BSc Thesis Applied Mathematics

Increasing the probability that an internet connection modelled using Poisson Point Process and STIRG model exists.

Kinga Renz

Supervisor: dr. Clara Stegehuis and Lotte Weedage

June, 2021

Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science

UNIVERSITY OF TWENTE.

PREFACE

I would like to thank my supervisors Lotte Weedage and Clara Stegehuis, for reading through all of the draft versions of this paper and for their patience. Without their guidance and helpful feedback, I would be lost.

Increasing the probability that an internet connection modelled using Poisson Point Process and STIRG model exists.

Kinga Renz*

June, 2021

Abstract

A necessity for the fifth generation of mobile communications to be deployed is the infrastructure of base stations to support the network. There are several reasons why a base station may fail to provide internet to a connected user, like a power outage or something is in the way of the connection. To assure that a base station will not experience any power outage, an additional power generator may be added. In this paper, we model the internet network with the usage of a Poisson Point Process. Furthermore, we model the reasons why a base station may fail to provide internet with the usage of the STIRG model. Then we use linear optimisation to find the optimal solution as to which base stations should get additional power generators. The general result is that it is more advantageous to add a power generator to a base station that has a higher probability to provide internet to a connected user. Moreover, no significant relation between the amount of protected base stations and the probability that an internet connection is interrupted has been noted.

Keywords: Poisson Point Process, STIRG, linear optimisation

*Email: k.renz@student.utwente.nl

CONTENTS

I INTRODUCTION

Since the first phones were used, mobile communication has changed a lot. The internet in the first generation (1G) had a speed of 2.4 kbps, poor voice quality and it was easy for a third party to eavesdrop. Then the second (2G), the third (3G) and the fourth generation (4G) followed. With every generation, significant improvements were made. Currently the 4G network is widely used. The internet in this generation is able to provide speed between 10 Mbps and 1 Gbps, it is possible to stream videos with high quality and the 4G network is much more secure than its predecessors. However, the fifth generation (5G) begins to be deployed. The 5G network offers improved functions from the previous generations as well as some new ones. It has much higher speed and capacity, the audio as well as video are in high resolution, also when calling, and it supports interactive multimedia, voice, streaming video, internet and many others. However, there are still some challenges facing the 5G network. One of them is the lack of infrastructure supporting the 5G network [8]. Some new antennas need to be built as the 5G network relies on higher-frequency bands. The development of the fifth generation can be advantageous for virtual reality, autonomous vehicles, healthcare equipment and many more, once all the challenges are addressed.

To create an infrastructure for the 5G network, we need base stations to which users can connect wirelessly. The connections between base stations and users can be viewed as a graph where base stations and users are nodes and the connections between them are edges. In order for a base station to work, it needs power supply. Moreover, to operate continuously, even when the power supply fails, the base station needs an external power source, a power generator for example. Such a base station with an additional power generator is called *protected*. Since it can be costly to protect a base station, we do not want to add a power generator to every one, but to the only ones whose protection cause the event of losing a connection between a user and a base station to be the most unlikely. Therefore we need to find a decision making process that chooses which base stations to protect.

In this paper the following research question will be addressed.

How to decide which base stations to protect to maximise the probability that an uninterrupted connection exists ?

This research question tackles a few aspects, like modelling of an internet network, and a decision process to decide which base stations should get power generators. Therefore, this research question can be viewed in terms of the following sub-questions.

How can the connections between internet users and base stations be mathematically modelled ?

How to decide which and how many base stations to protect ?

What is the relation between the amount of protected base stations and the probability that a internet connection is interrupted ?

The outline of this paper is such that first necessary theory is explained in Section II. Having that the detailed model can be introduced in Section III. In the same section relevant assumptions are explained, so that in Section IV simulations can be described as well as their results. In Section IV. A we perform one single simulation and we discuss and analyse the results. In Section IV. B we perform 500 different simulations and we discuss the average and extreme results, and we compare these results with the results from the single simulation. In the remainder of the paper, we discuss the model and the results, and formulate conclusions.

II THEORY

We model wireless networks by creating a graph where base stations and users are nodes and connections between them are edges. In the following sections, we will discuss the necessary theory to model a wireless networks. First, random and bipartite graphs will be discussed, then the Poisson Point Process (PPP) will be shortly explained, and lastly Signal-to-Interference Graph Ratio (STIRG) model will be introduced. All these terms will contribute to the mathematical model of the problem.

A RANDOM BIPARTITE GRAPHS

The concept of random graphs is rather new. There are several concepts that are referred to as random graphs. One of them considers that every possible edge in a graph appears independently of other edges with a probability p. It has been introduced in 1959 by Paul Erdös and Alfred Rényi to address the probability of a random graph being connected and the probability of other graph properties [6]. The concept of random graphs has since evolved and we can see random graphs often used to model random-like networks, such as internet networks, spread of human population, and neural networks [7].

A bipartite graph is a graph in which the set of vertices V can be split into two sets, V_1 and V_2, such that none of the vertices in V_1 is connected to any vertex in V_1. Similarly, none of the vertices in V_2 in connected to any vertex in V_2, but any vertex in V_2 is connected to some vertices in V_1, and any vertex in V_1 is connected to some vertices in V_2.

A random bipartite graph is a combination of both, random and bipartite graphs. In Figure 1 an example of such graph can be found. Consider $V_1 = [v_1, v_2, v_3]$ and $V_2 = [v_4, v_5, v_6, v_7, v_8]$. Then one can see that this is indeed a bipartite graph. Moreover, each vertex in V_1 is connected to vertices in V_2. Each of these edges exists independently of the other edges with probability p. Because of that the graph is called random. Since the graph is both bipartite and random, it is called a random bipartite graph. The probabilities for each edge to be present may differ, this is dependent on the assumptions one takes.

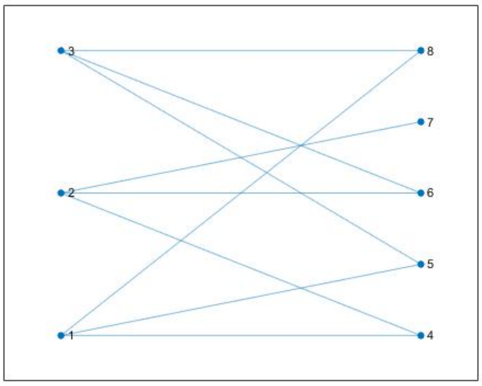


FIGURE 1: A simple example of a random bipartite graph.

In this paper, we assume that out of a set of points V, we choose a set of points B that represents base stations and another set, set $U = V \setminus B$, that represents users. Moreover, we assume that the users connect to at most one base station that is the closest. That is, neither base stations are connected to another base stations nor users are connected to other users. Furthermore, for each connection between a base station and a user, there exists a certain probability that indeed such connection exists. These are discussed in Section III. A. Therefore, the union of the sets of base stations and users, set V and the edges between them create a random bipartite graph.

B POISSON POINT PROCESS

In probability theory, the Poisson Point Process with density λ is a collection of random points distributed according to a Poisson process with density λ. The expected value and the variance of the amount of points within a given area in such process is $\lambda \cdot area$. Conditionally on the number of points within an area, these points are uniformly distributed over the area. The Poisson Point Process is often used when describing random-like networks [1]. The combination of two Poisson Point Processes is also a Poisson Point Process.

In this problem, we assume that the location of points representing the base stations and users comes from a Poisson Point Process. First, we create a set of points from a Poisson Point Process with density λ on a grid 10×10 m. Then out of all created points, a uniformly random number between 0.5 and 1.5 % is chosen which represents the amount of points from all created points that we take to create a set of points representing the locations of the base stations. The remaining points represent the locations of users.

C SIGNAL TO INTERFERENCE RATIO GRAPH

The Signal to Interference Ratio Graph model is often used in the context of wireless networks. This model is based on signal-to-interference-plus noise-ratio (SINR) and a link existence is dependent not only on the location of the user and the base station that we consider, but also on the location of all the other users and base stations in the network [5]. Furthermore, the STIRG model assumes that the location of the points is based on a Poisson Point Process and that the transmitted power of nodes representing base stations is independent of other base stations. The model considers that two nodes, i and j, are connected only if the ratio of the power from i to j to the total power received from all the other nodes and thermal noise is above a threshold [2]. Let P_i denote the transmitted power from base station i, $L(\mathbf{x}_i - \mathbf{x}_j)$ is an attenuation function, and let W be the thermal noise. Furthermore, let the interference of all other nodes be denoted as $\sum_{k \neq i,j} P_k L(\mathbf{x}_k - \mathbf{x}_j)$ and let $\gamma \in (0, 1)$ be the weight of the interference term. Then the condition that needs to be satisfied for the two nodes to be connected can be written as

$$f(\mathbf{x}_i, \mathbf{x}_j) = \frac{P_i L(\mathbf{x_i} - \mathbf{x_j})}{W + \gamma \sum_{k \neq i,j} P_k L(\mathbf{x_k} - \mathbf{x_j})} > T, \tag{1}$$

where T is a threshold.

For this problem, the attenuation function is chosen such that the larger the distance between a base station i and a user j, the smaller the probability that the connection ij exists. Let $\mathbf{x_i} = (x_{i_x}, x_{i_y})$, then the following attenuation function is considered.

$$L(\mathbf{x_i} - \mathbf{x_j}) = ||\mathbf{x_i} - \mathbf{x_j}||^{-2} = \frac{1}{(x_{i_x} - x_{j_x})^2 + (x_{i_y} - x_{j_y})^2}. \tag{2}$$

III MODEL

It is desired to know the relation between the amount of protected base stations and the probability that users lose internet connection, that is, lose the connection with a base station to which they were connected. We will do this by first mathematically modelling the problem using random bipartite graphs, Poisson Point Process and the Signal to Interference Ratio Graph model explained in Section II. C, and then finding an efficient decision-making process that decides which base stations to protect. In the remainder of this section, the model is compared to 0/1 knapsack problem.

