A question of aggregation

A comparative study of objective function creation methods applied to missile defense for helicopters

Master's thesis in Applied Physics

HANNAH DJURFELDT

Department of Applied Mechanics
Vehicle engineering and autonomous systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014

Master's thesis

MASTER'S THESIS IN APPLIED PHYSICS

A question of aggregation

A comparative study of objective function creation methods applied to missile defense for helicopters

HANNAH DJURFELDT

Department of Applied Mechanics
Vehicle engineering and autonomous systems
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2014

A question of aggregation
A comparative study of objective function creation methods applied to missile defense for helicopters
HANNAH DJURFELDT

Master's thesis
ISSN 1652-8557
Department of Applied Mechanics
Vehicle engineering and autonomous systems
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 1000

Cover:
A helicopter ejecting flares

Chalmers Reproservice
Gothenburg, Sweden 2014

A question of aggregation
A comparative study of objective function creation methods applied to missile defense for helicopters
Master's thesis in Applied Physics
HANNAH DJURFELDT
Department of Applied Mechanics
Vehicle engineering and autonomous systems
Chalmers University of Technology

Abstract

This report describes an investigation done about method one can use when creating a target function for optimization and the effects of those choices. Three simple and four advanced target function creation methods are described in detail and the results from optimizing with these target functions are analyzed. Lastly suggestions into possible continuations of this investigation is made.

Keywords: Multi-attribute desision making, Analytic hierarchy process, Evolutionary optimization, Target function creation, Missile protection

Preface

This is a report of a master thesis project carried out at the Swedish National Defence Research Agency and Chalmers University of Technology regarding the automatization of the optimization of missile defence for helicopters.

Acknowledgements

Many thanks to the wonderful people at the electronic warfare evaluation unit at FOI, especially my two mentors Peter Klum and Annelie Tonnvik. Thanks also to my mentor and examiner at Chalmers, Mattias Wahde, for waking my interest in evolutionary algorithms and helping me with this report.

Contents

1 Introduction

This report describes a project carried out at **Chalmers University of Technology** and the **Swedish Defence Research Agency, (FOI),** about target function creation methods for flare sequence optimization for helicopters under fire. The project was done as a master thesis during the summer and fall of 2013.

1.1 Background

When flying over hostile territory helicopters can be targeted by heat seeking missiles. In an attempt to mislead the software in the missile flares are ejected. The goal is to trick the missile into targeting the flare instead of the helicopter.

The success of the missile to keep, and retain, its target is affected by the flare sequences. Using different types of flares, different ejection ports and delays creates a complicated parameter space, and a perhaps more complicated output space. Development of new flare sequences demand much simulation and testing over long periods of time. The sequences are analyzed manually and are dependent on subjective judgements [12]. In an attempt to streamline this process the department of electronic warfare evaluation at FOI in Linköping has commissioned a master thesis on the effects of the method used to create the fitness function used when optimizing using an evolutionary algorithm to optimize the flare sequence.

1.2 Aim

The goal of this report is to describe an investigation done into the effects of the selection of method used to create the fitness function on the performance of the process of optimizing flare sequences for helicopters under fire by heat-seeking missiles. The time it takes to find the optimum, as well as the nature of the solutions determined to be optimal, is compared for seven different methods of aggregating 14 different simulation measurements into one fitness function. The solutions found are also compared with the set of possibly optimal solutions.

1.3 Method

To achieve the goal of this report the different parts of the investigation are described in detail. To help the reader understand the process a theoretical study is also presented.

A flare ejection has three parameters. The type of flare, the flare dispenser, and the delay in firing. An extra parameter is added to signal if the flare should be ejected or skipped. A flare sequence is a sequence of such ejections. The flare sequence is tested for missiles shot from 20 different positions evenly distributed in the sector between 120 and 150 degrees, counted from the nose of the helicopter, between one and five kilometers from the position of the helicopter. This sector is chosen due to the tail of the helicopter obstructing the view and making is more difficult to distract the missile. The simulations are carried out using two secret models developed at FOI [12].

The initial question is to decide which measurements to use as indicators of the real goal of optimizing the survival probability of the helicopter. These measurements are normalized and the normalized measurements are weighed together in different target functions created with different methods. Each target function construction method is evaluated by running an evolutionary optimization algorithm several times, noting the chosen optimum, the chance of finding that optimum and the time required to reach convergence.

The effects of the choice of target function are then investigated. Do the different methods choose the same optimum? How does the average convergence time vary? In addition the number of solutions that could be optimal is studied in reference to the different choices of target function.

2 Theory

This focus of part of the report can be summarized by a quote by W. Edwards and F. H. Barron [7]:

> That is, any outcome of a decision is most naturally described by a vector of numbers that relate to value. The task facing the analyst who wishes to use those numbers to guide decisions is to aggregate that vector into a scalar that the decision maker wishes to maximize - a single number measured at least on an interval scale.

If the situation addressed by this report is taken as an example, there is a model that gives out simulation data for different flare sequences. The objective is to aggregate these numbers to a one dimensional value function that can be used by an optimization algorithm.

Why such an aggregation is possible to accomplish is initially shown in section 2.1. Different methods of accomplishing the aggregation are then explained in sections 2.2 through 2.4. There are an indeterminate number of methods one can use in pursuit of this goal but only a small selection of common methods are presented here. To conclude the optimization method chosen is described and the reasons for which it was chosen are explained in section 2.5.

2.1 Multi-attribute preference theory

In order for there to be anything to optimize, one set of outcomes $X = (x_1, x_2, x_3, ...)$ must be preferred to another set of outcomes $Y = (y_1, y_2, y_3, ...)$. This existence of preference is the base of making improvements of the outcome, a necessary task in any method of optimization.

There exists a comprehensive theory to support the process of decision making, and as a result a notational standard. If the outcome X is preferred to the outcome Y this can be written as $X \succ Y$. The outcomes could also be equally preferred, perhaps more commonly described as the decision maker can be indifferent between the alternatives, written $X \sim Y$. Naturally there is also the case of X being equally or more preferred, $X \succeq Y$ seen as *either* $X \succ Y$ *or* $X \sim Y$ [6].

Assuming that the preferences are non-contradictory, e.g. not $X \succ Y$, $Y \succ Z$ and $Z \succ X$ simultaneously, one can place the possible outcomes in a strict order without any contradictions. From this it is natural to assign a value function $v(X)$ so that $v(X) \geq v(Y)$ *if and only if* $X \succ Y$. The problem of assigning such a function is non-trivial, especially when it is desired for the difference in value of the value function to reflect the strength of preference difference between the alternatives [6].

When considering value functions over several attributes one more assumption must be made, that of preference independence. A formal explanation is seen in equation (2.1.1). In short, the preference relation between two outcomes that differ only in one attribute depend only on the preference relation in that attribute and not on the other outcomes that remain unchanged [6].

$$(x_1, x_2, a, x_4, ...) \succ (x_1, x_2, b, x_4, ...) \Leftrightarrow a \succ b \quad \forall x_1, x_2, x_4, ... \tag{2.1.1}$$

The value functions are unique up to a linear transformation, therefore they are often normalized so that zero is the worst possible outcome and one the best possible outcome, but any scale would work equally well. The most common approach in choosing a value function is the additive form seen in equation (2.1.2) where the attributes are scaled to appropriately to what they represent and the relative importance of the attributes is controlled by use of weights [6].

Another type of utility function is the multiplicative utility function seen in equation (2.1.3). If the scaling constant λ is determined to be zero the equation reduces to the additive case [6].

$$v(X) = \sum_i \omega_i n_i(x_i) \tag{2.1.2}$$

$$1 + \lambda v(x) = \prod_i 1 + \lambda\lambda_i n_i(x_i) \tag{2.1.3}$$

There are several ways of calculating the weights ω_i and λ_i. Most often an additive utility function is used and one of the most common is called the **analytic hierarchy process**, henceforth abbreviated to AHP [6].

