4G Network Patch using Drones

Diogo Moura Gonçalves
diogo.m.goncalves@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

Novembro 2018

Abstract

The main goal of this thesis is to study and propose a new approach to network patching in mobile networks. The developed solution implements network patching making use of UAVs. The biggest challenge is to define the best positions to deploy the UAVs. A genetic algorithm was developed to tackle this problem. This algorithm takes advantage of the non-uniform distribution of data rate requirements among users to better position the UAVs. The UAVs make use of distant base stations to provide extra coverage or capacity in a given area. A simulator was developed to access the quality of the positions. The developed algorithm showed to be an effective way to compute the positions of the UAVs. Two metrics were analyzed to measure the effectiveness of the algorithm: the number of served users, and the total data rate served. Compared to other existing solutions, the algorithm was able to keep the number of served users in every scenario tested, and increase up to 24% the data rate provided. **Keywords:** Network Patching, UAV, LTE, Genetic Algorithm.

1. Introduction

If we consider the evolution of the telecommunications systems in the last years, we must acknowledge that this technology has grown at a very fast pace and has had a massive adoption by the general population all around the globe. In 2017 there were already 7.8 billion mobile subscribers worldwide, and that number is expected to grow up to 9.1 billion by 2023 [15]. If we realize that the world population is under the 8 billion mark, these numbers become even more staggering.

The demand for faster and better telecommunications systems does not come solely from the private market. Public security entities grow ever more dependent on communications to efficiently operate. So it is important to develop either more resilient systems to these kind of situations, or better solutions for replacing the network elements that are more exposed.

As this is a very competitive market, telecommunications operators are permanently working for increasing their coverage, their capacity and their availability. There are however a number of factors that can compromise these goals:

- ***Environmental Factors:*** Meteorologic/Natural events that destroy or compromise some parts of the network. These events can cause a temporary lack of coverage.
- ***Usage Peaks:*** Engineers design the network for a certain amount of users in a specific area. Sometimes there are special events that gather an unusual amount of people in a relative small space. That can cause lack of capacity in a cell.
- ***Components Breakdown:*** Some components of the network sometimes stop working properly causing disturbances on the telecommunications operation.

1.1. Solution Overview

There are several ways to correct network outages. The simplest techniques involve tilting the existing antennas of base stations in order to optimize the coverage and capacity of a certain cell. This method is rather limited in its effect, since there is not an increment to the network resources in that area.

Another way to compensate for outages is to deploy more base stations. This is a time and resource consuming operation.

Overall, these approaches are non-adaptive over time, the planning takes place without knowing where the users will be and they are time and resource consuming. Considering this, due to their capability to freely move and deploy on-demand, drones are expected to help alleviate some of these problems in the future. In this work, the terms drone and UAV are used interchangeably.

The use of UAVs in telecommunication patches allows engineers to re-frame a problem that has always existed. With this new tool, it is possible to radically change the position of the antennas in real time. Currently, only the positioning of the users is used to chose the positions of UAVs, but since this process will be taking place in real-time, there are other factors that could help achieve a bigger network efficiency. One of those cases is the required data throughput that users are demanding at a particular moment.

The driver of this work is to improve the signal of users that require a large amount of resources, specially in those situations where we do not need to compromise on the quality of service of users who need less resources.

The problem of finding the best position for the UAVs can be formulated in an optimization problem. Since there are various factors that we want to optimize, this work uses a genetic algorithm to tackle the problem.

2. Background

This section is going to address background and related work on three topics that concur for this thesis: LTE, UAVs and GAs.

2.1. LTE Fundamentals

This section was mainly based on [18], [6], [19], [10] and [12].

2.1.1 LTE Architecture

Figure 1 shows the High Level Architecture of the LTE system, composed of three main components: the UE, the E-UTRAN and the EPC.

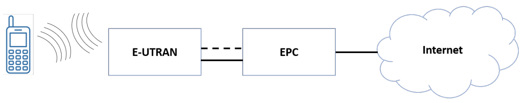


Figure 1: LTE architecture. Adapted from [6].

Figure 2 shows the E-UTRAN architecture which handles the radio communications between the mobile and the evolved packet core. This element has just one component, the eNB.

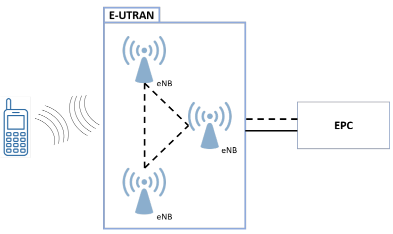


Figure 2: E-UTRAN architecture.

Each mobile is allocated to only one eNB at a time, and each eNB is responsible for controlling the mobiles in one or more cells. The eNB has two main functions: (1) sends radio transmissions to all its mobiles on the downlink and receives transmissions from them on the uplink; and (2) controls the low-level operations of all its mobiles that relate to those radio transmissions.

Figure 3 shows the main components of the EPC. The HSS is a central database that contains information about all the network operator's subscribers.

The P-GW is the EPC point of contact with the outside world. From here data packets are sent to the internet or the IP multimedia subsystem.

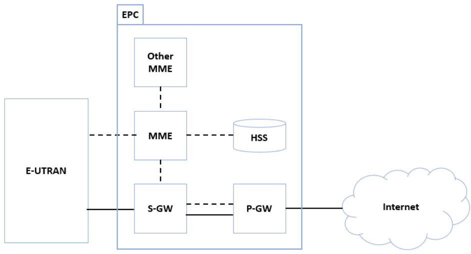


Figure 3: EPC architecture. Adapted from [6].