A DETAILED PROBLEM DEFINITION

In this paper, we assume that whether there is a connection between a base station i and a user j, such that the base station i is the closest base station, depends on the three events described as follows:

A_{ij} - user j in reach of base station i;
B_i - the base station i is working;
C_{ij} - the connection between a base station i and a user j is interrupted.

The probabilities of the events A_{ij}, B_i and C_{ij} as well as any other necessary assumption that are relevant to the probabilities of these events are described as follows.

The probability that a connection exists

The probability of the event A_{ij}, that user j is in reach of base station i can be modelled with the usage of STIRG model, and therefore, from Section II. C we have that condition as in (1) needs to be satisfied. Considering $f(\mathbf{x}_i, \mathbf{x}_j)$ from (1) and the attenuation function as in (2), we let the probability that user j is in reach of base station i to be

$$p_{ij}^{A} = \frac{P_i L(\mathbf{x_i} - \mathbf{x_j})}{W + \gamma \sum_{k \neq i,j} P_k L(\mathbf{x_k} - \mathbf{x_j})}.$$

For appropriate values of P_i, W and γ, it can be assured that p_{ij}^{A} is smaller or equal to 1. In this problem we take thermal noise (W) to be 1, and the weight of the interference term (γ) to be 0.001.

Transmitted power

We assume that the power of base station i can be described in two ways. The first setting assumes the transmitted power to be the same for all base stations while the second setting assumes that the transmitted power is dependent on the amount of users connected to a base station. In the first setting, we take $P_i = 10\ \forall i$. In the second setting, we assume that the power a base station transmits depends on the amount of connected users, such that the more users are connected to the base station, the lower the transmitted power. A base station needs to support all of the connected users which usually means that the transmitted power to each user decreases with every new joining user.

Let m be the amount of base stations in a Poisson Point Process, n the amount of users and a_i the amount of users connected to a base station i. Furthermore, let P denote the value of the transmitted power if it is described as in the first setting. Then the transmitted power of a base station i can be described as

$$P_i = \frac{n - a_i}{n} \cdot (P + \frac{P}{m}).$$

This way, the transmitted power differs per base station but the difference between transmitted powers among base stations is small. Furthermore, by adding the term $\frac{P}{m}$, we assure that the transmitted power is close to the value of P. Without this term, the value of P_i is smaller than P for all i.

The probability that a base station is working

The probability of the event B_i, that base station i is working can, again, be described in two ways. In the first setting, we assume p_i^B to be the same among all base stations while in the second setting, we assume p_i^B to be dependent on the amount of connected users to a base station i. In the first setting, we take the probability that a base station i is working to be $p_i^B = 0.9\ \forall i$. In the second setting, we assume that p_i^B is dependent on the amount of users connected. Since a base station needs to support all connected users which usually implies that the more connected users, the more likely the base station is to fail, i.e. the smaller the probability that a base station is working. In principle, the p_i^B is defined in the similar way as the transmitted power, and hence, letting p^B be the value of the probability that the base station is working if it is described as in the first setting, we have

$$p_i^B = \frac{n - a_i}{n} \cdot (p^B + \frac{p^B}{m}).$$

Again, the difference per base station is small and the value of each p_i^B is not much bigger or smaller than the value of p^B.

The interruption probability

The probability of the event C_{ij}, that the connection between base station i and user j is interrupted can also be described in two ways. In the first setting we assume p_{ij}^C to be the same for each base station and we take $p_{ij}^C = 0.15\ \forall i$. In the second setting, we assume that the interruption probability depends on the location of the base station and connected user. We consider a grid 10 x 10, and we introduce some regions in it. One can see it as a representation of a large area with dense city and rural areas surrounding the city. A dense city is busier, has higher buildings and more trucks drives there, in comparison to the rural areas. Therefore the interruption probability of a busier region is higher. The rural areas are less busy, they do not have many high buildings or

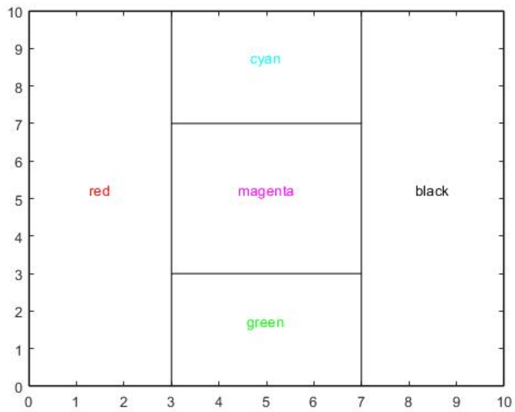


FIGURE 2: The assumed layout with the associated region colors.

trucks coming in, and therefore the interruption probability is lower. This assumption is represented in Figure 2. Each, out of the five regions has an interruption probability associated with them which is relevant for connections between base stations and users. If a user and a base station happen to be in two different regions their interruption probability is the average of the two. The magenta region is assumed to be the busiest region, and green to be the least busy region. With that in mind, the interruptions probabilities are assumed to be as represented in Table 1.

p_{ij}^C	red	green	magenta	cyan	black
red	0.15	0.10	0.20	0.125	0.15
green	0.10	0.05	0.15	0.075	0.1
magenta	0.20	0.15	0.25	0.175	0.2
cyan	0.125	0.075	0.175	0.10	0.125
black	0.15	0.10	0.20	0.125	0.

TABLE 1: The interruption probability for each region for the assumed layout.

If one wants to find out what is the interruption probability if a base station is in the red region and the connected user is in the magenta region, one looks for the value represented in row 1 (red) and column 3 (magenta).

The probability that an uninterrupted connection exists

The probability that the connection from base station i, that is working, to a user j exists and is uninterrupted is defined as p_{ij}. Since we assume the described events A_{ij}, B_i and C_{ij} independent, we have

$$p_{ij} = p_{ij}^A \cdot p_i^B \cdot (1 - p_{ij}^C).$$

B IMPLICATIONS OF ADDING A POWER GENERATOR

Adding a power generator to a base station i can be done at cost c_i and it causes that the probability that the base station i is working is 1, i.e. $p_i^B = 1$. We assume that the cost can also be described in two ways. It may be the same for each base station or it may vary depending on the amount of connected users to the base station. In the first setting, we take $c_i = 10\ 000\ \forall i$. In the second setting, we assume that the cost depends on the amount of users connected to a base station. The more users are connected to a base station, the more expensive power generator is needed. One can think of it in the sense that the more users is connected, a generator with more power is needed to support all of these users. Let m be the amount of the base stations in a Poisson Point Process, n the amount of users and a_i is the amount of users connected to a base station i. Furthermore, let c denote the value of cost as in the first setting. Then the cost for a power generator for a base station i is

$$c_i = \frac{a_i}{\frac{1}{m}\sum_{i=1}^m a_i} \cdot c.$$

This way, the cost differs per base station, but not significantly and are not much bigger than c. Furthermore, it is assumed that there is a budget C that can be spend on power generators to protect base stations.

Budget

The assumptions for budget can, in general, be done in several ways. Either amount of money can be given that one is willing to spend or one can decide how much money they are willing to spend depending on the amount of base stations, so at most how many base stations does one wish to protect. For this problem, it is assumed that the budget depends on the amount of base stations in a relevant Poisson Point Process. Let m be the amount of base stations in a Poisson Point Process and let c denote the value of cost for power generator if it is assumed to be the same for each base station. Then the budget for the power generators for the base stations, denoted as C, can be described as

$$C = \lfloor \frac{m}{2} \rfloor \cdot c.$$

The budget is always an integer, which makes computations a bit easier.

C LINEAR MODEL

In this section, we maximise the probabilities that there exist uninterrupted connections between base stations and users. The problem is constraint with the budget for the total cost for the additional power generators. This can be written as

$$\max \quad \sum_i \sum_j p_{ij} \qquad \text{s.t.} \sum_i c_i \leq C.$$

The problem written in this form is difficult to solve as the constraint impacts the objective function indirectly. To be more specific, the constraint changes p_i^B to be 1 for some base stations that are protected which changes $p_{ij} = p_{ij}^A p_i^B (1 - p_{ij}^C)$ to be $p_{ij} = p_{ij}^A (1 - p_{ij}^C)$. Therefore, we rewrite the model such that the constraints of the model impact the objective function directly.

Let $\alpha_i = \{0, 1\}$ be a decision variable that describes whether a base station i is protected. The value $\alpha_i = 0$ describes that base station is not protected and $\alpha_i = 1$ describes that base station is protected. The only thing that changes when a power generator is added to base station i is p_i^B. When base station i is not protected, so when $\alpha_i = 0$, $p_{ij} = p_{ij}^A p_i^B (1 - p_{ij}^C)$. When base station i is protected, so when $\alpha_i = 1$, p_{ij} changes to $p_{ij}^A (1 - p_{ij}^C)$. This implies that the objective function can be written as

$$\begin{aligned} &\sum_i \sum_j \left(p_i^B (1 - \alpha_i) + \alpha_i\right) p_{ij}^A (1 - p_{ij}^C) = \\ &\sum_i \sum_j \left[\alpha_i \cdot (1 - p_i^B) p_{ij}^A (1 - p_{ij}^C) + p_{ij}^A p_i^B (1 - p_{ij}^C)\right]. \end{aligned}$$

Therefore, the linear program becomes

$$\begin{aligned} \max \quad & \sum_i \sum_j [\alpha_i \cdot (1 - p_i^B) \frac{p_{ij}}{p_i^B} + p_{ij}] \\ & \text{s.t.} \sum_i c_i \alpha_i \leq C. \end{aligned} \tag{3}$$

Having the program in its linear form, we can solve it using a linear solver in MATLAB.