2.2 The analytic hierarchy process

One of the leading decision making tools is the analytic hierarchy method, (AHP), invented by Saaty in the late 1960s [8]. The rationale for introducing the AHP was to overcome the lack of a practical systematic

Table 2.2.1: Verbal statements about preference translated into numerical values [17].

Verbal statement	Corresponding number
Equally important	1
Marginally more important	2
Slightly more important	3
Moderately more important	4
Strongly more important	5
Quite strongly more important	6
Very strongly more important	7
Greatly more important	8
Extremely more important	9

approach for decision making and priority planning. Many of the worlds leading companies and institutions use AHP decision models, including CIA, NASA, IBM and the American department of defense. Examples of applications could be about which military bases to close, whether to introduce fish into the local ecology, environmental impact evaluations and portfolio management [8].

2.2.1 Principles and axioms of AHP

There are three basic principles of the AHP. The first is the **principle of decomposition**, allowing a complex problem to be subdivided into a hierarchy of clusters and subclusters of criteria. This is a very effective and intuitive way of dealing with complexity, almost universal in the structure of any large organism or organization [8].

The second principle is the **principle of comparative judgements**, enabling a decision maker to compare all the criteria in a subcluster pairwise as to their importance. These pairwise judgements are made to calculate the weights of the local criteria and need only be done in the local cluster, reducing the number of such comparisons to be made as well as simplifying for the person making the pairwise judgements. The weights constructed thus are referred to as local weights [8].

The final principle is the **principle of synthesis**, which allows one to multiply the local weights with the weight assigned to the cluster as a whole to get the global weights for each criterion. This allows one to evaluate each alternative to each criterion on the lowest order and still get the global fitness of the criteria without advanced mathematics [8].

The method is based on three simple axioms as well as the principles. First is the **reciprocal axiom** stating that if alternative A is x times better than B, eg. three times as long or heavy, than alternative B is $1/x$ times better than A. This makes very intuitive sense when talking about physical measurements but is extended to include subjective measurements such as beauty or durability [8].

The second axiom is the **homogeneity axiom** stating that the importance of the criteria being compared should be in the same order of magnitude. If this is not immediately fulfilled adjustments need to be made. When the difference is too great the decision maker generally makes judgements with less accuracy and greater inconsistencies [8].

The last axiom is the **synthesis axiom**, necessary for the hierarchy principle to apply. It states that that the priorities on one level do not depend on the levels below. An example of a violation could be that the importance of one criterion is lessened if all alternatives reach an above satisfactory level. If this axiom is violated an extension of AHP called the **analytical network process**, ANP, should be tried instead [8].

2.2.2 How to implement

Initially the different criteria are placed in a hierarchy, each criterion belonging to the subject defined by the category above it. On each level the decision maker compares all the criteria pair-wise, determining which one is more important and how much more important. This strength of preference can be given verbally, graphically of directly. The verbal assessment are translated into an integer from one to nine according to table 2.2.1 with intermediate values given to intermediate preferences if wanted [17].

These preference measurements are entered into a reciprocal matrix. Each diagonal element is equal to one, since each criterion is equally important as itself. If we call the matrix A and assume that criteria i is more important than criteria j then element A_{ij} becomes the integer from table 2.2.1 and element A_{ji} becomes the reciprocal value, or $\frac{1}{A_{ij}}$ [19].

The principal eigenvector of this matrix is then calculated and used as the local weights after being normalized so that the sum of the elements is equal to one. The process is repeated in each step of the hierarchy. The local weights are then multiplied by the local weights of their category in the level above, and so on. When there are no more levels left the lowest level criteria have gotten their global weights. All global weights should add up to one [8].

In each criteria the different alternative solutions are compared and given a score the same way the weights of the criteria are calculated. These criteria specific performance scores are then multiplied by the respective weights and added, forming the total score of the solution alternative. It should be noted that the total score of all alternatives will add up to one when the process is correctly done [8].

2.2.3 Modifications

The areas where AHP is primarily used is in fields where the alternatives are known beforehand and alternatives are neither added nor removed. The different alternatives are compared to access their qualities and the available points shared between them. This is not wanted when applying an optimization algorithm but very useful when considering unmeasurable qualities [8].

The methodology therefore needs to be adapted to suite an open set of alternatives. The simplest way of doing this is to compare the alternatives with some sort of predefined standard instead of each other. While not strictly needed for using an evolutionary optimization algorithm it facilitates to not have to make manual judgements in each iteration and tracking the progress becomes easier if the same solution gets the same score in each iteration.

A simple way of doing this is to give each solution a score in each criteria based on some sort of normalization function where zero is the worst possible and one the best possible. This makes the maximum score of an alternative equal to one, although it might not be possible to reach in reality.

2.2.4 Criticism

If one were to add or subtract an alternative when using an otherwise closed set of alternatives a phenomenon called **rank reversal** can occur. Simply put it is when two, or more, alternatives switch internal order when an alternative that is neither the best nor worst performing is added. This is the focal point of the criticism towards AHP compared with the **multi-attribute utility theory** method of assigning weights. Whilst catching more of the complexity of the decision and being fundamentally mathematically motivated the method is more strict, requires larger consistency and is much more complicated to complete. This makes the multi-attribute utility theory method rarely used in practice by decision makers [8].

If one assumes an open set of alternatives, and modifies the methodology as suggested in section 2.2.3, this phenomenon does not occur. Since this is the case in the application proposed in this report the phenomenon is not further discussed [8].

2.3 Direct point allocation

In direct point allocation the weights are assigned directly by the decision maker. An arbitrary number of points are assigned directly to the objectives or a predetermined number of points is assigned to one criterion and the others weighted freely from there. These values are normalized and become the weights. The method is simple and straightforward, an advantage in implementation, but lacks both the theoretical motivation and the abilities of more complex methods such as a hierarchical structure [17].

2.4 SMART

SMART stands for **simple multi-attribute rating technique** and was presented by Edwards in 1977 [7]. The motivation being that the indifference judgements between pairs of hypothetical options required by traditional multi-attribute decision theory seemed difficult and unstable. A simpler method with more direct assessments of the desired qualities was thought to be easier to use and less likely to result in logical errors [7].

The process starts by ranking all the criteria in order of importance. Then the decision maker is asked to make ratio estimates of the relative importance of each criterion compared to the one ranked lowest in importance. These ratios are then normalized and used as weights [17].

It should be noticed that Edwards himself, the inventor of SMART, recommends against using it. Instead it should be replaced by the SMARTS or SMARTER, methods that have evolved from SMART. The main objection is that the degree of importance of an attribute should depend on the spread of that attribute, something that is ignored in SMART. An example could be the situation of buying a new car. Price is generally a very important factor in this decision but if the price all considered vehicles were situated within a $1-2\%$ of the mean price, is that difference still important [7]?

2.4.1 SMARTS

SMARTS stands for **SMART using Swing**, the main difference from SMART being the method of calculating of the weights. The word swing referring to the operation of changing the score of an option in some dimension from one value to another [7].

There are two ways of calculating swing weights, magnitude estimates and indifference judgements. The question asked becomes: If a swing from the worst possible value to the highest possible value for the most important criterion is worth 100 points, what is the weight of the same swing of the performance in the second-most important attribute? This question is then repeated for all attributes, either in comparison with the top-ranked attribute or with the attribute closest above in importance [7].