The S-GW is responsible for routing and forwarding the data between the base station and the P-GW. Each S-GW serves all the mobiles of a certain geographic region.

The MME controls the high-level operation of the mobile such as security and the management of data streams that are unrelated to radio communications. As the S-GW, the MME is responsible for a certain geographic region.

2.1.2 LTE Radio Interface

OFDM works by dividing the available bandwidth in a group of orthogonal sub-carriers. These sub-carriers are generated in a way that in every peak of any of them in the frequency spectrum, the contribution of all the others sub-carriers is zero. This technique allows a dramatically reduction of the guard band between the sub-carrier's frequencies.

LTE uses OFDMA in the DL, a variation of OFDM, where users are assigned resources in the time and frequency domains.

In the UL, LTE uses SC-FDMA. We can see this as a special case of OFDMA where the signal is pre modulated before being modulated in OFDMA.

Another technique that is broadly used in LTE is MIMO. This means that there will be multiple antennas that can be used for different purposes.

Resource allocation in LTE is based on RB, which is the minimum allocation a UE can get. The smallest unit of a RB is the RE which translates to using one sub-carrier (15 kHz) to transmit one radio symbol. One RB uses 12 sub-carriers during 0,5 ms. In 0,5 ms it is possible to transmit 6 or 7 symbols, depending on the cyclic prefix, which leads to one RB having $7 \times 12 = 84$ symbols.

Depending on the modulation used, one symbol is equivalent to transmit two, four or six physical channel bits. The modulation is chosen taking into account several factors(e.g., the quality of the physical channel).

LTE can be deployed using various RF channel configurations. There are 27 bands specified for FDD and 12 bands for TDD. LTE can coexist with the previous 3GPP technologies. In Europe, mobile operators have over 600 MHz of spectrum available for LTE, including the 800 (LTE800), 900, 1 800 (LTE1800), 2 100 and 2 600 MHz bands (LTE2600).

2.1.3 LTE Scheduler

This section also considers the information in [8].

One important component of the LTE system is the scheduler. Since in LTE resources can be allocated both in frequency and time domains, this component is responsible for distributing the RBs by the UEs.

These are some examples of generic schedulers:

- ***Maximum Throughput Scheduler:*** The network tries to exploit channel variations to maximize the cell throughput. This scheduler starts by satisfying completely the users who have the best channel conditions.
- ***Blind Equal Throughput Scheduler:*** This policy distributes the resources in a Round Robin approach.
- ***Proportional Fair Scheduler:*** This scheduler tries to balance the two previous schedulers by allocating resources taking into account both the maximum rate achievable by a UE, and the average rate the UEs get.

2.1.4 Capacity

An approximate relation between LTE's bandwidth and the number of resource blocks available is given by the expression:

$$N_{RB} = \frac{B_{ch[kHz]}}{B_{RB[kHz]}} \times \frac{P_{Bch[\%]}}{100}, \qquad (1)$$

where:

- N_{RB}: number of RB;
- $B_{ch[kHz]}$: channel's bandwidth ;
- $B_{RB[kHz]}$: RB's bandwidth, which is 180 kHz;
- $P_{Bch[\%]}$: channel's bandwidth used percentage, $\sim$ 90%.

Once the number of RBs each user gets is known, we have to assess the amount of information each resource block can transport to compute the throughput. This quantity depends on the modulation used, which depends on the quality of the radio channel. This constrains will result in a different spectral efficiency for every user, being the theoretical maximum defined by Claude Shannon:

$$R_{b,max[bit/s/Hz]} = log_2(1+\rho_N), \qquad (2)$$

where:

- $R_{b,max[bit/s/Hz]}$: maximum data rate per hertz;
- ρ_N: SNR (in linear units).

2.1.5 Services and Performance Parameters

3GPP divided services into 4 classes: Conversational, Streaming, Interactive and Background. These classes are used to prioritize the data flows, allowing eNBs to treat data traffic differently.

In Table 1 there are examples of some services, their service class, and their approximate minimum, medium and maximum throughput.

Table 1: Throughput By Service. Extracted from [17].

Service	Service Class	Throughput [kbps/s]		
		min	med	max
VoIP	Conversational	8	32	64
E-mail	Background	10	100	1333
Smart Meters	Background	-	200	-
File Sharing	Interactive	200	1600	4 444
Video Streaming	Streaming	500	5000	10 000

2.1.6 Relaying

Relaying is the technique in which, by the addition of network nodes we complement the network of eNBs. This allows expansion of coverage or increase of capacity.

Modern relays in 4G designated RN, are network nodes connected wirelessly to a source eNB. Since they are under the full control of the radio access network, they allow similar monitoring and remote control capabilities as for an eNB.

This is the terminology relating RNs that 3GPP introduced:

- **Donor eNodeB/cell:** The source eNodeB/cell from which the NR receives its signal.
- **Relay cell:** The coverage area of the RN.
- **Backhaul link:** The link between the donor eNodeB and the RN.
- **Access link.** The link between the RN and a UE.
- **Direct link.** The link between the donor eNodeB and a UE.

- **Inband/outband.** An inband RN uses the same carrier frequency for the backhaul link as for the access link; otherwise, the RN is said to be outband.

- **Half/full duplex.** A half-duplex RN cannot receive on the backhaul link at the same time as transmitting on the access link, and vice versa, whereas a full-duplex RN has sufficient antenna isolation to be able to operate without this restriction. This distinction applies to inband RNs only, since outband RNs are always full-duplex.