D THE PROBLEM VS THE 0/1 KNAPSACK PROBLEM

The knapsack problem is a programming problem with one constraint. Consider a set of items $i = 1, 2, ..., n$, each with a weight w_i and value v_i. The goal is to determine the amount x_i each item can be included in a collection such that the total weight is under certain limit W and the value is as large as possible. The 0/1 knapsack problem restrict that the amount each item in the collection is either 0 or 1 [4]. Then this can be written as

$$\max \sum_{i=1}^{n} v_i x_i \qquad \text{s.t.} \sum_{i=1}^{n} w_i x_i \leq W \text{ and } \ x_i \in \{0, 1\}.$$

The 0/1 knapsack problem has many applications in decision-making problem and several algorithms have been created to solve it.

If one takes a closer look at the introduced model, (3), and the definition of 0/1 knapsack problem, one can note some similarities. The α_i in the model corresponds to x_i in the knapsack problem, $(1 - p_i^B)\frac{p_{ij}}{p_i^B} + p_{ij}$ in the model corresponds to v_i, the costs c_i in the model correspond to w_i and the budget C in the model corresponds to W.

Since the 0/1 knapsack problem is a NP-hard problem, no efficient algorithm exists [3].

IV SIMULATION AND RESULTS

As explained in Section III, we have that cost of power generators, transmitted power, p_i^B and p_{ij}^C can be described in two ways each. Therefore, there are 16 different assumptions that can be made. We consider the variables in that order: cost, power, p_i^B and p_{ij}^C, and we let 1 denote that a variable is as described in the first setting (the same among base stations and connections), and 0 to denote that a variable is as described in the second setting (differs among base stations and connections). One of the assumptions that can be made is that costs differ among base stations, the transmitted power is different for each base station, the probability that a base station is working, p_i^B, is the same for all base stations and that the interruption probability, p_{ij}^C, is different for all the connections. Such an assumption can be abbreviated to *0010*.

In this section, we first consider one Poisson Point Process and we analyse what is the impact of the different assumptions on the amount of protected base stations and what do the protected base stations have in common. Then, we perform a simulation of 500 different Poisson Point Processes for all 16 different assumptions and we analyse the results as well.

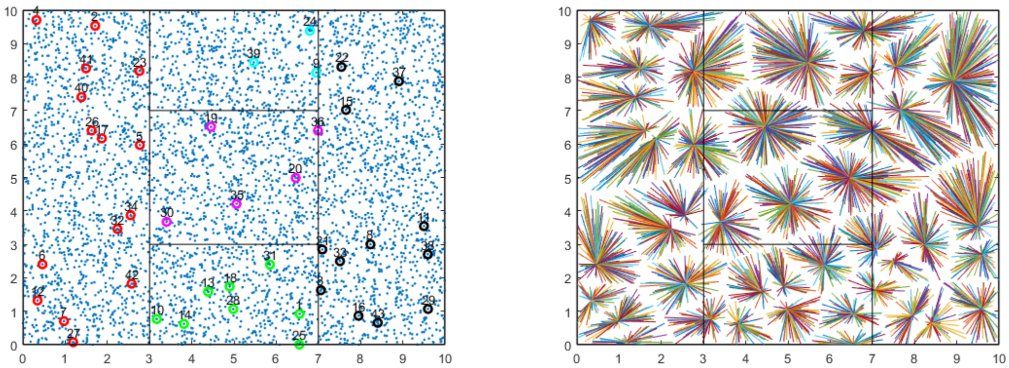
A ONE POISSON POINT PROCESS

In this part, we introduce a Poisson Point Process for which a simulation has been performed for all 16 cases. Afterwards, the results of these simulations are analysed and compared.

Poisson Point Process

Consider a set of base stations and users formed from a Poisson Point Process with $\lambda = 50$ in a grid $10 \times 10m$ as in Figure 3. This process has created 4875 points, out of which 43 are chosen to be base stations (0.88%). All of the points are spread in the grid and therefore have associated interruption probabilities, if applicable. The location of all points, specifically the base stations can be seen in Figure 3a. The base stations are also colored with the color corresponding to the region of the interruption probability for better visibility. In this Poisson Point Process, we have 15 base stations in the red region, 8 in the green region, 5 in the magenta region, 3 in the cyan and 12 in the black region.

In Figure 3b, it is represented which users connect to which base stations. As one can see, in the regions where the locations of the base stations are more dense, the amount of users to each base station is small while in the regions where there is not many base stations, more users are connected to each base station.



(A) The placement of users and base stations with the city layout.

(B) Links between users and base stations with the city layout.

FIGURE 3: The Poisson Point Process representation of the network with density $\lambda = 50$ with the city layout.

Amount of connected users

In Figure 4 one can see the histogram representing the amount of base stations in the Poisson Point Process that have the amount of connected users within specific intervals. Each bar has width 10, meaning that first bar represents the amount of base stations to which between 30 and 40 users connected. We note that this histogram is right-skewed, which makes sense as the amount of base stations that have certain amount of connected users must be at least 0 and at most the total amount of base stations created in the PPP.

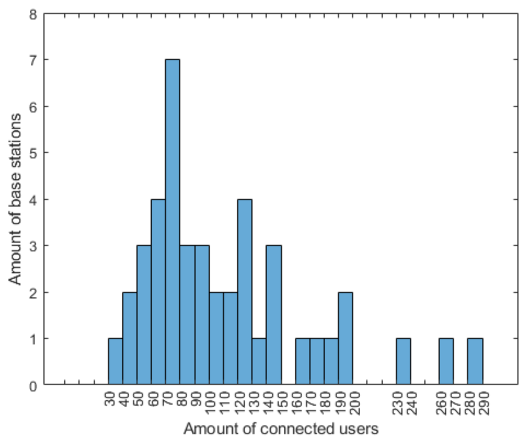
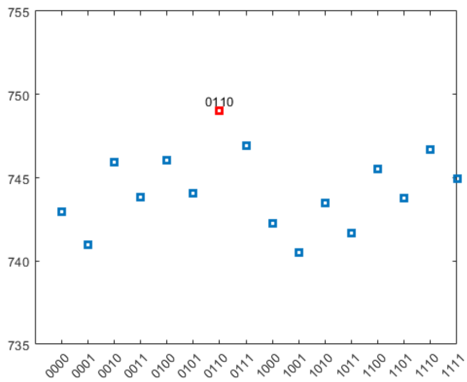


FIGURE 4: A histogram representing the amount of base stations that have amount of connected users within the intervals of length 10.

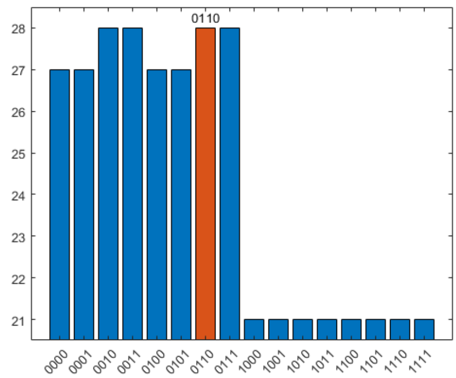
Most of the base stations have between 50 and 150 connected users (74.42%). The smallest amount of connected users was 37 while the largest was 282. The base station with the smallest amount of connected users is the base station which is located in the lower part of the black region. There are two base stations that are located closely to it that have 59 and 88 users connected to it. The base station that has the most users connected to it is the base station that is located in the upper part of the black region. Out of the base stations in the upper part of the grid, this one is located to the most right and therefore all the users that are located in the top right corner are connected to this base station, which explains the large amount of connected users.

Summary of results for all 16 cases

Considering the set of base stations and users as in Figure 3, 16 different runs have been made to see what happens if some of the variables are the same for base stations while other differ. In Figure 5 the comparison of the results of the 16 cases is shown. In Figure 5a the optimal values for each case is shown. Furthermore, the highest optimal value has been marked in red. This value has been achieved in the case 0110 when costs and p^C_{ij} differ per base station, and power and p^B_i stay the same. In Figure 5b the amount of protected base stations for each case is shown and the case for which the highest optimal value has been achieved is colored.



(A) The optimal value for $\sum_i \sum_j p_{ij}$ for each case.



(B) The amount of protected base stations for each case.

FIGURE 5: The comparison of results for all 16 cases.

Overall, the optimal values are not much different for all cases. The smallest optimal value is achieved for case 1001 (transmitted power and p^B_i differ among base stations while costs and p^C_{ij} are the same for all base stations and connections), and is 740.5127 while the highest value is achieved for case 0110 (cost and p^C_{ij} differ among base stations and connections while transmitted power and p^B_i are the same), and is 749.0134. The amount of protected base stations differs a bit more. The cases with the least protected base stations are all that have the same cost for power generators for all base stations. In these cases 21 base stations are protected, which is around 50% of all base stations, which makes sense as the budget in these cases only allows to protect half of the created base stations and it is advantageous to protect as much base stations as possible. The cases that protect the most base stations are 0010, 0011, 0110 and 0111, when 28 base stations are protected. The optimal values of these cases are 745.93, 743.85, 749.01 and 746.95, respectively. All of these values are within ten highest optimal values but three of them are one of the five highest optimal values.