The second method of extracting weights is through indifference judgements. A hypothetical option with the worst possible results in all categories is used as a template for comparison. Consider the least important attribute is changed from the worst possible to the best possible value. If instead the second least important attribute was to be changed, how much would it need to increase for the decision maker to be indifferent about the alternatives? The process is naturally repeated for all attributes but since comparison with the least important attribute can make the process unstable it is recommended to compare with the attribute raked just below in importance [7].

2.4.2 SMARTER

For a simpler elicitation of weights SMARTER, which stands for **SMART exploiting ranks**, was introduced in 1994. The weights are calculated using solely the rank of the criteria:

$$w_i = \frac{1}{n} \sum_{k=i}^{n} \frac{1}{k} \tag{2.4.1}$$

for n criteria sorted in descending order. Since calculation of the weights usually is very time-consuming and requires an intimate knowledge of both weight assessment methodologies as well as the potential outcomes this speeds up the process considerably and points out the same solution as optimal as SMARTS in 75-87 % of the cases studied by Barron. When the same optimum solution wasn't found the difference in value function using SMARTS between the two alternatives averaged at about 2% [7] [17].

2.5 Evolutionary algorithms

Drawing their inspiration from biological evolution, evolutionary algorithms are often used to solve optimization problems where traditional, often derivative based, methods are less suitable or even inapplicable. Developed by Holland in 1975 [11], the usage of evolutionary algorithms has grown with increased computational power and is one of the most popular categories in stochastic optimization [11].

2.5.1 Natural inspiration

Darwin's **theory of evolution** introduces the concept of a population of individuals from the same species. Same species is here defined by the ability to mate and produce fertile offspring. These offspring carry traits of their parents stored in their genes. An individual's genome determines, aside from environmental factors not considered in this report, the characteristics of the individual and is a combination of the genomes of the parents [22].

A chromosome is a long sequence of genes. The individual genes consist of sequences of base pairs. This makes the chromosomes double-stranded, allowing them to be replicated [22].

There are two types of reproduction in biology. The simplest is **asexual reproduction**. It involves one individual reproducing itself and is most commonly used by single celled organisms such as bacteria, but also more complex organisms such as certain plants and animals [22] [2]. The opposite of asexual reproduction is naturally **sexual reproduction**. Sexual reproduction if often a more complex procedure during which the genetic material of two individuals is combined [22].

The final concept to be introduced in this section is the concept of **mutation**. Mutations are random changes in the genome of an individual and are passed on to potential offspring. Most mutations are harmless, some are beneficial and others harmful. People often have a tendency to think of mutations as something negative causing death or disabilities but mutations are sometimes necessary for survival [9].

2.5.2 Simplifications and adaptations

Whilst drawing much inspiration and vocabulary from biological evolution, evolutionary algorithms adapt the mechanics to suit programming and optimization. Many simplifications are made, such as an individual being identical to its genome. Normally all individuals are replaced at the same time and the fitness landscape is usually non-stochastic and constant in time. The chromosomes are still sequences of genes, but the genes no longer consist of base pairs used as a template to create proteins. Instead they are sequences of numbers, binary or real-valued. They are used to calculate a fitness value indicating how good the individual performs [22].

There are two main way of selecting parents to an individual in the next generation, **tournament selection** and **roulette wheel selection**. In tournament selection a given number of individuals in the current population are chosen. These are ordered according to their fitness values. With a certain probability p the individual with the highest fitness value is chosen, if not the individual is removed from consideration and the process repeated until an individual is chosen or only one remains, which is then chosen. Notice that this method is based only on the rankings of the individuals, not the actual fitness values. Tournament selection is thought to be similar to animals, usually males, fighting over the chance to reproduce [22].

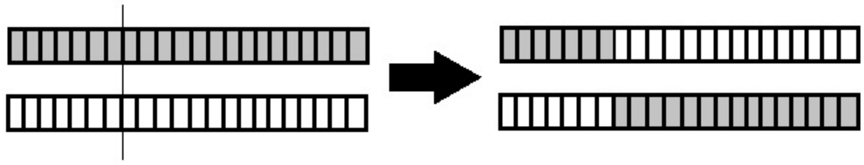
In roulette wheel selection the probability of an individual being chosen is in direct proportion to its fitness. This can be illustrated with spinning a wheel, each individual given a section according to its fitness, hence the name. This method makes individuals with a very low fitness less likely to be chosen than in tournament selection, but if there is an individual with much higher than average fitness it can easily come to dominate the populating, making the algorithm stuck in a local maximum. This can be overcome by only ranking the individuals and giving them a probability based on their position. This method has no near similarities with any biological selection mechanisms [22].

After two individuals in the current generation have been selected a process called crossover is carried out to produce one, or two, new individuals to use in the next generation. There is a pre-defined constant, the **crossover probability**, that determines the probability that a crossover is carried out. If not, one of the parents is copied into the next generation, seen as asexual reproduction. If this probability is set too high the genetic material of very successful individuals will spread very fast in the population, potentially trapping the algorithm in a local optimum [22].

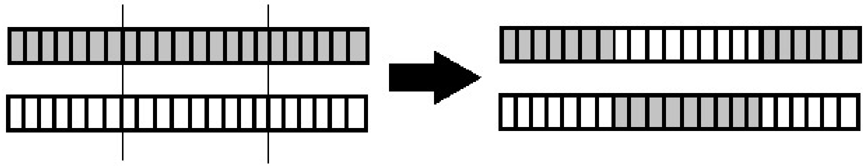
There are many different ways of implementing the crossover procedure. The first decision is the number of **crossover points**. A crossover point is where the genetic material from one parent stops and the other begins. Another choice is whether the crossover should be length-preserving or not. For many gene-encodings length preservation is given but it can be useful as it adds a larger diversity into the gene pool. The process is more easily illustrated by looking at figure 2.5.1 [22].

Before each new individual is entered into the next generation it undergoes a mutation operation. Each element in the chromosome is individually tested against a probability of mutation. An element can be mutated several ways depending on the encoding and application. When binary encoding is used a mutation often means that the element should be flipped, from 1 to 0 and vice versa [22].

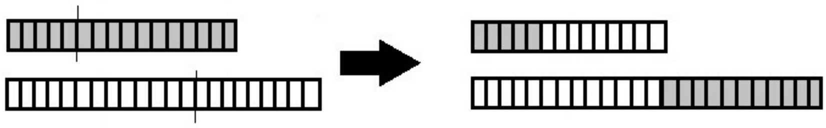
The final concept to be introduced is **elitism**. To ensure that the best solution so far is preserved one or a few copies of it is placed directly in the new generation, not passing through crossover and mutation. This does not significantly change the progress of optimization since the number of individuals typically is fairly large. However, it makes keeping track of the progress easier and termination more reliable not risking ending in a generation with lower fitness than a previous generation [22].



(a) Single point length preserving crossover



(b) Double point length preserving crossover



(c) Single point not length preserving crossover

Figure 2.5.1: Crossover examples

2.5.3 Pareto optimality

When optimizing over several objectives it is not always necessary to aggregate all objectives into a single value function. A genetic algorithm can, with a few modifications, arrive at a pareto-optimal set of solutions that the decision maker can choose between. This significantly lowers the number of solutions for a decision maker to consider but gives no instructions on how to consider them, only that each solution is optimal for at least one specific fitness function [18].

A solution is pareto-optimal if no other solution dominates it. A solution is dominated if another solution performs better in all criteria. In continuous optimization this creates a pareto-optimal front that can be approximated by a number of points in a subset of the outcome space. In discrete optimization the points directly define the pareto-optimal set [18].

There are many ways of carrying out the selection process when a pareto-optimal set is to be found. The simplest is to use new normalized randomized weights each time a selection is to be made and to use fitness proportional selection. In a way this methods illustrates the fact that the pareto optimal set is the set of solutions that could be optimal in at least one point in weight space [18].

