations are tested. The weapon-target pair with highest marginal return is selected, so that the firing unit is selected to the target, and the target value of the corresponding target is updated accordingly. After this, \(|W| - 1\) firing units are unallocated. In next iteration, the remaining \((|W| - 1) \times |T|\) weapon-target pairs are tested, and so on, until there does not remain any unallocated firing units. Hence, the time complexity of EMMR becomes \(O(|W|^2 |T|)\).

**Algorithm 3** Enhanced maximum marginal return algorithm (adapted from [16])

```
for all it such that 1 ≤ it ≤ |W| do
    highestValue ← −∞
    allocatedTarget ← 0
    allocatedWeapon ← 0
for all k such that 1 ≤ k ≤ |W| do
    for all i such that 1 ≤ i ≤ |T| do
        value ← \( V_i \times P_{ik} \)
        if value > highestValue then
            highestValue ← value
            allocatedWeapon ← k
            allocatedTarget ← i
    assign \( W_{allocatedWeapon} \) to \( T_{allocatedTarget} \)
    \( V_{allocatedTarget} ← V_{allocatedTarget} - highestValue \)
return allocation
```

As the standard MMR algorithm, EMMR is relying on target values. Hence, we calculate target values according to (6), solve the approximated target-based problem using EMMR, and return the solution as the solution to the asset-based problem.

4.3. A Genetic Algorithm for Static Weapon Allocation

In [13], we presented a genetic algorithm (GA) designed for real-time allocation of defensive weapon resources to targets. The original version of the algorithm was intended for the static target-based problem, but we have now with some modifications adapted it to also suit the static asset-based formulation of the problem.

The algorithm is described in pseudo code in Algorithm 4. First, an initial population consisting of \( nrOfIndividuals \) is created, through generation of a vector of length \(|W|\). In this vector each element \( W_k \) is assigned a random integer value in the interval \( \{1, \ldots, |T|\} \). In each generation we evaluate all individuals in the population and determine their objective function values in accordance with (1). Hence, each individual is assigned a fitness value that is used in the following phases of selection and recombination. After the evaluation phase, deterministic tournament selection is used as selection mechanism to determine which individuals in population \( Pop \) that should be used as parents for \( Pop' \), i.e., we pick two individuals at random from \( Pop \) and select the one with best fitness value. When two parents have been selected from \( Pop \), we apply one-point crossover at a randomly selected position \( k \in \{1, \ldots, |W|\} \), generating two individuals that become members of \( Pop' \). This is repeated until there are \( nrOfIndividuals \) in \( Pop' \). Thereafter, we apply mutation on a randomly selected position \( k \in \{1, \ldots, |W|\} \) in the first individual of \( Pop' \), where the old value is changed into \( i \in \{1, \ldots, |T|\} \). Hence, there is a probability of \( 1/|T| \) that the individual is unaffected of the mutation. The mutation operator is repeated on all individuals in \( Pop' \) and the resulting individuals become members of the new population \( Pop \). This loop is repeated until the termination criterion is fulfilled (the upper limit on
the computational time bound is reached). At this point, the individual with the best fitness found during all generations is returned as the allocation recommended by the algorithm.

**Algorithm 4: Pseudo code for our genetic algorithm**

```
fitness_{best} ← −∞
Pop ← GenerateInitialPopulation()
while termination criteria not met do
    for l ← 1 to nrOfIndividuals do
        J_l ← CalculateFitness(Pop(l))
        if J_l > fitness_{best} then
            g ← Pop(l)
            fitness_{best} ← J_l
        end
        Pop_l ← Crossover(Pop)
        Pop ← Mutate(Pop_l)
    end
    return g
```

Furthermore, we have implemented a variant of the genetic algorithm that is seeded with well-performing individuals. Instead of creating all individuals in the initial population at random, \( \kappa \) individuals are created based on the solution returned by the EMMR algorithm (a random swap between the targets of two of the firing units is first made for each of the seeded individuals in order to create some diversity among them). The remaining individuals are created randomly just as before. This seeded version of the genetic algorithm will in the following be referred to as GA-S.

### 4.4. A Particle Swarm Optimization Algorithm for Static Weapon Allocation

In [14], we developed a particle swarm optimization (PSO) algorithm for the static target-based weapon allocation problem. We have modified this algorithm to also suit the static asset-based weapon allocation problem.