0010 vs 0011 vs 0110 vs 0111

One might find it strange that the cases with the highest amount of protected base stations have different optimal values. Therefore, we compare the four cases: 0010, 0011, 0110 and 0111. All of them have 28 protected base stations while their optimal values vary from 743.8493 (for 0011) to 749.0134 (for 0110). All of these cases have the same values of the cost of power generators and probabilities that base stations are working. This can be seen as both, the costs and p_i^B are described in only two ways for every base station and for these four, these two variables are described in the same way. In this part, we will compare these four cases. Specifically the three cases, 0010, 0011 and 0111, are compared to 0110, as 0110 has the highest optimal value. In Table 2 one can see the differences between the cases.

variable	0010	0011	0110	0111
cost	differs	differs	differs	differs
power	**differs**	**differs**	same	same
p_i^B	same	same	same	same
p_{ij}^C	differs	**same**	differs	**same**

Table 2: Comparison for cases 0010, 0011, 0110 and 0111

Case 0010 considers different transmitted power for all base stations while 0110 considers that transmitted power is the same. The case 0011 also considers that the transmitted power differs among base stations, but also that the interruption probability is the same for all connections while the case 0110 considers that transmitted power is the same for all base stations and the interruption probability differs among connections. Lastly, the case 0111 is different from the case 0110 in the sense that it considers the interruption probability to be the same for all connections while, as already mentioned, the case 0110 considers it to differ. In Table 3 it is summarised how the variables that are impacted by the transmitted power behave for the four cases.

variables	0010	0011	0110	0111
optimal value	745.9291	743.8493	**749.0134**	746.9531
$\min_i P_i$	9.6354	9.6354	**10**	**10**
average power	9.9946	9.9946	**10**	**10**
$\max_i P_i$	**10.1542**	**10.1542**	10	10
$\min_i \sum_j p_{ij}$	**3.7434**	3.5658	3.7317	3.5546
mean $\sum_j p_{ij}$	15.8527	15.8135	**15.9213**	15.8825
$\max_i \sum_j p_{ij}$	93.9319	93.7448	**93.9936**	93.8057

Table 3: Comparison of several variables for cases 0010, 0011, 0110 and 0111.

On average, the case 0110 tends to achieve higher values for the transmitted power and for $\sum_j p_{ij}$. Furthermore, since the optimal values for cases with the same transmitted power among base stations, 0110 and 0111, are higher than in cases 0010 and 0011, it can be concluded that for this Poisson Point Process, if costs are different for all power generators and the probability that a base station is working is the same for all base stations, then the cases that consider the same value for transmitted power for all base stations perform better than the cases that consider different transmitted powers among base stations. Furthermore, since the optimal value for 0010 is higher than for 0011, and the optimal value for 0110 is higher than for 0111, it can also be concluded that cases that consider the interruption probability that differs among connections perform better.

One may wonder whether this behaviour also occurs for other cases that only differ by one of the variables. That is, whether cases that consider different costs among base stations, and whether cases that consider the interruption probability that differs among base station get higher optimal values. Moreover, if these variables impact the optimal value, maybe the other two variable can have an impact as well. Therefore cases that differ by only one of the variables are going to be compared and analysed.

Impact of the cost

In Table 4, such cases are compared that differ only by the variable of the cost. On top of the table, the optimal values of the cases that consider cost to be different for all power generators is shown and on the bottom the optimal values for cases that consider the costs to be the same. The higher values are marked for each comparison.

cases	0000 vs 1000	0001 vs 1001	0010 vs 1010	0011 vs 1011	0100 vs 1100	0101 vs 1101	0110 vs 1110	0111 vs 1111
differs	**742.96**	**740.95**	**745.93**	**743.85**	**746.04**	**744.06**	**749.01**	**746.95**
same	742.29	740.51	743.48	741.70	745.54	743.80	746.71	744.96

Table 4: Comparison of cases when the costs differ

As one can see, cases that consider different costs for power generators tend to have higher optimal values.

Impact of the transmitted power

The impact of the variable of the transmitted power is represented in Table 5. The higher optimal values are marked again and one can note that cases that consider same values for transmitted power of all base stations tend to have higher optimal values.

cases	0000 vs 0100	0001 vs 0101	0010 vs 0110	0011 vs 0111	1000 vs 1100	1001 vs 1101	1010 vs 1110	1011 vs 1111
differs	742.96	740.95	745.93	743.85	742.29	740.51	743.48	741.70
same	**746.04**	**744.06**	**749.01**	**746.95**	**745.54**	**743.80**	**746.71**	**744.96**

Table 5: Comparison of cases when the transmitted power differs

Impact of the probability that a base station is working

Now, the probability that a base station is working is considered. In Table 6, the optimal values of all cases that only differ by the variable p_i^B are compared and the higher optimal values are marked. Similarly, as in case of the transmitted powers, the cases which consider that the probability that a base station is working is the same have higher optimal values.

cases	0000 vs 0010	0001 vs 0011	0100 vs 0110	0101 vs 0111	1000 vs 1010	1001 vs 1011	1100 vs 1110	1101 vs 1111
differs	742.96	740.95	746.04	744.06	742.29	740.51	745.54	743.80
same	**745.93**	**743.85**	**749.01**	**746.95**	**743.48**	**741.70**	**746.71**	**744.96**

TABLE 6: Comparison of cases when p_i^B differs

Impact of the interruption probability

Lastly, the cases that differ by the variable of the interruption probability are compared in Table 7. The higher optimal values, which occur for cases with different interruption probabilities among connections are marked.

cases	0000 vs 0001	0010 vs 0011	0100 vs 0101	0110 vs 0111	1000 vs 1001	1010 vs 1011	1100 vs 1101	1110 vs 1111
differs	**742.96**	**745.93**	**746.04**	**749.01**	**742.29**	**743.48**	**745.54**	**746.71**
same	740.95	743.85	744.06	746.95	740.51	741.70	743.80	744.96

TABLE 7: Comparison of cases when p_{ij}^C differs

To summarise, for this Poisson Point Process cases that have different values of costs among base stations, same values for transmitted power and p_i^B for all base stations and that have values of interruption probability that differ among base station tend to have higher optimal values. As one can see, the case 0110 is the only case that considers the values in this way and is the case for which the highest optimal value has been achieved.

Cases with the 32 protected base stations

Considering again Figure 5b, one can note that the last 8 cases have the same amount of protected base stations. What these cases have in common is that the cost for a power generator is the same for each base station. Since the budget depends on the amount of base stations and the cost for the power generator for a base station, and in these cases the cost is the same for each base station, therefore the maximal amount of base stations that can be protected is the same for these 8 cases. Therefore, these last 8 cases all protect the maximal amount of base stations they can, which is 21.

Most often protected base stations

In Figure 6 three graphs are presented. The first one represents the average $\sum_j p_{ij}$ throughout all 16 cases for all base stations in a decreasing order while the second one shows how many times respective base stations were protected. The third graph represents the amount of users connected to each base station and 10 base stations that have largest amount of connected users are marked in orange. The first value of the first graph corresponds the highest mean $\sum_j p_{ij}$ among all base stations, the second graph concludes that the base station with the highest average $\sum_j p_{ij}$ was protected in all 16 cases, and from the third graph it follows that there were around 100 users connected to this base station.

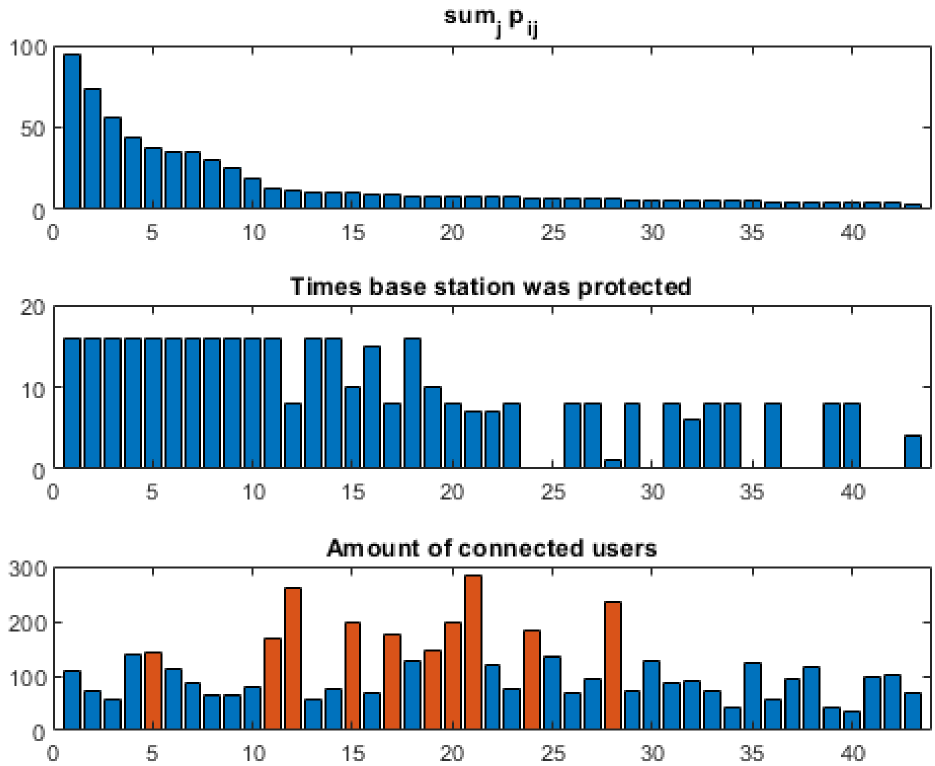


Figure 6: Average of $\sum_j p_{ij}$ for all base stations in a decreasing order, the amount of times the base stations were protected and amount of users of each base station.

