Analyzing Defense Capability of a Fleet of Military Aircraft Through Simulation

Eero Rantala

School of Science

Espoo August 31st 2018

Supervisor and advisor

Prof. Kai Virtanen

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Eero Rantala

Title Analyzing Defense Capability of a Fleet of Military Aircraft Through Simulation

Degree programme Mathematics and Operations Research

Major Systems and Operations Research **Code of major** SCI3055

Supervisor and advisor Prof. Kai Virtanen

Date August 31st 2018 **Number of pages** 68 **Language** English

Abstract

This thesis studies the air defense capability of a fleet of military aircraft with respect to the attributes of force fulfillment and engagement frontier. The force fulfillment measures the ability to deploy a number of assets consecutively to an operation area for a prolonged period of time. The analyzing of the force fulfillment is essentially an allocation and scheduling problem. Each asset has a limited amount of fuel, and the asset needs to return to base before fuel runs out. A replacing asset is needed at the operation area when the previous asset needs to return to base. After being refueled the asset is allocated again to some operation area. The replacement is done just in time whenever possible, and this sets up a problem where the flight schedules of assets are dictated by earlier allocations bit by bit.

The engagement frontier measures the capability to counter the first enemy attack using assets on alert and stationed at the bases. The airspace may be restricted by zones which must be circumvented. The assets carry missiles which reach far ahead of the asset and thus form a missile envelope. It might be wanted that the assets are gathered such that the engagement would happen with two assets present. The engagement frontier is formed where the opposing assets engage at the earliest.

This thesis presents two simulation models; one for computing measures of the force fulfillment, and one for computing the engagement frontier. The first model includes an allocation algorithm that uses a heuristic for controlling the use of ground resources and for allocating the aircraft to the operation areas. The allocation algorithm produces a flight and maintenance schedule from which the measures of the force fulfillment can be deduced. The simulation model for computing the engagement frontier is built upon network optimization where the earliest possible arrival times to each node in a grid are calculated.

The simulation models enable versatile analyses, e.g., the force fulfillment over a spatial area. With the simulation model for the engagement frontier, the use of assets and their standby times can be determined such that the desired engagement frontier is achieved without excessive high alert.

Keywords air defense capability, force fulfillment, engagement frontier, simulation, heuristic, network optimization

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Eero Rantala

Työn nimi Hävittäjälentokonelaivueen puolustuskyvyn analysointi simuloinnin avulla

Koulutusohjelma Matematiikka ja operaatiotutkimus

Pääaine Systeemi- ja operaatiotutkimus **Pääaineen koodi** SCI3055

Työn valvoja ja ohjaaja Prof. Kai Virtanen

Päivämäärä 31.8.2018 **Sivumäärä** 68 **Kieli** Englanti

Tiivistelmä

Tässä diplomityössä tarkastellaan lentokoneosastojen tuottaman ilmapuolustuskyvyn mittaamista kahdella eri attribuutilla; voiman riittävyydellä ja kohtaustasalla. Voiman riittävyys mittaa kykyä ylläpitää tietty määrä lentokoneosastoja peräjälkeen toiminta-alueella pitkällä aikavälillä. Voiman riittävyyden tarkastelu on pohjimmiltaan allokaatio ja aikataulutustehtävä. Kullakin lentokoneosastolla on rajallinen määrä polttoainetta, ja osaston täytyy palata tukikohtaan polttoaineen loppuessa. Korvaava osasto tarvitaan toiminta-alueelle, kun edellisen osaston täytyy kääntyä toiminta-alueelta pois. Tankkauksen jälkeen osasto allokoidaan jälleen jollekin toiminta-alueelle. Vaihto tapahtuu milloin mahdollista, ja tämä johtaa tehtävään, jossa osaston lennätysaikataulu määräytyy aiempien allokaatioiden perusteella pala palalta.

Kohtaustasa mittaa kykyä vastata ensihyökkäykseen lentokoneosastoilla, jotka ovat tukikohdissa määrätyssä valmiudessa. Ilmatilaa saattaa rajoittaa lentokieltoalueet, jotka täytyy kiertää. Lentokoneosastoilla on käytössään ilmatorjuntaohjuksia, jotka muodostavat ohjuskuoren lentokoneen edelle. Voimaa voidaan keskittää siten, että torjuntaan halutaan käyttää kahta lentokoneosastoa. Kohtaustasa muodostuu siihen kohtaan, jossa vastakkaisten osapuolien osastot kohtaavat aikaisimmillaan.

Tässä työssä esitellään kaksi erillistä simulaatiomallia; yksi voiman riittävyyden laskentaan, ja toinen kohtaustasan laskentaan. Voiman riittävyyden simulaatiomalli sisältää allokaatioalgoritmin, joka käyttää heuristiikkaa tukikohtaresurssien hallintaan ja lentokoneiden allokointiin toiminta-alueille. Allokaatioalgoritmi tuottaa lennätys- ja huoltoaikataulun, jonka perusteella voidaan mitata voiman riittävyyttä. Simulaatiomalli kohtaustasan laskentaan perustuu verkko-optimointitehtävän ratkaisemiseen, jossa lasketaan aikaisin mahdollinen saapumisaika kuhunkin hilapisteen solmuun.

Simulointimallit mahdollistavat erilaisia analyysejä liittyen esimerkiksi voiman riittävyyteen maantieteellisellä alueella. Kohtaustasan simulaatiomallin avulla pystytään määräämään lentokoneosastojen valmiuksia siten, että haluttu kohtaustasa saavutetaan, mutta osastot eivät joudu turhaan olemaan korkeassa valmiudessa.

Avainsanat ilmapuolustuskyky, voiman riittävyys, kohtaustasa, simulaatio, heuristiikka, verkko-optimointi

Preface

I want to thank the subject matter specialists with whom I worked with in close collaboration, and for whom this work is meaningful and of practical importance. Special thanks go to the chief specialist who had a passion for rockabilly culture and American vintage cars. It was also a great honor to work with one specialist who turned out to be a two time champion of the Jukola Relay.

I thank my supervisor Kai Virtanen for collaboration and guidance. I also thank my former co-workers for their peer support.

Finally, a very big thank you and apology to my wife and my children, from whom all this time has been off.

Espoo August 31st 2018

Eero Rantala

Contents

1. Introduction

1.1 Background and Motivation

In defensive military operations, a primary task of air assets is to secure the airspace of a nation. Fighter aircraft included in the assets are used for reacting to foreign aircraft that have intruded into the nation's airspace, or they can be employed in a preemptive manner to ensure the policing and surveillance of the airspace. The more aircraft are deployed both spatially and temporally, the better the *air defense capability*.

Air defense capability has many aspects that can be measured by different attributes. In this thesis, the air defense capability is measured by two different attributes, the *force fulfillment* and the *engagement frontier* [1]. The force fulfillment measures the ability to deploy a number of assets consecutively to an operation area for a prolonged period of time. The engagement frontier measures the capability to counter the first enemy attack using assets on alert and stationed at the bases. Surface to air weapons are not considered in this thesis and only the effect of fighter aircraft are taken into account when considering the air defense capability.

In order to get a holistic view of the air defense capability and to utilize the fleet of aircraft effectively, it is essential to analyze the performance of the fleet of aircraft and the ground resources in terms of the force fulfillment and the engagement frontier. A simulation model can be useful for planning pilot readiness, aircraft deployment and the utilization of both air and ground resources.

The aircraft operate in a group, and this group of aircraft is treated as an atomic asset in air defense. The flight paths of the aircraft may be affected by a restricted operating zone (ROZ), through which the aircraft are not allowed to fly, but which they must circumvent. The available ground resources also heavily affect the air defense capability. Ground resources include everything within a base; the runway, the plateaus where refueling and reloading weapons is done, the fuel and weapons stock, and the maintenance crew. The pilots have humane limitations in how long they are capable of being strapped in their seat in the cockpit and excess standby times are exhaustive for the pilots, and for this reason too, planning is required.

An essential part of air warfare is the strategy of choosing a suitable launch point for air to ground missiles. Air to ground missiles are different from air to air missiles.

Air to air missiles have their own propulsion and they strike a moving target, e.g., another aircraft [2]. Air to ground missiles are typically gliding bombs that do not have propulsion of their own and they are used to strike against a stationary target on the ground, e.g., a building [3]. Determining a suitable point of launch for such a gliding air to ground bomb is a complex problem of its own and it is not a topic in this thesis. The problem of finding a suitable launch point has been studied by, e.g., [4].

Deploying a number of assets consecutively to an operation area sets up an allocation and scheduling problem [5]. The allocation problem arises from allocating the assets to operation areas. The scheduling problem arises because each asset has a limited amount of fuel, and the asset needs to return to base before fuel runs out. A replacing asset is needed at the operation area when the previous asset needs to return to base. After being refueled the asset will be allocated again to some operation area. The replacement is done just in time whenever possible, and this sets up a problem where the flight schedules of assets are dictated by earlier allocations bit by bit.

Determining the engagement frontier requires calculating the minimum time in which the opposing assets will engage each other. This calculation is essentially a shortest path problem [6]. The solution would be trivial in the absence of any ROZ. The shortest path problems are a well studied field, and algorithms for solving them are readily available (see, e.g., [7]). However, the algorithms only solve explicitly defined problems and defining the problem in a analytical form is a unique issue.

An allocation problem would normally be formulated as an optimization problem where some cost function is minimized under constraints. One of the strongest candidates for solving a conventional allocation problem would be mixed integer linear programming (MILP) [8, p. 96]. In the case of force fulfillment, however, the constraints are not static as required in ordinary optimization but rather time-dependent. Also the total amount of the allocations is not known in advance. There are many other allocation and scheduling problems in the literature [9][10][11][12][13] that share some common features with the force fulfillment but none that could be used as such.

The basic setup of a scheduling problem is having a number of jobs that need to be all allocated for processing for a limited number of machines [14, Chap. 1]. Of course, the concept of a ‘machine’ is abstract and in the case of air defense capability, the ‘machine’ could be the operation area and the ‘jobs’ that are allocated could be the assets. Unfortunately, the fundamental purpose of scheduling problems in the literature is to get all the jobs processed. In the force fulfillment, the jobs, i.e., the assets, incur only cost and the fundamental purpose is to keep the ‘machine’, i.e., the operation area, running non-stop. No scheduling problem presented in the literature can be applied directly to solving the attribute of force fulfillment in air defense capability but ideas can be borrowed.

1.2 Research Objectives

Because the air defense capability is a multi-faceted issue, analyzing the performance of the fleet of aircraft is approached in this thesis from two perspectives which are the attributes of force fulfillment and engagement frontier. The attributes are unique and there exists no known problem formulation in the literature that could be applied as such to yield measures for either of the attributes. Thus, in this thesis, the attributes are fully defined and a problem formulation is done for determining a measure for each attribute. Two distinct simulation models are constructed for solving the problem set ups for each attribute and ultimately, to analyze the air defense capability.

The simulation model for the force fulfillment includes an allocation algorithm that uses a heuristic for controlling the use of ground resources and for allocating the aircraft to the operation areas. The allocation algorithm produces a flight and maintenance schedule which solves the problem. Different measures of how well the force fulfillment is achieved can be calculated from the schedule.

The simulation model for the engagement frontier is built upon a network optimization problem [15] where the earliest possible arrival times to each node are calculated. The earliest possible arrival times take into account the optimal way to circumvent any combination of ROZs. The engagement frontier is essentially the line where the assets from opposing sides have an equal minimum time to reach.

1.3 Structure

The thesis is organized as follows. In Chapter 2, the air defense capability is defined, two attributes of the air defense capability are presented as well as the components contributing to the overall air defense capability are discussed. In Chapter 3, the simulation model used for analyzing the force fulfillment is presented. Chapter 4 presents the simulation model for finding solutions of the engagement frontier. In Chapter 5, examples of the use of simulation models are presented. Finally, in Section 6, the contributions of this thesis are summarized and directions for future research are pointed out.

2. Air Defense Capability

The defense capability of a fleet of military aircraft is the ability of the fleet to be deployed both spatially and temporally. The fleet consists of a number assets, which are groups of military fighter aircraft. The assets rely on bases and they cannot operate without support. The bases serve two distinct functions. First, they provide maintenance such as refueling, rearming and repairing aircraft between missions. Second, aircraft on alert and ready to move out on short notice are also stationed at the bases. The bases are aircraft carriers, military airfields or road runways. The flight paths of the assets may be restricted by a restricted operating zone (ROZ). ROZs are areas through which the assets are not allowed to fly, and are forced to circumvent, if the ROZ lies in the assets' flight path.

In this thesis, the air defense capability is measured by two attributes, the force fulfillment and the engagement frontier. Both of these attributes are explained in detail in Sections 2.1 and 2.2. The factors affecting the air defense capability, as measured by the attributes, are explained in subsections of Section 2.3. The attributes measure different aspects of the air defense capability, and the attributes have considerable differences in which factors they use and how they use them.

2.1 Force Fulfillment

The force fulfillment measures the ability to deploy a number of assets consecutively to an operation area for a prolonged period of time. Aircraft are sent from fixed bases to fulfill temporal and spatial force requirements. These force requirements are given as operation areas. An operation area is some airspace with strategic importance where enemy air attack is anticipated, or where surveillance is needed. Each operation area requires a given number of assets for a given amount of time. Due to the assets having a restricted amount of fuel, they need to return to base for refueling. The goal is to rotate the aircraft in such a manner that the force requirement is maximized at all times.

The bases can only serve a few assets at a time and the base nearest to the operation area may not be able to serve an asset that is leaving the operation area. Naturally, the nearest base would be preferred as this minimizes unnecessary fuel consumption and time spent on commuting between the operation area and the base. Time spent at the operation area is what is wanted to maximize and all other time is wanted to minimize. The bases can offer hot refuel, which is faster than normal refueling, but there are limitations for the assets undergoing hot refuel. The assets may also consume weapons at the operation area, and the weapons load needs to be replenished at times. Weapons are not available at all of the bases, and in case the asset is in need of weapons, it must return to a base that offers weapons. There are an endless amount of ways how the assets can be rotated between the bases and the operation areas. The internal structure of a base and the commuting between the operation areas and the bases is presented in Figure 2.1.

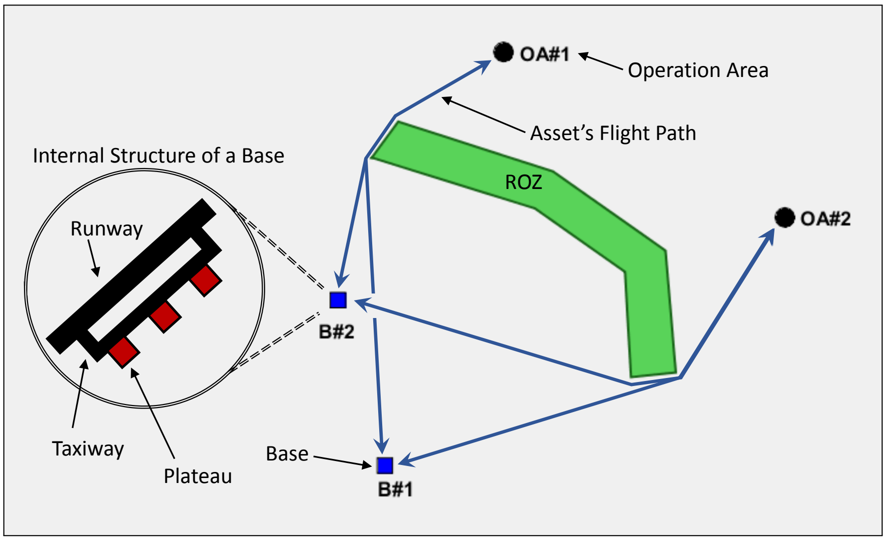

Figure 2.1. *The components relating to the air defense capability as measured by force fulfillment. The blue lines represent assets' flight paths circumventing a ROZ. The internal structure of a base is also illustrated. The base consists of a runway, taxiway and plateaus.*

2.2 Engagement Frontier

The engagement frontier measures the capability to counter the first enemy attack using assets on alert and stationed at the bases. This attribute is based on the concept of an engagement frontier, which is the frontier where approaching hostile aircraft are first intercepted. The aircraft fly a one time mission to intercept the foreign aircraft, and return back to their home base. In this thesis, the hostile counterpart will be referred to as Red, and the defending side as Blue. A crucial factor affecting the engagement frontier is the *advance warning*, which is the first moment when the Blue side becomes aware of any activity by the Reds. The earlier the defending aircraft are in flight, the further back the engagement frontier will be from the Blue bases. The advance warning can be triggered by the hostile aircraft cutting the *reconnaissance alert line*, or it can be gained by detecting activity directly from within a Red base. The components relating to the engagement frontier are presented in Figure 2.2.

The ROZ will greatly affect the shape and location of the engagement frontier as the ROZ needs to be circumvented and this adds to the flying time. Finding the shortest route is one issue of consideration, especially in the presence of multiple ROZs. Another factor affecting the engagement frontier is the use of missiles. The missiles can strike their target far ahead of the launching aircraft. This forms an *missile envelope* [16] for the aircraft, and the missile envelope may or may not want to be taken into account.

The engagement frontier has many different combinations to take into account. The intuitive case is the head-on engagement of the first assets. However, for a number of reasons, the engagement of the second asset on site may be of interest. Such reasons may be for example wanting air supremacy by numbers, or a fighter aircraft

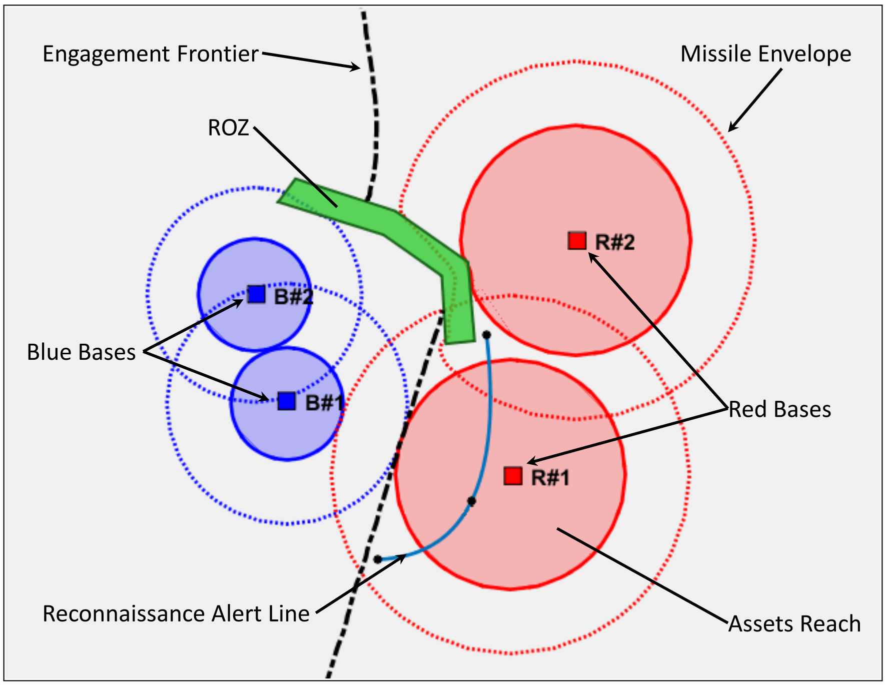

Figure 2.2. *The components relating to the air defense capability as measured by the engagement frontier. The solid red and blue areas are the reach of each asset and the dotted line around the reach is the asset's missile envelope. The engagement frontier is set to be the engagement of the first Blue missile envelope versus the first Red asset. The advance warning is gained from the Red asset from Base R#1 cutting the reconnaissance alert line.*

escorting a heavily armed bomber asset, and the interception of the bomber being the main interest. The combinations of first or second assets and missile envelope or no missile envelope results in a combination of a total of sixteen different alternatives for representing the engagement frontier. The combinations are illustrated in Figure 2.3.

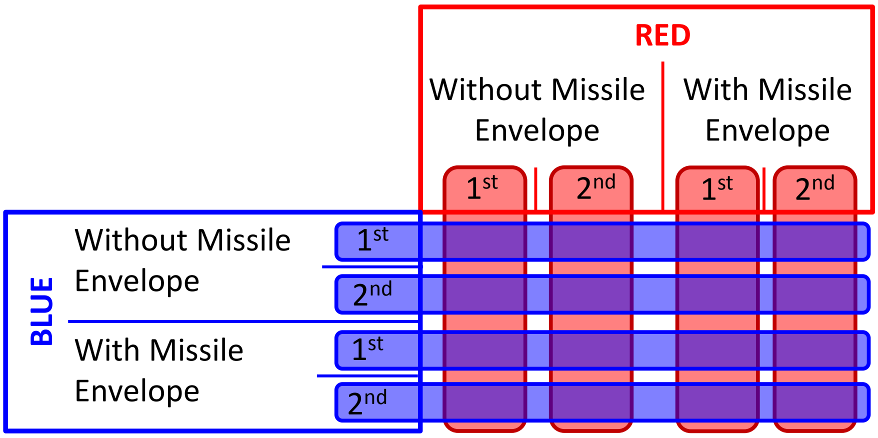

Figure 2.3. *The sixteen different alternatives for the engagement frontier. The engagement frontier can be determined with or without missile envelope and for the first or the second assets for both of the Blue and Red parties independently.*

2.3 Factors of Air Defense Capability

2.3.1 Asset

An asset is a single operating unit of military fighter aircraft. One asset consists of one or more aircraft, and the asset is treated as a single unit. Each aircraft is manned with one pilot. The pilots and the aircraft may be in different states of readiness. To take off from the highest state of readiness requires approximately only one minute of time, whereas to get the aircraft airborne from the most relaxed state may take a quarter of an hour. The highest level of readiness implies that the pilot is strapped in the cockpit with the aircraft engines running, and the aircraft is ready to take off on command. However, the pilots have humane limitations in how long they are capable of being strapped in their seat in the cockpit and excess standby times are exhaustive for the pilots. The aircraft has also limitations on how long its engines can be running on the ground without causing damage to the engines. Thus high levels of readiness come at a cost.

The individual aircraft within an asset have a fixed internal fuel tank and an optional external fuel tank. The aircraft have missile slots under their wings and bomb slots under their belly. Each slot can contain one weapon. Other specialized equipment is also possible to attach to the missile or bomb slots. Adding the external fuel tank is a trade-off, where the fuel capacity is increased, but he external fuel tank consumes one bomb slot and also considerably weakens the physical flight performance of the aircraft. The weakened flight performance and the reduced weapons load means that external fuel tanks are not used in air-to-air operations requiring agile maneuvering. In addition to the missile and bomb slots, the aircraft may carry some decoy weapons for deceiving enemy missiles or radars. Heat seeking missiles may be deceived by, e.g., torches.

