INFRAESTRUTURA DE ACESSO EM REDES SEM FIO OBSTRUÍDAS: DA INTRATABILIDADE À CONECTIVIDADE

MANASSÉS FERREIRA NETO

INFRAESTRUTURA DE ACESSO EM REDES SEM FIO OBSTRUÍDAS: DA INTRATABILIDADE À CONECTIVIDADE

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Orientador: Olga Nikolaevna Goussevskaia
Coorientador: Vinícius Fernandes dos Santos

Belo Horizonte
Fevereiro de 2017

MANASSÉS FERREIRA NETO

BACKBONE STRUCTURES IN OBSTRUCTED WIRELESS NETWORKS: FROM INTRACTABILITY TO CONNECTIVITY

Dissertation presented to the Graduate Program in Computer Science of the Federal University of Minas Gerais in partial fulfillment of the requirements for the degree of Master in Computer Science.

Advisor: Olga Nikolaevna Goussevskaia
Co-Advisor: Vinícius Fernandes dos Santos

Belo Horizonte
February 2017

Ferreira Neto, Manassés

D1234p Backbone Structures in Obstructed Wireless Networks: from Intractability to Connectivity / Manassés Ferreira Neto. — Belo Horizonte, 2017
xxiii, 53 f. : il. ; 29cm

Dissertação (mestrado) — Federal University of Minas Gerais
Orientador: Olga Nikolaevna Goussevskaia
Coorientador: Vinícius Fernandes dos Santos

1. Modelagem. 2. Simulação. 3. Redes Sem Fio Obstruídas. 4. NP-Completude. 5. Algoritmos. 6. Complexidade. I. Título.

CDU 519.6*82.10

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Infraestrutura de acesso em redes sem fio obstruídas: da intratabilidade à conectividade

MANASSES FERREIRA NETO

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROFA. OLGA NIKOLAEVNA GOUSSEVSKAIA - Orientadora
Departamento de Ciência da Computação - UFMG

PROF. VINÍCIUS FERNANDES DOS SANTOS - Coorientador
Departamento de Ciência da Computação - UFMG

PROF. ANTONIO ALFREDO FERREIRA LOUREIRO
Departamento de Ciência da Computação - UFMG

PROF. JAYME LUIZ SZWARCFITER
Programa de Engenharia de Sistemas e Computação - UFRJ

PROF. MARIO SÉRGIO FERREIRA ALVIM JÚNIOR
Departamento de Ciência da Computação - UFMG

Belo Horizonte, 22 de fevereiro de 2017.

Aos bons.

Acknowledgments

Agradeço àqueles que apoiaram e fizeram de mim o homem que sou: querida mãe Irene, mestre pai Vinicio, linda esposa Sarah, brava sogra Ângela, esperta *sister-in-law* Paula, eterno POVIM, ótima orientadora Olga, imprescindível coorientador Vinícius, incrível WISEMAP, genial CRC, forte STI-FALE, acolhedor DEST, surpreendente LSP, técnico CEFET, superior Campus-II, lúdico São Geraldo, chapante GFH, ninho ARNALDO, minhas filhas felinas: Little, Kiki e Jujuba, segura e terna Família, essenciais: Toninho, Clarice, Mônica, Agenor, Zézé, Marieta, Milude, Paulinho, Geraldo, Dinda, Eliane, Antônio Augusto, Edite, Sheila, Bia, Pum, Cyntia, Nadja, Murray, Tóia, inúmeros: Patrícia, Kelinho, Márcio, Cristina, Jane, Marta, Diorela, André, Diogo, Julio, Tati, Renatha, Samuel, Alice, Fabio, Bia, Aludra, Gabi, Rosalice, $(\cdots)$, Tê, Peu, Nick, Jojo, Uanderlai, irmãos de cora(cria)ção: Anderson, Marcelo, Guilherme, Arthur, Alan (os dois), Klávius, Kléver, Duds, Pedroca, Henry, Oui, Eureka, Juí, de longa: Bita, Naty, Natm, Jana, Tavo, Con, Fominha, Sart, Marisa, Ticiano, Carolina, Natalia, de média: bvc, Júlio, Mól, Alexandre, Rafael, Gladson, JP, Jonas, Fabiano, Roberto, Clay, JB, Ivão, Cotta, Pes, Maia, Fil, Ana, Tonin, Evandro, Paulo, Ronaldo, Fred, de nova: Murilo, László, Polly, Heitor, π, Jean, Armstrong, David, Ricardo, Alison, Cloves, Breno, GuiSaulo, eficiente Secretaria, inspiradores Professores e extraordinários Funcionários. À UFMG, ao Departamento de Computação, a todos, meu sincero agradecimento. Finalmente, muito obrigado aos envolvidos no desenvolvimento das preciosas aplicações: bash, atom, `R`, vim, ubuntu, `C++`, chrome, firefox, evince, gnome, xfce, lemon, java, dropbox, github, git, google, -inbox, -agenda, -scholar, gimp, xpdf e gnu-linux.