From the first of these graphs one can note how big is the difference between the smallest and the largest $\sum_j p_{ij}$. The maximum value is 93.87 while the minimum value is 3.66. From the second graph, it follows that the 11 base stations with the largest mean $\sum_j p_{ij}$ were protected in all 16 cases and most of the base stations that were protected at all, were protected in at least 8 cases. The last graph indicates that while 7 out of 10 base stations with the largest amount of connected users were protected in at least half of the cases, only 2 were protected in all of the cases.

In this Poisson Point Process, base stations with higher $\sum_j p_{ij}$ tend to be protected more often while a similar rule has not been observed for base stations with larger amount of users. Therefore, it can be concluded that for this Poisson Point Process, $\sum_j p_{ij}$ is more important than the amount of connected users when deciding whether a base station should be protected.

B COLLECTION OF POISSON POINT PROCESSES

In this part, we present and analyse results of 500 simulations, each performed on a different Poisson Point Process. These simulations have been performed for all 16 cases, resulting in total in 8000 different cases.

In general, we consider the same assumptions as explained in Section III, with the difference that the budget is set to 200 000. First, the overview of the Poisson Point Processes is presented. The amount of all created base stations and the location as well as the amount of connected users is discussed. Then the $\sum_j p_{ij}$ for all the base stations are discussed and some extreme values are shown and discussed. Finally, we present the results and we compare some of them to find out whether the impact of the variables as observed in Section IV. A can be supported.

Poisson Point Processes

There were 500 different Poisson Point Processes created for which the simulation for all 16 cases has been performed. Among these processes, the smallest amount of base stations that was created was 24, and the largest was 77. In total, there were 24 746 base stations created. In Figure 7 a box-plot of the amount of base stations in each region is shown. As one may note, the most base stations were placed in red and black regions. This make sense, as these regions have the largest areas. In the red region, there were at least 2 and at most 30 base stations while in the black region there were at least 3 and at most 31 base stations. The regions green and cyan had median of 6 base stations in these regions and there were at most 17 base stations in the green region and at most 15 base stations in the cyan region. The magenta region had a median of 7 base stations in this region, slightly higher than regions green and cyan. And its highest outlier amount of base stations was 19. The little amounts of base stations in each region correspond to runs with small amounts of base stations.

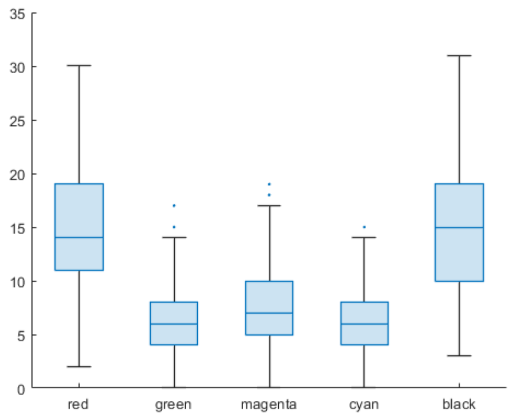
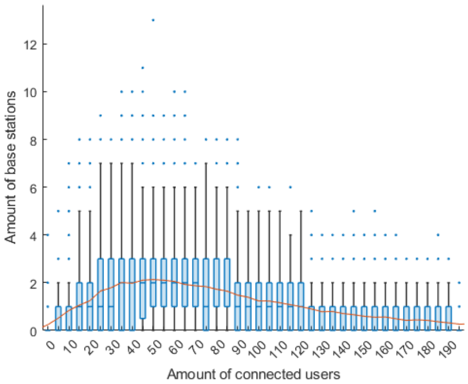


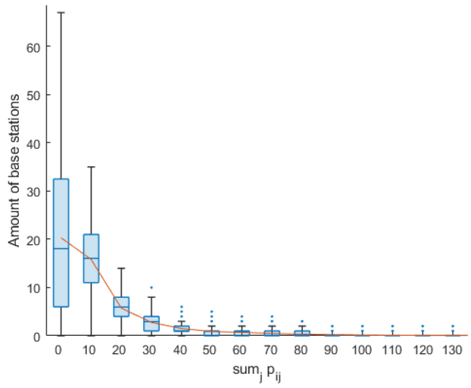
FIGURE 7: A box-plot representing the amount of base stations in each region throughout all PPP's.

Amount of connected users

In Figure 8a one can see the box-plot representing the amount of users connected to base stations that is within interval of length 5. For example, the first created box-plot represents the amount of base stations to which 0 to 5 users are connected. There were 200 different intervals, each of length 5, considered from 0 to 1000 connected users, but the intervals in the figure vary from $[0, 5]$ to $[195, 200]$. This is because the amount of base stations that have the amount of connected users within intervals larger than the interval $[195, 200]$ is considered in very little cases, at most 0.48% of all base stations are considered to have the amount of connected users to be in one of the intervals larger than $[195, 200]$. The largest amount of connected users is 910 which happened only for one base station throughout all the runs. To be specific, it happened for a run that had 30 base stations, most of which were in the bottom-right corner while this base station was the only one in the top-left corner. Therefore all users in the top-left corner connected to this base station. In the Figure 8a the average amount of base stations that have amount of connected users in respective intervals is shown as well, as the orange line. The largest average amount of connected users is in the interval $[55, 60]$ and is 2.132. That is, on average 2 base stations in a run had between 55 and 60 users connected. This interval is also the one with the largest value of the outlier. This outlier has value 14, that is, at most 14 base stations in one run had amount of connected users between 55 and 60.



(A) A box-plot representing the amount base stations to which amount of connected users is within intervals of length 5.



(B) A box-plot representing amount of base stations that have $\sum_j p_{ij}$ within intervals of length 10.

FIGURE 8: Summary of amount of base stations that have amount of connected users and $\sum_j p_{ij}$ within specific intervals.

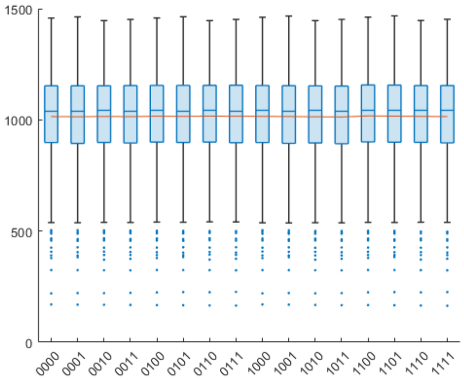
Overview of $\sum_j p_{ij}$

In Figure 8b, one can see a box-plot representing the amount of base stations that have $\sum_j p_{ij}$ within intervals of length 10. For example, the first box-plot represents the amount of base stations that have their $\sum_j p_{ij}$ between 0 and 10. The median of the amount of base stations for this interval is 18. The largest amount of base stations that had $\sum_j p_{ij}$ within that interval in one run was 67. In this run, there were in total 75 base stations. Three of them had their $\sum_j p_{ij}$ between 10 and 20, one of them between 30 and 40, two of them between 40 and 50 and the last two base stations had their $\sum_j p_{ij}$ between 70 and 80, and 180 and 190. From the figure, one can conclude that the larger the value of $\sum_j p_{ij}$, the smaller the amount of base stations and only 0.43% of all the considered base stations have their $\sum_j p_{ij}$ larger than 80.

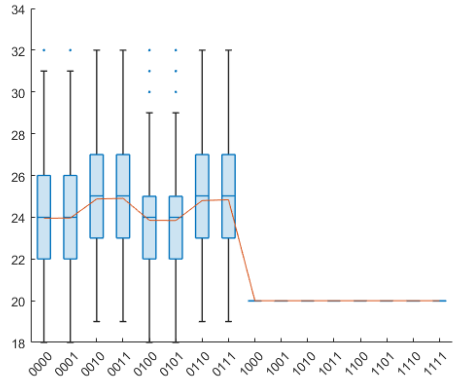
Summary of results for all 16 cases

For each of the 500 different Poisson Point Processes that were created, we consider 16 different cases, similarly as in Section IV. A, when one network created from one Poisson

Point Process was discussed. Recall that the order in which we consider variables is cost of additional power generator, transmitted power of base station, the probability that a base station is working, p_i^B, and the interruption probability, pC_{ij}. Furthermore, 1 denotes that a variable is the same among base stations and connections while 0 denotes that the variable differs. In Figure 9 the comparison of the results for all 16 cases is shown. In Figure 9a, a box-plot representing the optimal values for the cases is created as well as the mean value of the optimal values for each case is marked in orange. This value is always a bit lower than the median but the difference is not significant. In Figure 9b, a box-plot representing the amount of protected base stations for each case is shown and the average amount of protected base station for each case is marked in orange. This value tends to be the same or a bit smaller than the median.



(A) A box-plot representing the optimal value for $\sum_i \sum_j p_{ij}$ for each case.



(B) A box-plot representing the amount of protected base stations for each case.

FIGURE 9: The comparison of results for all 16 cases for 500 different runs.