The aircraft consume fuel and the fuel consumption is dependent on many factors: flight altitude, flight speed, the load of the aircraft, descent or climb, and the type of maneuvers the aircraft has to perform during battle engagement or other mission operations. The load of the aircraft means both the total weight of the weapons and fuel as well as the shape and dimension of the external weapons or fuel tank. The shape and dimensions of external weapons or fuel tank accounts for increased air resistance, which in turn results in increased fuel consumption. Flights to or from an operation are flown in cruising speed and these flight transitions consume a different amount of fuel than what is consumed at the operation itself. A fighter aircraft may use burst of after-burn, that give the aircraft a significant amount of extra thrust temporarily at the cost of rapid fuel consumption. These after-burn thrusts may partially account for a higher fuel consumption at the operation, compared to cruising flight.

Depending on the type of mission that the asset is performing, they spend a number of weapons. One of the main purposes of a military fighter aircraft is to launch missiles or to 'drop' bombs. Modern bombs are guided gliding bombs and they do not fall ballistically, hence 'dropping' in quotation marks. Weapons can only be loaded on the ground in a plateau of a base. Refueling is technically possible to do in mid-air from a tanker aircraft, but tanker aircraft need to be kept well within friendly territory. Hence, even in the case of mid-air refueling, the fighter aircraft need to retreat from the operation area or the engagement frontier and fly to the tanker, similar as they would retreat to a base.

Refueling can be done in a base as a normal refuel, where the aircraft engine is turned off and a mechanic has the chance to perform a standard check on the aircraft, while other ground crew reload weapons onto the aircraft. Alternatively refueling can be done as *hot refuel*, where the aircraft engines are not turned off. Reloading weapons is not done during hot refuel and there may be technical limitations as to how long the aircraft engine is allowed to run idle on ground. Hot refuel is much faster than normal refuel, but due to the aforementioned limitations, the aircraft needs to undergo normal refuel after one to three consecutive hot refuels.

The missiles are advanced and studying the launch strategy of missiles is a complex problem of its own. Missiles can be divided into Air-to-Air missiles (A/A) or Air-to-Ground missiles (A/G) depending on whether they have been designed for targets in the air or on the ground. The missiles are autonomous and can be launched on a fire-and-forget [17] basis, where the missile is capable of tracking its target and guiding itself to the target independently without further actions from the pilot. The flying speed of the missile is considerably greater than that of the aircraft. The missile reaches its target far ahead of the aircraft, and this forms an missile envelope for the aircraft. The missile envelope is visualized in Figure 2.2.

2.3.2 Base

Each base has a runway, plateaus, and taxiways that connect the plateaus to each other and to the runway. A plateau is a designated place where the aircraft can be held within the base and one plateau can only hold a limited number of aircraft at a time. Placing aircraft elsewhere than on a plateau, in the middle of the runway for example, will paralyze the operation of the base. However, some airfields may have excessively long runways in which case the ends of the runway may be allocated as plateaus. An asset that is placed in alert may occupy the runway such that the asset is placed at the end of the runway, ready to take off. In this case, no taxiing time is spent. The internal structure of an airbase is presented in Figure 2.1.

At some special bases, like aircraft carriers, there may not be any taxiways, as the plateaus are situated directly adjacent to the runway. At a normal land base, the aircraft move on the taxiways with their own engines, but at a slow speed, comparable to the speed of a heavy land vehicle. A taxiway may not necessarily connect to the end of the runway, and in the case of the taxiway connecting to the middle of the runway, the aircraft needs to taxi on the runway, adding to the total taxiing time and distance. In any case, if the asset is leaving from a plateau, there will always be at least a small delay for the aircraft to move to the end of the runway, nose pointed forward and the pilot being ready for take-off.

Only one way and ‘one lane’ traffic is allowed at a time on the taxiways and on the runway. Aircraft cannot pass each other on the taxiways or the runway and simultaneous take-off or landing is not possible. Only one aircraft at a time can use the runway and the individual aircraft within an asset have to take off consecutively one after another. In addition to the taxiing delay, there is also delay from this consecutive take-off, before all the aircraft within the asset are airborne.

Refueling and reloading weapons happens at a plateau, but not all plateaus have refueling and weapons reloading available. Some plateaus are simply places where the assets can be stowed away or where fully refueled assets may wait in standby. At some plateaus hot refueling is possible and at others not. In practice, nearly all the vehicles and equipment used in refueling or weapons loading is mobile and virtually any plateau can be equipped with refueling, hot refueling or weapons loading capability. It is a question of which plateaus have been prepared for each task. The available equipment is limited, so not all plateaus can offer every service.

2.3.3 Restricted Operating Zone

The flying routes of the assets may be altered by restricted operating zones. ROZs are areas through which the aircraft may not fly. This may be due to enemy radar, anti-aircraft weapons or for some geopolitical reasons. The assets need to go around the restricted operating zones which increases the flying distance. In the case of measuring the defense capability by force fulfillment, the effect of the ROZ is reduced to only a longer flying distance between a base and an operation area. When assessing the engagement frontier, the ROZ may greatly influence the achieved shape and location of the frontier.

2.3.4 Operation Area

An operation area is some area of airspace where the assets patrol and operate. Each operation area requires a certain number of assets to be present at certain times. These requirements are a function of time and the requirement may be zero at some times. Different operation areas may have different priorities such that force is primarily allocated to one operation area and only the idle or unused assets are then allocated to the secondary operation area. The number of different priorities is unlimited but in practice usually only primary and secondary priorities are used.

The force fulfillment is considered adequate when all operation areas have the required number of assets present at all times. A typical time span for the operation is several hours or more. One asset can stay at the operation area only for as long as it has sufficient amount of fuel remaining. When running low on fuel, the asset needs to return back to a base and a plateau that offers refueling. When one asset departs from the operation area, another asset should arrive to keep the force requirement fulfilled. At the operation area the assets' fuel consumption is dependent on the type of operation performed and especially whether after burn is used. If the asset engages in battle, missiles and decoy weapons may be used as well.

2.3.5 Advance Warning

When determining the engagement frontier, the advance warning plays a crucial role. The advance warning is the moment when the Blue side first becomes aware of any activity by the Red side that needs to be reacted upon. The advance warning may at best be before the Red assets have taken off. This would require a scout being positioned inside the premises of the Red base, and the scout would report on pilots boarding the aircraft.

A more common way of getting the advance warning is by radars. The radars have some view of sight and once an hostile aircraft comes into the view of sight, the advance warning is triggered. Gaining the advance warning by radar forms some frontier within which the Red assets may be detected. Such a frontier is called the *reconnaissance alert line* and it is presented in Figure 2.2.

3. Simulation Model for Force Fulfillment

The real life situation described in Chapter 2 is simplified into a simulation model. The model includes an allocation algorithm that uses a heuristic for controlling the use of ground resources and for allocating the aircraft to the operation areas. The allocation algorithm presents only one solution to each scenario, and the solutions are transparent and easy to justify. This makes what-if analysis easy, e.g., comparing the difference in air defense capability if the assets are equipped with an external fuel tank. The allocation algorithm includes a heuristic which has the common Earliest Due Date First (EDDF)-rule as its founding principle [18, p. 432]. The heuristic is also responsible for allocating the maintenance resources.

The solution produced by the allocation algorithm consists of a detailed master schedule which contains the flight and maintenance schedules for each asset. The master schedule also includes the allocations of the assets by determining which asset goes to which operation area and when. Other schedules can be derived from the master schedule, e.g., an aircraft maintenance schedule for each airbase. Also the tracking of fuel consumption at any given base can be done on the basis of the master schedule. Moreover, deriving a schedule for assets at the operation area gives the opportunity to calculate different measures for how well the force fulfillment was achieved. The measures defined in this thesis are *continuous force*, *full force* and *average force*. These measures are described in detail in Section 3.3.

With the simulation model, it becomes possible to analyze how the air and ground resources and the deployment of the aircraft affect the air defense capability. The most critical bases can be identified and extra maintenance resources can be allocated to those bases. Also a more holistic analysis is possible, e.g., analyzing the geographical coverage of the air defense capability rather than that of dedicated operation areas. The model is intended to be used for large scale planning rather than planning operations in detail; the schedule that is produced is not meant to be used as such. Therefore, it is justified to use a simple heuristic approach that is fast, easy to understand and that yields ‘good enough’ results.

The simulation takes place in a two dimensional Cartesian coordinate system. This means that factors such as the curvature of the earth is neglected. Time is an important element and it is measured on a continuous scale, but the parametrization is done using full minutes.

Table 3.1. *Parameters of the simulation model for force fulfillment.*

- **Asset**
  - Fuel capacity
  - Speed of cruising flight
  - Fuel consumption in cruise flight
  - Take-off and landing delay
  - Weapons capacity per type of weapon
- **Base**
  - Coordinates
  - Plateaus
  - **Plateau**
    - Refueling time
    - Hot refueling ability
    - Hot refueling time
    - Weapons reloading ability
    - Taxiing time
- **Operation Area**
  - Coordinates
  - Minimum staying time
  - Fuel consumption per asset
  - Weapons consumption per asset per weapon
  - Priority
  - Operation points
  - **Operation Point**
    - Activation time
    - Deactivation time
- **Restricted Operating Zone**
  - Coordinates of polygon vertices

3.1 Model Components

The parameters of the simulation model are listed in table 3.1. The components of assets, bases, operation areas and restricted operating zones are described in detail in the following subsections.

3.1.1 Assets

Assets are described as an atomic unit. The asset is never broken down into single individual aircraft and the consumption of fuel and weapons is for the whole asset. The asset has a constant flying speed and a constant rate of fuel consumption in cruising. Effects of ascending or descending are not considered. While at the operation area the asset has a constant fuel consumption that differs from the consumption during cruising. Weapons may be used at the operation area and weapons usage is given as a fixed amount of weapons per visit to the operation area. The asset has a fuel capacity and a maximum weapon carrying capability per type of weapon. The fact that assets will not fly their fuel tanks completely empty in real life has to be taken into account in determining the assets fuel capacity. The fuel capacity is the maximum amount of fuel that the asset can safely use, thus the amount of fuel spared for safe landing needs to be subtracted from the assets absolute capacity of fuel tanks. The asset is always fully refueled when refueled. No consecutive hot refuels are allowed; the asset has to undergo normal refuel at least every other time.

If an asset has used its weapons such that there is some operation area which consumes more weapons than what the asset has left, then the asset needs to go to a plateau that provides weapons. The maximum length of stay at the operation area is determined by the amount of fuel remaining. When the asset has just the amount of fuel remaining that is needed for the flight back to base, it must leave the operation area.

A take-off delay is added to the flight time between the base and the operation area. The take-off delay compensates for the aircraft taking off one at a time, accelerating and climbing to flight altitude. Flight altitude is otherwise ignored. A fixed landing delay is added to compensate for the limitations of the runway and the slowed down speed of approach. As with taking off, only one aircraft at a time can use the runway and thus the individual aircraft land one after another. This landing delay is assumed to be the same as the take off delay, so a combined take off and landing delay is used. The phases of one flight operation for an asset are illustrated in Figure 3.1.

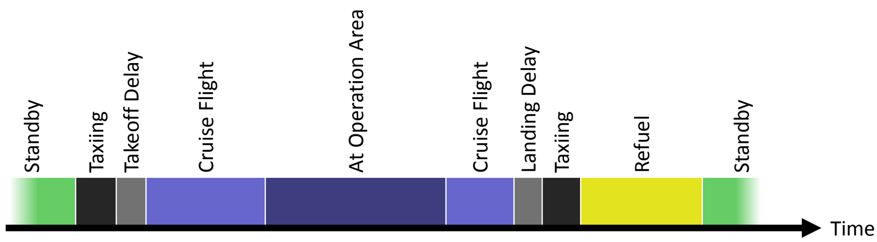

Figure 3.1. *The phases of one flight operation for an asset. The taxiing time is a parameter of a plateau at a base. The takeoff and landing delays are the same and are a parameter of an asset. The speed of cruising flight is a parameter of an asset, and the time spent depends directly on the distance flown. The fuel consumption at the operation area is a parameter of the operation area, and fuel is spent such that there is just enough for safe return to base. The refuel time is a parameter of a plateau at a base.*

3.1.2 Bases

The base consists of a runway, taxiways and plateaus. The base is represented as a point on the coordinate axes. The capacity of the runway is accounted for only in the take off and landing delay of the assets. The simultaneous landing of two individual assets is not restricted in the model. The same goes for the taxiways - the simultaneous use of the taxiway by multiple assets is not restricted.

Each base has a number of plateaus, and each plateau can hold exactly one asset. Each plateau offers the capability for normal refuel. Hot refuel and weapons reload are optional. The duration of the normal refueling, hot refuel and weapons reload is assumed to be fixed and is a property of the individual plateau. The ground crew is not taken into account but each plateau is assumed to be able to offer its service whenever an asset arrives at that plateau. Taxiways are not modeled as such, and only the approximated taxiing time is a parameter of the model. The taxiing time to a plateau is a property of that individual plateau.

3.1.3 Operation Areas

The operation areas are broken down into one or more fragments which each require one asset present for a single continuous period of time. These fragments are referred to as *operation points*. The splitting of an operation area into operation points is illustrated in Figure 3.2. The operation area is defined as a point on the coordinate axes. Operation areas have a minimum length of stay so that if the asset is not able to stay for the minimum time, then it is not allocated to the operation area at all. The assets' fuel consumption at the operation area is defined per operation area per asset. This fuel consumption is different from the assets' fuel consumption in cruise flight. The total requirement of assets at an operation area is defined as a time-dependent staircase function, as seen in Figure 3.2. The operation points are a reduction of that stair function and each operation point has a window of activation. The assets are required to be present from the beginning of the activation time to the end of the deactivation time.

The operation points inherit their priority from the operation area. Hence, all the operation points that are descendants of the same operation area will have equal priorities but operation points from two different operation areas may have different priorities. Priorities are treated such that any operation points with the highest level priority are treated exactly as described in the allocation heuristic in Section 3.2.3. Lower level priorities are treated iteratively so that operation points with a higher priority are treated before any operation points with a lower level priority.

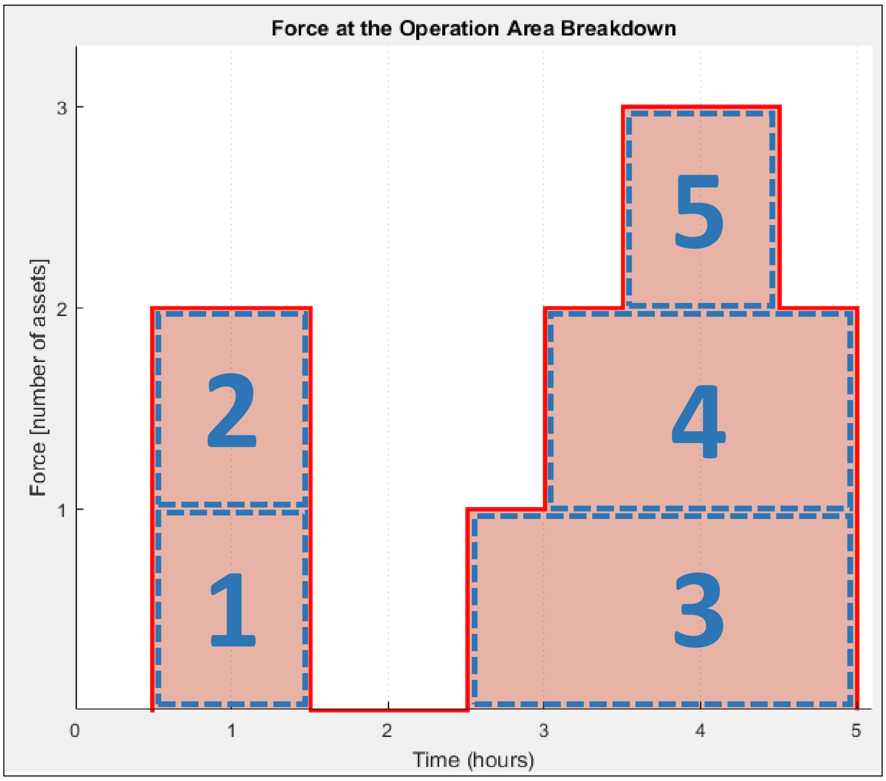

Figure 3.2. *Breakdown of an operation area into operation points. In this example, an operation area with a varying and discontinuous force requirement is split into five operation points. An operation point is always continuous and requires exactly one asset present.*

Secondary and lower level priorities have an extra limitation in the force allocation. If an individual asset's schedule is appended, in other words the asset has completed all the tasks allocated to it so far, then any further allocation will not interfere with the earlier schedule, and the extra limitations do not apply. This is the case in Example 5.1.3 in Figures 5.6 and 5.7, where asset A#5 has an allocation to OA#2 appended right after the allocation to OA#1. In the example, by chance, the asset returns back to the base from which it started off but it is not restricted to do so by the heuristic. However, if the asset has a long period of idle time before it is due for an prioritized operation area, as asset A#5 in Example 5.1.3 in Figure 5.6, then limitations on allocations to the lower level operation area apply.

In Figure 5.7, the chronologically first allocation of A#5 is done so that the asset has already been fixed to OA#1 on the higher priority run of the algorithm. In this case, extra limitations apply. The asset must return to exactly the same base and the same plateau, and if the asset was hot refueled in the original schedule, then hot refueling will not be allowed in order not to violate the rule of not having two consecutive hot refuels. It is also clear that the asset must complete the whole allocation cycle of cruise - operation point - cruise - refueling, within the dedicated time window.

The arriving and departing as well as the maintenance of the assets are presented in Figure 3.3. These just-in-time swaps do not happen at predefined intervals or at the same time for all assets. If it is not possible to get an asset to arrive when one asset is departing from the operation area, then a deficit is formed in the force requirement of that operation area and the force fulfillment is weakened.

Figure 3.3. *Example of a schedule. The operation area has been reduced into two operation points. The colors are the same as in Figure 3.1, except that the take off and landing delays have been embedded into the cruise flight. At the operation points green marks an asset present and pale red marks a deficit.*

3.1.4 Restricted Operating Zones

Restricted operating zones are polygons or polylines, whose number of vertices is not restricted. The shape of the ROZ is not restricted. The enclosed polygons or the polylines act the same; assets flight path may not intersect any part of the ROZ. The ROZ only affects the distance that an asset needs to travel between a base and an operation area. The ROZs are thus reduced to just an increase in flight distances and they result in no added complexity in the rest of the simulation. The distances between bases and operation areas can be solved using network optimization [15] and the result acts the same as if simply the straight line distances were longer.

The modeling of the ROZ is described in more detail in Subsection 4.1.4. Unlike in the simulation model for the engagement frontier discussed in Section 4, only a custom graph is formed which contains nodes for the bases, the operation areas and a node for each vertex of each ROZ. The connections in the graph are built with a ‘brute force’ approach, since the total number of nodes will be significantly lower than in the simulation model for the engagement frontier. All pairs of nodes are investigated whether a straight line connection exists in between them or whether they cut through a ROZ at any point. The connections of ROZ vertices may or may not self intersect the ROZ that they belong to or intersect another ROZ. Traveling along the edge of a ROZ is allowed, i.e., adjacent ROZ vertices are always connected.

Once the graph defining all the allowed straight line connections is built, the minimum distances from each base to each operation area are calculated with the same external algorithm as in Section 4.3. The distances are stored and in the rest of the simulation only the value of the distance is relevant.

3.2 Allocation Algorithm

The overall allocation schedule of assets to operation areas is constructed with an algorithm that repeats a subproblem of allocating one asset to one operation point. The algorithm is continued until there are no assets that could be further allocated to any operation points. The operation points are treated in order of their priorities as discussed in Section 3.1.3, with the extra limitations applying that any formerly constructed schedule must not be broken. The algorithm is responsible for feeding the heuristic with a list of operation points that are to be allocated next, according to their priorities.

The subproblem of allocating one asset to one operation point is solved using a novel heuristic. The heuristic is deterministic and myopic. Deterministic, because it will always make exactly the same decisions starting from the same initial conditions. Myopic, because it does not take into account any possible consequences of an allocation. The heuristic's logic is founded on the Earliest Due Date First (EDDF) -principle [18, p. 432]. Rationale in the EDDF principle is a greedy heuristic that tries to extend the period of continuous full force for as long as possible. By allocating a feasible asset to the operation point which has the earliest departure time, the period of continuous full force gets extended. If such an extension can be carried along until the deactivation of the operation area, a feasible solution of the force fulfillment problem has been found. The algorithm and the heuristic is presented graphically in Figure 3.4. The heuristic consists of four stages:

- Choosing an operation point where to allocate the next asset.
- Choosing an asset which is allocated to that operation point.
- Maximizing the length of stay of the the asset at the operation point.
- Choosing a return base and allocating the asset to a plateau for refueling.

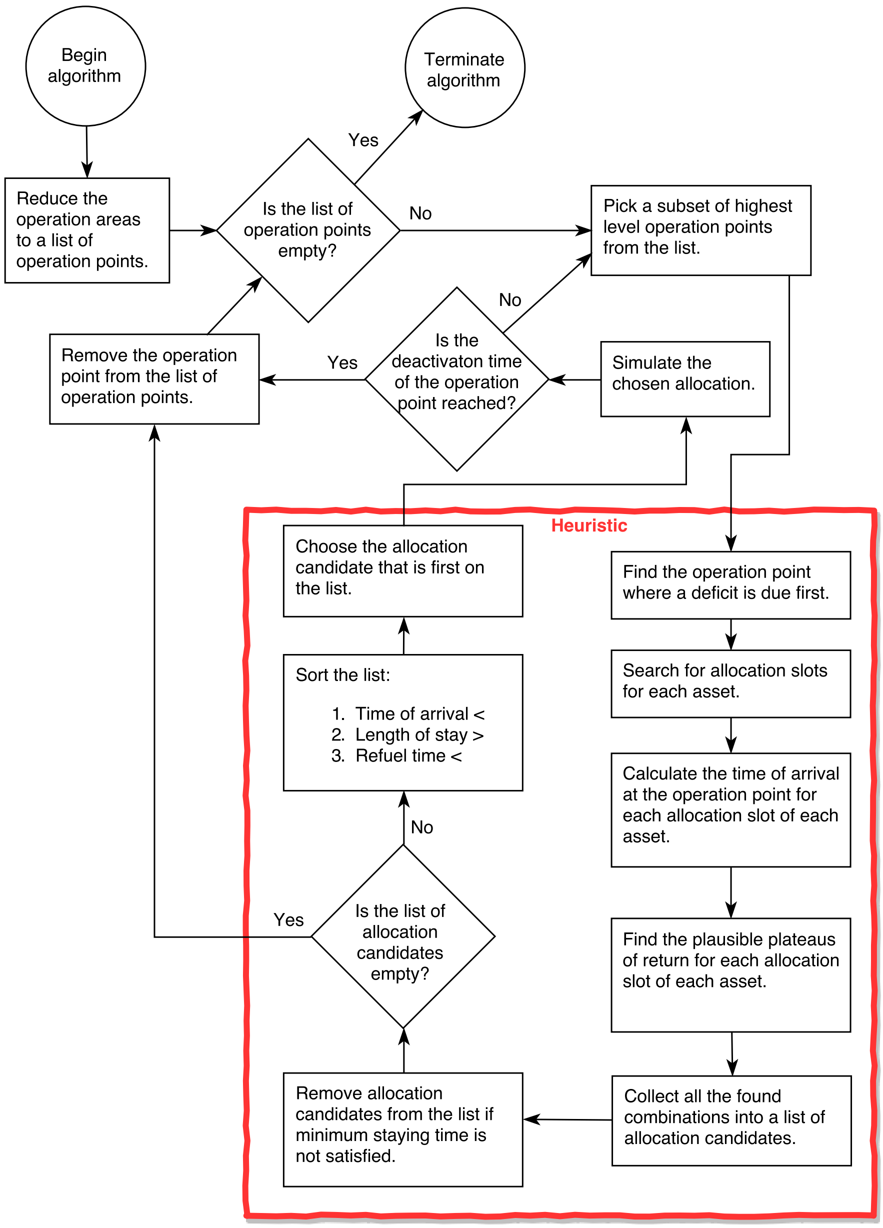

Figure 3.4. *A flowchart of the algorithm and the heuristic. The heuristic is iterated within the algorithm. The '<' and '>' characters in the sorting refer to sorting smallest first and greatest first, respectively.*

3.2.1 Principles of Heuristic

The heuristic seeks for the operation point which has the earliest due date. In this case, the earliest due date is the earliest departure of an asset. This operation point is chosen for the next allocation of an asset. Then, the heuristic tries to allocate an asset to that operation point. An allocation consists of choosing an asset, an allocation slot for that asset, and a plateau of return. The heuristic does a 'brute-force' comparison of all the possible combinations of assets, allocation slots and plateaus of return and then chooses the best rated of these. The composition of possible allocation combinations is presented in Figure 3.5.