- **Donor and coverage antennas.** At the RN, the donor antenna(s) are used for the backhaul link, while the coverage antenna(s) are used for the access link. In some cases, the physical donor and coverage antennas may be the same.

Backhaul's link on inband RNs consume radio resources, thus reducing the capacity of the RN. When RNs operate in half-duplex, there are also other complications for the system design. Outband relaying with full-duplex operation increases the relay cell capacity and simplifies the system design. This increases the system cost since the last configuration requires two isolated antennas, and the availability of a second carrier frequency.

RNs can belong to one of three categories, depending on the functionality layers they provide. Layer 1 RNs are simple repeaters and may include some baseband processing such as FEC. Layer 2 RNs provide Medium Access Control (MAC) functions such as scheduling. Layer 3 RNs have their own PCI signalled by the PSS/SSS, and all of Layer 1 and Layer 2 functions are supported by the RN.

2.2. UAVs

UAVs are adapted to the specific task at hands, which makes them highly specialized and efficient tools.

Almost all the control functions of an UAV can be delegated to an auto-pilot and modern UAV controller systems are turning possible the collaboration between several drones. Some centralized controller may be needed.

Applying UAVs to telecommunications seems to be a logical step, in [11] the authors studied the use of UAVs in emergency scenarios. UAVs are generally a good fit for these scenarios, however, the authors raise relevant questions about power consumption in these peculiar situations where basic infrastructures may be compromised.

In [16] the authors studied the relaying in tactical situations, proclaiming increases by a factor of two in throughput, and by 67% in connectivity when compared to the ad hoc ground network without using UAVs.

In [1] and [5] the authors focused on studying the impact of altitude in the UAV performance assuming the drone-cells are isolated from other base stations. The altitude of the drone is responsible for an important trade-off between achieving a LoS radio channel with the users, which is more probable with higher altitudes, and achieving a lower path loss, which is lower at smaller distances, hence, lower UAV's altitudes.

In [4] the integration of drone-cells in the existing LTE network is addressed, making the case that this is not a simple task, and ending up suggesting a multi-tier approach to this problem where high level UAVs, flying at a higher altitude, give support to lower level UAVs. They also envision a decoupling of this activity (UAV deployment) from the current network providers, and becoming themselves infrastructure providers and mobile virtual network operators providing services to the current network operators.

In [7] a clustering algorithm of type K-Means is used for addressing the problem of UAV localization in a scenario where UAVs and fixed based stations coexist.

2.3. Genetic Algorithms

A GA is a meta-heuristic algorithm, which is generally used to solve optimization problems. They use a combination of random choices and knowledge of previous results to address the search space. This section is based in [14] and [13].

2.3.1 Components

The higher level component of genetic evolution is the generation. Generations are composed of individuals that have genetic material which determines their characteristics/traits. This genetic material is organized into chromosomes and genes, one individual has one or a group of chromosomes, which are composed of genes. The gene is the most basic component of genetics, this structure stores the information about one specific trait of the individual.

Developing a specific genetic algorithm, requires being able to frame our problem in a way that we can represent it with these basic components. This task should not be underrated since it is known that the chosen representation will directly impact the performance of the algorithm. Different authors will advocate for different representations for the sake of two things: convergence speed and the ability to avoid locally optimal solutions.

2.3.2 Evolution

In nature, the fitness of an individual is defined as the level of adaptation to a certain environment. The more fit an individual is, the more likely it is to have a longer life, to reproduce and to generate descendents. It is called natural selection to the process of elimination of lesser fit individuals. In the genetic algorithms world, the fitness value is a representation of how well a certain solution responds to the initial problem. The design of this ranking tool, is one of the most difficult challenges in designing GAs.

The main genetic operators are the following:

- **selection**, how we choose individuals to reproduce;
- **crossover**, how does the mixing of genetic material between two parents happen;
- **mutation**, the random process of changing gene values.

Figure 4 shows the genetic evolution process which starts with a random group of individuals, the first generation. Then we iterate the following process to generate the subsequent generations. Choose two individuals from the previous generation, crossover their genetic material to generate a new individual, randomly mutate its genes and add this individual to the new generation. This process is iterated until the next generation has the same number of individuals as the previous.

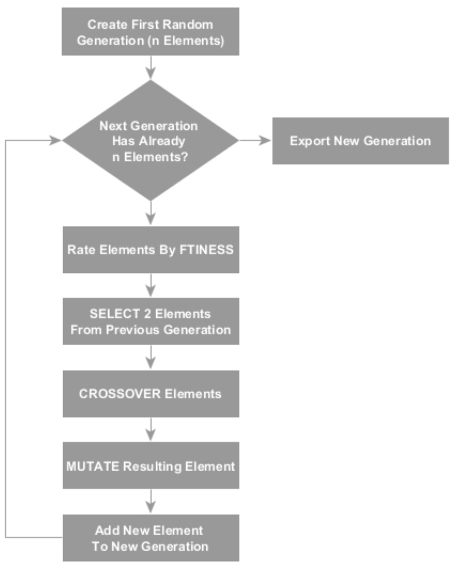


Figure 4: Evolution Process.

In the GA context, convergence speed is the required number of iterations/generations we need to reach a stable value of fitness. In theory, one can use this parameter to know when to stop iterating generations. In every generation one can search for the best adapted solution, and if its value of fitness is unchanged for a large number of iterations one can assume the optimal solution has been reached and stop iterating. The time needed for this to happen is directly connected to the search space being addressed. The larger the number of possible options for an individual, the longer it will take to reach this optimal solution.