A particle swarm consists of \( nrOfParticles \) particles, in which each particle is associated with a position \( \vec{x}_l \), a velocity \( \vec{v}_l \), and a memory \( \vec{b}_l \) storing the particle’s personal best position. Moreover, we also store the swarm’s global best position in a vector \( \vec{g} \). Each particle corresponds to a solution, given by the particle’s position.

**Algorithm 5: Pseudo code for our particle swarm optimization algorithm**

```
Initialization()
while termination criteria not met do
    for l ← 1 to nrOfParticles do
        J_l ← CalculateFitness(\vec{x}_l)
        if J_l = CalculateFitness(\vec{g}) then
            p_l ← Reinitialize()
        else
            if J_l > CalculateFitness(\vec{b}_l) then
                \vec{b}_l ← \vec{x}_l
            end
            if J_l > CalculateFitness(\vec{g}) then
                \vec{g} ← \vec{x}_l
            end
        end
    end
    for l ← 1 to nrOfParticles do
        \vec{v}_l ← UpdateVelocity(\vec{p}_l)
        \vec{x}_l ← UpdatePosition(\vec{p}_l)
    end
    return \vec{g}
```

The algorithm is described in pseudo code in Algorithm 5. In an initialization phase, each particle is assigned an initial position \( \vec{x}_l^0 \) (where the elements in the initial position vectors are integers randomly distributed between 1 and \( |T| \)), and an initial velocity \( \vec{v}_l^0 \) (a vector of real numbers randomly distributed from the uniform distribution \( U[-0.5|T|,0.5|T|] \)). A fitness value is calculated for each particle, given by the objective function value \( J \) (see (1)) that is obtained for the solution corresponding to the particle’s position. The new fitness is compared to the personal best and the global best to see whether these should be updated accordingly. After this, the velocity and position is updated for each particle, according to (7) and (8).

\[
\vec{v}_l^{t+1} = \omega \vec{v}_l^t + c_1 \vec{r}_1^t \cdot (\vec{b}_l^t - \vec{x}_l^t) + c_2 \vec{r}_2^t \cdot (\vec{g}^t - \vec{x}_l^t) \quad (7)
\]

\[
\vec{x}_l^{t+1} = \vec{x}_l^t + \vec{v}_l^{t+1}. \quad (8)
\]

In (7), \( \omega \) is a parameter referred to as inertia or momentum weight, specifying the importance of the previous velocity vector, while \( c_1 \) and \( c_2 \) are positive constants specifying how much a particle should be affected by the personal best and global best positions (referred to as the cognitive and social components, respectively). \( \vec{r}_1^t \) and \( \vec{r}_2^t \) are vectors with random numbers drawn uniformly from the interval \([0, 1]\). Moreover, the \( \cdot \)-operator denotes the Hadamard product, i.e., element-by-element multiplication of the vectors. In order to avoid that particles gain too much momentum, a \( V_{max} \) parameter that constrains the velocities to stay in the interval \([-V_{max}, V_{max}] \) has been introduced.

After the position update specified in (8), we round off the particles’ positions to their closest integer counterpart. In next iteration we calculate the particles’ new fitness values, whereupon the velocities and positions are updated, and so on. This is repeated until a termination criterion is met, i.e., that no more time remains. When this happens, the best solution obtained so far is returned as output from the algorithm.

A problem that must be handled is particles moving outside the bounds of the search space. When this happens, we reinitialize the position and velocity values of the coordinate for which the problem occurred. Moreover, in order to avoid premature convergence to local optima (stagnation), we reinitialize the velocity vector for particles rediscovering the current best solution. For a more thorough explanation of the problem of stagnation in particle swarm optimization, see [34].

In addition to the described particle swarm optimization algorithm, we have also included a variant in which we seed the starting position for \( \kappa \) particles in the initial population in the same way as with the GA-S algorithm (while their initial velocities are randomized in the same manner as for the remaining particles). This seeded particle swarm optimization algorithm will in the following be referred to as PSO-S. To the best of our knowledge, the use of seeded particles is novel for the weapon allocation problem.
5. EXPERIMENTS

In the experiments reported here, we have used the open source testbed SWARD\(^\dagger\) (System for Weapon Allocation Research and Development) which we have developed in order to allow for systematic comparison of various weapon allocation algorithms [11, 15]. The testbed is implemented in Java, and we have been running the experiments on a computer with a 2.67 GHz Intel Core i7 CPU and 8 GB RAM. By using SWARD, we make sure that the experiments presented here are easily reproducible, so that researchers can test other algorithms on the same problem instances.