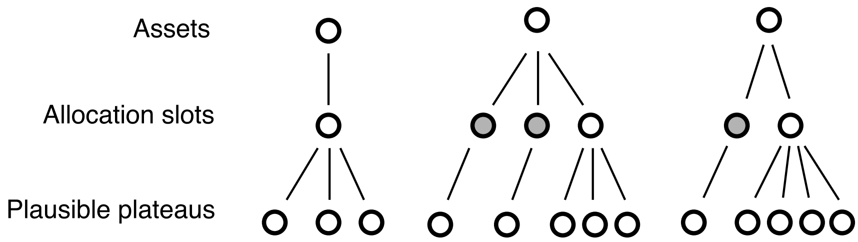

Figure 3.5. *An example of possible allocation combinations. The gray filled allocation slots are* interior slots *and hence they have only one plausible plateau each.*

3.2.2 Allocation Slots

The allocation slot is either a period of long enough idle time in before a previously fixed allocation, or at the end of the allocation schedule of the asset. The allocation slots are illustrated in Figure 3.6. The allocation slots in before allocations that have been done on earlier iterations of the heuristic will be referred to as *interior slots*. The allocation slots that are at the end of the so far cumulated allocation schedule of the asset will be referred to as *exterior slots*. The interior slots are marked with a pale blue border in the figure and the exterior slots are marked with a purple border. The exterior slots have an ‘open end’, meaning that there are no limitations dictating when the asset should be returned and refueled. Many of the interior slots in Figure 3.6 are so small that an allocation will not be possible.

For an allocation to be possible into an interior slot, the slot needs to be large enough to fit a complete flight operation cycle of an asset from standby to standby with at least minimum staying time spent at the operation area. This complete cycle includes all the phases illustrated in Figure 3.1. In the case of an interior slot, the extra limitations discussed in Section 3.1.3 apply, and the asset is forced to return to the same base and the same plateau from where it left.

The primary criterion in selecting the asset and the allocation slot is the time when the asset can reach the operation point. The secondary criterion is the time that the asset is able to spend at the operation point. The third criterion is the refueling time at the plateau where the asset will be returning to. All the three criteria are strictly hierarchic so that the second criterion is only used if two or more assets are exactly equal according to the first criterion and third criterion is only used if the assets are exactly equal according to the first and second criteria.

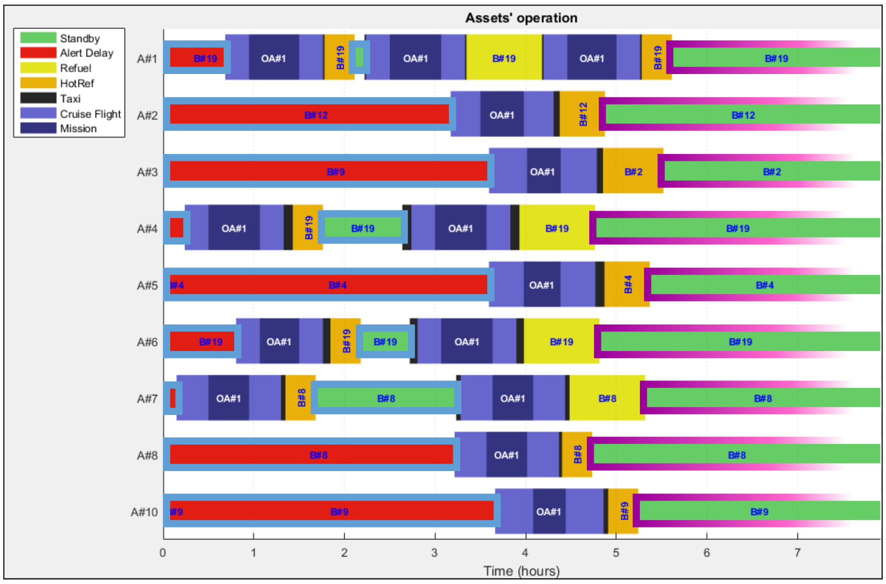

Figure 3.6. *An example of possible allocation slots. The* interior slots *are marked with a pale blue border and the* exterior slots *are marked with a purple border.*

The primary criterion aims at filling the force requirement without deficits, or minimizing the deficit caused by the current allocation. If there are no assets available that would reach the operation point on time, an asset is chosen which is able to arrive earliest to the operation point. The secondary criterion greedily maximizes the time spent at the operation point, without paying any attention to possible future deficits. The third criterion aims at effective use of the base's resources, using for example a plateau capable of hot refuel whenever possible.

Assets always stay at the operation area for as long as possible so that when the asset comes to land, it has no fuel to spare. This is again greedy; returning an asset for refueling early may be forethought and perhaps prevent a deficit from happening later on. The heuristic tries to extend the period of continuous force at any cost for the current allocation.

The maximum time that the asset is able to spend at the operation point is dependent on the base to where the asset will return. The heuristic goes through all the possible return bases and effectively the rule for choosing the 'best' return base is simple - choose the one nearest to the operation point. The heuristic does a 'brute force' search within each iteration, going through all the possible combinations of allocation slots, assets and return plateaus and builds a list of allocation candidates.

For each return base, the return time can be calculated explicitly and likewise the arrival time to a plateau at the base can be calculated. When the return time is known, the vacancy of the plateau can be determined. If the plateau is vacant for some allocation combination of allocation slot and asset, this allocation candidate is added to the list of allocation candidates.

There is a special phenomenon of plateaus getting 'reserved' for the asset that initially started off from the plateau. This unnecessary reservation of the plateaus may lead to a stall position where the asset and the maintenance place get paired up so that each asset uses the same plateau each time and that the plateau serves only the one asset that it is 'paired with'. For the sake of example, assume that the number of plateaus is exactly equal to the number of assets. In this case, whenever an asset will be returning for refueling then all the other plateaus will be reserved. This is because the allocation algorithm works such that it always ends an iteration to the phase where the asset will be refueled. The only plateau appearing to be vacant will be the plateau where the asset initially started off from.

It is however predictable and expected that any asset will not stay for long at the plateau after maintenance has completed. To counteract this problem, a virtual queue has been introduced. A virtual queue is an allocation option to where an asset can return in a base instead of a real plateau. Only one asset at a time may be at the queue at each base. The asset in the queue is allocated to a plateau immediately when one becomes free. If the freeing of a plateau happens chronologically earlier than when the asset entered the queue, then the actual queuing time for the asset will be zero. This virtual queue allows for an asset to come to land without certain knowledge when a plateau will become vacant. This ensures that the assets move freely and efficiently from one airbase to another, making the use of plateaus more efficient.

The asset will be refueled at the plateau and its weapons stock will be replenished, if the plateau has weapons available. The plateau needs to be vacant at the time when the asset would taxi to the plateau. If the asset has consumed weapons so that there exists an operation point to which the asset would not have enough weapons remaining, then the plateau needs to provide weapons. A plateau is a *plausible plateau* if it is vacant and if it satisfies the possible need for weapons refill. All the combinations of allocation slot, asset and plausible plateaus are added to a list of allocation candidates. It may turn out that the 'best' asset would, e.g., require weapons but any plausible plateau was not found for the asset and hence no allocation candidate is added to the list for the 'best' asset.

Any allocation candidates for which the minimum staying time is not met are removed from the list of allocation candidates. If the list is completely empty, it means that no plausible allocation combination was found and the operation point will be left with a deficit. The list of allocation candidates is sorted according to the first, second and third criteria. The first criterion is the time of arrival, that at best is the same as the previous asset departing. The second criterion is the maximum length of stay that which essentially favors returning to the nearest possible base. The third criterion is the refuel time at the plateau of return which is the only place where the heuristic tries to improve the odds of finding a good allocation on the next iteration. A faster refueling makes the asset ready for next flight operation earlier. The heuristic sorts all the found allocation candidates in order according to the three criteria and chooses the allocation combination that came first in the sort.

3.2.3 Heuristic

The steps of the heuristic in full are as follows:

1. Find an operation point which has due the earliest deficit. The earliest deficit is determined by the time that the last asset is leaving the operation area.

2. For each asset search for allocation slots. An allocation slot is either a long enough period of standby in between two consecutive allocations of an asset or the end of the allocation schedule for the asset. The former is referred to as an interior slot and the latter an exterior slot.

3. Calculate the time of arrival at the operation point for each allocation slot of each asset.

   (a) If the asset can arrive on time, then the time of arrival is the same as the earliest deficit in Step 1.

   (b) If the asset cannot arrive on time, then the arrival time will be the earliest possible time of arrival for the asset.

4. For each allocation slot of each asset, find the plausible plateaus of return. The time of stay at the operation point is calculated by subtracting the amount of fuel spent in flight to and from the operation area from the asset's fuel capacity and dividing by the fuel consumption at the operation area. Once the maximum stay at the operation point is known, the time of arrival at the plausible plateau of return can be determined by summing together the time of arrival from Step 3, the maximum stay, the return flight and the taxiing. The plausible plateaus of return are:

(a) For an interior slot, only the same plateau from where the asset left from. If the asset needs weapon refill and the plateau does not offer it, then no plausible plateau exists.

(b) For an exterior slot, any plateau that is vacant at the time of taxiing to the plateau and that offers weapons refill, if the asset needs weapons refill. For each base include the virtual queue as an plausible 'plateau', if any plateau at the base satisfies the possible need for weapons refill.

5. Collect a list of all the allocation slots for all the assets for which a plausible plateau of return was found. There may be more than one plausible plateau of return for an allocation slot for an asset and these are all collected in the list.

6. Remove any entries in the list for which the minimum staying time is not satisfied.

7. Sort the list in order according to the hierarchic sorting rules. The hierarchic sorting means that consecutive rules are only used if the previous rule resulted in a tie. Having a tie in sorting happens, e.g., when more than one asset is able to arrive on time at the operation point. The sorting rules are:

   (a) Time of arrival at the operation point from Step 3, sorted earliest to latest.

   (b) The maximum stay at the operation point from Step 4, sorted longest to shortest.

   (c) The refuel time at the plateau for the asset, sorted fastest to slowest.

8. Choose the combination of allocation slot, asset and plateau of return that appears at the top of the list.

3.3 Measures for Achieved Force Fulfillment

Three different measures for the achieved force fulfillment have been recognized. In all the three measures, time periods where the force requirement is zero are omitted. All the three measures can be expressed as such or as a percentage 0 - 100 % of the maximum possible value representing a perfectly fulfilled force requirement.

Continuous Force measures for how long can the force requirement be fully met before the first deficit. In Figure 3.7, the continuous force is presented as the orange line. The continuous force is broken at the first instant when there is deficit in the achieved number of assets present at the operation area. Should it happen that the very first asset arrives late at the operation area, the continuous force would be trivially zero.

Full Force measures all the periods of time when the force requirement at the operation area is fully met. In contrast to the continuous force measure, the full force measure does not stop at the first deficit but rather continues on to sum up all the periods of time when the force requirement is fulfilled. The measure is presented in Figure 3.7 as the addition of the orange line and the yellow lines.

Average Force measures the average of all the achieved force, even if the force requirement is not fully met. The average force is calculated by dividing the 'blue' surface area by the total time span of non-zero force requirement.

In the example in Figure 3.7, the total time span of force requirements is $60 + 150 = 210$ minutes. The continuous force is the time from the first instant of force requirement (time 30.0 minutes in Figure 3.7) to the time instant of the first deficit (time 81.0 minutes in Figure 3.7), i.e., 51.0 minutes. The maximum value in the example is the total time of 210.0 minutes. The relative percentage value is $51.0/210.0 \approx 24.29\,\%$.

The full force continues on from the continuous force adding to the sum all the other time spans where the required force is fulfilled. The time period $1\frac{1}{2}$ - $2\frac{1}{2}$ hours where the force requirement is zero in Figure 3.7 is ignored. Thus, the full force in the example is calculated as $51.0 + 30.0 + 30.0 + 32.3 + 20.0 = 163.3$ minutes. As with the continuous force, the maximum value is the total time of 210.0 minutes and the relative percentage value is $163.3/210.0 \approx 77.76\,\%$.

The average force in Figure 3.7 is calculated by summing together the surface areas of the areas a - g. Each individual area is a rectangle whose sides are time by number of assets. The number of assets is an integer and each area a - g represents the force of exactly one asset. Hence, the calculation reduces to summing up all the widths of the areas a - g. The unit of the surface area is 'asset hours'. The average force is acquired by dividing the total 'asset hours' by the total time span. In Figure 3.7, the total 'asset hours' is $(60.0+51.0+106.9+75.6+32.3+28.0+23.0) = 376.8$, which divided by the total time yields the average force value of $376.8/210 \approx 1.79$. The maximum value for 'asset hours' in the example is $2 * 60 + 150 + 120 + 60 = 450$ and thus the

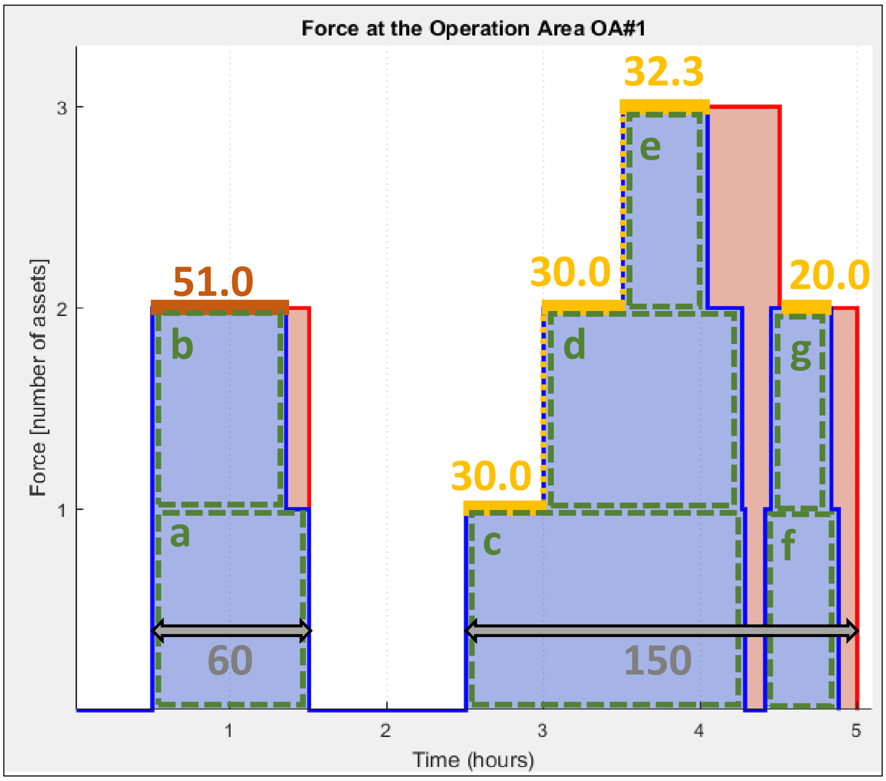

Figure 3.7. *The three measures of achieved force. Continuous force is* 51 *minutes in orange. Full force is* $30 + 30 + 32 + 20$ *minutes in yellow in addition to the continuous force, totaling at* 163 *minutes. Average force is the sum of surface areas* a *-* g *divided by the total time span. The total time span is* $60 + 150 = 210$ *minutes.*

Table 3.2. *Summary of the values of the measures of force in Figure 3.7.*

Continuous Force	51.0	24.29 %
Full Force	163.3	77.76 %
Average Force	1.79	83.73 %

relative percentage value of the achieved average force is $376.8/450 \approx 83.73\,\%$. The different measures and their relative percentage values are summarized in Table 3.2.

All the three measures treat the extreme points exactly the same. This means that if the achieved force at the operation area is zero at all times, then all the three measures give the value of zero and the relative value of 0 %. If at the other extreme the achieved force fulfills the required force at all times, all the three measures give the relative value of 100 %. The behavior between the extremes varies. The continuous force is the most sensitive measure, as it will get a value of zero if there is even an arbitrarily short deficit right at the beginning of the force requirement. The full force will always get a value that is at least equal to the continuous force, as

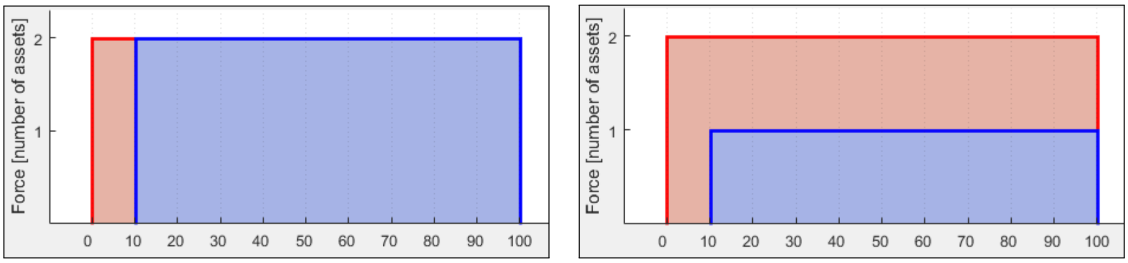

Figure 3.8. *The special cases where the measures of force fulfillment differ drastically. On the left, the values of continuous force (CF), full force (FF) and average force (AF) are CF 0 %, FF 90 % and AF 90 %. On the right, the values are CF 0 %, FF 0 % and AF 45 %.*

by definition the period of continuous force is included within the full force. The average force is the most robust of the measures, not getting 'deceived' by spreading the achieved force differently.

However, the usefulness of the measures is dependent on the real life scenario. It may be that some mission is considered a failure once the required force cannot be fulfilled any more. In this case, the continuous force will tell for what portion of time can the mission be sustained successfully. It may be that a mission is considered a failure whenever the achieved force drops below the required force. In this case, the full force will give a measure of what portion of the mission is a success and what portion a failure. If the momentary amount of force is irrelevant for whatever reason, the average force will give the most robust measure for the total 'integral' of the achieved force. The special cases where the measures differ significantly are presented in Figure 3.8.

3.4 Spatial Batch Run

Simulating the force fulfillment for one operation area can be extended to a batch run where a large spatial region is discretized to a uniform rectangular grid. The force fulfillment is then calculated for each point in the grid at a time, placing the operation area at that grid point. Here is where the three measures of force fulfillment come in useful. For a single operation area, it is possible to analyze the schedule yielded by the allocation algorithm, but for a relatively coarse grid of. e.g., 20 by 30 grid points, analyzing the resulting 600 schedules becomes unpractical. The measures of force fulfillment present the result with one single value and these values are easy to color code and present visually, giving the observer a holistic view of the potential force fulfillment for a large spatial region. It must be noted that the potential force fulfillment applies to only one operation area at one grid point at a time, not for the whole spatial region simultaneously. An example result of a spatial batch run is presented in Figure 3.9.

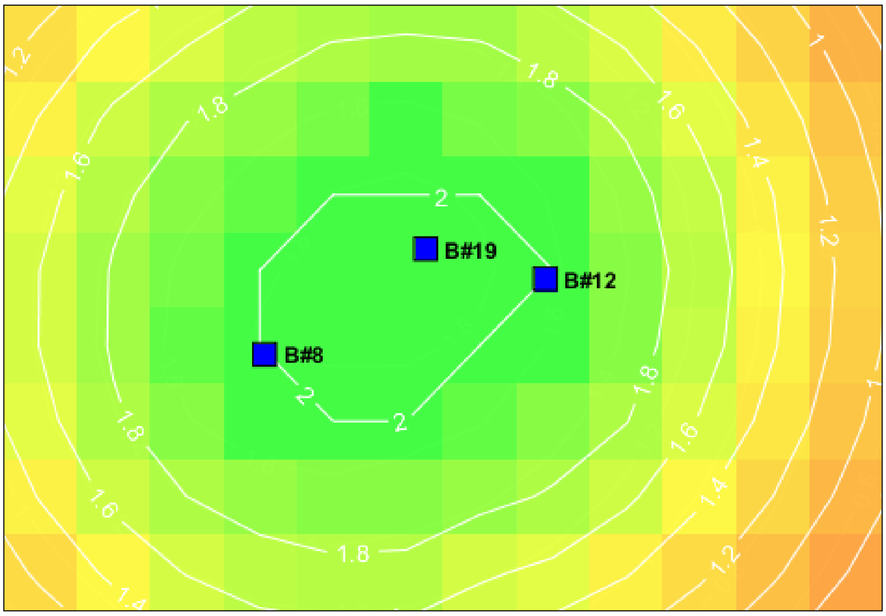

Figure 3.9. *A simple example of a spatial batch run. There are three bases and six assets and the operation area has a force requirement of two assets. The color scaled squares depict the values of the achieved average force. Contour lines are approximated by interpolation and drawn to clarify the achieved values and their approximated boundaries. The actual point by which the value of each square is calculated is situated at the center of the square.*

It is also possible to determine the excess defense capability by the measure of force fulfillment with the spatial batch run. The allocation algorithm can be run for a predetermined operation area and the spatial batch run can then be applied such that some of the resources have already been spent and fixed for the higher priority fixed operation area. Again the same three measures for force fulfillment are available to analyze the potential spatial force fulfillment.