There is no certainty that a genetic algorithm will find the best solution. By definition this is a stochastic algorithm, and so, it is possible that the algorithm converges to a local optimum. However, there are ways to improve the odds of finding this perfect solution. One of them is to make an effort to design the genetic operators in a way that the generations are kept with a fair amount of diversity, making sure that solutions with a low value of fitness are not deleted as soon as they appear because they may be, globally, a bad fit for the problem, but they may contain some of the genes present in the global optimum.

2.3.3 Operators

The next subsections will address the various operators present in the a genetic algorithm.

Selection Operator

The selection operator intends to simulate the mating behavior of the individuals. This process is designed to distribute the selection probabilities to the different individuals.

To create a new generation, we need to choose pairs of individuals of a previous generation to crossover and generate a new individual. According to [3] it is possible to name some categories for this selection methods.

A selection method can be *dynamic*, meaning the selection probabilities are proportional to the actual fitness-values; or it can be *static* meaning the individual is selected according to its position in the fitness ranking.

A selection method can be *preservative*, which attributes a non-zero probability to every element in a generation; or *extinctive* which dictates that some elements are excluded right away from the reproduction operation.

When a selection method is *extinctive*, it can be *right extinctive*, if the elements with zero probability are the low performers, or *left extinctive*, if the elements with zero probability are the top performers. Although this last category may seem counter productive, it is very useful if one wants to make sure that the generations are kept with a certain degree of variability, thus preventing the algorithm to converge too soon to a local optimum.

Aside from the distribution of probabilities, it is possible to define a selection method as being *elitist* or *pure*. In the elitist method, some of the top performers skip the reproduction process and are directly included in the next generation.

Fitness Function

In genetic algorithms, the fitness function is the ranking tool used to distinguish the individuals that better answer to the initial problem, from those who do it worse. The more accurate the fitness function is, the best our final solution will be.

Crossover Operator

When a binary representation is used, the simplest way to implement the crossover is to choose a random point in the chromosome and all of the genes before that point are copied from one parent, and all the others are copied from the other parent. However, it is possible to implement a two point crossover operation. In this case, two points of the chain are chosen and the segments between them are exchanged. Lastly, it is important to check if it makes sense, in a specific implementation, to allow a gene to be cut in half in this operation.

Mutation Operator

Along with crossover mutation is the main instrument of disruption and innovation in an evolution process. The balance of amount of mutation and crossover is extremely important to the convergence speed/global optimum trade off.

2.3.4 Stop Conditions

Firstly, one can impose a limit to the number of iterations the algorithm runs. Then, one can impose a limit to the time the algorithm is running. And finally the algorithm can stop when the chance of achieving a significant change in the fitness values is very low.

2.4. K-Means Drone Disposition Algorithm

K-Means is a well known clustering algorithm, which works by minimizing the distance of all the points/UEs to the clusters' centroids coordinates. This work implements a variation of K-Means called K-Means++. This algorithm will be used as a baseline for the GA, and the specific implementation details were based in [9], [7] and [2].

The time complexity of the K-Means algorithm is given by $\mathcal{O}(MNKI)$, where M is the number of points, N is the number of dimensions, K the number of clusters, and I the number of iterations.

3. Implementation

This section describes the implementation decisions of this work.

3.1. System's Architecture

Figure 5 shows the general architecture of the proposed solution.

In order to control the positioning of the UAVs the network would need an UAV Controller which would be located in the donor eNB. This element is the one that knows the UAV current coordinates,

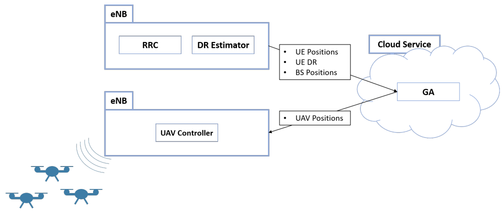


Figure 5: Architecture of the system.

and, upon receiving the desired UAV Positions, can instruct the UAVs to move accordingly.

The decision about the UAV coordinates would take place in a cloud based service. The GA is the algorithm chosen to perform the decisions about the UAV coordinates. UEs connect either to the assigned UAV or, if the UAVs run out of RBs, to the nearest BS. The RRC and DR Estimation for a certain user is located in its corresponding servicing eNB, and the UAV Controller is in the eNB that is coordinating the UAV operation.

3.2. Genetic Drone Disposition Algorithm

As discussed in 2.3.1 the genetic algorithm is divided into 4 structural components: genes, chromosomes, individuals and generations.

In this work, the genes implement the coordinates of a single drone. The coordinates are represented in bits, and the implementation makes it possible to choose the level of granularity we want for our coordinates. This translates into changing the number of bits used to represent each coordinate.

Every chromosome has the same number of genes, and since each gene encodes one drone, the number of genes will be equal to the number of drones. In our problem, each individual would have only one chromosome. This is why there is no distinction between these two structures.

The generation represents the population at any given moment and the number of individuals in each generation is also configurable.

The selection method implemented is *dynamic*, *preservative* and *elitist*. The number of champions is an entry parameter.

The selection process is achieved by assigning to each individual a value of relative fitness:

$$f_i = \frac{F_i}{\sum_{i=0} F_i}, \tag{3}$$

where:

- F: absolute fitness value;
- f: relative fitness value.

After the selection process, the two chromosomes are crossedover and generate descendants.

The fitness function is used by the GA to classify the individuals. The algorithm runs the Telco

Module for every individual and then uses these results to classify the UAV configurations. The Telco Module will be further explored.