" The mathematical theory of information came into being
when it was realised that the flow of information can be expressed numerically
in the same way as distance, time, mass, temperature."
(Alfréd Rényi, A diary on information theory*)*

"A great discovery solves a great problem
but there is a grain of discovery

in the solution of any problem.

Your problem may be modest;
but if it challenges your curiosity and
brings into play your inventive faculties,

and if you solve it by your own means,

you may experience the tension and
enjoy the triumph of discovery. "
(George Pólya, How to Solve It)

Resumo

Neste trabalho é considerada uma rede sem fio ad hoc realizada em uma grade regular quadrada, na qual a comunicação entre os dispositivos é influenciada por obstáculos regularmente espaçados. O raio crítico de transmissão para obter conectividade nesse tipo de rede cresce com o tamanho da grade, o que pode prejudicar a viabilidade em larga escala de tecnologias sem fio de baixa potência. Avalia-se portanto como introduzir eficientemente uma infraestrutura conectada de pontos de acesso em cenários subcríticos, nos quais o raio de transmissão é insuficiente para estabelecer a conectividade. Formula-se o problema de posicionar o menor número de pontos de acesso, de tal modo que todo componente conectado seja coberto por pelo menos uma estação base, e denomina-se esse problema de Obstructed Wireless Network Backbone Cover Problem (OWN-BC). Prova-se que OWN-BC é NP-Completo e propõe-se um algoritmo 2-aproximativo para obter soluções com garantia de qualidade. Realiza-se simulações para ilustrar o desempenho do algoritmo em diferentes cenários. Além disso, é feita uma caracterização de cenários para os quais o algoritmo proposto obtêm soluções ótimas em tempo polinomial com alta probabilidade.

Palavras-chave: Redes de Comunicação Sem Fio Obstruídas, NP-Completude, Algoritmos Aproximativos, Modelagem Analítica, Validação de Modelos, Limites Fundamentais, Escalabilidade.

Abstract

In this work we consider a wireless ad-hoc network deployed on a finite street grid, where communication between devices is disrupted by regularly spaced obstacles. The critical transmission range for connectivity in such networks grows with the size of the grid, which might impair the feasibility of low-power wireless technologies in large networks. We therefore analyze how the connectivity of such networks in sub-critical scenarios, where the transmission range is insufficient to establish connectivity, can be improved by introducing a global backbone with a set of access points. We formulate the problem of positioning a minimum number of access points, such that every connected component is covered by at least one access point, and refer to it as the Obstructed Wireless Network Backbone Cover Problem (OWN-BC). We prove that OWN-BC is NP-complete and present a 2-approximation algorithm to find solutions with guaranteed quality. Furthermore, we derive a lower bound on the probability of finding optimal solutions in random network scenarios. Finally, we perform a series of simulations to illustrate the performance of the approximation algorithm and characterize scenarios in which the proposed algorithm obtains optimal solutions in polynomial time.