4. Simulation Model for Engagement Frontier

The *Engagement Frontier* represents the anticipated level of where two opposing sets of assets engage each other. These opposing sets are referred to as Blue and Red. The Blue assets are reacting with defense against the hostile activity of the Red assets. The crucial factors determining the engagement frontier are the standby readiness of the Blue assets, the advance warning time, and the flying speed of the assets. The advance warning time is gained from reconnaissance or airspace monitoring. The advance warning time is the first instant of making an observation of activity by the Red assets.

The engagement frontier is considered a one flight job for each asset and circulation of assets is not considered as in the force fulfillment. However, the engagement frontier and force fulfillment are linked. If the bases are very far away from the desired engagement frontier and the assets are not able to reach the desired frontier in time, then Combat Air Patrol (CAP), might come into consideration, where assets are kept in 'standby' in mid-air.

A basic engagement frontier is presented in Figure 4.1, where the black dash-dotted line is the engagement frontier. The engagement frontier can also be calculated for the second assets, i.e., such that it is required that two assets from the same side are present. This 'second asset' engagement frontier is presented in Figure 4.2. The second asset engagement frontier can be solved for either of the two parties.

The assets also have a missile envelope that may or may not be taken into account in the engagement frontier. The missile envelopes can likewise be taken into account for each party individually, giving four different alternatives. An engagement frontier with a missile envelope taken into account for only the blue party is presented in Figure 4.3. Together there a total of sixteen possible choices for the engagement frontier, as clarified in Figure 2.3.

4.1 Model Components

The parameters of the simulation model used in assessing the engagement frontier are listed in Table 4.1. The components of assets, bases, advance warning time, restricted operating zones and 2D grid are described in detail in the following subsections. The simulation takes place in a two dimensional Cartesian coordinate system that is discretized to form the 2D grid.

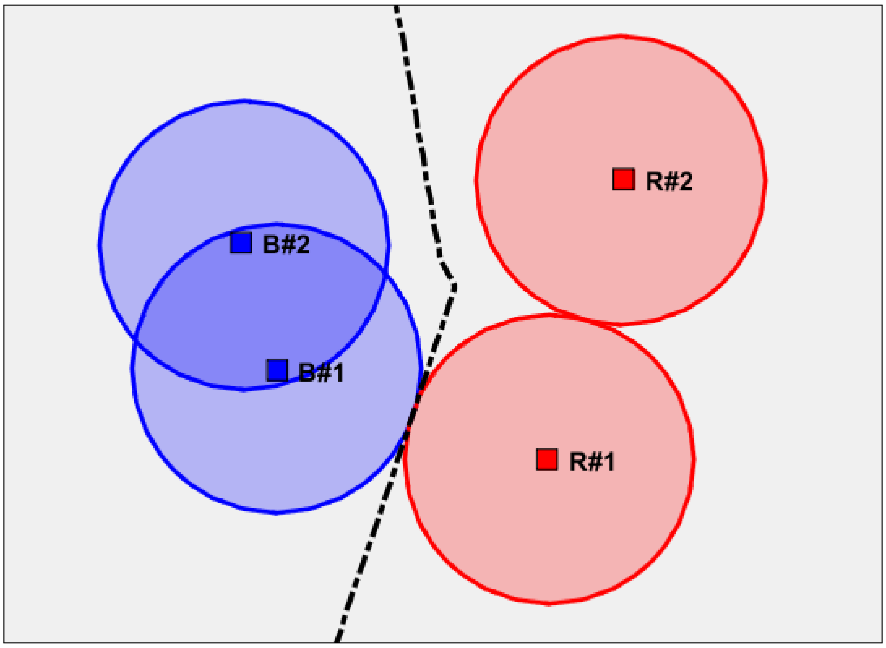

Figure 4.1. *A basic engagement frontier. The black dash-dotted line is the engagement frontier and the blue and red circles represent the reach of each asset at a specific moment in time. The moment of time is such that the earliest possible engagement has just happened.*

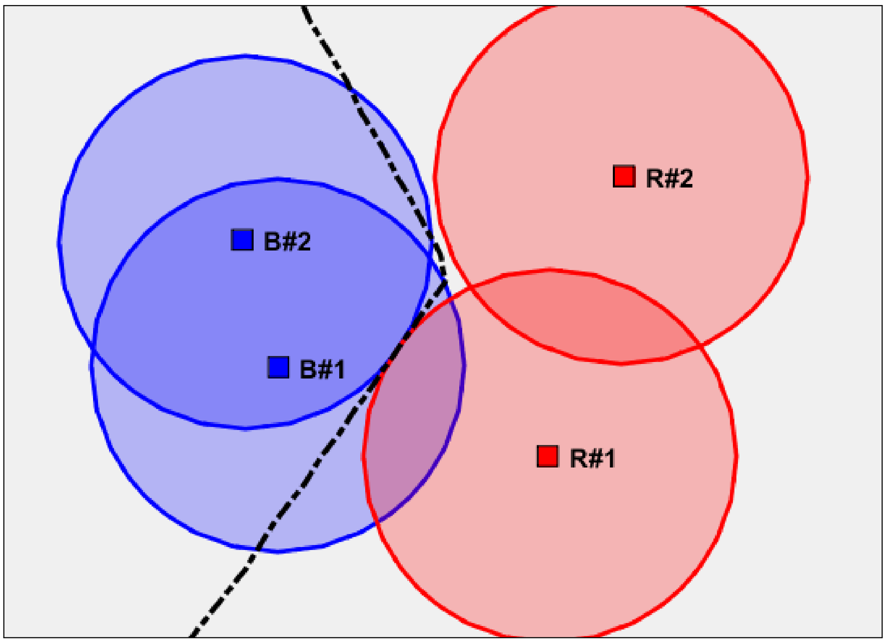

Figure 4.2. *An engagement frontier with the requirement of having two Blue assets.*

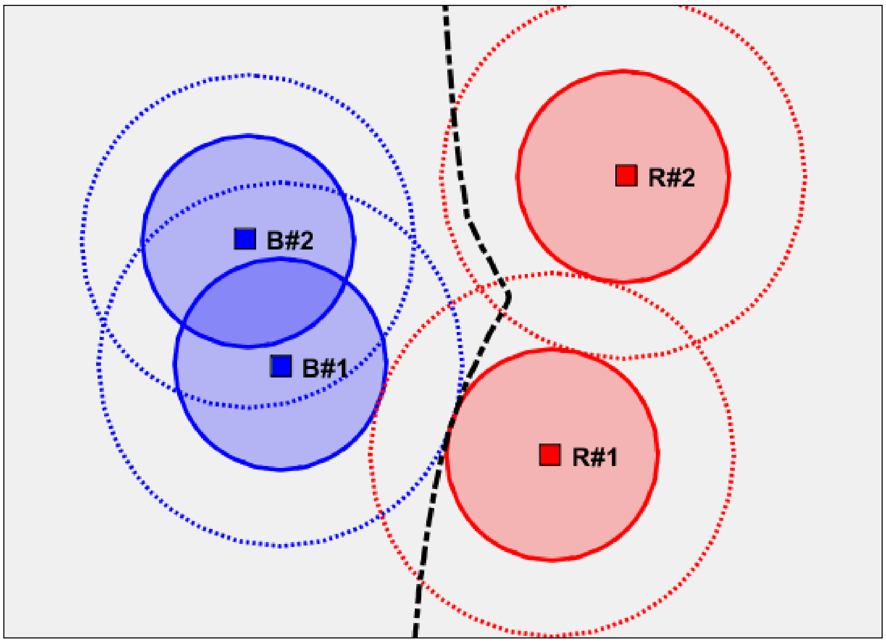

Figure 4.3. *An engagement frontier with a missile envelope for the Blue assets.*

Table 4.1. *Parameters of the simulation model for engagement frontier.*

Blue Asset
- Standby time
- Take-off delay
- Speed of flight
- Missile envelope

Red Asset
- Airborne instant
- Speed of flight
- Missile envelope

Base
- Coordinates

Advance Warning Time
- Instant of alert

Reconnaissance Alert Line
- Coordinates of line segment vertices

Restricted Operating Zone
- Coordinates of polygon vertices
- Parties under influence

2D Grid
- Adjacency matrix

4.1.1 Assets

Assets are described as an atomic unit. The asset is never broken down into single individual aircraft. The most important property of the asset is the standby time that it has been commanded. After the standby time, the take-off delay is added and after adding up the times of both, the asset ‘jumps’ into full speed flight directly on top of the base, with all weapons ready for launch. The take-off delay compensates for the aircraft taking off one at a time, accelerating and climbing to flight altitude.

The asset’s fuel capacity is assumed not to be a limiting factor and thus the fuel capacity and fuel consumption are ignored. The asset has a missile envelope that tells the distance between the asset and a missile at the moment that the missile would be able to strike the opposing aircraft. Launching a missile includes making an observation of an opposing asset, locking the missile to the target and launching the missile. At the time of launch, the distance between the opposing assets will be significantly greater than the distance of the missile envelope. The speed of the missile is considerably greater than the speed of either asset, but during the flight of the missile, the opposing asset flies closer and the missile launched asset continues flying in the direction of the missile. The missile envelope distance is thus the anticipated time to strike multiplied by the difference in speed of the missile and the missile launching asset.

The opposing assets are given an airborne instant. The airborne instant is an arbitrary moment of time when the Red asset will start traversing at full speed from the location of its departing base. The Red assets have likewise a missile envelope and a speed of flight.

4.1.2 Bases

A base in the engagement frontier simulation is essentially only a starting point for an asset. The limitations in the use of the runway, that only one aircraft at a time is able to use for take-off, is taken into account in the asset’s take-off delay. Taxiways are ignored and the standby time of the asset embeds any possible taxiing on the ground. A very short standby time essentially means that the asset would be situated at the end of the runway. The base is thus represented simply as a point on the coordinate axes. A base could also be used in the simulation to represent a CAP point by setting the assets standby time and take-off delay both to zero. Altitudes are not considered in the simulation model.

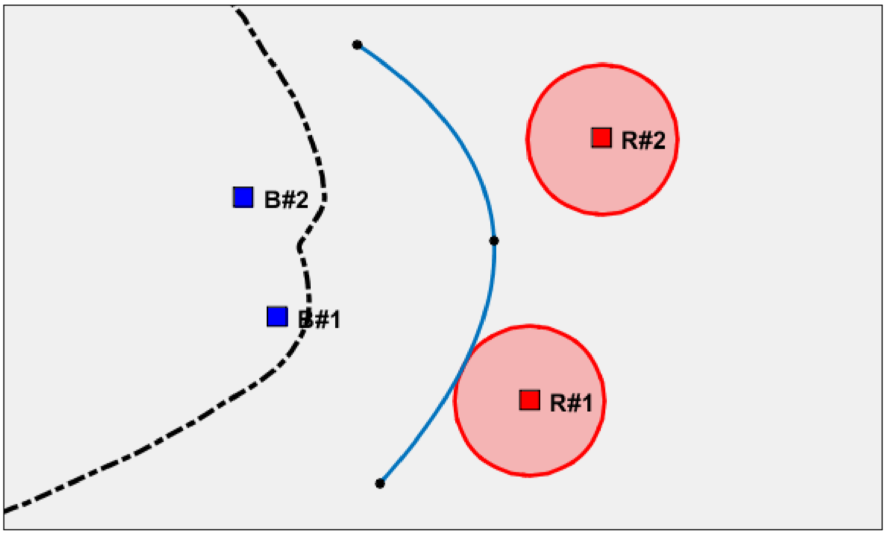

Figure 4.4. *A reconnaissance line that activates the blue assets is shown with a solid blue line. The Red asset R#1 has just reached the reconnaissance line and an advance warning is gained by the Blue side.*

4.1.3 Reconnaissance Alert Line

A reconnaissance alert line, as depicted in Figure 4.4, is a polyline that once cut by the opposing assets gives an advance warning to the defending assets. After the advance warning, the defending assets will spend their standby time plus their take off delay, before advancing at full speed from their starting bases. The reconnaissance alert line is always a polyline, composed of straight line segments joined together. However, the reconnaissance alert line can be defined as a spline going through dedicated points, but for matters of calculation, this spline is discretized into a hundred points between any two dedicated points.

The advance warning can also be set manually to simulate a situation where the advance warning is gained directly by detecting activity of opposing assets on ground. The reconnaissance alert line is only a convenience for setting the advance warning time.

4.1.4 Restricted Operating Zones

Restricted operating zones are polylines that may or may not be enclosed to form a polygon. If not enclosed, the ROZ will not form a 'zone', but rather a border that may not be crossed. As with the reconnaissance alert line, the ROZ lines may be defined as curves passing through designated points. The curves are formed as splines and discretized tightly enough to form a series of straight lines that seem curved. Different types and shapes of ROZ are presented in Figure 4.5. Assets going around a ROZ is visualized in Figure 4.6.

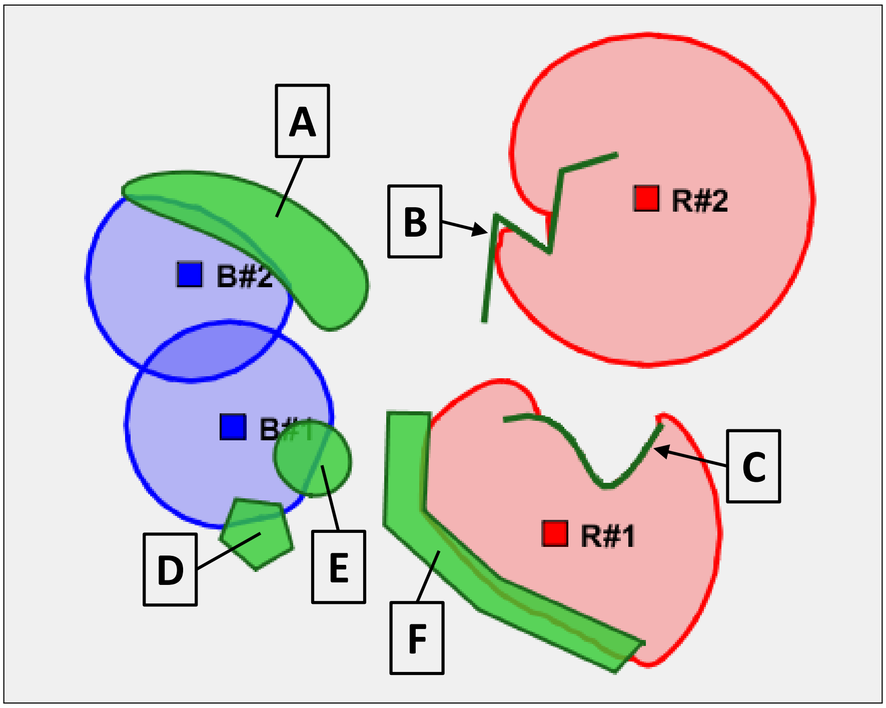

Figure 4.5. *Assets being limited by ROZs of various shapes and sizes. The assets' reach areas penetrate the enclosed ROZs for a clearer visualization. A is an enclosed ROZ formed of a spline passing through five dedicated points. B is a simple polyline with five vertices. C is a curve defined by four points. D is a simple pentagon. E is a circle or oval defined by a curve passing through five points. F is a polygon.*

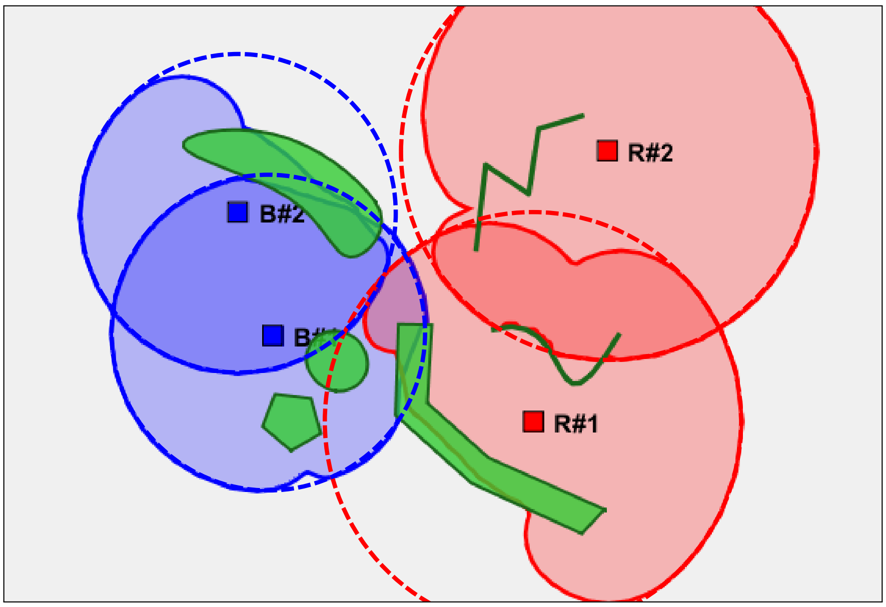

Figure 4.6. *Assets going around ROZ with a dotted ring indicating the advancement of each asset if advancing without ROZ.*

4.1.5 2D Grid

The ROZ complicates the calculation of the *reach areas* of the assets. The reach area of an asset is the enclosed spatial area that an asset is capable of reaching within a given time. The reach area of an asset is the elementary factor in determining the engagement frontier. Without the ROZ it would be possible to determine the engagement frontier using analytical methods. For example, two opposing assets flying at the exact same speed without any hindrances in flight paths form an engagement frontier that is a straight line. Flying at slightly different speeds the engagement frontier would be a symmetrically curved line.

Because of the increased complexity caused by the ROZ, the reach areas of assets are simulated with a tightly discretized evenly spaced 2D grid. The grid is formed to cover the whole area of observation. Each point in the grid thus represents a coordinate point, and all locations and movements of assets are described in this discretized grid. The grid is interpreted as an *undirected graph*. A graph is a structure that contains *nodes* connected by *edges*. For an undirected graph, the edges have no orientation, meaning that the connection is always a two way connection.

An example of a graph is presented in Figure 4.7. The edges are represented by an *adjacency matrix* where the nodes are collected as a list and the list of nodes is mapped against itself to obtain a square matrix with dimensions equal to the total number of nodes. The adjacency matrix for the example graph is presented in Figure 4.8.

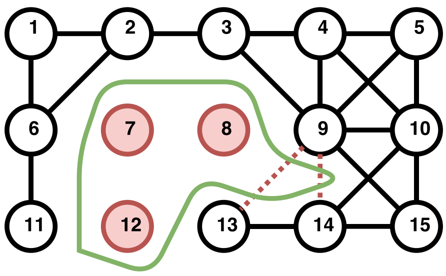

Figure 4.7. *An example of a simple undirected graph with 15 nodes and 18 edges. The nodes are the circles and the edges are the lines connecting the nodes. The green enclosed line is a ROZ and the nodes inside the ROZ have no edges. None of the edges may pass through a ROZ and the ROZ blocks two edges that are colored red and drawn with a dotted line.*

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	·	1	·	·	·	1	•	•	·	·	·	•	·	·	·
2	1	·	1	·	·	$\sqrt{2}$	•	•	·	·	·	•	·	·	·
3	·	1	·	1	·	·	•	•	$\sqrt{2}$	·	·	•	·	·	·
4	·	·	1	·	1	·	•	•	1	$\sqrt{2}$	·	•	·	·	·
5	·	·	·	1	·	·	•	•	$\sqrt{2}$	1	·	•	·	·	·
6	1	$\sqrt{2}$	·	·	·	·	•	•	·	·	1	•	·	·	·
7	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
8	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
9	·	·	$\sqrt{2}$	1	$\sqrt{2}$	·	•	•	·	1	·	•	·	·	$\sqrt{2}$
10	·	·	·	$\sqrt{2}$	1	·	•	•	1	·	·	•	·	$\sqrt{2}$	1
11	·	·	·	·	·	1	•	•	·	·	·	•	·	·	·
12	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
13	·	·	·	·	·	·	•	•	·	·	·	•	·	1	·
14	·	·	·	·	·	·	•	•	·	$\sqrt{2}$	·	•	1	·	1
15	·	·	·	·	·	·	•	•	$\sqrt{2}$	1	·	•	·	1	·

Figure 4.8. *The adjacency matrix for the example graph in Figure 4.7. Notice the symmetry of the matrix and the fact that nodes 7, 8 and 12 have no edges and the rows and columns referring to those nodes are empty. The rows and column of nodes 7, 8 and 12 are highlighted with red bullets.*

4.2 Forming 2D Grid

The 2D grid is interpreted as a graph. The connections in a graph are defined by an adjacency matrix. In the simulation, the connections to the neighboring nodes are as presented on the right hand side in Figure 4.9. The connections are two-way, meaning that the connection can be traversed both ways. The connections are designed such that the distribution of angular directions would be as even as possible. As seen in the figure, the red square has straight line connections to all the nodes on its perimeter. Inside the square however lie 16 nodes that cannot be reached by a straight line path. Leaving out straight connections to these nodes is a conscious decision, with the aim of keeping the model as simple as possible while keeping the angular distribution of the connections as even as possible.

Using the connections defined in Figure 4.9, a network of connections is achieved. This criss-crossing of connections is depicted in Figure 4.10 where the connections leaving from one node are drawn with dark blue, and all the connections leaving from other nodes are drawn with a lighter blue. The figure shows the network of connections within one step away from the node in the center. The figure is perfectly symmetrical diagonally, horizontally and vertically.