Three different fitness functions are implemented:

- Served Users: number of UEs being serviced at the same time;
- Total DR: that returns the sum of DR obtain by all UEs;
- Proportion: the sum of the proportion, obtained by each user, between the required DR and the obtained DR:

$$F_i = \sum \frac{DR_{obt}}{DR_{req}}, \tag{4}$$

where:

- DR_{obt}: Obtained Data Rate by a user;
- DR_{req}: Required Data Rate by a user.

For the crossover the algorithm randomly chooses a point in the gene sequence to perform the crossover. All the genes before this point will come from one parent and all of the following will be from the other parent.

There are two ways of implementing this operation: (1) two parents are used to generate two descendants, and (2) two parents are used to generate one descendant. This work implements the second option.

Mutation is divided into two sub-parameters: the *mutation rate* and *the mutation span*.

The mutation operation is applied right after a new individual is created. The *mutation rate* is the probability of that new chromosome being mutated. The *mutation span* is how different should the mutated gene become.

This work implements a stop condition that counts the number of iterations performed, and when a pre-defined threshold is achieved the algorithm stops.

3.3. Telecom Model

Figure 6 shows the architecture of the algorithm that implements the telecom model. This algorithm starts by receiving the coordinates of all the UEs, all the UAVs and all BSs. Then it iterates through all UEs and assigns them to the available UAVs in order to allocate all of their capacity. The remaining UEs will be serviced by the BSs.

Figure 7 (a) shows the architecture of the UAV scheduler. This scheduler is a *Maximum Throughput Scheduler*.

Figure 7 (b) shows the architecture of the BS scheduler. This scheduler is a *Blind Equal Throughput Scheduler*.

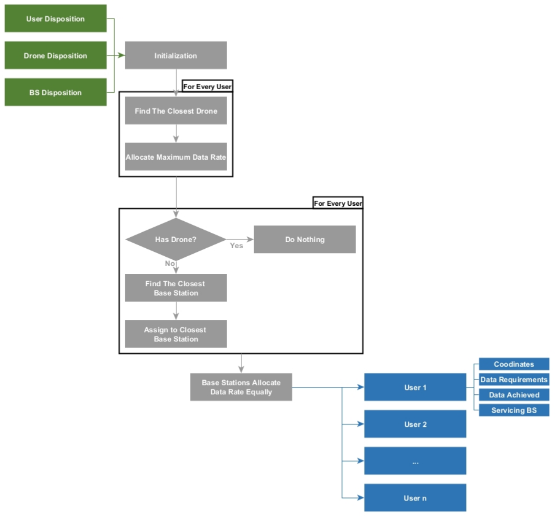
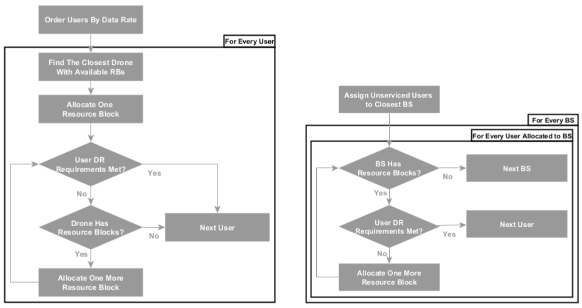


Figure 6: Telecom Model Architecture.



(a) Drone Scheduler (b) Base Station Scheduler

Figure 7: Resource Allocation Schedulers.

It is important for both of these schedulers to make a prediction of how much data rate the UEs are going to obtain. There are mainly three links that need to be addressed for this purpose: (1) the Backhaul Link, which will be in LoS, and will be assumed to have enough capacity to offload the UAVs' necessities, (2) the Direct Link in LoS by an UAV and (3)the Direct Link in NLoS by the BS. The following subsection will discuss both implementations.

3.3.1 Propagation Models

The goal of this sub-section is to describe how to arrive at the data rate a UE obtains. This sub-section is based in [12] and is implemented in the Telecommunications Module. To arrive at this value one needs to calculate some quantities, the first one is the available power at the receiver, which can be obtained as:

$$P_{r[dBm]} = P_{t[dBm]} + G_{t[dBi]} + G_{r[dBi]} - L_{p[dB]}, \tag{5}$$

where:

- $P_{t[dBm]}$: transmit power;
- $G_{t[dBi]}$: gain of the transmitting antenna;
- $G_{r[dBi]}$: gain of the receiving antenna;
- $L_{p[dB]}$: path loss.

The used gains in the simulator were respectively 15 dBi and 0 dBi. We assumed BSs transmitted at 48 dBm and the UAVs transmitted at 43 dBm.

The path loss for the LoS links, considering the average power decay, is given by:

$$L_{p[dB]} = 32.44 + 20\ log(d_{[km]}) + 20\ log(f_{[MHz]}) + 10\ a_{pd}\ log(d_{km}), \tag{6}$$

where:

- $d_{[km]}$: distance;
- $f_{[MHz]}$: frequency;
- a_{pd}: average power decay, 2.