The optimal values vary a lot throughout the runs, but do not vary much throughout the cases. If the outliers are included in the consideration, the smallest optimal value is 164.2071 while the largest one is 1467.71. The maximum optimal value has occurred in the run 450 for the case 1101, that is all the variables, except p_i^B, are the same for all base stations and connections. In this run, there were only 24 base stations, out of which 20 were protected. The amount of protected base stations varies a bit less among all the runs and cases. The smallest amount of protected base stations was 18 while the largest one was 32. Furthermore, the box-plot in Figure 9b suggests that the cases 0010, 0011, 0110 and 0111 tend to protect more base stations than all the other cases. Interestingly, these are the same cases as in Section IV. A which also protected the most base stations. Moreover, the cases that had the smallest amount of protected base stations also had costs of power generators to differ among base stations. These cases also were the ones which, on average, performed worse compared to other cases that had the cost for power generators set to differ among base stations. Therefore it cannot be concluded that the varying costs of power generators yields more protected base stations.

0010 vs 0011 vs 0110 vs 0111

We again compare these four results as they all protect the largest amount of base stations. The goal of this comparison is to find out whether 1) the results from the single Poisson Point Process support the average results drawn from the several runs and 2) whether it can be concluded which variables impact the optimal value and amount of protected base stations.

variables	0010	0011	0110	0111
min optimal value	**168.4**	167.6	165.6	164.8
median optimal value	1042.3	1039.2	**1044.7**	1040.6
max optimal value	1446	**1452**	1446	**1452**
mean $\sum_j p_{ij}$	12.1766	12.1656	**12.1967**	12.1856
max $\sum_j p_{ij}$	343.7298	343.9614	349.0883	**349.3302**

Table 8: Comparison of several variables for cases 0010, 0011, 0110 and 0111 resulting from several runs.

In the single Poisson Point Process, we had that the optimal values for cases with the same transmitted power among base stations, 0110 and 0111, are higher than in the two other cases, as can be seen in Table 3 in Section IV. A. However, the average results from several Poisson Point Processes suggest that the optimal values for cases which consider the interruption probability that differs among connections is highest among the four cases and the impact of the variable for transmitted power cannot be concluded. Considering the average $\sum_j p_{ij}$, the results of the single Poisson Point Process support the average results. In both, the highest average $\sum_j p_{ij}$ was achieved for case 0110, then for the case 0010, then for 0111 and the smallest $\sum_j p_{ij}$ was achieved for case 0011.

Despite that the results for optimal values for these four cases for the single run do not support the average results for these four cases, we can note that the impact of the interruption probability is the same, at least for these four cases. Therefore, to find out whether such conclusion can be made among all cases, the median optimal values for all cases that only differ by one variable will be compared. Since such comparison was also made for a single Poisson Point Process, we will see whether the results from Section IV. A are supported by the average results.

Impact of the cost

In Table 9, such cases are compared that differ only by the variable of the cost. The table has three rows. The first rows explains which cases are compared in each column while the second and third row represent median optimal values and amount of protected base stations. The second row represents these values for the case on top, so the case in which the costs are different among base stations. The third row represents the values for the cases on the bottom, so the cases in which the costs are the same for all base station. The higher values are marked for each comparison.

In the single Poisson Point Process, it was clear that the cases that consider different costs for power generators tend to have higher optimal values, see Table 4. From the average such conclusion cannot be drawn. Moreover, the cases that consider different costs for power generators have higher median optimal value in only three out of eight comparisons. What these three comparisons have in common is that p_i^B is the same for all base stations. However, no specific conclusion can be drawn about the impact of p_i^B here as in the cases

cases	0000 vs 1000	0001 vs 1001	0010 vs 1010	0011 vs 1011	0100 vs 1100	0101 vs 1101	0110 vs 1110	0111 vs 1111
differs	1040.7 **24**	1038.0 **24**	**1042.3** **25**	**1039.2** **25**	1041.4 **24**	1039.3 **24**	**1044.7** **25**	1040.6 **25**
same	**1042.2** 20	**1040.1** 20	1040.1 20	1039.0 20	**1043.0** 20	**1041.6** 20	1042.3 20	**1041.1** 20

Table 9: Comparison of median optimal values and median amount of protected base stations for cases when the variable of the cost differs.

0111 and 1111, which also have set the probability that a base station is working to be the same among base stations, the median optimal value is higher for the case which has the cost of power generators set to differ.

The amount of protected base stations, on the other hand is clearly always higher for the cases which consider that costs of the additional power generators differ among base stations. This makes sense as the budget only allows to have at most 20 base stations protected, if the costs are the same, but if they differ, they often can be smaller than the cost if they are the same.

Impact of the transmitted power

The impact of the variable of the transmitted power is represented in Table 10. The table has three rows again. The first rows explains which cases are compared in each column while the second and third row represent median optimal values and amount of protected base stations. The second row represents these values for the case on top, so the case in which the transmitted powers differ among base stations, and the third row represents the values for the cases on the bottom, so the cases in which the transmitted powers are the same for all base station. The higher values are marked for each comparison.

cases	0000 vs 0100	0001 vs 0101	0010 vs 0110	0011 vs 0111	1000 vs 1100	1001 vs 1101	1010 vs 1110	1011 vs 1111
differs	1040.7 24	1038.0 24	1042.3 25	1039.2 25	1042.2 20	1040.1 20	1041.4 20	1039.0 20
same	**1041.4** 24	**1039.3** 24	**1044.7** 25	**1040.6** 25	**1043.0** 20	**1041.6** 20	**1042.3** 20	**1041.1** 20

Table 10: Comparison of median optimal values and median amount of protected base stations for cases when the variable of the transmitted power differs.

In the single Poisson Point Process, see Table 5, as well as in the average results, it is clear that the cases that have the same transmitted power among all base stations achieve higher optimal values. Moreover, the amount of protected base stations in cases that only differ by the variable of the transmitted power, is the same. Therefore it can be concluded that the variable of the transmitted power impact the behaviour of the median optimal values, it performs better for the same transmitted powers, and has no impact on the median amount of protected base stations.

Impact of the probability that a base station is working

Now, the probability that a base station is working is considered. In Table 11, the median optimal values of all cases that only differ by the variable p_i^B are compared and the higher optimal values are marked. Moreover, the median amount of protected base stations is compared for such cases.

cases	0000 vs 0010	0001 vs 0011	0100 vs 0110	0101 vs 0111	1000 vs 1010	1001 vs 1011	1100 vs 1110	1101 vs 1111
differs	1040.7 24	1038.0 24	1041.4 24	1039.3 24	**1042.2** 20	**1040.1** 20	**1043.0** 20	**1041.6** 20
same	**1042.3** **25**	**1039.2** **25**	**1044.7** **25**	**1040.6** **25**	1041.4 20	1039.0 20	1042.3 20	1041.1 20

TABLE 11: Comparison of median optimal values and median amount of protected base stations for cases when the variable of the probability that a base station is working differs.

In the single Poisson Point Process, we concluded that the impact of the variable of the probability that the base station is working is similar as the one of the variable of the transmitted power, see Table 6. However, Table 11 indicates that if the cost of additional power generator differs among base stations, then the cases with same p_i^B among base stations perform better considering median optimal value. These cases also have higher median of the amount of protected base stations. On the other hand, if the cost of additional power generator is the same among base stations, then the cases with different p_i^B perform better considering median optimal value. Since the costs are the same, the median of the amount of protected base stations is 20 for all such cases.

Impact of the interruption probability

Lastly, the cases that differ by the variable of the interruption probability are compared in Table 12. The higher median optimal values as well as the larger median amount of protected base stations are marked.

The results of the comparison of the cases when only the variable of the interruption differs for the single Poisson Point Process, see Table 7, are consistent with the median results for the several Poisson Point Processes. The median optimal values tend to be higher for the cases that consider that the interruption probability depends on the location of the base stations and users. Moreover, the median amount of protected base stations for the cases is exactly the same. Also, it can be seen that the box-plots in Figure 8a behave in the same way for the compared cases that differ only by the variable of the interruption probability. Therefore, it can be concluded that considering the interruption probability that is dependent on the location of the base stations and users tends to achieve higher

cases	0000 vs 0001	0010 vs 0011	0100 vs 0101	0110 vs 0111	1000 vs 1001	1010 vs 1011	1100 vs 1101	1110 vs 1111
differs	**1040.7** 24	**1042.3** 25	**1041.4** 24	**1044.7** 25	**1042.2** 20	**1041.4** 20	**1043.0** 20	**1042.3** 20
same	1038.0 24	1039.2 25	1039.3 24	1040.6 25	1040.1 20	1039.0 20	1041.6 20	1041.1 20

Table 12: Comparison of median optimal values and median amount of protected base stations for cases when the variable of the interruption probability differs.

optimal values but does not impact the amount of protected base stations.