Because of the symmetry, the algorithm used for forming the adjacency matrix for the graph is simplified by first forming an asymmetric adjacency matrix and later mirroring the matrix to form two-way connections. The picture on the left

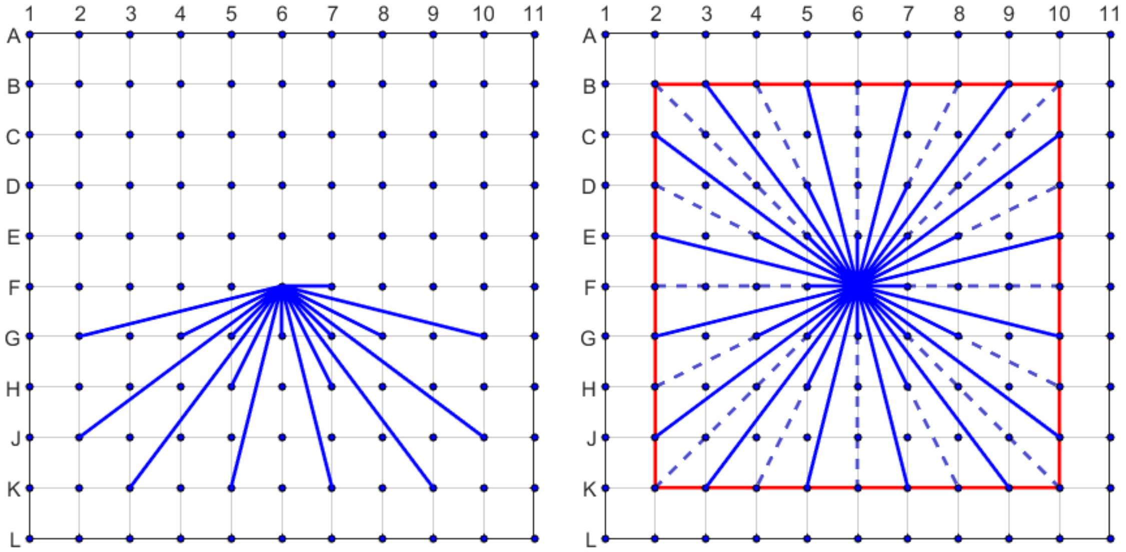

Figure 4.9. *The edges of the grid shaped graph used in the simulation. On the left only asymmetric arcs are shown, i.e., the ones that connect to nodes with a higher index number and that form an upper triangular part of the adjacency matrix. On the right the the full arcs are shown. The red box shows the outermost perimeter that has straight line connections to each node on the perimeter. This distribution gives a fairly even angular distribution of the edges.*

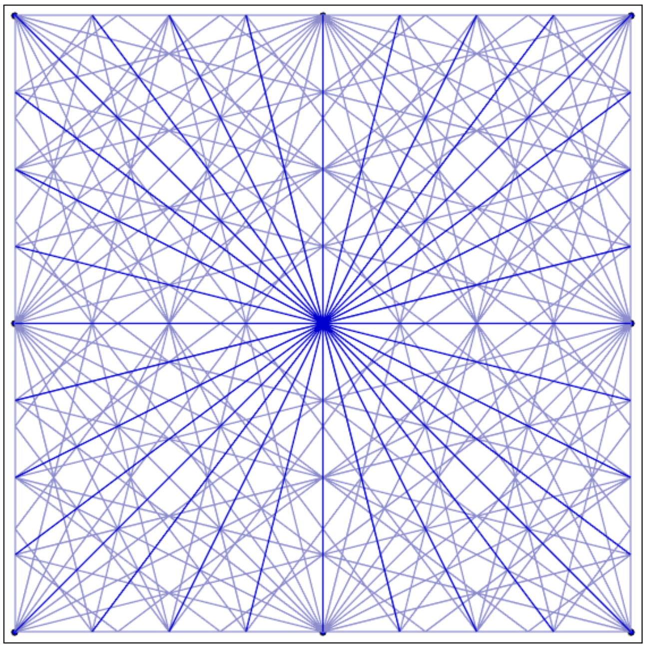

Figure 4.10. *The network of connecting edges shown around one node. The edges connecting to the central node are shown in darker blue and the network of other edges are shown in lighter blue. The area presented is a 3 by 3 grid with the nodes laying on the perimeter of the area.*

of Figure 4.9 shows such an asymmetric case that is used in the algorithm that forms the adjacency matrix. All the nodes are enumerated left to right and top to bottom as on the area of observation. The asymmetric connections in Figure 4.9 define the edges connecting to nodes with a greater enumeration than the current node in question. The algorithm proceeds left to right and top to bottom on the grid, creating edges according to the pattern. A full adjacency matrix including two-way connections for a very simple 6 by 6 grid is illustrated in Figure 4.11 for the purposes of example. Only the positions of non-empty elements are displayed and not the actual values. The values of the edges are trivially determined by Pythagoras Theorem.

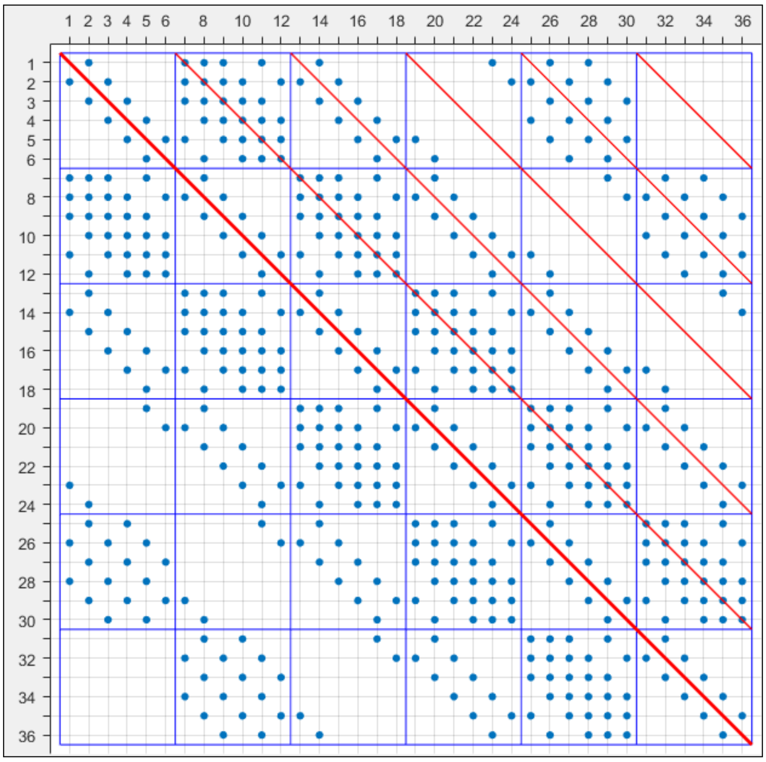

Figure 4.11. *The adjacency matrix for a 6 by 6 grid with edges as defined on the right hand side in Figure 4.9. Only the edges are indicated, not the actual values. Each edge has a value that is the straight line distance between the connected grid points. The thick red line is the central diagonal and it acts as a line of symmetry. The thinner red diagonals indicate nodes that are directly below in the grid. No arcs reach five rows lower and hence there are no edges from the first row (nodes 1-6) to the sixth row (nodes 31-36) and no edges the other way.*

4.2.1 Removing Connections

ROZs are areas through which the arcs may not pass and any nodes inside a ROZ have no edges connecting to them. As seen in Figure 4.8, the rows and columns referring to a node inside a ROZ are trivially empty. The nodes that lie inside a ROZ are determined by a Matlab function `inpolygon`. This is the reason why even circular shaped ROZ are actually described as a polygon. ROZ that are not enclosed but only a restricting line are ignored in this phase. The rows and columns referring to nodes inside a ROZ are simply emptied.

Nodes that are not inside a ROZ may still have their connections restricted by a ROZ as illustrated in Figure 4.7. Determining whether or not a connection cuts a ROZ is computationally a relatively heavy operation and thus the algorithm for forming the adjacency matrix is optimized not to investigate connections from nodes that lie a safe distance away from the nearest ROZ. Figure 4.12 illustrates the situation where the influence of the ROZ is only considered at nodes close to the ROZ. The nodes are traversed left to right and top to bottom and the minimum distance to the ROZ is calculated. Five units are deducted from the minimum distance to a ROZ and any nodes within this radius are guaranteed not to be influenced by a ROZ.

Points which are a safe distance from the nearest ROZ are shown in Figure 4.12 as colored squares. These points are treated as a batch and the algorithm continues traversing the nodes left to right and top to bottom skipping any nodes that have been determined safe in a batch. The white squares in the figure are nodes which are a safe distance away from the ROZ but which do not result in a batch of nodes. The black dots are points that lie below the ROZ, and using the asymmetric edges in Figure 4.9, the connections are guaranteed.

The red crosses are nodes that lie inside a ROZ and these nodes have no connections either way and the corresponding rows and columns of the adjacency matrix have been emptied. The red hexagrams in the figure are nodes for which the heavy computation, of exactly which connections are plausible, needs to be performed. The computation involves investigating all the 16 asymmetric connections from the current node. Each connection forms a segment of a straight line and this line segment is investigated against each line segment of the ROZ.

The pentagon shaped ROZ labeled D in Figure 4.5 has only five segments. This would result in a maximum of 80 line intersection calculations. An oval or round ROZ consists of many more line segments, and an arbitrarily shaped ROZ can have even more. Obviously the investigation can be terminated when the first intersection

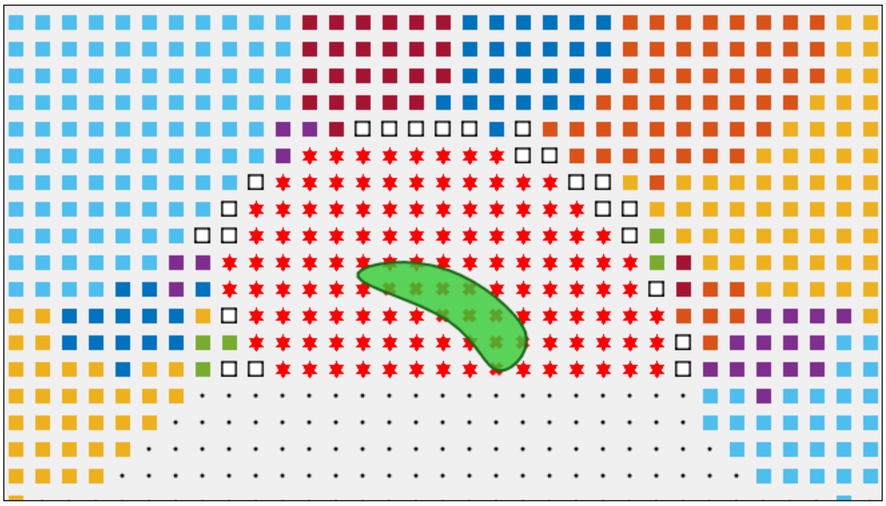

Figure 4.12. *Optimizing the calculation of connecting arcs around a ROZ. The red hexagrams (stars) around the ROZ are nodes that are within five units of length from the ROZ and for which an expensive operation of removing some arcs needs to be performed. The white squares are nodes that are far enough from the ROZ, so that it is guaranteed that no arcs will need to be removed. The colored squares are blobs of nodes that are within a safe radius and are guaranteed not to be influenced by any ROZ. The black dots are all below the ROZ and thus guaranteed not to be influenced by the ROZ.*

of line segments is found and the corresponding element of the adjacency matrix is removed.

4.2.2 Approximation Error

The maximum error in distance in the grid is found on nearby nodes that are three steps straight and one step to the side. Such a node, e.g., G3 in Figure 4.9, has a straight line distance of $\sqrt{(3^2+1^2)} = \sqrt{10}$, but to reach it via the defined edges a path via G4 to G3 needs to be taken. This path via the edges has a distance of $\sqrt{(2^2+1^2)}+1 = \sqrt{5}+1$. The relative error in the distance is $(\sqrt{5}+1)/\sqrt{10} \approx 2.3\%$.

From the group of nodes just outside the red box on the right hand side in Figure 4.9, the largest relative error is on the nodes five steps straight and two steps to the side. Such a node is, e.g., the node H1 in Figure 4.9. This node has a relative error of 1.6%.

The directions of the edges have a maximum angle of separation of

$$\theta = \arctan(1/4) \approx 14.0\,^\circ. \quad (4.1)$$

This occurs on either side of the horizontal or vertical directions. An example is illustrated in Figure 4.13. Theoretically the greatest error caused due to discretized directions occurs exactly in the middle of the two directions L_1 and L_3, i.e., in the direction L_2 with an angle of $\theta/2$. Scaling the situation such that the length l_1 of L_1 is exactly one, the length of L_3 is

$$l_3 = \sqrt{17/16}. \tag{4.2}$$

The length of the midway line L_2 is

$$l_2 = \frac{1}{\cos\left(\frac{\theta}{2}\right)}, \tag{4.3}$$

and its vertical displacement is

$$y_2 = \tan\left(\frac{\theta}{2}\right). \tag{4.4}$$

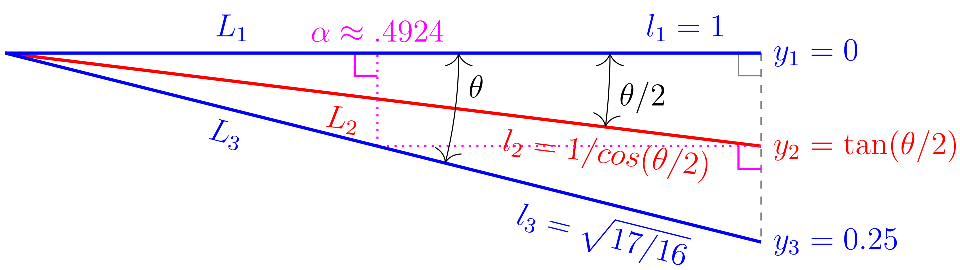

Figure 4.13. *Calculating an error in the direction exactly in between the two directions L_1 and L_3 having the greatest angle of separation in between.*

The ratio of the vertical displacements of line L_2 and L_3 is

$$\alpha = \frac{y_2}{y_3} = \frac{\tan(\theta/2)}{0.25} \approx 0.4924. \tag{4.5}$$

Because the problem is linear, the ratio of α is also the ratio of L_3 that needs to be traveled in order to reach the vertical displacement of y_2. Traveling along L_3 for a portion of α likewise linearly results in a horizontal displacement with a value exactly equal to α, because the length of L_1 is exactly one.

In order to reach the end of line L_2 via the grid directions, the line L_3 is traveled for a portion of α and then the line L_1 for a portion of $1 - \alpha$. The relative error can now be calculated as

$$\Delta L = \frac{\alpha \cdot l_3 + (1-\alpha) \cdot l_1 - l_2}{l_2} \approx 0.75\ \%. \tag{4.6}$$

When long enough distances are traveled, the relative error in distance will fall below the maximum error found for the shorter distance. It can be safely approximated that the relative error caused by moving in the discretized grid will be below 1 %. Such an error in distance does not at all jeopardize the credibility of the evaluation of the engagement frontier.

4.3 Calculating Minimum Reach Times

Once the adjacency matrix is constructed containing the information which nodes are connected and the lengths of those connections, different minimum distance calculations can be performed. The simulation model in this thesis uses *MatlabBGL* (Matlab Boost Graph Library) [19], a Matlab compatible version of the Boost Graph Library (BGL) [20]. The whole original *Boost Libraries* are implemented in `C++`.

The adjacency matrix is passed onto the `shortest_paths` function of the MatlabBGL. In addition to the adjacency matrix, the function takes as an input a single node from which the shortest distances to every other node are calculated. These shortest paths are calculated for the starting positions of each asset from each side. The function returns the minimum distance to each node from the given starting node. Once the minimum distance from the starting node of an asset to every other node is known, the minimum flying time for the asset to reach each node is calculated by dividing the distance by the flying speed of the asset.

Figure 4.14 presents a simple scenario where two Red assets start at time $t = 0$ and two Blue assets are in 5 minute standby with a 3 minute take off delay. The Blue assets start advancing at time $t = 8$. There is a simple polygon shaped ROZ between assets B#2 and R#2. The engagement frontier in the example is determined by the first assets and without missile envelope. As seen in the figure, the engagement frontier is undefined if it would lie inside a ROZ.

The reach times for the Red assets are illustrated in Figure 4.15 and the reach times for the Blue assets in Figure 4.16. The nodes inside the ROZ cannot be reached and this results in a hole in the figures. The reach times are color coded from dark red through orange-yellow to white and the reach time is drawn as a surface where the z-component represents the minimum time to reach the point on the x-y -plane.

The minimum reach times are calculated for each asset and the result is stored as a matrix where the rows and columns refer directly to those of the 2D grid and the matrix elements store the corresponding minimum reach time of that node. The standby time, the take-off delay and the advance warning time are added to the minimum flying time for the Blue assets. For the Red assets, the airborne instant

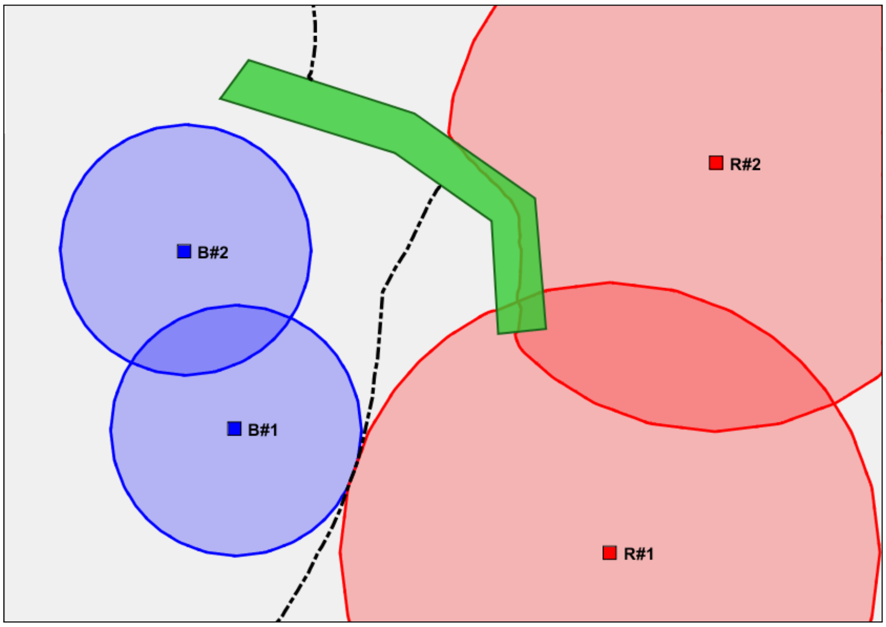

Figure 4.14. *The scenario used in Figures 4.15 to 4.19. Blue has two assets, B#1 and B#2, and Red has likewise two assets R#1 and R#2. There is a simple polygon shaped ROZ in green that effects both the Blue and the Red assets. The engagement frontier is the black dash-dotted line and it is formed by the first assets and without missile envelope.*

is only added. The advance warning time and the airborne instant may both be negative to simulate different real-life scenarios.

Minimum reach times are also calculated for the missile envelope for each asset. The minimum reach time matrix of the assets' missile envelope is based on the assets' minimum reach time. The missile envelope is parameterized as a fixed distance. The distance of the missile envelope is divided by the corresponding asset's flying speed and this amount of time is subtracted from the asset's minimum reach time matrix to form the missile envelope's minimum reach time matrix.

The reach areas of the assets and the missile envelopes at a given moment of time can be visualized by drawing a filled contour line at the value of the time instant. To tidy up the visualization, the undefined reach times of nodes inside a ROZ can be interpolated from surrounding nodes. Such a technique of visualizing the assets reach areas and tidying the visualization of assets flying 'around' ROZ is used throughout this thesis.

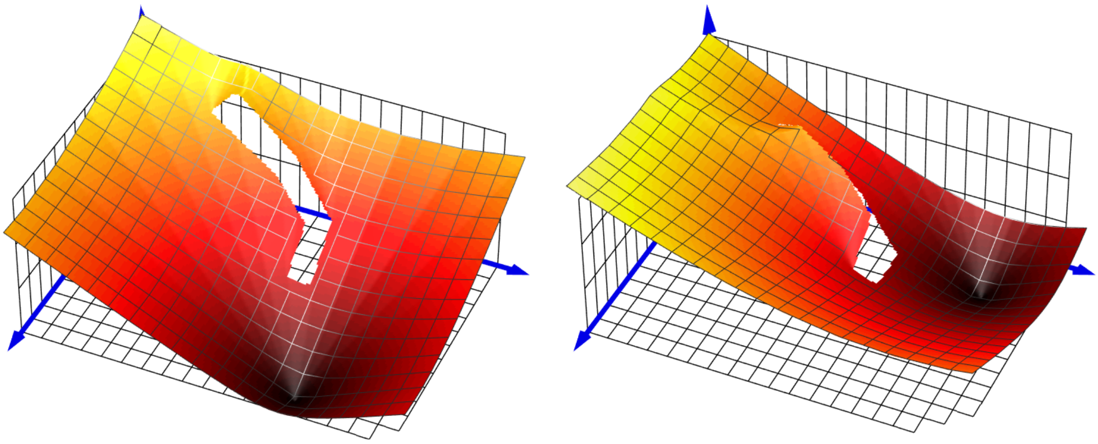

Figure 4.15. *The reach times of assets* R1 *and* R2 *as a 3-D surface. The height and the color both represent the reaching time. A lower height and a darker color mean a smaller reaching time.*

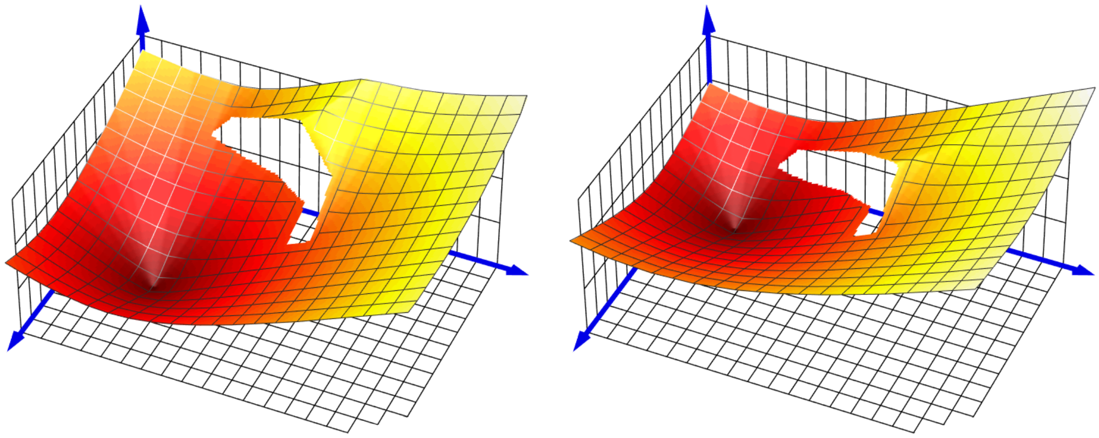

Figure 4.16. *The reach times of assets* B1 *and* B2 *similar to Figure 4.15.*

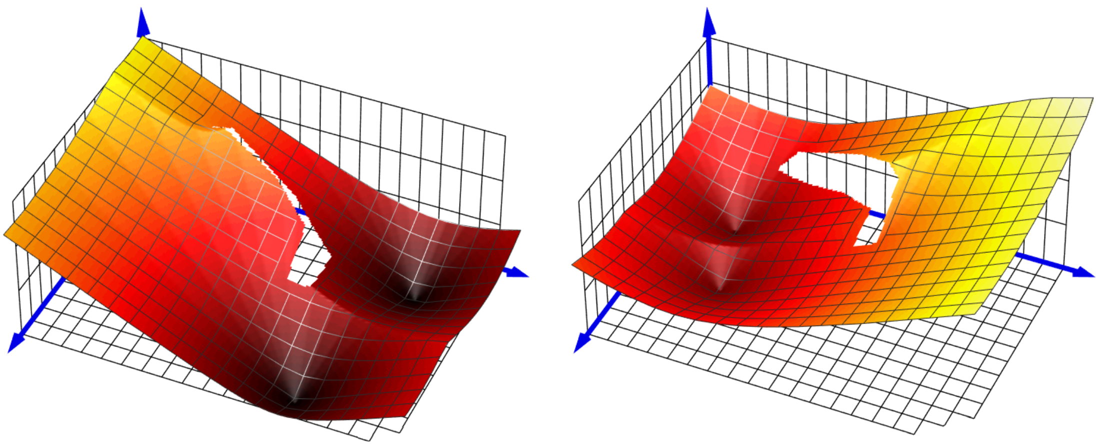

Figure 4.17. *The minimum reach times of the Red assets on the left and the Blue assets on the right. The values are obtained by taking the minimum of the reach times in Figures 4.15 and 4.16.*

4.4 Forming Engagement Frontier

Once the minimum reach time matrices have been calculated for each asset and each missile envelope for both parties, all the 16 different combinations of the engagement frontier can be calculated. All the individual reach time matrices of assets and missile envelopes for both parties are combined into a 3D matrix of minimum reach times, where the first two dimensions of the 3D matrix represent the nodes of the 2D grid and the third dimension holds the different values for each asset or missile envelope. There are a total of four such 3D matrices formed; Blue assets, Blue missile envelopes, Red assets, and Red missile envelopes.