For the NLoS links the path loss can be calculated using the Okumura-Hata model. The deployment of drones intends to increase capacity, so it is expected that operators choose to use the frequency bands that give the most capacity, the 1800 MHz and the 2600 MHz. Since Okumura-Hata only models scenarios with frequencies ranging from 150 MHz to 2000 MHz, we opted to use the extension COST 231 to Hata's model which models scenarios with frequencies ranging from 1.5 GHz to 2 GHz. According to this model, the path loss can be calculated from:

$$\begin{aligned} L_{p[dB]} = {} & 46.30 + 33.90\ log(f_{[MHz]}) \\ & - 13.82\ log(h_{be[m]}) \\ & + [44.9 - 6.55\ log(h_{be[m]})]\ log(d_{[Km]}) \\ & - H_{mu[dB]} + C_{m[dB]} \\ & - \sum correction factors, \end{aligned}$$

where $h_{be[m]}$ is the effective height of BS antenna;

$$C_{m[dB]} = \begin{cases} 0, & smallcity; \\ 3, & urbancentres; \end{cases} \tag{7}$$

$$H_{mu[dB]} = \begin{cases} [1.10\ log(f_{[MHz]}) - 0.7]h_{m[m]} \\ -[1.56 \log(f_{[MHz]}) - 0.8], \\ smallcitty; \\ 8.29\ log^2(1.54\ h_{m[m]}) - 1.0, \\ f \leq 200 MHz, \\ large\ city; \\ 3.20\ log^2(11.75\ h_{m[m]}) - 4.97, \\ f \geq 400 MHz, \\ large\ city; \end{cases} \tag{8}$$

where $h_{m[m]}$ is the UE's height.

To apply this model to an area that is somehow open spaced, we applied the correction factor for quasi open areas K_{op}, which for our transmitting frequencies yields 23 dB.

3.3.2 Link Capacity

An estimation of the data rate obtain by the UE can be obtained by:

$$r_b = B_{av[kHz]} \times R_{b,max[bit/s/Hz]}, \tag{9}$$

where:

- $B_{av[kHz]}$: is the available bandwidth for a specific UE;
- $R_{b,max[bit/s/Hz]}$: maximum data rate per hertz.

As mentioned before the R_b, max can be calculated by the expression 2, and the B_{av} is a result of the amount of resource blocks allocated to the UE. Manipulating the expression 1 we arrive at the UE's available bandwidth given by:

$$B_{av[kHz]} = N_{RB} \times B_{RB[kHz]} \times \frac{100}{P_{Bch[\%]}}, \tag{10}$$

where $P_{Bch[\%]}$ is the maximum data rate per hertz.

The Signal to Noise Ratio (ρ_N), is a function of the signal power received and the noise and interference:

$$\rho_N = \frac{P_{r[W]}}{N_{[W]} + I_{[W]}}, \tag{11}$$

where:

- $P_{r[w]}$: power received;
- $N_{[W]}$: thermal noise;
- $I_{[W]}$: interference.

We will assume a very low interference in comparison with the thermal noise due to the LoS and an efficient management of radio resources.

Finally, the thermal noise is given by:

$$N_{[W]} = K_{[m^2\ kg\ s^{-2}\ K^{-1}]} T_{[K]} B_{av[Hz]}, \tag{12}$$

where:

- $K_{[m^2\ kg\ s^{-2}\ K^{-1}]}$: boltzmann constant, $1.38064852x10^{-23}\ m^2\ kg\ s^{-2}\ K^{-1}$;
- $T_{[K]}$: temperature = 25.

3.4. Simulator

The Figure 8 illustrates the general architecture of the developed simulator. Initially, the simulator creates the UE disposition inside the scenario dimensions. Once that is done, the UE disposition is used by two algorithms which output a drone disposition each. The end goal is to compare the performance of these two drone dispositions.

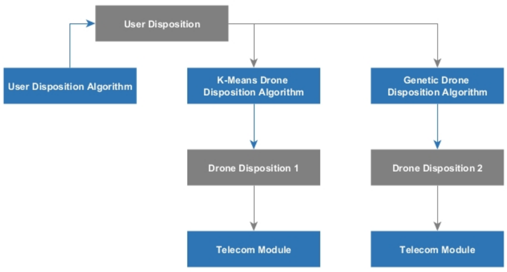


Figure 8: Simulator's architecture.

In order to compare the two drone dispositions, the simulator calls the Telco Module.

3.4.1 Simulator's General Components

The simulator's scenario is composed of a rectangle with customizable dimensions. This is the area where UEs are deployed. A UE has a set of coordinates and a data rate requirement. The drones will also be deployed inside this window. It is also possible to include a representation of multiple Base Stations (BS)/macrocells.

The simulator assumes the existence of fixed based stations that will provide the Backhaul link for the UAVs. The relays are Outband and Full Duplex since they can communicate with the UEs and the Donor BS at the same time.

We have decided to implement three categories of data rates for UEs: low demand, medium demand and high demand, and they will respectively be represented by 32 kbps/s, 200 kbps/s and 5 Mbps/s.

3.4.2 UE Disposition Algorithm

The simulator uses a Poisson distribution to generate the UE's positions. It assumes some UEs will be agglomerated in small groups, or hotspots, and then distributes the remaining UEs uniformly throughout the scenario. Figure 9 depicts the architecture of this process.

The number of UEs in each hotspot will be the result of a Poisson distribution with the average provided as an input argument.