In order to recreate the problem instances used in the experiment presented in Section 5.1.1, the following settings should be used in SWARD:

\[
\begin{align*}
T_{\text{start}} & = 5, \quad W_{\text{start}} = 5, \\
T_{\text{end}} & = 9, \quad W_{\text{end}} = 9, \\
T_{\text{step}} & = 1, \quad W_{\text{step}} = 1, \\
\text{iterations} & = 10, \quad DA\text{s} = 5, \\
\text{seed} & = 0, \quad \text{timeLimit} = 1000 \text{ ms}.
\end{align*}
\]

Similarly, for recreating the problem instances used in the experiment presented in Section 5.1.2, the following settings should be used:

\[
\begin{align*}
T_{\text{start}} & = 10, \quad W_{\text{start}} = 10, \\
T_{\text{end}} & = 30, \quad W_{\text{end}} = 30, \\
T_{\text{step}} & = 10, \quad W_{\text{step}} = 10, \\
\text{iterations} & = 100, \quad DA\text{s} = 5, \\
\text{seed} & = 0, \quad \text{timeLimit} = 1000 \text{ ms}.
\end{align*}
\]

The time limits make sure that no algorithms are allowed to run for more than a total time of one second (including seeding). For the genetic algorithms we have used the parameter setting: \(nrOf\text{Ind}\text{ividuals} = \text{max}(|T|, |W|)\). Additionally, we have for the seeded version used \(\kappa = 0.5 \times nrOf\text{Ind}\text{ividuals}\). For the particle swarm optimization algorithm we have used \(nrOf\text{Particles} = 50, \quad c_1 = 2.0, \quad c_2 = 2.0, \quad \omega = 0.8, \quad \text{and} \quad V_{\text{max}} = 0.5 \times |T|\). The same settings have been used for the seeded version, with the additional parameter setting \(\kappa = 25\).

5.1. Heuristic Algorithm Performance

For scenarios that demand solving the static asset-based weapon allocation problem faster than is possible with optimal algorithms, we have to rely on heuristic algorithms. In Section 5.1.1, we present experimental results obtained with the suggested heuristic algorithms on small-scale problem instances, while we in Section 5.1.2 present results on large-scale problem instances.

5.1.1. A comparison against the optimal solution for small-scale problems

We have in order to investigate the quality of the solutions generated by the suggested algorithms compared their obtained objective function values to the optimal objective function values obtained by exhaustive search for relatively small-scale scenarios between \(|T| = 5, |W| = 5\) and \(|T| = 9, |W| = 9\).

The average percentage deviation from the optimal solution is a common metric to use for evaluating heuristic algorithms on small-scale optimization problems where the optimal solution can be calculated, and therefore it also has been used here. The percentage deviation \(\Delta_{\text{alg}}\) for a specific algorithm on a specific problem instance has been calculated as:

\[
\Delta_{\text{alg}} = \frac{|J_{\text{alg}} - J_{\text{opt}}|}{J_{\text{opt}}} \times 100
\]

(9)

where \(J_{\text{alg}}\) is the objective function value for the tested algorithm and \(J_{\text{opt}}\) is the optimal objective function value. In the tables, we use **bold** to show which obtained objective function value that is the best for each tested problem size.

Looking at Table II, the algorithms’ percentage deviations from the optimal solution show that most of the algorithms are able to find optimal or very near-optimal solutions for the smallest tested problem sizes. The MMR algorithm is by far the worst of the algorithms on the tested small-scale scenarios, but when allowed to improve its initial solution by local search (i.e., the MMR-LS algorithm), the quality is improved. The EMMR algorithm produces solutions that are better than both the MMR and MMR-LS algorithms. However, as can be seen, all these greedy heuristics are outperformed by the nature-inspired metaheuristics. Of the nature-inspired metaheuristics, the PSO algorithm performs somewhat better than the others. In fact, it produces optimal solutions to all problem instances of size \(|T| = 5, |W| = 5\)–\(|T| = 8, |W| = 8\), and for \(|T| = 9, |W| = 9\) it is in one second able to generate almost optimal solutions to problems consisting of \(9^6 = 387,420,489\) feasible solutions.