Palavras-chave: Obstructed Wireless Communication Networks, NP-Completeness, Approximation Algorithms, Analytical Modeling, Model Validation, Fundamental Limits, Scalability.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

In this work we explore the limits of connectivity of wireless networks performed in environments with obstacles – the so-called Obstructed Wireless Networks (OWN). Obstacles are present in a variety of networking application scenarios, such as vehicular networks operating in urban street grids (Bian et al. [2015b]; Florian et al. [2013]), home networks (Kawsar and Brush [2013]), or other networks deployed indoors, tunnels, or underground mines (Forooshani et al. [2013]). Let consider, for instance, the formation of a network only between the cars located on the streets of a city with many buildings, like New York City (see Figure 1.1). The described vehicular network does not rely on a pre-existing infrastructure, being called *ad hoc*.

We start this work by analyzing the scalability of ad hoc wireless networks obstructed by regularly spaced obstacles, using the model proposed in Almiron et al. [2013]. In this model, the street blocks of the city of Figure 1.1 are represented as squares on a two-dimensional grid and the devices, as points on uni-dimensional streets (see Figure 1.2a). A device has the communication range dependent on the wireless technology applied. A pair of devices are connected if they meet specific communication rules, roughly, a function of distance and the absence of obstacles between them. A set of connected devices forms the so-called connected component, or just, component. The components are represented in Figure 1.2a as the points (devices) connected by lines (communication links).

The main contribution of this work is the study of sub-critical communication scenarios in the model proposed in Almiron et al. [2013], i.e., scenarios where either the number of devices or its communication range are not large enough to ensure overall connectivity and the network is partitioned into several components, exactly what happens in Figure 1.2a.

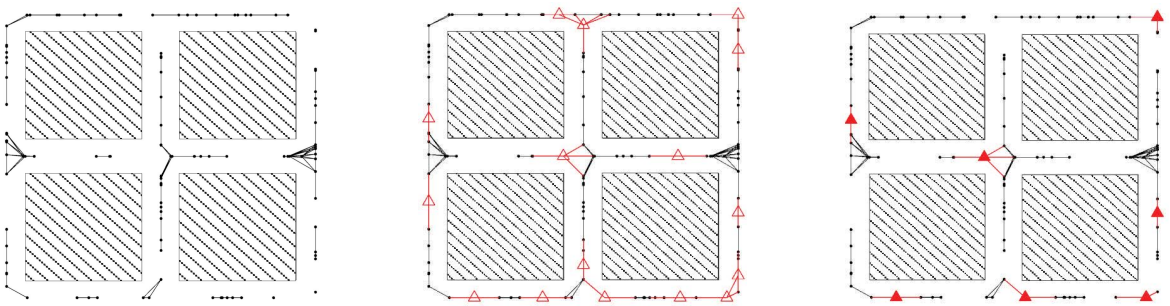
We assume that a global communication infrastrucure exists, e.g. the Internet,

Figure 1.1: Application scenario. Photo available at http://www.topik.in/post/amazing-new-york-city-above-photos. Access on April/2017.

which can be accessed through one or more access points by any device, as long as there is a path of communication links from that device to at least one access point to the infrastrucure. This infrastructure, also referred in this work as backbone, provides connectivity between the ad hoc components. We represent in Figure 1.2b, with empty triangles, the candidate positions for access points. In Figure 1.2c, we represent, with solid triangles, the minimum set of access points, selected from the candidates, to establish connectivity. We define the Obstructed Wireless Network Backbone Cover (OWN-BC) problem:

> Given an OWN, what is the minimum number of access points that need to be deployed on the grid, such that every component is covered by at least one access point and, consequently, every device has a communication path to the backbone?

We prove that, given an arbitrary device distribution on the grid, OWN-BC is



(a) Components (b) Candidates (c) Access Points

Figure 1.2: OWN-BC instance.