Another way of use the pareto rank as a selection fitness. There are two ways of accomplishing this. Simplest is to count the number of solutions an individual is dominated by. Another way is to determine which solutions are pareto-optimal in the current set, give them a predefined maximal value and then remove them from consideration. The procedure is repeated, with a new fitness that is lower than the previous, until no solutions remain. The second method is computationally more expensive but less dependent on the clustering of different solutions [18].

2.5.4 Advantages and drawbacks

The largest advantage of evolutionary algorithms are their versatility. They can be applied to problems with both continuous and discrete variables and work with numerically generated data, experimental data as well as analytical functions. The algorithms are well suited to search complex multidimensional surfaces due to the fact that they can jump out of local minima, do not require even the potential existence of derivatives and can easily handle a large number of variables. Since they simultaneously search a wide sampling of the fitness landscape the algorithms are also well adapted to parallel computing [11].

Traditional optimization methods are very quick to find the optimum of well-behaved convex analytical functions, making evolutionary algorithms much slower, and more complicated and time-consuming to construct. Analytical methods also sometimes can prove that the global, or at least local, optimum is reached, something no stochastic optimization algorithms can guarantee. The problem targeted by this report is far from well-behaved and convex, making a genetic algorithm a natural choice. [11].

3 Implementation

For any conclusions and results to be depended upon the implementation through which they were gathered must be understood. The setup of the simulation, what values to take into account and how to normalize them are decisions to make before any optimization can be thought of. The actual parameters of optimization, the construction of weights with different methods as well as the setup of the evolutionary algorithm that carries out the actual optimization process are also important to consider if one is to stand any chance of defending or duplicating the results.

3.1 Simulation set-up

Three different missile types are used in order to make the helicopter capable of defending itself against a variety of threats. Each missile is fired from twenty different positions between one and five kilometers from the target. The positions are in a sector of 120-150° counted from the nose of the helicopter and all three missile types are shot from the same positions.

To make it harder for the algorithm to optimize a response the helicopter is hovering at a height of 100 m. This makes the separation between the flares and the target smaller than if it had been for a moving helicopter. This makes it harder to achieve large miss distances. The low height reduces the time the helicopter has to react to the threat as well as the distance between the shooter and target.

There is also one restraint that needs to be taken to reality, if not to the model of itself. Two flares of different types can not be ejected simultaneously from the same ejection port. Two flares of the same type can be ejected, essentially doubling the intensity of one flare since, at least in the model, they will have identical trajectories. This is handled by postponing the remaining sequence one time step of length 0.1s.

3.2 Choice of criteria

Selecting what indicators to optimize in one of the first steps in an optimization process. The target goal is to maximize the survival chance of the pilot of the helicopter under fire, but what indicators can one use to describe a solution that is effective and still robust, versatile and cost-effective?

The two most intuitive indicators are the mean and minimum of the miss distances. A miss distance of zero means a hit. Another measurement is the dispersal of miss distances. If the distances are closely gathered it indicates that missiles fired from similar positions might be effectively combated as well. The dispersal is counted as the number of shots, with the same type of missile, shot from within a distance of one kilometer and whose miss distance differs more than ten meters from the shot considered. This is calculated for all shooting positions and the average over all positions is taken as an indicator.

$$I_{\text{dispersal}} = \frac{\sum_{i=1}^{N} \sum_{i \neq j} \chi(|\mathbf{r}_i - \mathbf{r}_j| < 1000)\chi(|d_i - d_j| > 10)}{N} \tag{3.2.1}$$

The helicopter is equipped with a warning system that tells the pilot from which sector the treat is approaching. This makes it possible to have different responses to threats from different sectors. However, this system is not perfect and missiles shot from positions very close to the edges of these sectors are especially important to handle. The mean miss distance from shots fired within 10° of the sector limits are therefore taken as a separate indicator.

In order for the helicopter to survive multiple attacks it is important that the number of flares used remains relatively small. The supply of flares is limited during a mission and in order to withstand attacks from different sectors it is beneficial if the consumption of the two different flare types is even. The difference in the number of flares used of the two types is used as a indicator.

$$I_{\text{even\_usage}} = |n_1 - n_2| \tag{3.2.2}$$

3.2.1 Normalization functions

The different indicators must be normalized into the same scale, for practicality chosen so that zero is the worst possible value and one the best possible. The chosen functions are the gaussian and hyperbolic tangent

Indicator	Ideal point	Halfway point
Number of used flares	1	3
Even usage of flares	0	2
Dispersal of miss distances	0	3

Table 3.2.1: Halfway point and ideal points for normalization of different countable measurements

function, but in both cases a parameter controlling the size of the function is necessary. These parameters are chosen by defining a value for the indicator that is OK but not good and deserve a score of 0.5. This indicator value is called a **halfway point** for both functions and was determined by A. Tonnvik [21].

For the different distance measurements the hyperbolic tangent function was used to illustrate that an increase in miss distance is most important for small distances, where the gradient of the hyperbolic tangent function is greatest, but no matter how large the distance becomes a larger distance is always preferable to a smaller. The chosen halfway point was 15 m for all the distance indicators, making the normalization function as seen below for the simulated distance indicator, both minimum and average, I_d.

$$v(I_d) = \tanh(\frac{\tanh^{-1}(0.5)}{15} I_d) \tag{3.2.3}$$

For the indicators where elements were counted, the gaussian function was used. The ideal point, α, indicating what value could be seen as optimal and is different from indicator to indicator. The variance σ was also calculated using a value $n_{0.5}$ as a halfway point determined by A. Tonnvik [21]. The halfway points as well as the ideal points vary between the indicators as in table 3.2.1.

$$v(I) = \exp \frac{-(I-\alpha)^2}{4\sigma^2} \tag{3.2.4}$$

$$\sigma = \frac{I_{0.5} - \alpha}{\sqrt{-\ln 0.5}} \tag{3.2.5}$$

3.3 Construction of weights

The different weights were constructed according to the methods in sections 2.2 to 2.4. The AHP hierarchy, with local weights, look as in figure 3.3.1. For SMARTS and SMARTER the sorted list of priorities can be seen in table 3.3.1. All the judgements made in the process of weight construction were done by A. Tonnvik [21]. The final weights for all target functions can be seen in table 3.3.2.

The process of constructing SMARTS-weights was adapted to better suit the situation at hand. If all missiles hit the helicopter the number of flares used becomes irrelevant, therefore the baseline scenario was a scenario where all measurements get a normalized score of one half. Also the measurements from different missiles were not compared, all missiles therefore receive equal weight. This was done to speed up and facilitate the estimation process. The criteria are also not independent. A situation with a higher minimum than average miss distance is dificult to visualize and weigh.

Some simpler target functions are also used for comparison. The first two are maximizing the mean and minimum miss distance with all three missiles given equal weight. These target functions are intuitive, easy to motivate and might be seen as the logical choices of target function if one had no knowledge of more complex aggregation methods. As a last resort, to ensure that all the effort was spent for a reason, the target function with all attributes weighed equally important is also tested. These three target functions are so simple that they are not listed in table 3.3.2.

3.4 Evolutionary algorithm

Since the individual parameter values and implementation choices done in an evolutionary algorithm can affect the overall results such as convergence speeds and the possibility to duplicate the results it is important to clearly state what choices were made and why. For readers unfamiliar with evolutionary algorithms reading chapter 2.5 is recommended.