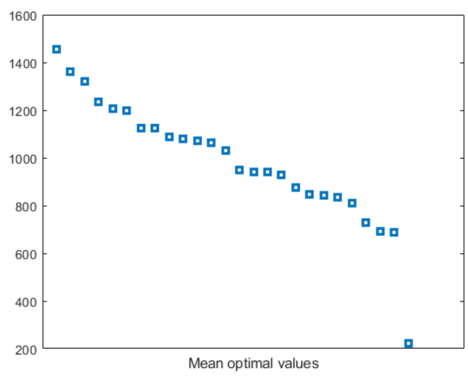
For the single Poisson Point Process, it was clear that cases that have different values of costs, same values of transmitted powers and of p_i^B as well as different values of p_{ij}^C tend to have higher optimal values. After considering many more Poisson Point Processes it can be concluded that it is not clear how the cost of power generators impact the median optimal value, but the cases with costs of power generators that differ among base stations allow to protect more base stations. The cases which consider the same transmitted power among base stations have higher median optimal values, but no impact on the amount of protected base stations have been noted. Furthermore, it seems that the impact of the probability that a base station is working is linked to the impact of the cost of power generators. The cases which consider different costs for power generators have higher median optimal values when the p_i^B was the same among base stations while the cases which considered the costs for power generators to be the same have higher median optimal values when the p_i^B differs among base stations. Lastly, the impact of the interruption probability for the single Poisson Point Process supports the impacts of the p_{ij}^C on the median optimal values, the cases which consider the interruption probability to be dependent on the location of the base stations and users to have higher median optimal values.

Considering impacts of the variables on the median optimal values, two cases, 0110 and 1100, are the ones that have their variables considered in this way and these cases were also the ones that achieved higher median optimal values. Out of these two values, 0110 had higher median optimal value and higher median amount of protected base stations but the case 1100 has achieved the highest optimal value in all the runs and cases.

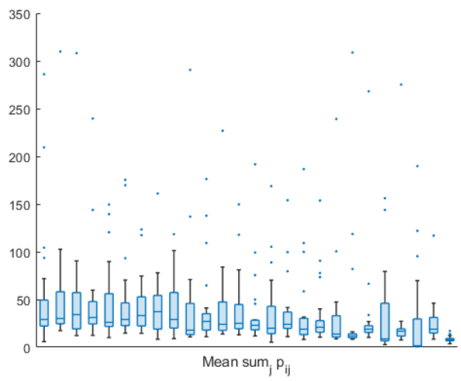
Comparison of runs that had 24 base stations

There were 26 runs that had created 24 base stations. For one of these runs the highest optimal value has been reached. We've seen that the smallest amount of protected base stations among all runs was 18. That is 75% of the base stations in this run. Therefore, one may wonder whether it is a rule that the smaller the amount of base stations in a Poisson Point Process, the higher the optimal value. We compare the optimal values and $\sum_j p_{ij}$ in the 26 runs that had 24 base stations.

The run that had reached the highest optimal value was the run 450. The highest optimal value was reached only in one case, 1101. This case considers the same cost for all power generators, therefore it was possible to protect at most 20, out of 24, base stations in this run. That is 83% of all base stations created in this Poisson Point Process. Since it is advantageous to protect base stations, the maximum amount of base stations that can be protected is indeed protected. The smallest optimal value has been obtained in case 0000 for run 30. In Figure 10a the median optimal value for all the runs with 24 base stations is represented while in Figure 10b a box-plot representation of the average $\sum_j p_{ij}$ for all base stations in each run that had 24 base stations. This average is taken over the 16 different cases. The values of the median optimal values are presented in the decreasing order and the values of the average $\sum_j p_{ij}$ are plot respectively. For example, the first value in Figure 10a represents the highest median optimal value and the box-plot in Figure 10b represent the average $\sum_j p_{ij}$ among 16 different cases for all base stations.



(A) A box-plot representing the optimal values for runs that have 24 base stations.



(B) A box-plot representing the mean $\sum_j p_{ij}$ for all base stations in each run that have 24 base stations.

FIGURE 10: Optimal values and mean $sum_j p_{ij}$ in each run that have 24 base stations.

Consider the runs 450, the run with the highest, and the run 30 with the smallest optimal value among runs with 24 created base stations. The $\sum_j p_{ij}$ of these two runs differ a lot. The average for the run 450 is 54.1486 while for the run 30 it is 8.3248 and overall, the optimal values among all 16 cases for these two runs are not far different from each other. Therefore, this implies that the set-up of the Poisson Point Process causes the difference in the optimal value.

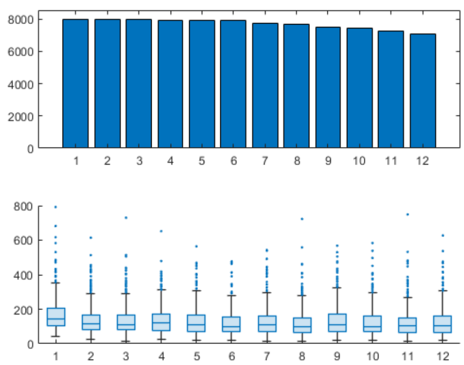
Consider the base stations that have the largest amount of users connected in both runs. Such base station in run 450 has 513 connected users and its average $\sum_j p_{ij}$ over the 16 cases is 209.36 while the base stations with the largest amount of connected users in run 30 has 462 connected users and its average $\sum_j p_{ij}$ is 6.79. Since the median of amount of connected users are in general larger for run 30, which indicates that this run has more base stations with larger amount of connected users, we also consider the average $\sum_j p_{ij}$ for the base stations with the smallest amount of connected users in each run. For the run 450, it is 5.79, for a base station with 6 connected users, and for the run 30, 7.67 for a base station with 19 connected users. This result indicates that the amount of connected users to a base station does not have to impact the value of $\sum_j p_{ij}$ for a base station. Moreover, the results for the runs with 24 base stations indicate that the location of points from a

Poisson Point Process that are chosen to be base stations, impact the $\sum_j p_{ij}$.

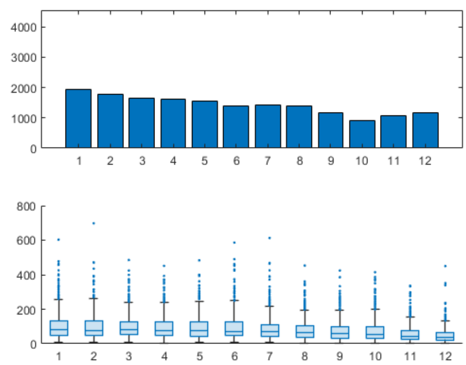
If the amount of connected users had a clear impact on $\sum_j p_{ij}$, we would expect the average $\sum_j p_{ij}$ over the 16 case for a base station with the smallest amount of connected users to be smaller or larger than the average $\sum_j p_{ij}$ over the 16 case for a base station with the largest amount of connected users. However, for the run 450, the base station with the largest amount of connected users had larger $\sum_j p_{ij}$ and for the run 30, the base station with the smallest amount of connected users had larger $\sum_j p_{ij}$.

Most often protected base stations

In Section IV. A, we have found that in the single Poisson Point Process, the base stations with highest $\sum_j p_{ij}$ tend to be protected. In Figure 11 the amount of times the base stations with largest and smallest $\sum_j p_{ij}$ were protected out of 8000 cases possible. Furthermore, a box-plot representing amount of connected users to such base stations is represented. In Figure 11a, these values for 12 largest $\sum_j p_{ij}$ are presented while Figure 11b shows these values for the 12 smallest $\sum_j p_{ij}$. These are plot in such a way that on the most left, one can see the values for the largest, and to the most right the values for the smallest $\sum_j p_{ij}$.



(A) Amount of times base stations with 12 largest $\sum_j p_{ij}$ were protected and box-plot of amount of connected users to respective base stations.



(B) Amount of times base stations with 12 smallest $\sum_j p_{ij}$ were protected and box-plot of amount of connected users to respective base stations.

FIGURE 11: Summary of amount of times base stations with largest and smallest $\sum_j p_{ij}$ were protected and amount of connected users to the respective base stations.

As one may notice, the base stations with the largest $\sum_j p_{ij}$ are protected much more often. The base stations that had the largest $\sum_j p_{ij}$ in each run and each case, were protected in 7997 out of 8000 times (99.96%) while the base stations that had the smallest $\sum_j p_{ij}$ in each run and each case, were protected 1155 out of 8000 times (14.44%). Furthermore, it has been noted that the base stations with larger $\sum_j p_{ij}$ tend to have more users connected to it. The median of connected users of base stations with 12 largest $\sum_j p_{ij}$ varies between 97 and 144.5 while for base stations with 12 smallest $\sum_j p_{ij}$ it varies between 36 and 83. Also the outlier values tend to be higher for the base stations with larger $\sum_j p_{ij}$. Therefore we can conclude that base stations with larger $\sum_j p_{ij}$ tend to be protected more often, which is in line with the results from the single Poisson Point Process. Moreover, the base stations with larger $\sum_j p_{ij}$ tend to have more connected users, on average.

V CONCLUSION

An infrastructure of the 5G network can be advantageous to our world. It can revolutionise, among other, the virtual reality, autonomous vehicle but also the healthcare equipment. The development of these fields can have impact on the improvement of life but also on the environment, or poverty. But to create such infrastructure, base stations are needed to which users can connect wirelessly. Moreover, the creation of an infrastructure and optimal placement of the base stations are not the only issues that need to be addressed. There are many challenges when it comes to assuring that a wireless connection is stable and uninterrupted. In this paper we have focused on increasing the probability that a wireless connection from a base station that is working to a user within the range of the base station exists and is uninterrupted.

Assumptions
To model the internet network, we have made use of Poisson Point Process. In this paper, we have set it up in such a way that from all created points from a Poisson Point Process, we choose $0.5 - 1.5\%$ points to represent base stations which support users represented by the remaining points. We have assumed that a user connects to only one base station that is the closest and we have introduced three events that impact the existence of such connection, events A_{ij}, B_i and C_{ij}. The probability of the event A_{ij}, that a user j is within range of the base station i has been described with the usage of the STIRG model while the probabilities of the events B_i and C_{ij} has been described in two ways. In the first setting we assume B_i and C_{ij} to be the same among base stations and connections. In the second setting, we assume the probability of the event B_i, that base station i is working, to differ among base stations, and we described it in such a way that the more users connected to the base station, the less likely is the base station to be working. Still in the second setting, we assume the probability of the event C_{ij}, that a connection between base station i and user j is interrupted, to also differ among connections, and we assume it to be dependent on the location of the base station and of the connected user.