NP-complete and propose a 2-approximation algorithm to compute close-to-optimum solutions in polynomial time.

We simulated diverse OWN scenarios and observed that the obtained solutions were remarkably close to the optimum, which motivated us to analyze the complexity of typical problem instances. Because of this, we turn our attention to the randomly-generated device distributions. By analyzing the structure of the problem in random scenarios, we showed that it can be solved optimally in polynomial time, as long as certain criteria are satisfied at the intersections of the grid structure. In particular, we present an analytical lower bound on the probability of the approximation algorithm to compute optimum solutions in polynomial time.

We validate this analytical result by simulating the empirical probability of finding optimum solutions in polynomial time in sub-critical network configurations. The simulations corroborated the good quality of the analytical result.

We believe that, from a practical point of view, these are potentially valuable results, since they allow (a communication engineer) to ensure connectivity in OWN with arbitrary device deployment, with or without a backbone, for any network size, communication technology, or geometry of a particular obstacle grid. Moreover, it is possible to identify scenarios, where optimal backbone solutions can be quickly computed.

1.1 Dissertation Outline

We started this chapter by making an informal introduction to the subject of work. The pillar of this work is an ad hoc wireless network carried out in a square grid, in which communication between devices is influenced by regularly spaced obstacles. The rest of the dissertation is organized as follows.

In Chapter 2, we describe the deterministic and probabilistic versions of OWN model. We highlight useful results demonstrated in Almiron et al. [2013] in the scope of the Random OWN model.

In Chapter 3, we analyze the scalability properties of ad hoc OWN. We study the scalability limits of connectivity in OWN. The critical range of transmission for connectivity in this type of network grows with grid size, which can impair the viability of the large-scale use of low-power wireless technologies in obstructed environments.

In Chapter 4, we formalize the OWN-BC problem and study its complexity. We study how to efficiently introduce a backbone of access points in subcritical scenarios, in which the range of transmission is insufficient to establish connectivity. The problem

consists of positioning the minimum number of access points so that every connected component is covered by at least one base station. We demonstrate that OWN-BC is NP-Complete, by showing that: (i) we can evaluate a given solution in polynomial time with a deterministic machine and (ii) any instance of a known NP-hard problem, Exact Cover by 3-Sets, can be mapped into a specific instance of OWN-BC, in such a way that the latter has a solution if and only if (iff) the former does.

In Chapter 5, we propose a constant approximation algorithm for OWN-BC. We propose a 2-approximation algorithm to obtain solutions with quality guarantee. Our algorithm is based on the solution of minimum edge cover problem, consequently, it has a maximum matching core, which explains why it fits best in scenarios where base stations assist at most two connected components. We prove correctness and the approximation factor.

In Chapter 6, we discuss the probability of "polynomial connectivity with a backbone". We show analytically the necessary conditions for polynomial complexity instances of OWN-BC. A geometrical feature of the model, observed in the simulations, is used to analyze the conditions that produce, with high probability, instances for which the proposed algorithm finds optimal solutions in polynomial time. In addition, we show the connectivity gap in OWN between the two modes of device deployment: ad hoc *versus* "with backbone structures".

In Chapter 7, simulations are performed to illustrate the performance of the algorithm in different scenarios. We present simulation data generated to compute minimum backbones with our algorithm (and to compare it with a baseline). We also use simulations to validate our analytical "polynomial connectivity with a backbone" procedures.

Finally, in Chapter 8, we discuss related work, pointing out interesting aspects of the maturing capabilities of our work, and, in Chapter 9, we present our conclusions and discuss possible future directions.