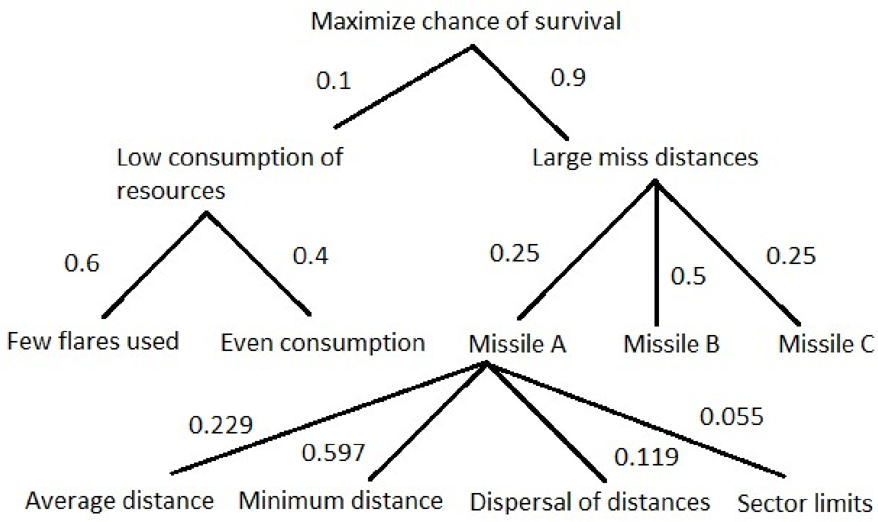


Figure 3.3.1: AHP hierarchy. Criteria weighted the same under all missile types.

Table 3.3.1: Criteria sorted in descending order of importance

Minimum miss distance missile B
Minimum miss distance missile A
Minimum miss distance missile C
Mean miss distance missile B
Mean miss distance missile A
Mean miss distance missile C
Dispersal of miss distances missile B
Dispersal of miss distances missile A
Dispersal of miss distances missile C
Mean miss distances close to sector limits missile B
Mean miss distances close to sector limits missile A
Mean miss distances close to sector limits missile C
Number of flares used
Even usage of flares

Table 3.3.2: Weights for different target function construction methods

Criterion	AHP	DPA	SMARTER	SMARTS
Number of flares used	0.06	0.0348	0.0106	0.0477
Even usage of flares	0.04	0.0174	0.0051	0.0286
Mean miss distance missile A	0.0515	0.0870	0.0834	0.0789
Minimum miss distance missile A	0.1343	0.1304	0.1608	0.0986
Dispersal of miss distances missile A	0.0268	0.0522	0.0471	0.0709
Sector limits missile A	0.0124	0.0417	0.0230	0.0596
Mean miss distance missile B	0.1031	0.0957	0.1013	0.0789
Minimum miss distance missile B	0.2687	0.1391	0.2323	0.0986
Dispersal of miss distances missile B	0.0536	0.0557	0.0573	0.0709
Sector limits missile B	0.0248	0.0435	0.0302	0.0596
Mean miss distance missile C	0.0515	0.0783	0.0692	0.0789
Minimum miss distance missile C	0.1343	0.1304	0.1251	0.0986
Dispersal of miss distances missile C	0.0268	0.0522	0.0381	0.0709
Sector limits missile C	0.0124	0.0417	0.0165	0.0596

Each optimization was carried out with fifty generations of fifty individuals each. In the first generation each chromosome consisted of four genes. In each generation the individual with the highest fitness was copied once into the next generation so that the solution was kept.

Each gene was encoded with four variables. The first position was a flag indicating if that particular gene should be active. Having such a functionality meant that a larger quantity of genetic material could be saved without penalty. The second parameter indicated which of the two dispensers on the helicopter should be used and the third parameter indicated which of the two types of flares should be used. The first flare type had a narrower emission spectrum but a higher intensity than the second flare type. The final parameter to the set was the delay with which the flare was to be ejected. This was done in steps of 0.1s up to 3s. During initialization all parameter values were equally likely.

Tournament selection was carried out with a tournament size of two and a probability of 0.7 of choosing the individual with the highest fitness. The two individuals in the tournament were chosen with replacement from the population and every individual had equal chance of being chosen to participate in the tournament.

Crossover was done with two crossover points as in a combination of figures 2.5.1(b) and 2.5.1(c) on page 7 with each individual cell representing a single gene. The difference being that only one individual was generated. The probability of crossover was 0.85.

Each individual underwent a mutation operation where each variable in each gene was mutated with a probability of 0.05. If a mutation was performed a new value was drawn from a uniform distribution of all possible values for that parameter.

3.5 Experimental set-up

There were seven different target functions to be tested. Each target function was tested ten times through optimization with an evolutionary algorithm as described above. Each iteration a new, random, starting population was used. The best solution found in the end of each iteration, as well as the average and maximum fitness for each generation was retrieved.

Each solution took approximately one minute to test and to avoid unnecessary computation time a database was kept over all tested solutions with both the raw data from the test simulation and the normalized indicators. This database was used to look up if the sequence had been tested before and if so to collect the simulation results directly from there.

4 Results and discussion

In this chapter the results are presented and discussed. The results are grouped into two categories, First the simple objective functions that only consider the mean and minimum miss distances respectively. Second the more complex objective functions where all fourteen objectives are considered. Which flare sequences that were declared optimal and how they performe is also discussed.

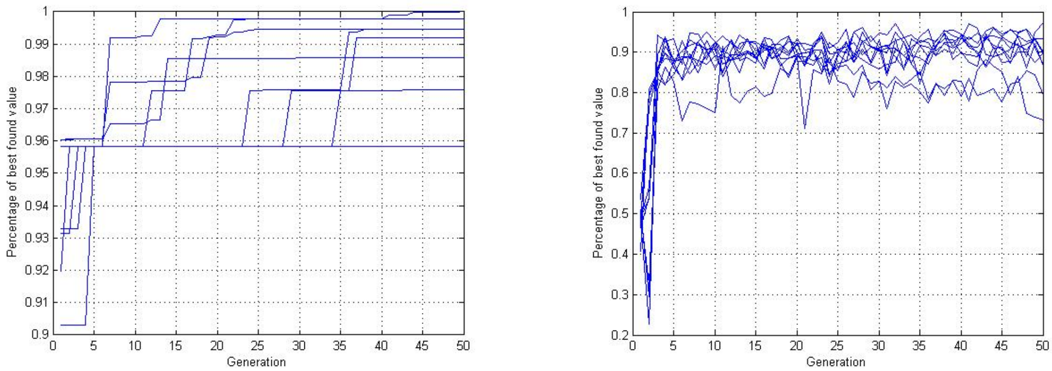
4.1 Simple objective functions

Looking at figures 4.1.1 (a) and 4.1.2 (a) one can easily see two tendencies. First the best individual in the first generation is relatively close to the solution with the highest fitness. This can be taken as a sign that if 50 random sequences with maximum four flares at most of them will probably give decent miss distances.

There are also important differences between figures 4.1.1 (a) and 4.1.2 (a). figure 4.1.1 (a) shows larger jumps and almost all trials appear to have found different maxima. However the fitness begins lower for the mean fitness function, and therefore there is more to be done.

The fitness also tends to make rather large jumps in increasing fitness. The large jumps could be takes as indications of the irregularities in both the input and output space. The maximum fitness never goes down due to elitism which significantly affects the appearance of the data.