Each 3D matrix is sorted in ascending order on the third dimension, i.e., the minimum reach time of each asset or missile envelope is sorted per each node such that a Blue or Red minimum reach time is obtained on the 'top layer' of the matrix. The second layer of the matrix contains the minimum reach times of the second assets. Such a 3D matrix is presented in Figure 4.18. The data is the same as in Figure 4.15 but sorted node wise by the reach time of the individual assets. The top layer of the matrix presents the same data as the minimum reach time of the Red assets in Figure 4.17. The 3D matrix in Figure 4.18 holds only a coarse sample of the data to better illustrate the discrete values.

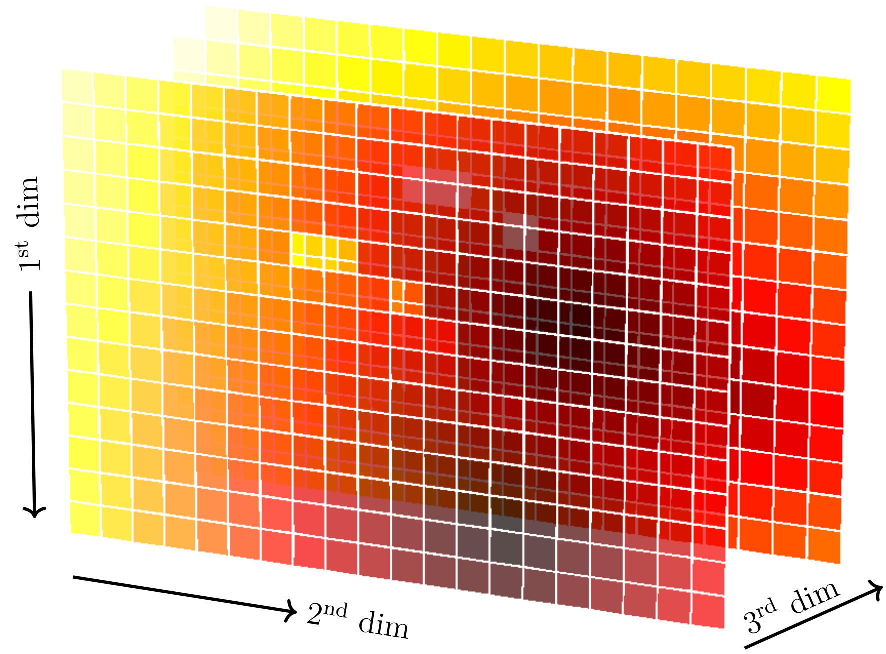

Figure 4.18. *3-D sorted matrix with the 1^{st}, 2^{nd} and 3^{rd} dimensions labeled.*

The engagement frontier is the intersection of the chosen Blue and Red minimum reaching times. Such an intersection is illustrated in Figure 4.19. Technically the intersection is easiest to determine by subtracting the Blue minimum reach times from the Red minimum reach times. The intersection has a value of zero. The engagement frontier is visualized by drawing a contour line at the zero value.

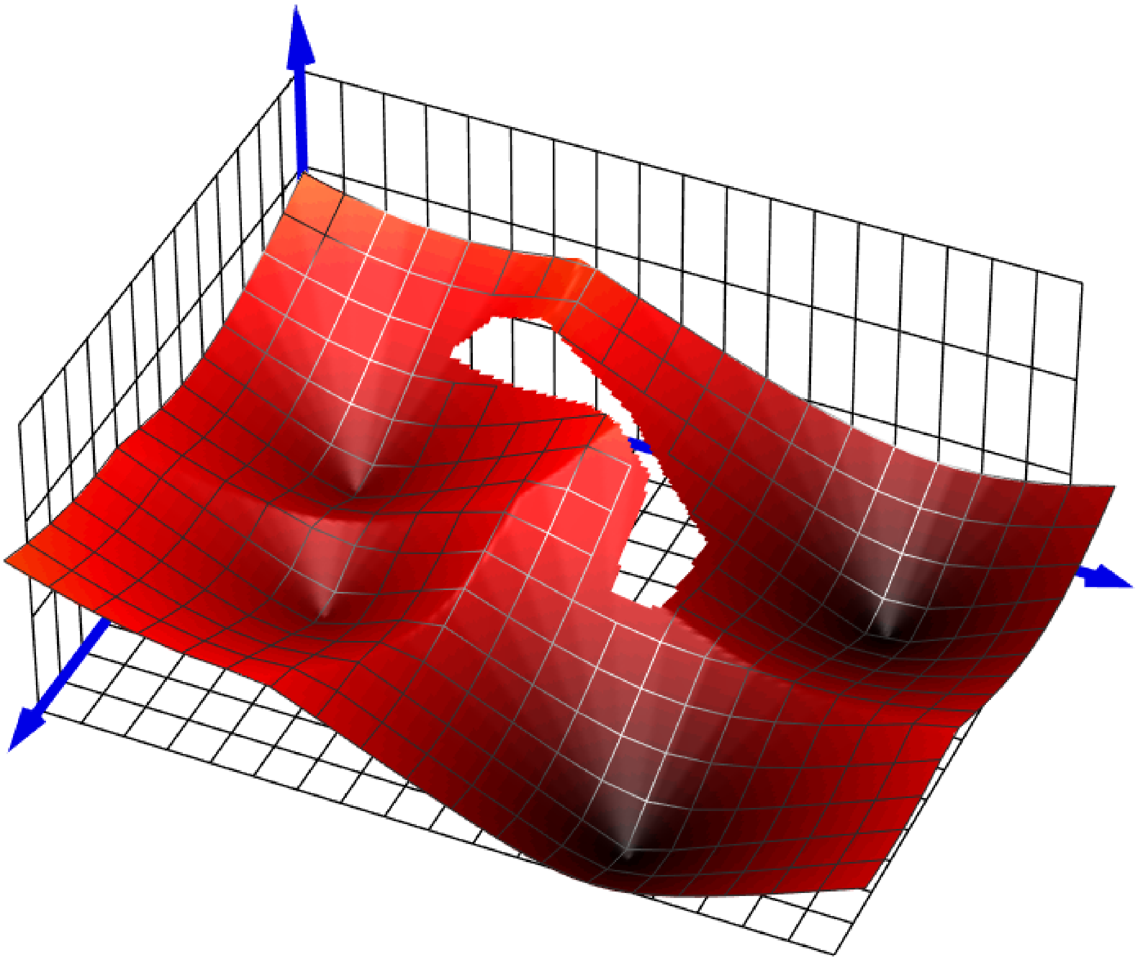

Figure 4.19. *The engagement frontier is formed at where the surfaces in Figure 4.17 intersect. The engagement frontier is the ridge running vertically in the middle.*

5. Examples

5.1 Force Fulfillment

5.1.1 Operation Area with a Continuous One Asset Force Requirement

Scenario 5.1.1 presents the simplest possible case of force fulfillment where the target of force allocation is one operation point, i.e., an operation area with one asset force requirement for a continuous period of time. The force requirement is visualized in Figure 5.1. The requirement is one asset for a continuous time period of seven and a half hours. There are three assets that are supported by a single base. Hot refueling is allowed only once after each normal refuel. Figure 5.2 presents the use of the assets in trying to fulfill the force requirement at the operation area.

As can be seen from Figure 5.2, the assets are in active use and there is little standby time in between any two consecutive missions for an asset. This active use of the assets combined with the restriction of every hot refuel being followed by a normal refuel, leads to a deficiency in the achieved force at the operation area. This deficiency is seen in Figure 5.1 as the two narrow red strips at around 4 hours and 8 hours. Asset A#4 is still being refueled when it already would be needed to replace Asset A#7 at just before 4 hours. Asset A#4 flies to the operation area immediately after refueling and taxiing but the deficiency has already happened.

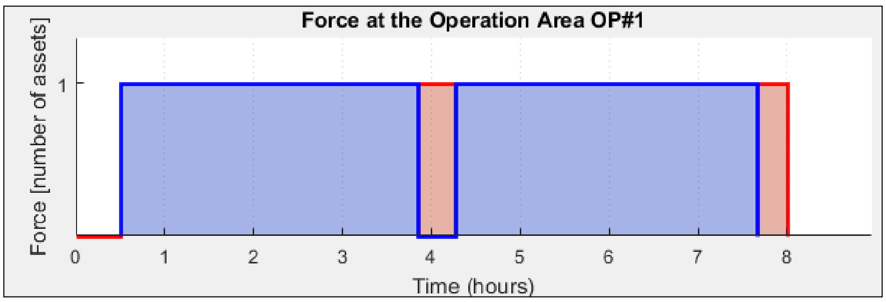

Figure 5.1. *The achieved use of force in the single operation point in Scenario 5.1.1. The red color presents the force requirement and the actual achieved force is presented on top of the requirement in blue color. Thus, any red color that is showing presents a deficiency in the achieved force.*

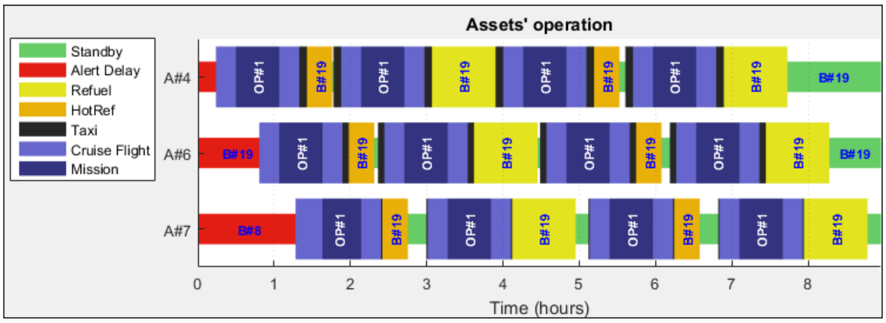

Figure 5.2. *The use of the three assets supported by one base in Scenario 5.1.1.*

5.1.2 Operation Area with a Varying Force Requirement

In Scenario 5.1.2, one operation area is served by nine assets and six bases. The operation area has a varying force requirement. The force requirement begins with an initial short 'burst' of two asset force requirement, then a period of idle time. After the idle time the force requirement builds up incrementally to a maximum of three assets, before reducing down to two and then ending completely. The force requirement along with the achieved force fulfillment is presented in Figure 5.3. The operation area is technically broken down into five individual operation points but the breakdown is not visible in the figure.

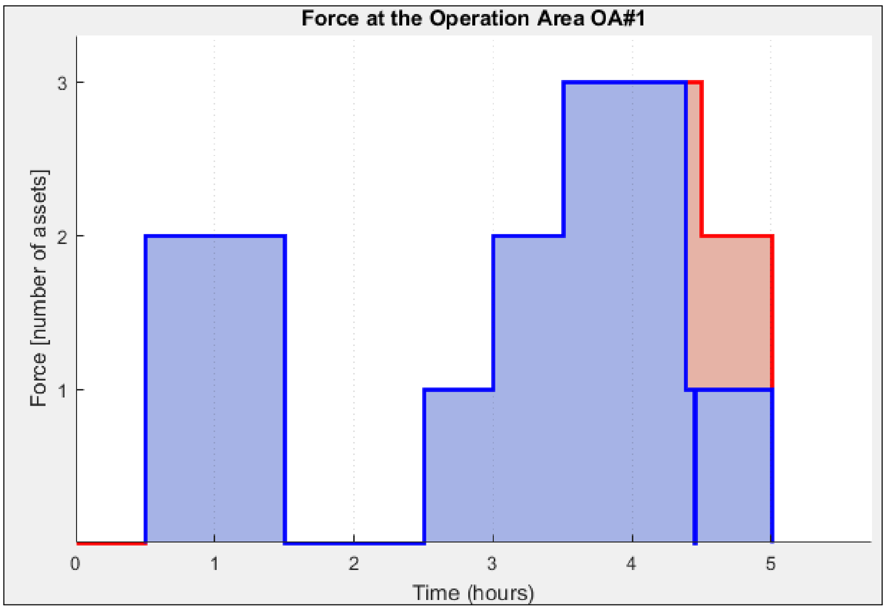

Figure 5.3. *The achieved use of force in Scenario 5.1.2 where the operation area has a varying force requirement.*

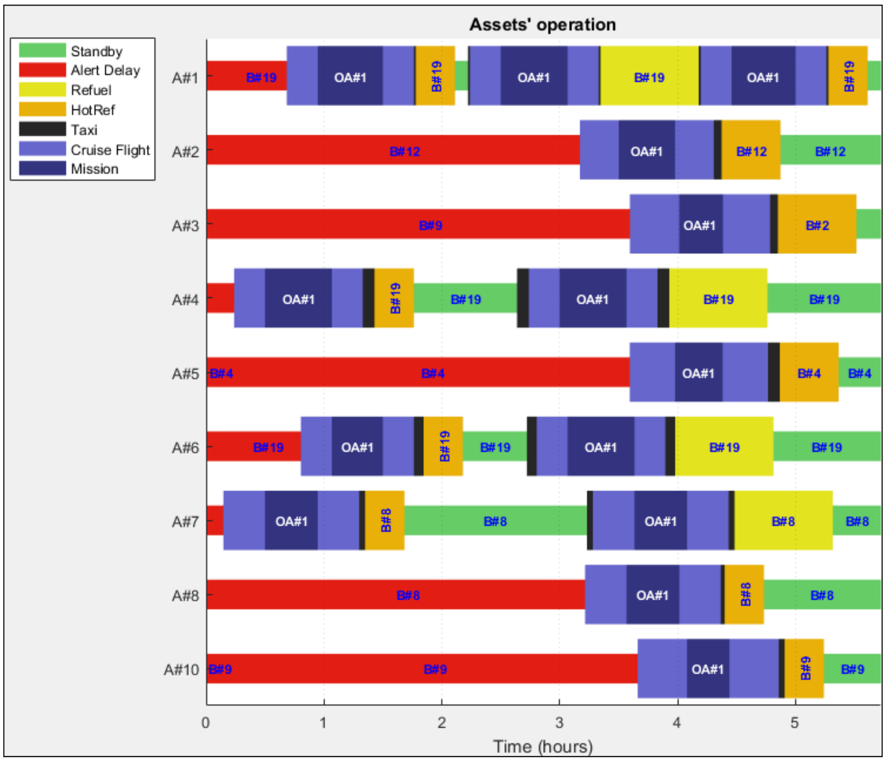

Figure 5.4. *The use of the nine assets in Scenario 5.1.2.*

The use of the assets is presented in Figure 5.4. The first 'burst' is handled by Assets A#4 and A#6 which are then replaced by Assets A#1 and A#7. Asset A#1 does a hot refuel and after a brief standby period it is allocated to the operation area at 2.5 hours to satisfy the one asset force requirement active at that time. The maximum force requirement of three assets is achieved but then all the assets are in active use. The three asset force cannot be sustained for the whole period that was required, and a deficiency is formed. For the last half an hour, only asset A#1 is able to serve the operation area, the other assets being refueled or returning to base.

5.1.3 Two Operation Areas with Different Priorities

Operation areas can also be prioritized. In Scenario 5.1.3, there are two operation areas with Operation Area OA#1 prioritized over Operation Area OA#2. Both operation areas have only a one asset force requirement for a continuous period of time. The operation areas and their achieved force are presented in Figure 5.5. The prioritization can be seen in the figure as OA#1 has an almost full achievement of force but OA#2 is severely lacking.

The principle in prioritizing is that the primary operation area is resolved first and then the secondary operation area is resolved with any leftovers from the primary area. When making allocations to the secondary operation area, the allocations made

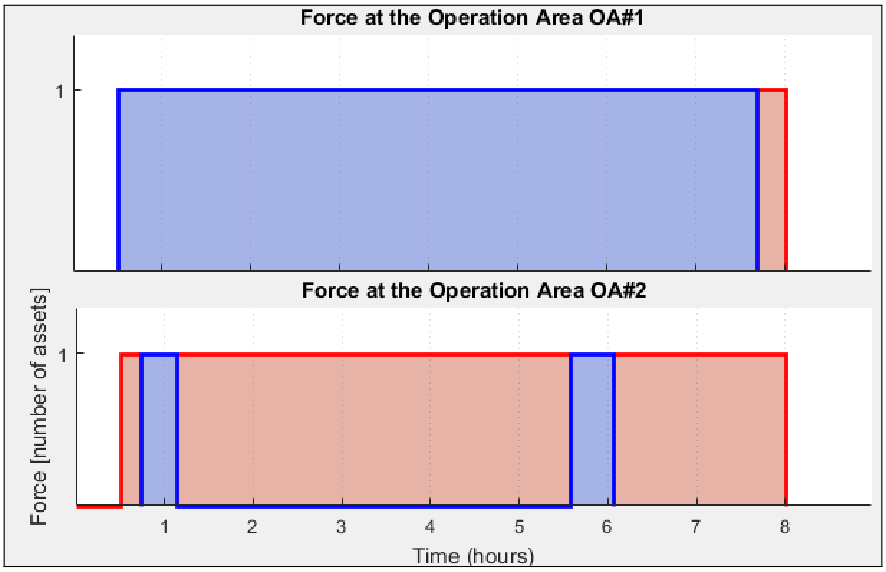

Figure 5.5. *The combined achieved use of force for the operation areas in Scenario 5.1.3. Operation Area OA#1 is prioritized over Operation Area OA#2.*

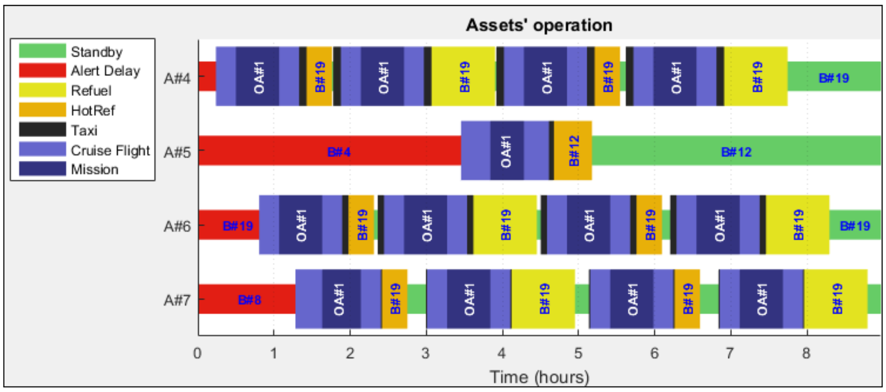

Figure 5.6. *The use of assets for the prioritized Operation Area OA#1 in Scenario 5.1.3.*

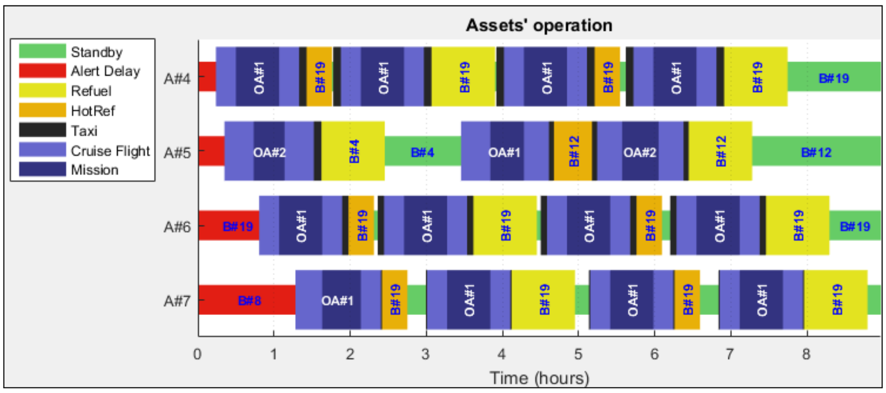

Figure 5.7. *The combined use of assets for the Operation Areas OA#1 and OA#2 in Scenario 5.1.3. Asset A#5 has two additional flight operations compared to the schedule in Figure 5.6.*

for the primary operation area are untouched. This means that only assets that can be returned in time back to the base where they left, and assets that have no more flight operations for the primary operation area, may be used. This is clarified by comparing Figure 5.6, which shows the operations of the assets after resolving the primary operation area, with Figure 5.7, which shows the operations of the assets after resolving the secondary operation area as well. The only change between these figures is that Asset A#5 has gotten two additional operations, one before and one after the operation in Figure 5.6. These two operations can be seen in Figure 5.5 as the two spikes of force fulfillment at the Operation Area OA#2.

5.1.4 Spatial Force Fulfillment Potential with a Prioritized Operation

The force fulfillment can be calculated in a batch run where the calculation for one operation area is iterated over a grid of different positions. The values at each grid point are calculated independently of each other. The batch run calculation can be combined with a prioritization of operation areas. The ‘floating’ operation area is always calculated with the lowest prioritization.

In Scenario 5.1.4, there are four bases and six assets. Base B#4 has only one plateau, B#8 has two, B#19 has three, and B#18 has an abundance of five plateaus. The fixed and prioritized Operation Area OA#1 has a force requirement of one asset and a period of duration of seven and a half hours. The use of assets for the Operation Area OA#1 is presented in Figure 5.8.

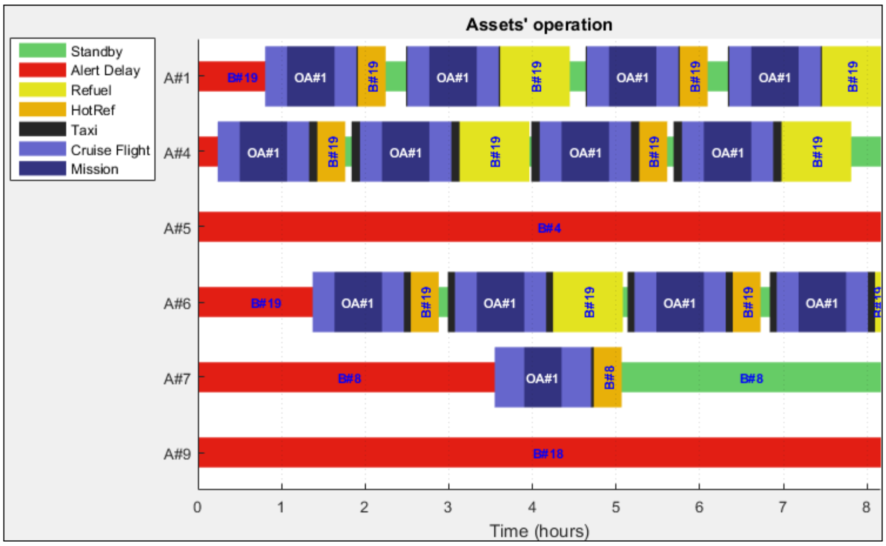

Figure 5.8. *The use of assets for only the prioritized Operation Area OA#1 in Scenario 5.1.4.*

The floating operation areas have the same force requirement of one asset and the same duration of seven and a half hours as the fixed Operation Area OA#1. The result of the batch run is presented in Figure 5.9. The values in the figure present the measure of average force in the floating operation area. The contour lines are approximations based on the actual calculated values in the grid. The figure also shows the spatial positions of the bases and the fixed Operation Area OA#1. Two example points are picked from the batch run. These example points are labeled as OA#F1 and OA#F2 and are shown with a purple dot in Figure 5.9.