Linear Model
Having that, we have introduced a linear model which maximises the probability that an uninterrupted connections between base stations and connected users exist. At first, it was done in such a way that the constraint impacted the objective function indirectly, therefore, it has been rewritten. We introduced a decision variable α_i that describes whether base station i is protected, and in the new linear model, the constraint impacts the objective function directly.

One Poisson Point Process
Once the linear model has been introduced and solved using the linear optimisation, we have presented a single Poisson Point Process which has been analysed. We have seen that base stations that are located far from other base stations tend to have more users connected to it while base stations that are located more closely tend to have less users connected. Furthermore, for this PPP, the difference between the smallest and largest amount of connected users is very big. Later, the optimal values and amount of protected base stations have been compared for all 16 cases for the Poisson Point Process. It turns

out that the optimal values do not differ much between the cases but the amount of protected base stations does. We have compared four cases, one of them had the highest optimal value while all of them had the largest amount of protected base stations, among all 16 cases. Since we were considering four cases, the impact of the variables which made them different has been analysed. We have noted that for this Poisson Point Process, cases that consider the same values of transmitted power and probability that a base station is working for all base stations achieve higher optimal values. Lastly we checked what do the cases with the same amount of protected base stations have in common as well as what do protected base stations have in common. We concluded that the cases with the same amount of protected base stations can at most protect 21 base stations as the cost for each base station in all of these cases is the same. It has also been checked that the base stations with the higher mean total probability of existence of uninterrupted connections, $\sum_j p_{ij}$, are protected more often than the other base stations. We have also seen that the amount of users connected to base stations does not impact the mean $\sum_j p_{ij}$ or the amount of times the base station was protected.

Collection of Poisson Point Processes
After the simulation for a single Poisson Point Process has been performed, we have created a collection of 500 different Poisson Point Processes. For this collection, a simulation has been performed and the results have been collected and analyses. In this collection of Poisson Point Processes, we have seen that most of the base stations had between 5 and 190 connected users but that there was a base station which had 910 connected users. Therefore, the amount of connected users varies a lot, which has also been observed in the single Poisson Point Process. Later, we have observed that most of the base stations had their $\sum_j p_{ij}$ to be at most 80. Furthermore, we concluded that the median optimal values did not differ much between all 16 cases, and that the highest optimal value occurred for the case 0110 which considered transmitted power and p_i^B to differ among base stations and costs and p_{ij}^C to be the same among base stations and connections. A similar conclusion has been made for the single Poisson Point Processes. The highest recorded optimal value throughout all runs and cases was for the case 1101 which considers all variables except p_i^B to be the same among base stations and connections. The median optimal value for this case is only the fifth largest one, and yet for this case the highest optimal value has been recorded.

Impact of the four variables
After the simulations were discussed, we have focused on the impact of the four variables, costs of power generators, transmitted power, p_i^B and p_{ij}^C. In this paper we have described the assumptions for these four variables in two ways each. We have found out that for the simulation of the single Poisson Point Process, the cases with the same values for transmitted power and p_i^B among base stations and different values for costs and p_{ij}^C among base stations and connections tend to have higher optimal values. As we saw, the only case that fit exactly in all of these four descriptions is the case 0110, for which we have recorded the highest optimal value in the single Poisson Point Process, performed in this paper. A similar analysis was performed for the median optimal values of 500 different Poisson Point Processes. Each of these assumptions did not have a very significant impact on the results but we have seen that the cases with the same transmitted power among base stations and the interruption probability dependent on the location of a base station and a user tend to result in higher optimal values. Moreover, we have seen that the assumptions whether costs are the same or differ among power generators are linked to the assumptions whether

the probability that a base station is working is the same or differs among base stations. We have seen that the cases with the same costs among base stations tend to have higher optimal values if the probability that a base station is working differs among base stations while the cases with the costs that differ among base stations tend to have higher optimal values if the probability that a base station is working is the same among base stations.

Final conclusions
In this paper, we have seen that the connections between users and base stations can be modelled with the use of STIRG model and to decide which base stations to protect, we use linear optimisation. We have concluded that the total probability of existence of uninterrupted connections, $\sum_j p_{ij}$, for a base station i is a very important factor when deciding whether to protect the base station. We have seen that the base stations with highest $\sum_j p_{ij}$ have been protected the most often. Moreover, from the results for the runs with 24 base stations, it follows that the $\sum_j p_{ij}$ is impacted by the location of base stations and users, but not so much by the amount of connected users to base stations.

VI DISCUSSION

In this paper we have been working towards a conclusion which base stations to protect such that the probability that a connection from a working base station to a user exists and is uninterrupted is maximised. We have done that by modelling the connection between base stations and users, and by finding a decision process about which base stations to protect. Lastly, we have been looking into whether the amount of connected users to a base station had impact on the amount of times base stations that was protected and their $\sum_j p_{ij}$, and overall, how these three are correlated.

In the single Poisson Point Process created in this paper, it was clear that overall, the higher the $\sum_j p_{ij}$, the more likely is the base station to be chosen to be protected. In the same simulation, it could not have been concluded that the amount of connected users impacts the value of $\sum_j p_{ij}$ of a base station, or the amount of times the base station is protected. When performing a similar analysis for the collection of the Poisson Point Processes, we could see that, again, the most often the base stations with higher $\sum_j p_{ij}$ tend to be given a power generator to support them and that these base station had, overall, more connected users. We have seen, when discussing the runs with 24 base stations, that the amount of connected users does not impact $\sum_j p_{ij}$. What has not been analysed in detail is whether the base stations with the largest amounts of connected users are protected more often. This has been analysed for the single Poisson Point Process, but not for the collection. From the single simulation it actually follows that the base stations with largest amount of connected users are not always protected. In fact, only 2 out of base stations with 10 largest amounts of connected users have been protected in all 16 cases but there was also a base station that had the amount of connected users in the top 10 that was not protected at all.

In this paper we have focused a lot on the four introduced variables for cost for power generator, transmitted power of base stations, the probability that a base station is working and the interruption probability. The way in which these variables were described was rather symbolic and was not meant to have a severe impact on the results. Despite that, we could have noted several repeated behaviours regarding these variables, like that the cases with the same transmitted power among base stations and the interruption probability dependent on the location of the base station and connected user were resulting in higher optimal values. However, to make the impact of these variables more visible, the variable of which impact one would be interested, should be changed significantly. Moreover, the introduced variables give a lot of freedom to redefine them.

For example, one could make the transmitted power or the probability that a base station is working very small when large amount of users is connected. Furthermore, a limitation of a power base station could be added, that is, that a base station is not able to support more than x amount of connected users, and that its transmitted power (or the probability that a base station is working) is 0 for all connected users when the amount of connected users is larger than x.

Considering the interruption probability, one could introduce much more complicated or much larger regions for which the interruption probability is described. In general, the Poisson Point Processes from which we take the locations of base stations and users can be made much larger, either in the sense of the area or the amount of points. This can be done by manipulating the grid on which a Poisson Point Process is defined and by enlarging the density λ. Since there is a lot of freedom in redefining the variables, one could describe the interruption probability in a totally different way. One could also introduce different interruption probabilities for a different situation that interrupts the connection. In this paper, we considered the interruption as a high building or a truck being in a way of a connection. These two events could be separated and each could get their own defined interruption probability. Moreover, more interruptions could be added.

To summarise, there are still several factors that can be analysed from the simulations. The impact of the amount of connected users, or perhaps the location of the base stations could be analysed in more detail. Furthermore, the way the problem is modelled makes it easy to redefine the assumptions, which gives a lot of flexibility if one was interested in maximising the probability that an internet connection is uninterrupted in a specific region.

REFERENCES

[1] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke. *Stochastic Geometry and Its Applications*, chapter Point Processes I - The Poisson Point Process. John Wiley Sons Ltd, 2013.

[2] Olivier Dousse, François Baccelli, and Patrick Thiran. Impact of interferences on connectivity in ad hoc. 09 2004.

[3] Absalom E. Ezugwu, Verosha Pillay, Divyan Hirasen, Kershen Sivanarain, and Melvin Govender. A comparative study of meta-heuristic optimization algorithms for $0-1$ knapsack problem: Some initial results. *IEEE Access*, 7:43979–44001, 2019.

[4] Arnaud Freville and Gérard Plateau. An efficient preprocessing procedure for the multidimensional 0–1 knapsack problem. *Discrete Applied Mathematics*, 49(1):189–212, 1994. Special Volume Viewpoints on Optimization.

[5] Martin Haenggi, Jeffrey Andrews, François Baccelli, Olivier Dousse, and M. Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *Selected Areas in Communications, IEEE Journal on*, 27:1029 – 1046, 10 2009.

[6] Mihyun Kang and Zdenek Petrášek. Random graphs: Theory and applications from nature to society to the brain. 2014.

[7] Mihai Tesliuc. Random graphs and their applications.

[8] Lopa J. Vora. Evolution of mobile generation technology: 1g to 5g and review of upcoming wireless technology 5g. *International Journal of Modern Trends in Engineering and Research*, 2:281–290, 10 2015.