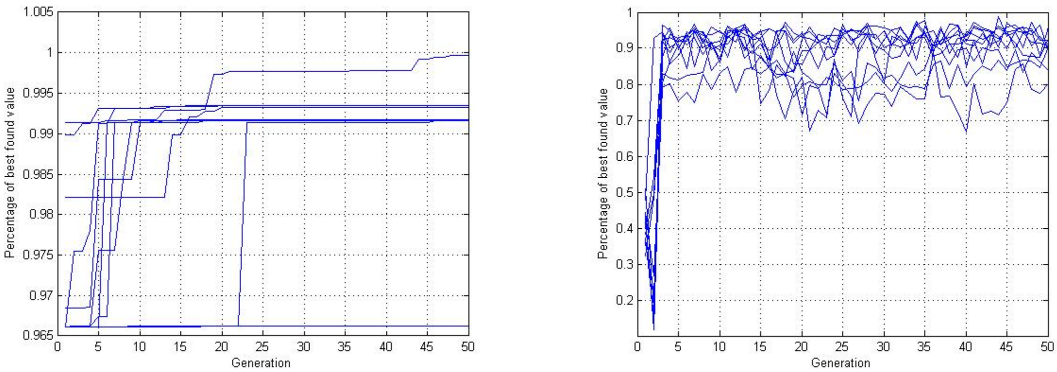
It seems to be a matter of chance what solutions are found as best and with a large probability the global maximum is not found. Considering that these solutions use very many flares, up to 10 which is the maximum in the model used, the relevant volume of the input space becomes very large and more difficult to search. A larger population size could probably help but considering that each flare sequence takes about a minute to



(a) Fitness of the individual with the highest fitness in each generation

(b) Average fitness of the population in each generation

Figure 4.1.1: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using only the average miss distance as fitness function.



(a) Fitness of the individual with the highest fitness in each generation

(b) Average fitness of the population in each generation

Figure 4.1.2: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using only the minimum miss distance as fitness function.

evaluate, and if a larger portion of the input space is to be searched thoroughly more sequences need to be evaluated.

When looking at figures 4.1.1 (b) and 4.1.2 (b) one can see that the average fitness is rather noisy, but there are a few interesting tendencies. Firstly the initial sequences start with a relatively low fitness of about half the maximum. Together with the high maximum fitness this indicates that there is a large spread of fitness values, which is positive.

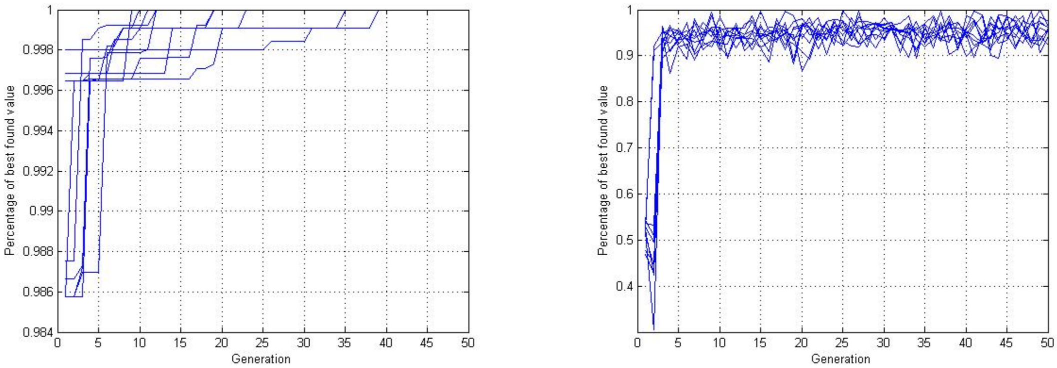
Many, but not all, iterations also make a large dip in the second generation with fitness values of around 20 - 30% of the maximum achieved fitness. Other iteration remain relatively close to the first average fitness and others increase the average but in the third generation all iteration achieve around 80 - 90% and remain there for all subsequent generations.

4.2 Advanced objective functions

Looking at Figure 4.1.1 (a) one can immediately see that when using the AHP fitness function all iterations reach the same optimum solution, a solution not superseded by anything found by using another fitness function. AHP is alone in this behavior.

Table 4.2.1: Average convergence time for four target function

Target function	Convergence time
AHP	19.9 ± 10.6192
Equal weights	14.25 ± 7.5166
DPA	19.625 ± 15.8919
SMARTS	17.8889 ± 6.1531



(a) Fitness of the individual with the highest fitness in each generation

(b) Average fitness of the population in each generation

Figure 4.2.1: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using the AHP fitness function.

Looking at figures 4.2.3 (a) and 4.2.5 (a) one can see that the fitness when using the SMARTS fitness function as well as the equal weights fitness function share half of this property. The solutions found also do not become superseded by solutions found using other fitness functions. A majority, 9/10 when using the SMARTS fitness function and 8/10 when using the equal weights fitness function, find the optimum solution, but not all.

In a comparison between these three figures one can see that the initial fitness values using AHP are very high and after a few generation the best soultion found in every iteration are within 0.4% of the final fitness. When using equal weights the initial fitness values on the other hand start relatively low but then rises very quickly. The same cannot be said about Figure 4.2.3 (a) where the convergence time is relatively long with big and sudden improvements being made midway in one iteration.

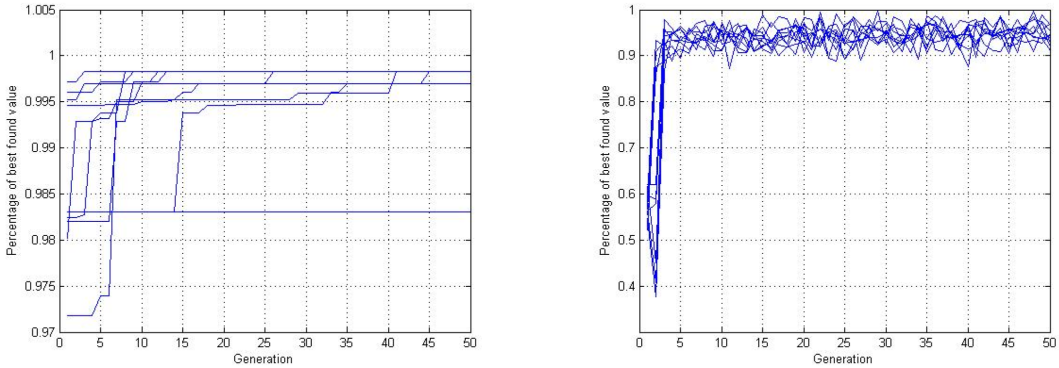
Figures 4.2.2 (a) and 4.2.4 (a) show the best solution found by using the corresponding fitness function was superseded by a solution found using another fitness function. The solutions found are however very close to the highest scoring solution. Both show good convergence with 8/10 iterations converging on the same solution. However, in the case of SMARTER the converged upon solution is trumped by the solution found in another iteration .

In table 4.2.1 one can see the convergence time for the four fitness function AHP, DPA, SMARTS and equal weights. One can see that equal weights has the shortest average convergence time but all are fairly close, well within the standard deviations. The convergence time is naturally only applied to iterations landing in the same solution, excluding the simple objective function that found different solutions in every iteration. SMARTS is also excluded due to the fact that only one iteration found the solution with the highest fitness.

Looking at the mean fitness in each generation instead in figures 4.2.1 (b), 4.2.2 (b), 4.2.3 (b), 4.2.4 (b) and 4.2.5 (b) not as much can be gained. Most interesting is the dip in average fitness that occurs in the second generation. This is shared by all fitness functions, including the simpler fitness functions seen in figures 4.2.1 (b) and 4.2.2 (b). In figure 4.2.3 (b) only one iteration goes down in the second generation but it seems to be the most common alternative for the remaining fitness functions.