For the distribution of DR requirements, two possible implementations were considered: Uniformly Distributed and Non-Uniformly Distributed. In the

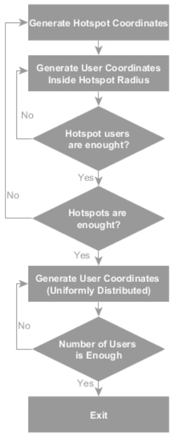
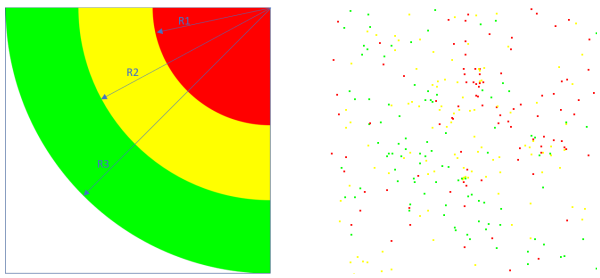


Figure 9: Poisson UE Disposition Algorithm Architecture.

first case, the three possible levels for DR are distributed uniformly throughout the scenario. In the second case, a random point is chosen from one of the corners of the scenario, then three circles with different radii are drawn around that point, as it is depicted in Figure 10 (a). These will represent areas of a certain type of data rate requirement. When a user falls in one of those areas, it will be more probably assigned the respective data rate. If a UE is inside one of those areas it has 50% probability of getting the level corresponding to that area and 50% probability of getting a random level of DR. If a user is not inside any the colored areas, it has equal probability of getting one of the three levels.

The output of this process is a list of UEs, each containing two coordinates and a data rate requirement in bit/s. Figure 10 (b) shows an example of an output from the UE Disposition Algorithm Non-Uniformly Distributed. This configuration has 300 UEs, 5 hotspots and, on average, each hotspot has 20 UEs. The color of the points represents the different data rate requirements of each UE: red UEs have the higher data rate requirements, the green UEs have the lower data rate requirements and the yellow UEs have a medium data rate requirements.



(a) Data Rate Areas (b) User Disposition Example

Figure 10: User Disposition.

4. Results

This section starts by presenting some of the possible configurations of the simulator and presents the simulation results of this work.

4.1. GA Experimental Parameters

The results of a GA are dependent on the right configuration of its parameters. Several experiments were performed to determine these values. These experiments yielded that at generation 300 the fitness value was unlikely to significantly change, the mutation rate should be set to 30 %, the population size to 50 individuals and the number of bits per coordinate at 7.

4.2. Simulations Description

In all the simulations both the GA and the K-Means++ were deployed. The references to comparing the two algorithms, should be understood as the comparison between these two algorithms.

The scenario had a 300m x 300m dimensions, yielding a 90 000 m2 area. All simulations deployed two macrocell BS at the coordinates (150, 400) and (400, 150) and had 5 UAVs available.

Two performance metrics were evaluated:

- **Served DR**: this is the sum of all the DR obtained by each UE;
- **Served Users**: this is the sum of all UEs which were allocated at least one RB.

The first goal of the simulations was to understand the impact of the two possible UE's disposition regarding how DR is distributed throughout the UEs. Both the Uniform and Non-Uniform Distributions were tested.

The second goal of the simulations was to understand how different amounts of UEs affect the effectiveness of the two algorithms. Two different amounts of UEs were deployed:

- **Limit Case**: the number of UEs is expected to fully consume, or even marginally exceed the network resources;
- **Excess Case**: the number of UEs clearly exceeds the limit that the network is able to serve.

The limit case had 150 UEs in total, 5 hotspots, each hotspot had a 40 meters radius and on average 15 UEs. The Excess Case had 230 UEs, 5 hotspots, each hotspot had a 40 meters radius, and on average 20 UEs.

The third goal was to compare the performance of the two algorithms. The K-Means++ worked as a baseline. The last goal of the simulations was to find the fitness function that yields the best results, and compare those results with the baseline.

Every combination of these possibilities was tested over 1000 Monte Carlo trials.

The results provided by the Telecom Module were analyzed, and the averages were computed with a level of significance of 95%.

4.3. Served Users Comparison

Tables 2 and 3 show the results of the simulations for the number of served users.

Table 2: Served Users by Algorithm in the Non-Uniform distribution.

Users	Algorithm		Served Users
			Average
Limit	K-Means++		136.91
	GA	Served Users	150 (+8.73%)
		Served DR	150 (+8.73%)
		Proportion	150 (+8.73%)
Excess	K-Means++		225.4
	GA	Served Users	230 (+3.07%)
		Served DR	219.51 (-3.92%)
		Proportion	229.98 (+3.06%)

Table 3: Served Users by Algorithm in the Uniform distribution.

Users	Algorithm		Served Users
			Average
Limit	K-Means++		141,56
	GA	Served Users	150 (+5.63%)
		Served DR	150 (+5.63%)
		Proportion	150 (+5.63%)
Excess	K-Means++		215,59
	GA	Served Users	229,98(+6.26%)
		Served DR	210,64 (-2.15%)
		Proportion	229,9 (+6.22%)

4.4. Data Rate Comparison

Tables 4 and 5 show the results obtained for the served data rate.

Table 4: Served Data Rate by Algorithm in the Non-Uniform distribution.

Users	Algorithm		Served DR	Served Users
			Average [bit/s]	*Average*
Limit	K-Means++		2427205440	73.96
	GA	Served Users	2437485884	74.27 (+0.31%)
		Served DR	3238186945	98.67 (+24.71%)
		Proportion	3216277558	97.99 (+24.04%)
Excess	K-Means++		2808253150	50.92
	GA	Served Users	2856064736	51.79 (+0.87%)
		Served DR	3560569954	64.56 (+13.64%)
		Proportion	3087416760	55.98 (+5.06%)

4.5. Results Analysis

Comparing the results between the K-Means++ and the Genetic Algorithm, these results show that the GA has a better performance for almost every scenario. It is however important to consider the trade-off between having a higher global data rate being consumed, and having every user with access to the telecommunications service.