It should be noted that the results obtained on small-scale problems do not necessarily extends to large-scale problems. For small instances of any combinatorial problem, it is likely that algorithms such as PSO algorithms and GAs are able to search a large fraction of the solution space in a short period of time, making it more probable to find a high quality solution, while

---

\(\dagger\)The open source testbed SWARD can be downloaded from http://sourceforge.net/projects/sward/.
over 100 problem instances) in Tables III—V. We also plotted the objective function values obtained (averaged from the optimal solution, we have here simply allowed to run for one second on each problem instance. The algorithms have also in this experiment been allowed to run for one second on each problem instance. The optimal solutions are hard to obtain for large-scale problem instances, so instead of calculating the deviation from the optimal solution, we have here simply plotted the objective function values obtained (averaged over 100 problem instances) in Tables III–V. We also show the associated standard deviations within parentheses. As before, bold is used to indicate the best obtained objective function value on each problem size.

5.1.2. A comparison between algorithms on larger-scale problems

In a second experiment with the heuristic algorithms, we have tested them on larger-scale problems ranging in between ([|T| = 10, |W| = 10] and ([|T| = 30, |W| = 30]). The algorithms have also in this experiment been allowed to run for one second on each problem instance. The optimal solutions are hard to obtain for large-scale problem instances, so instead of calculating the deviation from the optimal solution, we have here simply plotted the objective function values obtained (averaged over 100 problem instances) in Tables III–V. We also show the associated standard deviations within parentheses. As before, bold is used to indicate the best obtained objective function value on each problem size.

A note to make is that the standard deviations shown in many cases are larger than the differences in mean values among the algorithms. However, this should not be interpreted as that there are no significant differences among the algorithms. Rather, the largest part of these standard deviations are due to the differences between various problem instances. In some problem instances the optimal objective function values are lower, while they in others are higher (as a natural result of the random fashion in which the problem instances are generated). As a consequence of this, also optimal algorithms would obtain large standard deviations.

When analyzing the obtained results, it can be seen that the use of local search significantly improves the quality of the solutions found using MMR also on large problem sizes. A comparison of the solutions generated by MMR-LS with the ones returned by the EMMR algorithm shows that the performance of these are approximately equally good (although EMMR is significantly faster than MMR-LS). This indicates that it in the future may be worth studying if it would be beneficial to apply simple local search also to EMMR.

It can be seen that the seeded particle swarm optimization algorithm (i.e., PSO-S) is performing best relative to the other algorithms on all tested problem sizes except the largest, on which the seeded genetic algorithm (GA-S) performs slightly better. We have in earlier work [14] shown that PSO runs into some trouble when applied to large target-based problem instances under tight real-time constraints, and this trend can be seen also for the large asset-based problem instances tested here. However, when combined with the seeding mechanism, particle swarm optimization seems to work very well. It can be seen that the obtained objective function values for the greedy algorithms MMR-LS and EMMR are reasonably close to the best algorithms’ objective function values for many of the tested problem sizes, while they for problem instances where |T| > |W| are much worse. These results are in line with the analytical arguments in [7], predicting that it will work well to approximate the static asset-based weapon allocation problem with its target-based counterpart on problem instances involving a strong defense (a large number of firing units compared to the number of targets), while the approximation will work bad in cases of a weak defense (i.e., problem instances where there are more targets than firing units). Although the differences between e.g., EMMR and PSO-S or GA-S and PSO-S are not very large for problem instances involving a strong defense, the differences should not be ignored, since such small but significant differences can have severe