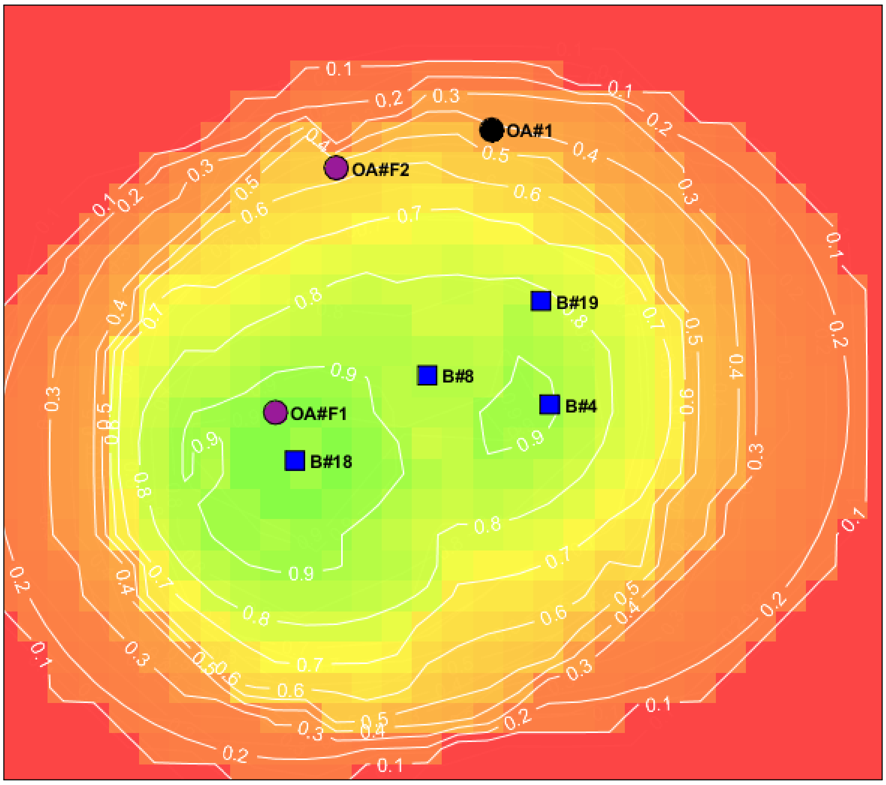

Figure 5.9. *The spatial force fulfillment potential with a prioritized operation area in Scenario 5.1.4. The black dot represents the fixed and prioritized operation area and the purple dots represent two sample points. The squares are color coded by the average force measure of the force fulfillment. Force fulfillment is simulated individually for each square.*

The fixed operation area is prioritized and hence the simulations for each single grid point start off with the fixed schedule that is presented in Figure 5.8. The use of assets for the sample operation areas OA#F1 and OA#F2 are presented in Figures 5.12 and 5.13. Note how both of these schedules for the use of assets include the schedule in Figure 5.8 as such and only add additional flight operations to the assets. The force fulfillment of the two sample points OA#F1 and OA#F2 are presented in Figures 5.10 and 5.11 respectively. OA#F1 has an average force of 0.94 assets and OA#F2 has an average force of 0.57 assets.

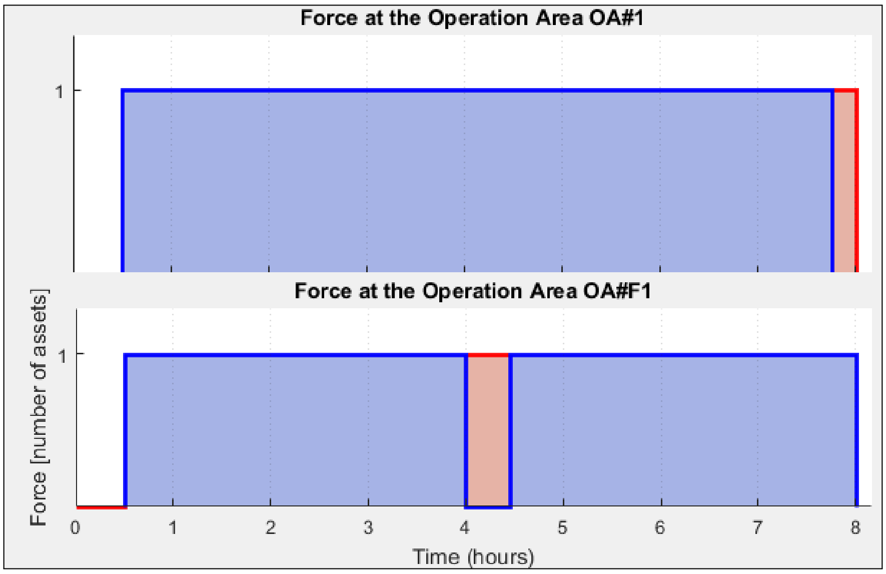

Figure 5.10. *The force fulfillment for OA#F1 together with the prioritized OA#1 in Scenario 5.1.4. The average force at OA#F1 is 0.94 assets.*

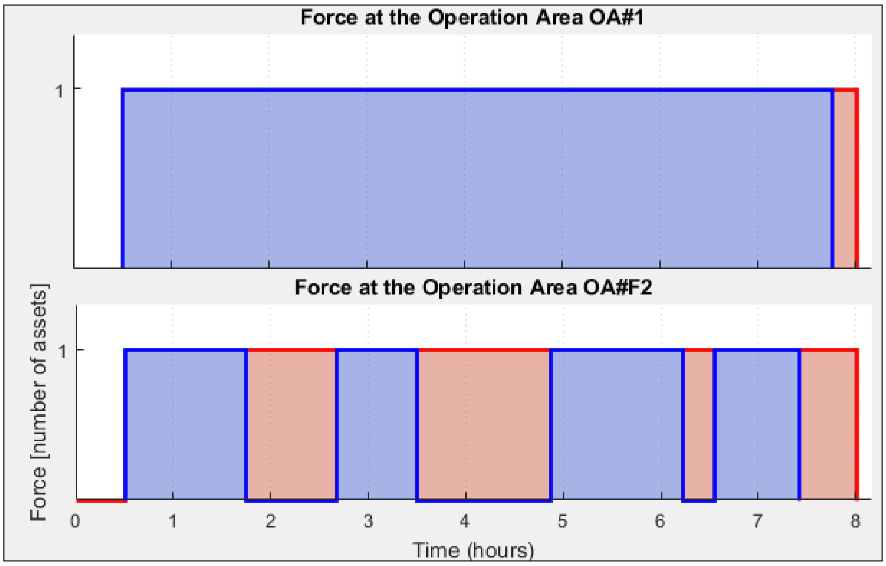

Figure 5.11. *The force fulfillment for OA#F2 together with the prioritized OA#1 in Scenario 5.1.4. The average force at OA#F1 is 0.57 assets.*

5.1.5 Discussion

In Scenario 5.1.2, the situation has a chance to 'reload' during the one hour gap between the two sets of operations. All assets are in standby when the operation area again requires assets at two and half hours and Asset A#1 has had time to undergo a hot refuel before being allocated to the operation area on time. By the time the force requirement builds up to three assets at the time instant of 3.5 hours, the earlier allocated assets need to return to base. There are enough assets to initially satisfy

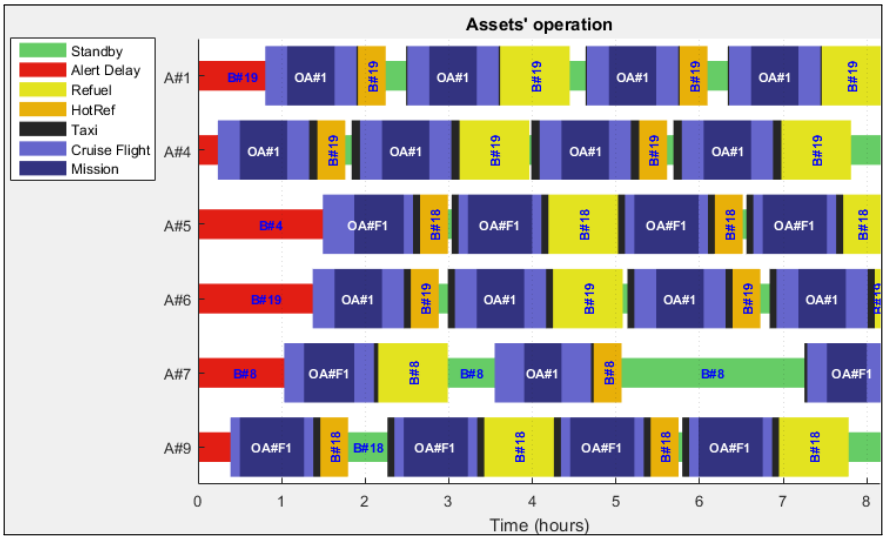

Figure 5.12. *The combined use of assets for the prioritized Operation Area OA#1 and the floating operation area at sample point OA#F1 in Scenario 5.1.4.*

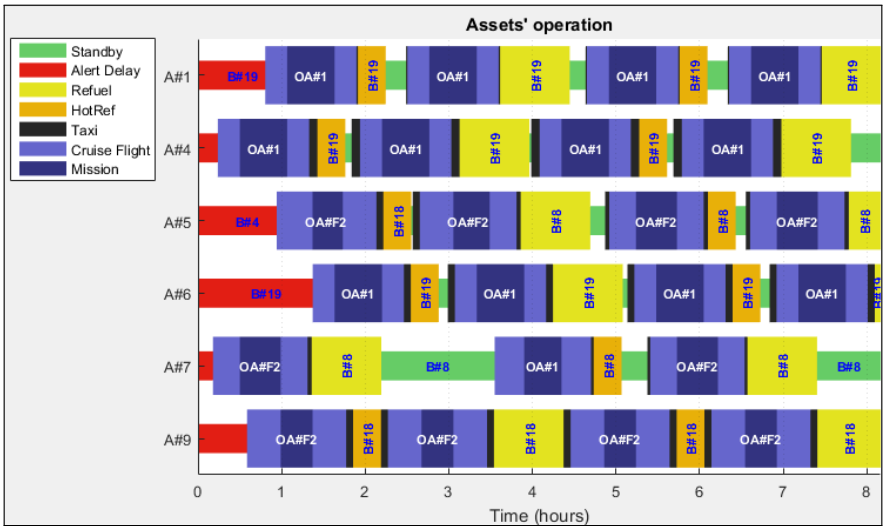

Figure 5.13. *The combined use of assets for the prioritized Operation Area OA#1 and the floating operation area at sample point OA#F2 in Scenario 5.1.4.*

the maximum force requirement but the force cannot be sustained and a serious deficiency is formed at the operation area. There are no replacing assets available for those returning back to base at the time instant after 4 hours.

The heuristic is always greedy and tries to get the assets back into standby as fast as possible. In Scenario 5.1.2, Assets A#4, A#6 and A#7 would have time for a normal refuel, but instead the heuristic chooses for the hot refuel. As in this scenario hot refueling is allowed only once after each normal refuel, using the hot refuel and going into standby is a waste. Not being allowed for hot refuel in the subsequent flight operation weakens the achieved force at the operation area. A more clever solution would be to do a normal refuel for Assets A#4, A#6 and A#7 after their first flight operation and spare the hot refuel option for later when the demand is the greatest. This would be an improvement to the heuristic presented in this thesis.

In Scenario 5.1.4, the batch run of geographical coverage yields an even larger ‘big picture’ of the air defense capability than simulating the force fulfillment for a set of fixed operation areas. The batch run does not only answer the question whether force fulfillment can be achieved at a dedicated single point. The batch run gives a picture of how the force fulfillment can be achieved in the neighborhood of different bases and how the air defense capability weakens as the operation area is further away from the bases.

In Scenario 5.1.4, the intuitive result is confirmed, that force fulfillment is at its strongest near the bases and weakens as the distance between the nearest base and the operation area grows. This is sensible as with a longer distance more and more fuel and time is spent ’commuting’ to and from the operation area, and less fuel and less time will remain available to be spent doing an effective operation. As the fraction of time spent at the operation area with respect to the whole flight operation grows, the more assets would be needed to achieve the same level of force fulfillment.

The prioritization of the operation areas works in a predictable manner, despite the limitations applying to interior allocation slots, that an asset needs to return in time back to the plateau from where it left. Solving the combined schedule within one iteration would allow for a slightly better solution but would require a considerably more complex algorithm. In the implementation used in this thesis, the schedule for the prioritized operation area is calculated first, and based on that the lower prioritized operation areas are allocated. This asynchronous approach allows for the reuse of the schedule which results in less computing and a faster solution time than if solving the combined schedule from scratch within one iteration.

5.2 Engagement Frontier

5.2.1 Head-On Engagement

Scenario 5.2.1 illustrates a case where there are multiple ROZs. The shapes and the positioning of the ROZs are purely for demonstrational purposes to illustrate how the assets bypass them. Furthermore, the ROZs illustrate the point that the engagement frontier is undefined if it would coincide with an enclosed ROZ. The Blue asset starting from Base B#2 has a slightly faster flying speed than assets from Bases B#1 and R#1 and the asset from Base R#2 has respectively a slightly slower flying speed. The different flying speeds result in an engagement frontier that is curved. Figure 5.14 presents the engagement frontier for Scenario 5.2.1 at the moment of time when first contact has happened. The engagement frontier is without missile envelopes and for the first assets, i.e., an I vs I engagement.

Note how the engagement frontier is practically straight where Assets B#1 and R#1 engage. This is a special case where the assets are flying at the same speed and have taken to flight at the same. Only the round shaped ROZ in front of Base B#1 causes a notch in the otherwise straight engagement frontier. The engagement frontier where assets B#2 and R#2 engage is curved.

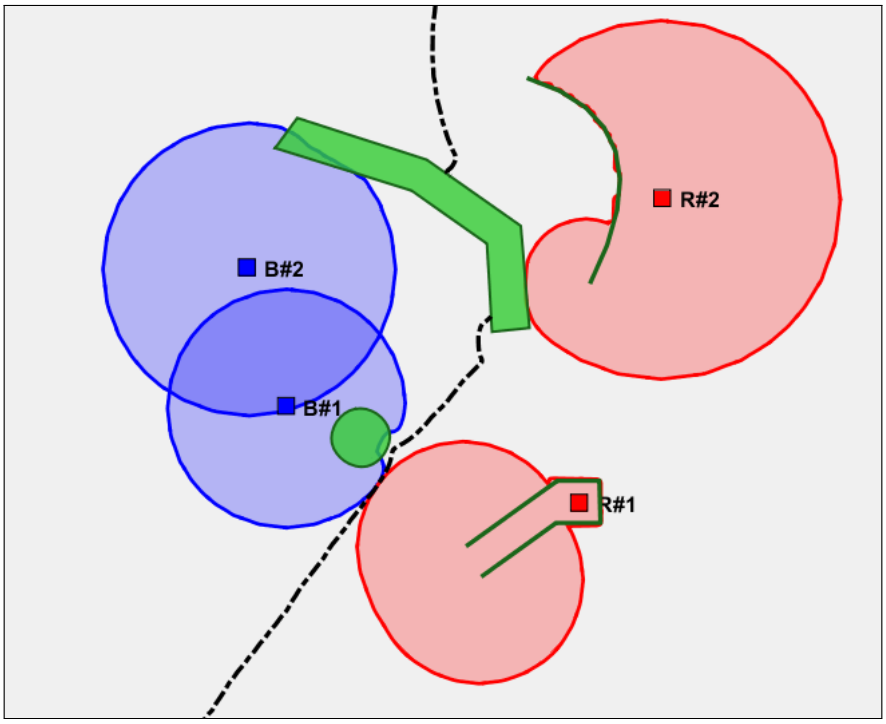

Figure 5.14. *The first contact in a I vs I engagement frontier without a missile envelope in Scenario 5.2.1.*

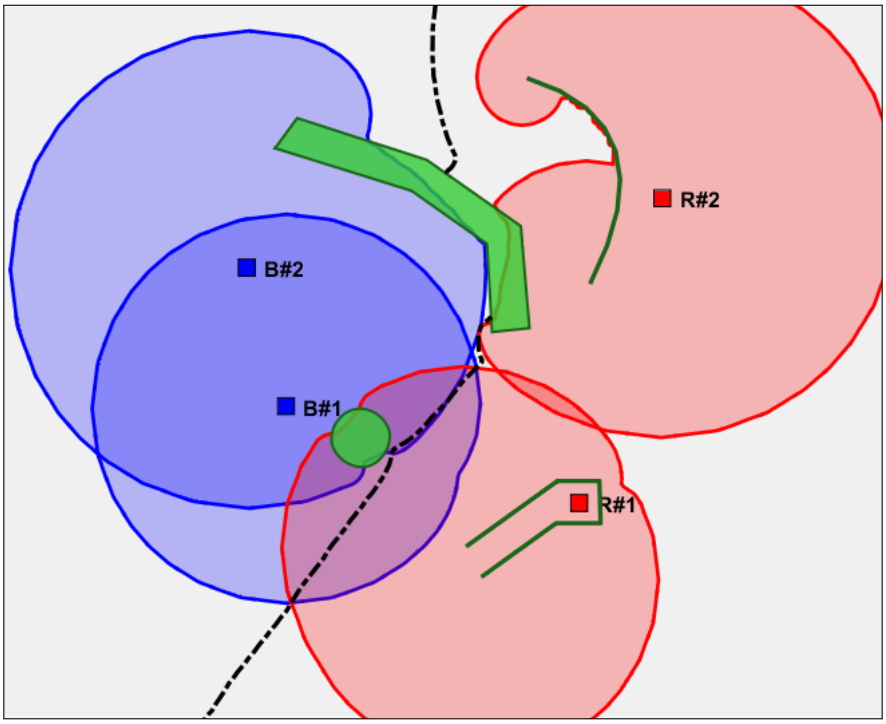

Figure 5.15. *A later contact in a I vs I engagement frontier without a missile envelope in Scenario 5.2.1 with multiple ROZs.*

Figure 5.15 presents the same scenario but where the assets reach areas is visualized at a later instant of time. The instant of time is the first contact of assets from Bases B#2 and R#2. Now the engagement frontier is curved such that the center of curvature is on the side of Asset R#2. This is because Asset B#2 has a faster flying speed and thus it 'wraps itself' around the reach area of Asset R#2. The engagement frontier in general lies always at the intersection of two opposing reach areas at some moment in time.

5.2.2 Interception with Missile Envelope

Scenario 5.2.2 presents a case with a simple polygonal ROZ and an engagement frontier with missile envelopes. The bases and the assets are identical to those in Scenario 5.2.1 but the engagement frontier is determined taking into account the missile envelopes from either sides separately.

Figure 5.16 presents the engagement frontier with the blue assets' missile envelope and Figure 5.17 presents respectively the engagement frontier with the red assets' missile envelope. It is obvious that the engagement frontier is closer to the blue bases, when observing the red assets' missile envelope intercepting the blue assets than other way round.

If missile envelopes from both sides are considered at the same time, the engagement frontier would lie somewhere in the middle compared to Figures 5.16 and 5.17. It is however questionable what the interception of two missile envelopes would reflect in real air combat.

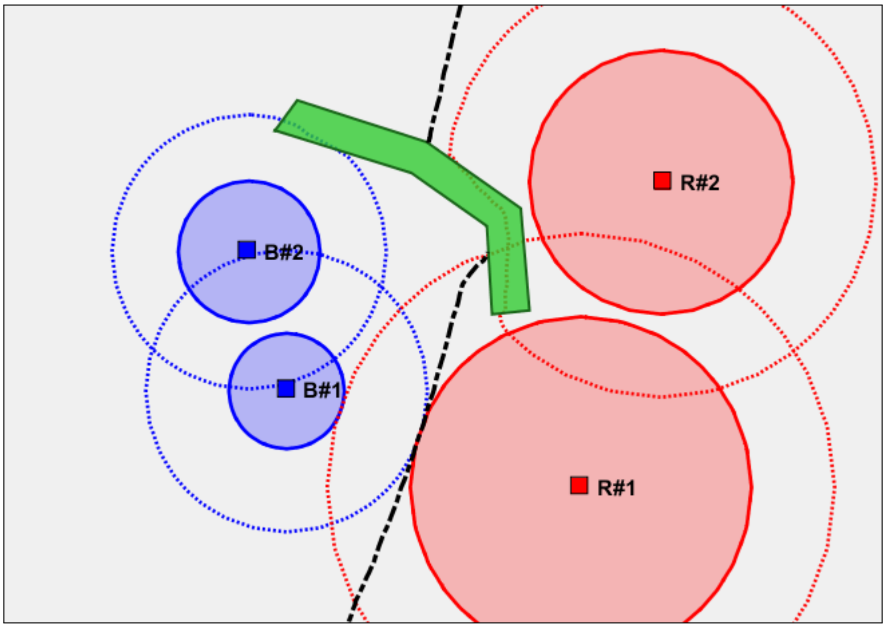

Figure 5.16. *The first contact in a I vs I engagement frontier taking into account only the blue assets' missile envelope in Scenario 5.2.2.*

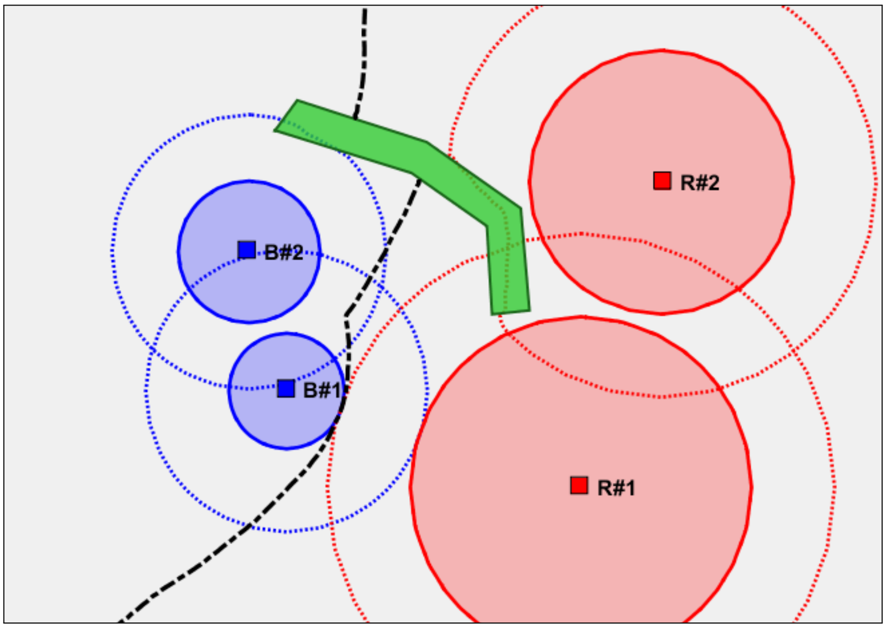

Figure 5.17. *The first contact in a I vs I engagement frontier taking into account only the red assets' missile envelope in Scenario 5.2.2.*

5.2.3 Protecting a Target

Scenario 5.2.3 presents a case with a higher level goal of protecting a target from a hostile attack. This scenario combines together all the features of the engagement frontier simulation model. Figure 5.18 shows the overall situation where the Blue side has two bases B#18 and B#19, and the Red side has bases R#3 and R#4. The target, named Target #1, is marked with a dot. The available airspace is restricted by a polygon shaped ROZ.