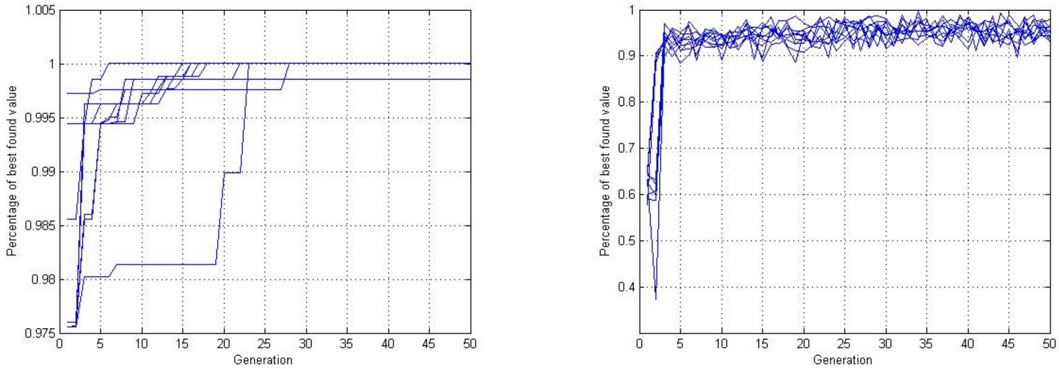
4.3 The different winning solutions

It is also important to study the individual winning solutions and how they perform in the different criteria. Remembering how to decode a chromosome from section 3.4 only the encoded sequences will be used.



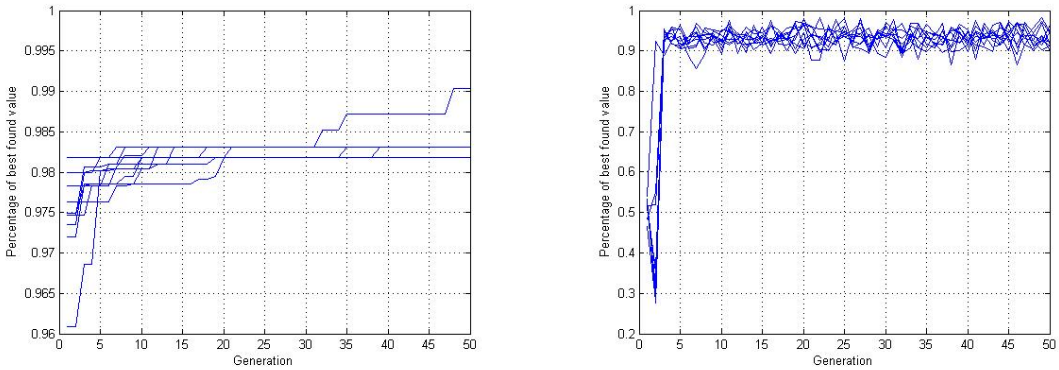
(a) Fitness of the individual with the highest fitness in each generation (b) Average fitness of the population in each generation

Figure 4.2.2: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using the DPA fitness function.



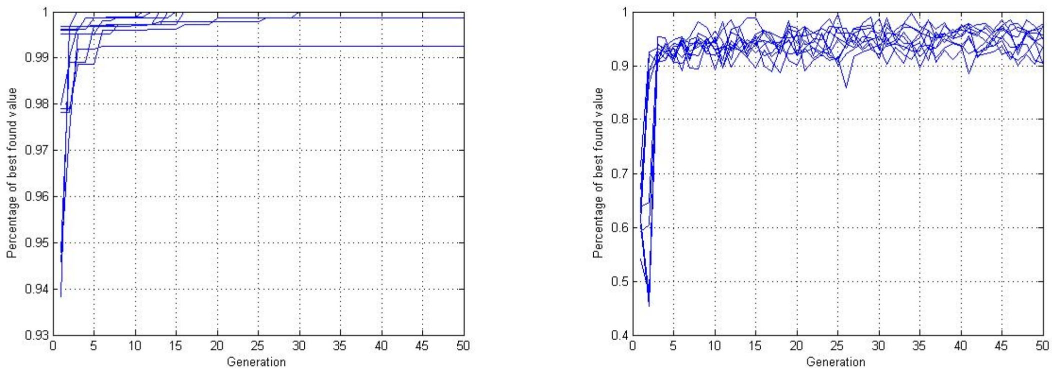
(a) Fitness of the individual with the highest fitness in each generation (b) Average fitness of the population in each generation

Figure 4.2.3: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using the SMARTS fitness function.



(a) Fitness of the individual with the highest fitness in each generation (b) Average fitness of the population in each generation

Figure 4.2.4: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using the SMARTER fitness function.



(a) Fitness of the individual with the highest fitness in each generation

(b) Average fitness of the population in each generation

Figure 4.2.5: The average fitness of each generation, as well as the fitness of the best individual of the generation measured as percentage of solution with the highest fitness-value found using the fitness function with all weights equal.

Table 4.3.1: Winning flare sequence when using SMARTER target function

1 2 2 0
1 2 1 1
1 2 2 1
1 2 1 3
1 1 1 7
1 2 1 0
1 2 2 1
1 2 2 2
1 2 1 1
1 2 2 1

The most commonly victorious sequence was: 1 2 2 0 1 1 1 11. It was the winning solution when using AHP, DPA, SMARTS and equal weights. it was also the most commonly found solution when using SMARTER. For the case of DPA the sequence was not the best found overall but the best one found when optimizing with that fitness function. The optimal solution with DPA was: 1 2 2 0 1 2 1 3 1 2 1 3.

The best solution found using SMARTER was: 1 2 2 0 1 2 1 3 1 2 1 5, but the solution with the highest fitness is for simplicity not given in the text but seen in table 4.3.1 and is, as seen, very long. Not surprising since SMARTER is the target function giving the smallest weight to the criterion to use few flares, see table 4.4.1.

4.4 Sequence specific performance

Seeing the winning flare sequences for different fitness functions in table 4.4.1 some things become clear. Firstly that missile C is much harder to protect against than missiles A and B and secondly that the dispersal criteria is easily maxed out. The increased difficulty of protecting against missile C is the reason the missile type was added, though in hindsight perhaps the difficulty was a bit to great, especially considering that the minimum miss distance for missile C is zero for all tested flare sequences. Interesting to note is that the sector limit measurements are identical to zero. Since this is the average miss distance from shots fired close to the sector limits it tells one that these shots are never protected against, a serious deficiency if the goal of this report was to conclude the optimal protection strategy.

The dispersal criteria is, as mentioned in section 3.2, the number of shots fired from within one kilometer, and of the same missile type, whose miss distance differs more than ten meters from the shot in question. It is maxed out to zero neighbors in all but three cases, all for missile C. Perhaps most logically the cases of mean and minimum miss distance target functions that ignore the criteria in question. The final case is the DPA

Table 4.4.1: Normalized and unnormalized performance in different categories for the optimal solutions