<table>
<thead>
<tr>
<th>Table III</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Average Objective Function Value for</td>
</tr>
<tr>
<td>Averaged Over 100 Static Asset-Based Problem Instances</td>
</tr>
<tr>
<td>(higher objective function values are better)</td>
</tr>
<tr>
<td>10 × 10</td>
</tr>
<tr>
<td>GA</td>
</tr>
<tr>
<td>GA-S</td>
</tr>
<tr>
<td>PSO</td>
</tr>
<tr>
<td>PSO-S</td>
</tr>
<tr>
<td>MMR</td>
</tr>
<tr>
<td>EMMR</td>
</tr>
<tr>
<td>MMR-LS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Average Objective Function Value for</td>
</tr>
<tr>
<td>Averaged Over 100 Static Asset-Based Problem Instances</td>
</tr>
<tr>
<td>(higher objective function values are better)</td>
</tr>
<tr>
<td>20 × 10</td>
</tr>
<tr>
<td>GA</td>
</tr>
<tr>
<td>GA-S</td>
</tr>
<tr>
<td>PSO</td>
</tr>
<tr>
<td>PSO-S</td>
</tr>
<tr>
<td>MMR</td>
</tr>
<tr>
<td>EMMR</td>
</tr>
<tr>
<td>MMR-LS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table V</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Average Objective Function Value for</td>
</tr>
<tr>
<td>Averaged Over 100 Static Asset-Based Problem Instances</td>
</tr>
<tr>
<td>(higher objective function values are better)</td>
</tr>
<tr>
<td>30 × 10</td>
</tr>
<tr>
<td>GA</td>
</tr>
<tr>
<td>GA-S</td>
</tr>
<tr>
<td>PSO</td>
</tr>
<tr>
<td>PSO-S</td>
</tr>
<tr>
<td>MMR</td>
</tr>
<tr>
<td>EMMR</td>
</tr>
<tr>
<td>MMR-LS</td>
</tr>
</tbody>
</table>
impact on the end result if such algorithms are applied in a real-world C-RAM system.

6. CONCLUSIONS AND FUTURE WORK

We have in this article presented the static asset-based weapon allocation problem, which is an optimization problem that needs to be solved in a short amount of time in air defense situations involving RAM threats such as rockets and mortars. We have also presented the static target-based weapon allocation problem, but the focus has been on the asset-based case. We have implemented two versions of a genetic algorithm, two versions of a particle swarm optimization algorithm, and various versions of the greedy maximum marginal return algorithm. Such algorithms have earlier been used for the static target-based weapon allocation problem, but as far as we know, it is previously unknown how they perform on the asset-based version of the problem. Our experiments have shown that optimal or very near-optimal solutions are obtained in real-time by the genetic algorithms and the particle swarm optimization algorithms on small-scale problems. The standard maximum marginal return algorithm yields worse solutions, but these can easily be improved upon by local search, or by using an enhanced version of the algorithm. However, the quality does not become as good as that of the genetic algorithms or the particle swarm optimization algorithms.

For larger problem instances the optimal solutions are not known, and can therefore not be used for comparison. Instead, the objective function values produced by the algorithms have been compared to each other. It has been shown that the greedy algorithms create solutions of good quality (compared to the other algorithms) for scenarios with a strong defense, but that they perform bad on scenarios involving a weak defense, i.e., where there is a larger number of targets than there are firing units.

The algorithm that has been performing the best on large-scale problem instances is a novel improvement on the particle swarm optimization algorithm where the initial population is seeded with individuals based on small variations of the solution returned by the enhanced maximum marginal return algorithm. For the problem instances where there is a strong defense, the algorithm is not able to improve very much on the solution returned by the enhanced maximum marginal return algorithm, but for the problem instances involving a weak defense, the difference is dramatic. For problems of quite small scale, the difference in solution quality is very small between the particle swarm optimization algorithm and its seeded version. However, as the problem size increases, the difference in solution quality becomes very evident.

6.1. Future Work

The obtained results can be used as benchmarks for other heuristic algorithms. Hence, it is our hope that the used data sets (problem instances) will be used by other researchers as well, so that a better understanding of which algorithms that work well for static asset-based weapon allocation is obtained. Moreover, in the current research on static asset-based weapon allocation, it is assumed that kill probabilities, lethality probabilities, and target aims are known with certainty. Obviously, these estimates will in real-world systems be associated with uncertainty, and it would therefore be interesting and useful to know how sensitive the solutions produced by the algorithms are to such uncertainties.