If one considers only the Limit Scenario, the fitness function that yields the best results is the

Table 5: Served Data Rate by Algorithm in the Non-Uniform distribution.

Users	Algorithm		Served DR Average [bit/s]	Served Users Average
Limit	K-Means++		2551410483	83,94
	GA	Served Users	2363379351	77,80 (-6.14%)
		Served DR	3035138097	99,08 (15.14%)
		Proportion	3000172608	98,83 (14.88%)
Excess	K-Means++		2883848357	61,83
	GA	Served Users	2830697170	60,71 (-1.11%)
		Served DR	3411952986	73,27 (11.45%)
		Proportion	3017640540	64,97 (+3.14%)

Served DR, which maximizes the total data rate serviced by the network. This happens because the performance of the three fitness functions are similar in the number of serviced users metric and the Served DR fitness function is slightly better than the Proportion in the Served DR metric.

However, if one considers the Excess Scenario, the Served DR fitness function is no longer the obvious choice since it yields now the worst results in the Served Users metric, worse than the K-Means++ algorithm. For this scenario, the wisest choice would be the Proportion fitness function, which compromises a little bit in Served DR, but keeps a high number of Served Users.

Taking this into consideration, the best fitness function would be the Proportion.

These results show that in the best case scenario, the GA outperforms the K-Means by 24%. Another relevant conclusion is that when the amount of unserviced DR increases too much (the Excess cases), the difference between the two algorithms also becomes less expressive.

The big conclusion of all these simulations is that if the fitness function is properly chosen, even in the worst case scenarios, the GA outperforms the K-Means.

5. Conclusions

In this thesis two different user dispositions were used to compare the two algorithms. The GA showed to be better in both dispositions, in particular in the Non-Uniform disposition. Two different quantity of UEs were tested. The GA remains better in both cases, and is particularly good in the Limit case.

Two metrics were used to measure the quality of the algorithms: the number of Served Users and the Served DR. The GA was capable to keep the Served Users and still outperform the K-Means++ in the Served DR.

As for future work, it would be interesting to perform some extensions to this analysis. First, it would be useful to develop a better method to go from one state of UAV positions to another. This could result in lower quality UAV positioning, trading off with the amount of time the UAVs could stay in the air by reducing the effort needed to reallocate.

It would also be interesting to evaluate the performance of other network indicators besides the number of served users and the average data rate per user. For example how would the delay be affected by a deployment of this type? And what would be the lifetime of UE's batteries when they are closer to the antennas?

Finally, it is important to better study the link between the UAVs and its serving BS. It is important to access the feasibility and trade-off of using the same frequencies as the ones used in the normal telecommunications operation, or if it would be better to use specific frequencies.

References

[1] A. Al-Hourani, S. Kandeepan, and S. Lardner. Optimal lap altitude for maximum coverage. *IEEE Wireless Communications Letters*, 3(6):569–572, 2014.

[2] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.

[3] T. Bäck and F. Hoffmeister. Extended selection mechanisms in genetic algorithms. 1991.

[4] I. Bor-Yaliniz and H. Yanikomeroglu. The new frontier in ran heterogeneity: Multi-tier drone-cells. *IEEE Communications Magazine*, 54(11):48–55, 2016.

[5] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu. Efficient 3-d placement of an aerial base station in next generation cellular networks. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–5. IEEE, 2016.

[6] C. Cox. *An introduction to LTE: LTE, LTE-advanced, SAE and 4G mobile communications*. John Wiley & Sons, 2014.

[7] B. Galkin, J. Kibilda, and L. A. DaSilva. Deployment of uav-mounted access points according to spatial user locations in two-tier cellular networks. In *Wireless Days (WD), 2016*, pages 1–6. IEEE, 2016.

[8] O. Grøndalen, A. Zanella, K. Mahmood, M. Carpin, J. Rasool, and O. N. Østerbø. Scheduling policies in time and frequency domains for lte downlink channel: a performance comparison. *IEEE Transactions on Vehicular Technology*, 66(4):3345–3360, 2017.

[9] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[10] H. Holma and A. Toskala. *LTE for UMTS: Evolution to LTE-advanced*. John Wiley & Sons, 2011.

[11] S. Kandeepan, K. Gomez, L. Reynaud, and T. Rasheed. Aerial-terrestrial communications: terrestrial cooperation and energy-efficient transmissions to aerial base stations. *IEEE Transactions on Aerospace and Electronic Systems*, 50(4):2715–2735, 2014.

[12] L.M.Correia. Mobile communications systems – lecture notes. Instituto Superior Técnico, Lisbon, Portugal, 2017.

[13] B. L. Miller and D. E. Goldberg. Genetic algorithms, selection schemes, and the varying effects of noise. *Evolutionary computation*, 4(2):113–131, 1996.

[14] M. Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.

[15] W. OBILE. Ericsson mobility report, 2017.

[16] R. C. Palat, A. Annamalau, and J. Reed. Cooperative relaying for ad-hoc ground networks using swarm uavs. In *Military Communications Conference, 2005. MILCOM 2005. IEEE*, pages 1588–1594. IEEE, 2005.

[17] S. Pedrinho. Análise do uso do lte-a mcc entre os militares em ambientes operacionais. Master Thesis, Instituto Superior Técnico, Lisbon, Portugal, 2017.

[18] S. Sesia, M. Baker, and I. Toufik. *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[19] T.-T. Tran, Y. Shin, and O.-S. Shin. Overview of enabling technologies for 3gpp lte-advanced. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):54, 2012.