Objective	AHP Equal , SMARTS	DPA	SMARTER	Mean	Minimun
Nr flares	0.8350 (2)	0.5 (3)	0 (10)	0.0002 (8)	0.0131 (6)
Even consumption	1.0000 (0)	1.0000 (0)	0.2102 (3)	0.2102 (3)	0.5 (2)
Mean miss distance Missile A	0.8590 (35.2159)	0.8761 (37.1047)	9031 (40.6491)	0.9138 (42.3358)	0.9138 (42.3358)
Min miss distance Missile A	0.8555 (34.8478)	0.8722 (36.6557)	0.9002 (40.2293)	0.9111 (41.8961)	0.9112 (41.8994)
Dispersal Missile A	1.0000 (0)	1.0000 (0)	1.0000 (0)	1.0000 (0)	1.0000 (0)
Sector Limits Missile A	0.8555 (34.8478)	0.8722 (36.6557)	0.9002 (40.2293)	0.9111 (41.8961)	0.9112 (41.8994)
Mean miss distance Missile B	0.8591 (35.2245)	0.8591 (35.2207)	0.8600 (35.3199)	0.8641 (35.7557)	0.8641 (35.7573)
Min miss distance Missile B	0.8553 (34.8356)	0.8553 (34.8328)	0.8560 (34.8988)	0.8596 (35.2730)	0.8596 (35.2730)
Dispersal Missile B	1.0000 (0)	1.0000 (0)	1.0000 (0)	1.0000 (0)	1.0000 (0)
Sector limits Missile B	0.8553 (34.8356)	0.8553 (34.8328)	0.8560 (34.8988)	0.8596 (35.2730)	0.8596 (35.2730)
Mean miss distance Missile C	0.0635 (1.7373)	0.2476 (6.9057)	0.2021 (5.5947)	0.3893 (11.2218)	0.3892 (11.2198)
Min miss distance Missile C	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Dispersal Missile C	1.0000 (0)	0.9395 (0.9)	1.0000 (0)	0.0625 (6)	0.0625 (6)
Sector limits Missile C	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)

fitness function. The criteria is not under-appreciated, in fact only SMARTS weighs it heavier, but the score is after all very close to maximum and outperforms in other categories, perhaps most noticeably the mean miss distance for missile C.

5 Conclusions

None of the tested methods of constructing target function can be described as bad or meaningless. The general conclusion is that it pays of very well to consider a more complex target function. The discussion leading up to determination of the criteria and the process of calculating the weight is also helpful making one think through what really is important and how things correlate.

One can in general say that the analytic hierarchy process gives very good results. The solution is found fast and reliably. That the solution is found to be optimal with other target functions also indicates that the solution is good.

The biggest advantage of AHP over the other methods is the construction of a hierarchy in which the objectives are organized. When many objectives are used, such as in this case 14 objectives, it simplifies thinking and lessens the amount of necessary comparisons. In this case of fourteen criteria 91 comparisons are needed to do a full evaluation when not using hierarchies. Comparing two very different objectives, where the difference in importance also is great, is also very difficult to do accurately and can lead to instabilities and inconsistencies.

SMARTS was considered very complicated to carry out with strange situations to judge. The original version with the baseline scenario with all indicators at their worst possible value was instinctively hard to judge because if the minimum miss distance is zero, which is the same at the helicopter is shot down, one does not care about the other parameters. The objectives are also not entirely independent. The situation with a higher minimum than mean is difficult to visualize.

5.1 Reservations

The sequences presented as optimal in this report can not directly be said to be optimal in the real world. First the scenario was to simple and the heat emissions of the helicopter had to be changed so that it became more difficult to trick the missile. This was done for missile C and explains the very poor results for protection against that missile. Secondly no amount of optimization can optimize more than the model it is given. Certain factors are not incorporated in the model and therefore aren't considered.

5.2 Further study

The dispersal indicator, because of its easy maximization, might be subject to revision were another experiment to be performed. Either the criteria could be sharpened by lessening the allowed variations of neighborly miss distances or subject to removal. In the case of removal one should be careful because even if the performance in this criteria is uninteresting in the final result there is no way to tell how big a part it plays in the progress of finding the optimal solution. Also the criteria describes an important consideration and in removed should be replaced with another measure of estimating the variations in miss distances.

Given more time more computations could be done and the convergence comparisons could be done with larger data sets. Also bear in mind that this target function creation method comparison is done with only one situation. A choice that appears good here might not be as good when another situation is looked at.

Bibliography

[1] A. F. Abd El-Wahed, ”Intelligent fuzzy multi-criteria decision making: Review and analysis”, *Fuzzy Muliti-Criteria Decision Making*, USA: Springer, 2008, ch. 2, pp 19-50.

[2] B. O. Bengtsson, ”Asex and evolution: A very large scale overview” in *Lost Sex*, Netherlands: Springer, 2009, ch. 1, pp. 1-19.

[3] T. Demiriel, N. Demiriel, C. Kahramen, ”Fuzzy analytic hierarchy process and its application”, *Fuzzy Muliti-Criteria Decision Making*, USA: Springer, 2008, ch. 3, pp 53-83.

[4] J. S. Dyer, R. K. Sarin, ”Measurable Multiattribute Value Functions”, Operations research, vol. 27, no. 4, pp. 810-822, Jul-Aug 1979.

[5] J. S. Dyer, ”Remarks on the Analytic Hierarchy Process,” *Management Science*, vol 36, no 3, pp. 249-258, Mar, 1990.

[6] J. S. Dyer, ”MAUT - Multiattribute Utility Theory,” in *Multiple Criteria Decision Analysis: State of the Art Surveys*,2005 ed. New York, USA: Springer, 2005, ch. IV, pp 265-292.

[7] W. Edwards, F. H. Barron, ”SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement” ,in *Organizational Behavior and Human Decision Processes*, vol. 60, no. 3, pp. 306-325, 1994.

[8] E. H. Forman, S. I. Gass, ”The Analytic Hierarchy Process: An Exposition,” *Operations Research*, vol. 49, no. 4, pp. 469-486, Jul-Aug 2001.

[9] D. R. Forsdyke, ”Mutation” in *Evolutionary bioinformatics*, New York, USA: Springer, 2011, ch. 7, pp. 131-151.

[10] M. Gershon, ”The role of weights and scales in the application of multiobjective decision making”, in *European Journal of Operational Research*, Vol. 15, No. 2, pp. 244-250, Feb 1984.

[11] R. L. Haupt, S. E. Haupt, *Practical Genetic Algorithms*, 2nd ed. Hoboken, New Jersey, USA: John Wiley & Sons, 2004.

[12] M. Hø”rberg et al., ”Automatiserad framtagning av reactiv fackelfällsekvens,” FOI, Linköping,FOI-R–1367–SE ,Jan. 2011.

[13] C. Kahraman, ”Multi-criteria decision making methods and fuzzy sets” , *Fuzzy Muliti-Criteria Decision Making*, USA: Springer, 2008, ch. 1, pp 1-18.

[14] C. Kahraman, S. Birgün, V. Z. Yenen, ”Fuzzy multi-attribute scoring methods with applications”, *Fuzzy Muliti-Criteria Decision Making*, USA: Springer, 2008, ch. 7, pp 187-208.

[15] B. Malakooti, ”Systematic decision process for intelligent decision making,” *Journal of Intelligent Manufacturing*, vol. 22, no. 4, pp. 627-642, Aug 2011.

[16] P. Mendez Jr, ”Analytic Hierarchy Process,” in *Demand Driven Supply Chain*,. Berlin Heidelberg, Germany: Springer, 2011, ch. 6, pp149-155.

[17] M. Pöyhönen, R. P. Hämöläinen, ”On the convergence of multiattribute weighting methods” in *European Journal of Operational Research*, vol. 129, no. 3, pp. 569-585, Mar 2001.

[18] G. Rohling, ”Multiple Objective Evolutionary Algorithms for Independent, Computationally Expensive Objective Evaluations,” Ph.D. dissertation, School of Electrical and Computer Engineering, Georgia Institute of Technology, Nov. 2004.

[19] T. L. Saaty, L. G. Vargas, ”The Seven Pillars of the Analytic Hierarchy Process” in *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*, New York, USA: Springer, 2012, ch. 2, pp 22-40.

[20] A. Salo, P. Hämöläinen,"Preference programming - Mulitcriteria weighting models under incomplete information" in *Handbook of multtcriteria analysis*,vol 103, Heidelberg, Berlin, Germany: Springer, 2010, ch. 5,pp 167-187.

[21] A. Tonnvik, private communication, Sep 2013.

[22] M. Wahde, *Biologically inspired optimization methods*, Southampton, United Kingdom: WIT Press, 2008.