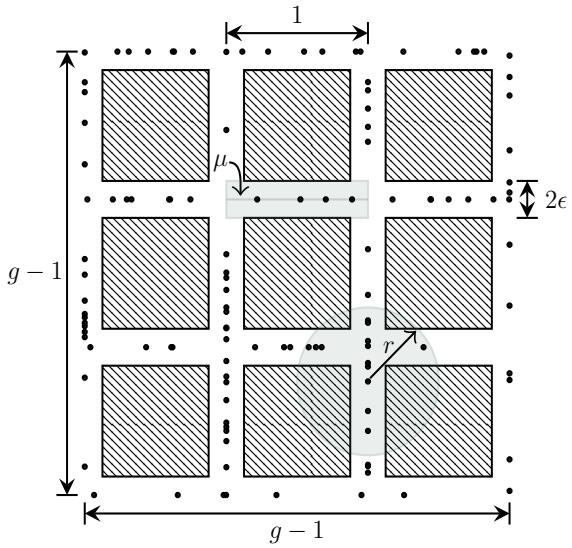
Chapter 2

Model

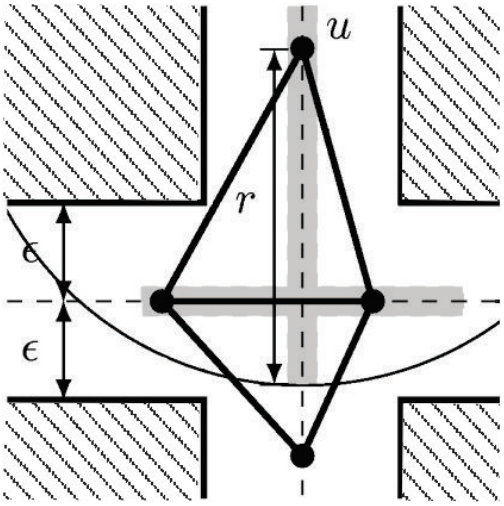
Let consider networks in which devices participates in routing by forwarding data for other devices, so the determination of which devices forward data is made dynamically on the basis of network connectivity. Also let assume that communication is wireless, such ad hoc wireless networks in practice often have to operate in obstructed environments, where the propagation of the signal cannot be described by communication models that assume open-space and omnidirectional signal propagation. For instance, geometric graphs studied by Clark et al. [1990], the protocol model proposed by Gupta and Kumar [2000] and the physical models addressed in Goussevskaia et al. [2013], are interesting models that do not consider the influence of obstacles.

In general, analytical work considering obstacles in wireless networks is scarce, most work has been focused on randomly positioned devices (Almiron et al. [2013]; Frieze et al. [2009]; Penrose [2016]). In particular, in Almiron et al. [2013], a two-layer model was proposed, in which the deployment region consists of a regular grid structure, overlaid with a set of randomly-positioned wireless devices. Communication links are formed between devices if distance and line-of-sight conditions are met. An instance is defined by four parameters: grid size, device density, street width and transmission range. For a given combination of grid size and device density, the so-called Critical Transmission Range (CTR) for connectivity is computed, above which value the network becomes connected with high probability (w.h.p.).

It is not unusual to have technological and/or market restrictions on the settings of wireless devices, making such analytical bounds on the CTR not useful in certain practical situations. Alternatively, we can think about the transmission range as a constant and analyze the scalability of the network in terms of size. When the size of the grid goes to infinity, an infinite amount of device components emerges, independently of the transmission range. As a consequence, we have a restriction on the grid size and



(a) Grid layout.



(b) LoS: The gray region is the coverage of device u

Figure 2.1: OWN model: deployment and connectivity.

device density. More specifically, larger grids require higher device densities.

In order to study obstructed wireless networks, we need a model that captures some essential characteristics of obstructed environments encountered in practice and is simple enough to provide analytical tools for network properties, such as connectivity. The model proposed in Almiron et al. [2013] meets these requirements, so we decided to use it in our work. We refer to this model as (Random) OWN and describe it below.

The deployment area is defined as a regular grid structure, which can be viewed as a Manhattan-style street map[1]. In this way, street segments are represented by the grid edges, street intersections by the grid vertices, and obstacles are represented by the city blocks (see Figure 2.1a). This grid layout is then overlaid by a set of communication devices, which are deployed arbitrarily over the street segments. Pairs of devices are able to communicate with each other as long as they are close enough, considering their wireless signal strength, and there is no obstacle (city block) between them. This two-layer model provides us with enough abstraction to analyze many desirable network properties. More formally, the OWN model is defined as follows.