In the experiments presented in this article, we have been generating problem instances in which there are no dependences among the values of the parameters. As an example, there is no correlation between any of the kill probabilities involving a specific target (or rather, there might be such correlation, but if so, this is by pure chance). This is consistent with how weapon allocation algorithms have been evaluated earlier in reported literature, but it can be discussed whether this lack of structure really would be seen in estimated kill probabilities from real-world air defense scenarios. Thinking of such a scenario, two targets, \( T_1 \) and \( T_2 \), of the same type, approaching a firing unit \( W_1 \) from the same direction and on the same altitude, would most likely result in kill probabilities \( P_{11} \) and \( P_{12} \) being quite similar. Likewise, two firing units \( W_2 \) and \( W_3 \) would obtain kill probabilities of approximately same magnitude, given that the firing units were positioned close together and being of the same type. Hence, the random fashion in which problem instances have been generated here (and in previous reported experiments with weapon allocation algorithms) may not necessarily create the same kinds of search spaces that would be experienced in real-world air defense situations. An idea that could be of interest for the future is therefore to create problem instances with an inbound structure that better reflect reality.

ACKNOWLEDGMENT

We would like to express gratitude to the reviewers for their constructive comments and suggestions that have helped to improve the article. This work was supported by the Information Fusion Research Program (University of Skövde, Sweden) in partnership with Saab AB and the Swedish Knowledge Foundation under grant 2003/0104.

REFERENCES

[1] R. Ahuja, A. Kumar, K. Jha, and J. Orlin

Exact and heuristic methods for the weapon target assignment problem. 


Evolutionary decision-makings for the dynamic weapon-target assignment problem. 


[10] K. C. Jha


[12] F. Johansson and G. Falkman


[16] B. A. Julstrom

[17] O. Karasakal

[18] S. E. Kolitz

[19] Z-J. Lee and C-Y. Lee


[21] Z. J. Lee and W. L. Lee


[23] S. P. Lloyd and H. S. Witsenhausen


[26] S. Matlin

[27] W. A. Metler and F. L. Preston

[28] R. A. Murphey
Target-based weapon target assignment problems. In P. M. Pardalos and L. S. Pitsoulis (Eds.), Nonlinear assignment problems: algorithms and applications, 2000, 39–53.

[29] S. Paradis, A. R. Benaskeur, M. Oxenham, and P. Cutler


[32] B. Suman and P. Kumar
Fredrik Johansson obtained his M.Sc. in computer science from University of Skövde, Sweden, in 2005 and his Ph.D. in computer science from Örebro University, Sweden, in 2010.

During his Ph.D. studies he was a member of the Skövde Artificial Intelligence Lab (SAIL) and the Information Fusion Research Program at the Informatics Research Centre (IRC) in Skövde. Currently, he works as a scientist at the Swedish Defence Research Agency (FOI) in Kista.

His research interests are applied artificial intelligence and high-level information fusion, and the application of probabilistic techniques such as Bayesian networks for decision support. He is also interested in the use of techniques such as social network analysis, natural language processing, and web harvesting for supporting the work of intelligence analysts.

Göran Falkman obtained his Ph.D. in computing science from Chalmers University of Technology, Sweden, in 2003.

He holds a position as an Associate Professor of Computer Science, with a specialty in Interactive Knowledge Systems, at University of Skövde, Sweden, where he works as a researcher and senior lecturer within the Skövde Artificial Intelligence Lab (SAIL) at the Informatics Research Centre (IRC). He has been a project leader for three applied research projects within the area of information fusion, focusing on algorithms for threat evaluation and weapon allocation, visual analytics and maritime domain awareness, and anomaly detection for surveillance applications, respectively. He has also been the leader for the Situation Awareness scenario within the Infusion research program at University of Skövde. Currently, he is one of the principal investigators of the Uncertainty Management in High-Level Information Fusion (UMIF) research project. Since 2009, he is an elected member of the Executive Board of the Swedish Artificial Intelligence Society (SAIS).

The research interests lie in the intersection of applied artificial intelligence, knowledge systems, interaction design, and information fusion. This includes work on the design, implementation and use of formal knowledge representation and knowledge-based systems (especially, case-based reasoning, ontology engineering, and the Semantic Web), as well as the use of interactive visualization for supporting knowledge-based reasoning processes (especially, situation analysis and decision-making).