It is wanted that the blue assets gain air supremacy and that they would intercept the first red asset with a two against one supremacy. Thus, the 2^{nd} Blue vs. 1^{st} Red engagement frontier is of interest where the second blue asset intercepts the first red asset. The Blue side gets the alert from the reconnaissance alert line. Figure 5.18 presents this moment of time when the red asset from Base R#3 has just cut the reconnaissance alert line. Notice how any blue assets are not airborne yet. Cutting the reconnaissance alert line triggers an alert and the standby time comes into effect. The moment of triggering the alert is called the advance warning time. After the

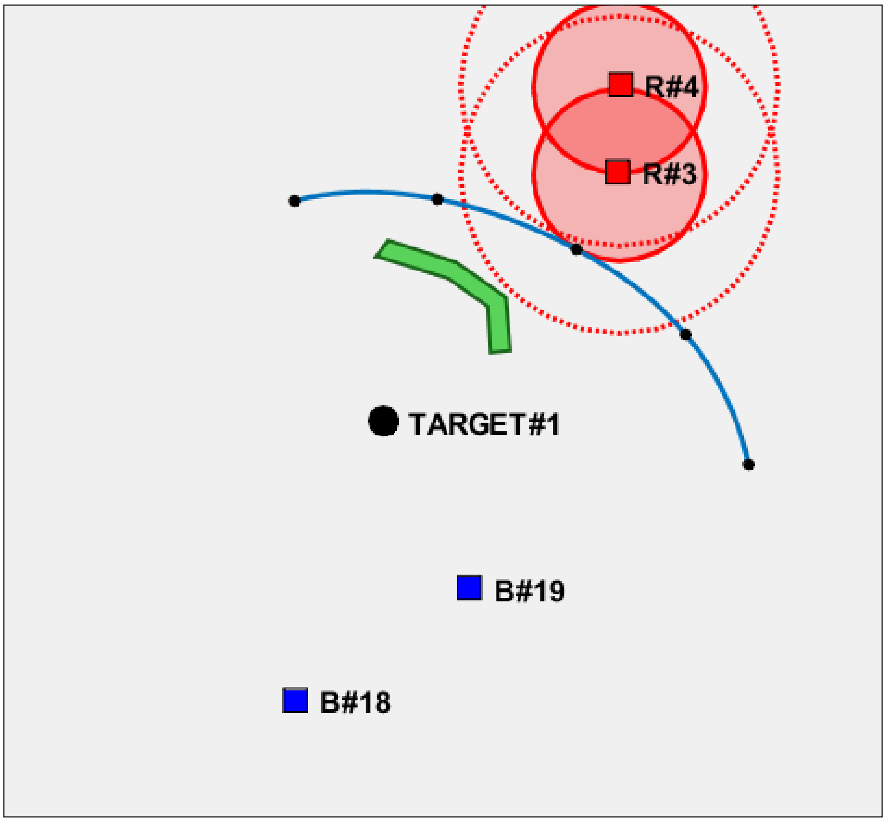

Figure 5.18. *The overall situation in Scenario 5.2.3. The red asset from base R#3 has just cut the reconnaissance alert line and triggered the advance warning time. Blue assets are still in standby in their bases.*

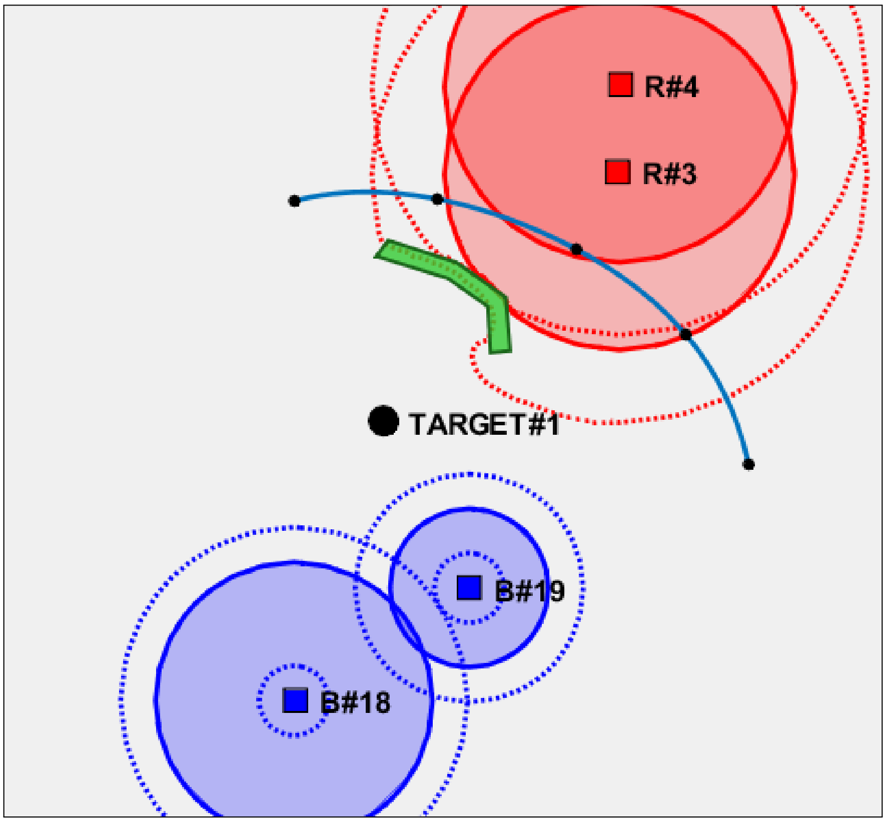

Figure 5.19. *The reach of assets at the moment in time when all the assets are airborne in Scenario 5.2.3. The first blue asset from Base B#18 has started off six minutes before the first blue asset from Base B#19. The second assets from either base start off at the same time, eight minutes after the first asset from Base B#19.*

standby time has elapsed, the concerned asset will take off from the runway. The take off delay is further added to the standby time before the asset is considered to be in full flight altitude and in full speed on top of the base.

Both of the Blue bases have two assets present. The other one of the assets at each base is in raised alert and the other asset is in standard 15 minute alert. The asset at Base B#18 is in one minute alert and the asset at Base B#19 is in seven minute alert. All assets have a three minute takeoff delay. Figure 5.19 presents the moment in time when all the blue assets have just gotten airborne. Note how the asset from Base B#18 has a larger reach area because it has started six minutes earlier than the asset from Base B#19. The missile envelope is considered active immediately after the take off delay has elapsed.

Figure 5.20 presents the moment in time when the first interception by the second blue asset missile envelope happens against the first red asset. The Red missile envelope does not reach Target #1 which was the purpose in this scenario. The engagement frontier is also well beyond Target #1. It can be concluded that the assets from Base B#18 are not as crucial as the assets from Base B#19. If the number of assets that can be afforded to keep in a high state of alert is limited, a better air defense capability could be reached by placing both of the high alert assets in Base B#19.

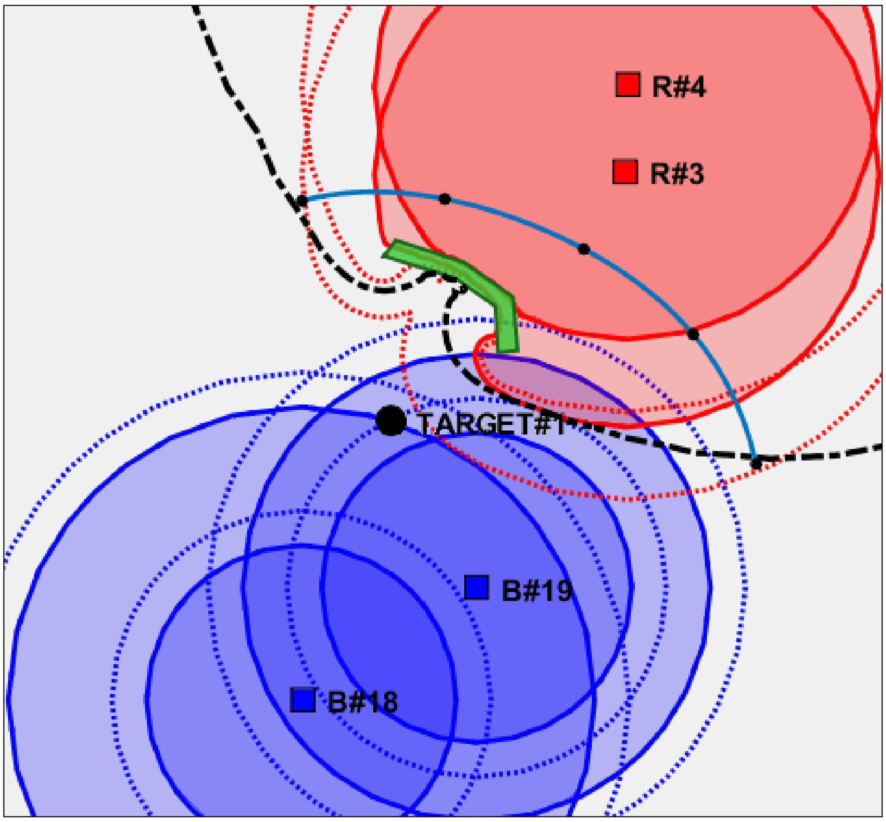

Figure 5.20. *The moment of contact of the blue* second *asset's missile envelope with the* first *red asset in Scenario 5.2.3. The red missile envelope does not reach Target #1 and the engagement frontier is well beyond Target #1.*

5.2.4 Discussion

It can be criticized that missiles need to bypass ROZs the same way that the assets need to in the simulation model for engagement frontier. It is reasonable that a missile would be able to fly through a ROZ and avoid, e.g., anti-aircraft fire from the ground, supporting the view that the missiles should ignore any ROZs. On the other hand, missiles are launched by the pilot of an aircraft and the pilot has to have the target in sight, e.g., in radar. There are limitations and certain preconditions for launching missiles that go beyond the scope of this thesis. Modeling the missile envelope as accurately as possible would require considerably more input from the environment, such as sight of radar or the angle of incidence of the approaching hostile aircraft.

Scenario 5.2.3 was carefully constructed by hand such that the higher level goal of protecting the target was achieved. This is how the simulation model is intended to be used. There are many different alternatives as to how to achieve the goal in the scenario. Improving the radar coverage would push the reconnaissance alert line further back and gain crucial extra minutes for the blue assets to get airborne. Placing the assets at a base that is closer to the target would allow for a lower standby time and yet the same level of protection. Increasing the flight speed of the assets or changing the weapons setup would likewise affect the engagement frontier.

There are many different alternative ways to achieve the same engagement frontier and the pros and cons of each alternative are subject to debate. Thus, it is better to provide a simulation that does predictably what it was asked to do. It is not good for the simulation model to go too deep into detail to collide with different nuances of, e.g., a missiles probability to hit its target. Such probabilistic models are a topic of their own.

6. Conclusions

6.1 Contributions

In this thesis, two simulation models for analyzing the *air defense capability* of a fleet of military aircraft were presented. The air defense capability is a multifaceted issue, and two different attributes regarding the air defense capability were identified and defined in this thesis. These two attributes are the *force fulfillment* and the *engagement frontier*. Both of these attributes relate to a different aspect of the air defense capability.

The force fulfillment considers the case where aircraft are sent from fixed bases to fulfill temporal and spatial force requirements, defined as operation areas. The assets (aircraft) have a limited time that they can spend at the operation area before they need to return to base for refueling. When an asset returns to base, a replacing asset is needed just in time at the operation area. This sets up a problem where the assets need to be circulated between bases and operation areas to maintain the required number of assets at the operation area at any moment of time.

The engagement frontier considers the case where the aircraft need to react to hostile aircraft by approaching and intercepting them. The aircraft are stationed at bases in different states of readiness. Once the hostile activity is detected, e.g., by a radar, the assets get ready, take off and fly towards the approaching hostile aircraft. The airspace may be restricted by ROZs and the assets need to bypass them. Bypassing a ROZ and especially a bundle of them requires finding an optimal route for the bypass. The interception may be done head-on or with missiles. Missiles have a considerable reach and they form a missile envelope for the assets. The problem is to find the spatial intersection of the minimum reach times of the opposing assets or their missile envelopes.

The simulation models share a common domain model, but some parameters of the model for the force fulfillment are irrelevant for the model for the engagement frontier and vice versa. Despite the similarities in the inputs of the models, their purpose and implementation are totally different.

The simulation model for force fulfillment contains an allocation and scheduling algorithm that has a novel heuristic at its decision making core. Actual flight paths are irrelevant and only the flight time between bases and operation areas matter. The number of plateaus at a base and their service time for hot refuel or normal refuel are crucial. The assets' fuel capacity and fuel consumption in cruising flight and at the operation area are essential for the simulation. The outcome of the simulation model for force fulfillment is a flight schedule for each asset, from which the achieved force fulfillment at each operation area can be deduced.

The simulation model for the engagement frontier focuses on flight paths. Fuel consumptions are not of interest, as the interception is a one time mission and fuel can be assumed not to be a limiting factor. Likewise the plateaus and refueling are not included in the simulation model for the engagement frontier. The core of the simulation is the reach times of assets and their missile envelopes. In order to calculate the reach times the whole spatial area is discretized into a large grid of nodes. One problem was to determine how the nodes should be connected in order to keep computational time at a minimum while not jeopardizing accuracy of the model. Any connections that would cut through a ROZ are removed. Explicitly testing whether a single connection cuts through a ROZ is computationally expensive, and testing every single connection would be a waste. Hence, different tuning methods were used to reduce the computational load.

The set of grid points together with their connections, form a graph. The graph is passed to an external optimization algorithm that solves the shortest paths. Once the paths are solved for each asset and each assets' missile envelope, the minimum reach times for the set of assets for both the blue and the red sides can be determined. Finally, the engagement frontier is found by the intersection of the minimum reach times from each opposing side.

6.2 Pragmatic Value

The purpose of this thesis was to build simulation models for analyzing defense capability of a fleet of military aircraft. The simulation models provide a holistic view of the air defense capability and help to utilize the fleet of aircraft effectively by providing analyzes of the capabilities of the fleet.

The simulation for force fulfillment analyzes the performance of the fleet of aircraft and the ground resources. Taking into account multiple operation areas, a large fleet, hot refuels, and limited ground resources (plateaus) is not trivial but with this simulation model, analyzing the force fulfillment is quick and easy. As presented in the examples in this thesis, a spatial batch run is possible where an operation area is iterated over a grid of different positions. This batch run gives a 'big picture' of the air defense capability with respect to force fulfillment. Other types of batch runs are also possible, e.g., investigating the effect of hot refuel or normal refuel times to the force fulfillment. It is possible to run the simulation in a loop such that the refuel time is increased, e.g., by five minutes in each iteration and to compare the outcomes.

The simulation model for engagement frontier is useful for planning pilot readiness and aircraft deployment. Especially in the presence of ROZ in the airspace, determining the engagement frontier is not trivial. It is possible to try out different allocations of assets to bases with different standby times, and then one can compare and analyze the results to choose the alternative that is most appropriate.

6.3 Topics for Future Research

The simulation model for the force fulfillment could be further improved by modeling logistics within the bases. The logistic would include the use of ground personnel and the fuel distribution. In the real world, there is a limited amount of ground personnel and if they need to serve several assets simultaneously, the turnaround times of these assets are increased. The internal fuel distribution may be carried out by fixed pipes or by mobile tanker vehicles. Once the tanker vehicle has refueled an asset, it needs to head back to the base's central fuel reservoir to refill itself. This tanker refill consumes time and thus a plateau may not be able to immediately refuel the next asset, as is assumed in the current model.

The possibility to vary not only the duration of maintenance but also the capacity of the base would add to the usability of the simulation model. It is a plausible real life scenario that the focus of a large military operation shifts according to a plan. Resources will then be moved from one base to another, and the capacity to refuel aircraft would increase at some bases at the cost of reducing at some others. Further considerations would be to change the role-equipment of a multi-role fighter aircraft. This would involve a dedicated asset to be sent to a certain base, and after the role change, the asset would be capable of performing some other type of mission.

A different approach to the allocation algorithm would be to 'throw all the assets in air and start juggling'. In the current model, an allocation involves a complete flight operation of an asset, starting from being fully refueled in a base and ending in being fully refueled in a base. It would be interesting to investigate whether an allocation algorithm where each iteration begins and ends with an asset being at the operation area could be made to work. Assets would thus be allocated to operation areas without prior knowledge to which base they return to. There is constant battle in the current model with the plateaus being reserved just because the next allocation of an asset at the plateau is not known at that particular instant of time. If the situation was turned the other way round that the next allocation to an operation area would be known when the asset is refueled, then the exact moment of time when the plateau will be vacant again will be known and no unnecessary reservation of a plateau will happen.

Some further nuances that could be taken into account would be, e.g., the effect of the weather in taxiing, take-off and landing times. In icy winter conditions, all movement on ground is slower than in clear and dry summer weather.

The major improvement potential in the simulation model for the engagement frontier would be to describe the missile launch and missile strike in more detail. This is however in controversy with the discussion in Section 5.2.4. Taking into account the missile launch and strike realistically would involve a whole another simulation model. There are issues such as getting a radar sighting of the opposing aircraft which is dependent on the specific aircraft model and the angle of approach of the two aircraft. A fighter aircraft has a considerable smaller radar reflection head-on than from the side. Should the aircraft do evasive maneuvers, the strike probability of the missile comes down to aerodynamic properties and physical inertia of the missile. A plausible solution could be to integrate an external simulation model to the model presented in this thesis.

Adding 3D modeling to the current simulation model would allow for a more accurate description of ROZs. Most anti-aircraft missile systems that are launched from the ground have a cone shaped area of influence. Flying lower would thus allow for the asset to be out of sight for anti-aircraft missiles. It would simply be a matter of optimization whether to drop altitude and make a smaller sideways detour, or to keep the altitude and make a larger sideways detour. The flying speed of the aircraft are altitude dependent and this could also be taken into account in the simulation.

References

[1] M. Harju, J. Liesiö, and K. Virtanen. *Spatial Multi-Attribute Decision Analysis: Axiomatic Foundations and Incomplete Preference Information*. Submitted for publication. October 27, 2017.

[2] I. Orenstein. *Guidance System for Air-to-Air Missiles*. U.S. Patent US5938148A, Israel Aircraft Ind Ltd, 1996.

[3] F. P. Oglesby and W. L. Wuster. *Method of Precision Bombing*. U.S. Patent US4281809A, US Secretary of Navy, 1979.

[4] H. Puustinen. *Military Aircraft Routing with Multi-Objective Network Optimization and Simulation*. Aalto University, Department of Mathematics and Systems Analysis, Master's Thesis, 2013.

[5] I. Akgün and B. Ç. Tansel. *Optimization of Transportation Requirements in the Deployment of Military Units*. Computers & operations research, Volume 34, Issue 4, April 2007, pp. 1158-1176

[6] J. O. Royset, W. M. Carlyle, and R. K. Wood. *Routing Military Aircraft With A Constrained Shortest-Path Algorithm*. Military Operations Research, Vol. 14, No. 3 (2009), pp. 31-52

[7] A. K. Nemani and R. K. Ahuja. *Shortest Path Problem Algorithms*. In Wiley Encyclopedia of Operations Research and Management Science, John Wiley & Sons, New Jersey, U.S.A., 2011.

[8] C. A. Floudas. *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford University Press, Oxford, U.K., 1995.

[9] N. K. Jaiswal. *Military Operations Research: Quantitative Decision Making*. Vol. 5. Springer Science & Business Media, New York, U.S.A., 2012.

[10] R. Gopalan and K. T. Talluri. *Mathematical Models in Airline Schedule Planning: A Survey*. Annals of Operations Research, Volume 76(1998), pp. 155-185.

[11] G. Kozanidis, A. Gavranis, and G. Liberopoulos. *Heuristics for Maximizing Fleet Availability Subject to Flight & Maintenance Requirements*. Proceedings

of the International Conference on Applications of Advanced Technologies in Transpotation, Athens, Greece, 2008.

[12] J. Kokkala. *Optimal Allocation of Defensive Fighter Force.* Aalto University, Department of Mathematics and Systems Analysis, Master's Thesis, 2010.

[13] G. Konzadinis, A. Gavranis, and E. Kostarelou. *Mixed Integer Least Squares Optimization for Flight and Maintenance Planning of Mission Aircraft.* Naval Research Logistics, Volume 59(2012), pp. 212-229.

[14] P. Brucker. *Scheduling Algorithms.* No. 4, Springer, Berlin, Germany, 2004.

[15] B. Golany, et al. *Network Optimization Models for Resource Allocation in Developing Military Countermeasures.* Operations Research, Volume 60, Number 1, pp. 48-63, 2012.

[16] F. Neuman. *Methodology for Determination and Use of the No-Escape Envelope of an Air-to-Air-Missile.* Proceedings of the Guidance, Navigation and Control Conference, Minneapolis, U.S.A., 1988, p. 4137.

[17] A. Davidovitz and J. Shinar. *Two-Target Game Model of an Air Combat with Fire-and-Forget All-Aspect Missiles.* Journal of Optimization Theory and Applications, Volume 63(1989), Issue 2, pp. 133–165.

[18] S. Nahmias and Y. Cheng. *Production and Operations Analysis.* Volume 6. New York: McGraw-Hill, 2009.

[19] D. Gleich. *MatlabBGL.* https://github.com/dgleich/matlab-bgl. Referenced 23rd March 2018.

[20] J. Siek, L. Lee, and A. Lumsdaine. *The Boost Graph Library: User Guide and Reference Manual.* New Jersey: Pearson Education, 2001.