Definition 1. *Obstructed Wireless Network (OWN): An instance of an OWN is defined by the following parameters: grid size g, street width ϵ, communication range r, and device set D. As illustrated by Figure 2.1a, device deployment is done over the $g \times g$ grid. Each street is comprised of $g-1$ blocks, to which we refer as segments and g crossroads, to which refer as intersections. The normalized length of a segment is set*

[1]The grid structure can also be viewed as a tunnel or indoor corridor lattice.

to one, and the width a street is set to $0 < 2\epsilon < 1$. *The communication devices in* D *are positioned at arbitrary coordinates, but always on uni-dimensional and centralized lines along street segments. The communication links are established between any pair of devices* $(u, v) \in D$, *that meet two criteria: (1) Euclidean distance:* $d(u, v) \leq r$; *and (2) Visibility: there is no obstacle between* u *and* v, *i.e., there is Line-of-Sight (LoS) (see Figure 2.1b).*

Given an OWN, several questions can be formulated concerning its connectivity. If the network is disconnected, partitioned into several components, then the connectivity might be provided by increasing the communication range. Alternatively, if the communication range cannot be altered due to technological restrictions, then one might introduce additional devices or access points, to connect the network components to a backbone infrastructure. The former approach was studied in Almiron et al. [2013], which we briefly discuss below. The latter approach we formulate as the OWN-BC problem, which we define and study in Chapters 4 and 5.

The deployment of devices over the grid structure can be performed arbitrarily in a deterministic or probabilistic manner. Below, we define a probabilistic version of the OWN model, which uses a uniformly distributed device deployment and was first introduced in Almiron et al. [2013].

Definition 2. *Random OWN: In addition to the parameters defined in Definition 1, the Random OWN uses the parameter node density* μ. *The deployment of devices in set* D *is performed in the following manner: in each street segment of the grid,* μ *communication devices are deployed uniformly at random (on uni-dimensional and centralized lines). Therefore,* $|D| = 2g(g-1) \cdot \mu$.

In Almiron et al. [2013] discrete percolation theory elements were applied to study global network properties, such as the formation of a giant connected component and the CTR for connectivity.

Definition 3. *CTR for Connectivity: Suppose* n *devices are distributed in a deployment region. The CTR for connectivity can be defined as the minimum transmission range, denoted by* r_c, *which induces a communication graph with a unique connected component, including all* n *devices.*

Therefore, the CTR is a threshold, beyond which connectivity is warranted w.h.p. (*super-critical scenarios*), and under which the network becomes fragmented, or partitioned into several connected components (*sub-critical scenarios*).

Theorem 1. *(Almiron et al. [2013]) Given a Random OWN, the CTR for connectivity w.h.p., denoted by r_c, is*

$$r_c = \frac{\ln\left(g^{a+1/2}\right) + \ln\left(\mu - 1\right)}{\mu} \tag{2.1}$$

for $a > 0$, whenever $\epsilon \geq \epsilon_c$.

The value ϵ_c is the so-called *critical width*, i.e., the minimal value of ϵ that guarantees the probability of connectivity at street intersections is higher than the probability of connectivity at street segments. Any positive value of parameter a induces connectivity when $g \to \infty$, and higher values can be used for faster convergence according to the specific value of g.

In this work we study both the deterministic and the probabilistic versions of the OWN model. Note that both models have potential applications in practice. From the network engineering perspective, the probabilistic network can provide insights into scenarios, where the exact location of the communication devices cannot be chosen by the engineer but is rather imposed by the, e.g. urban, environment. The deterministic version, on the other hand, is useful for situations where devices can be placed anywhere, for example, inside a building or a tunnel. Moreover, the deterministic model is useful for theoretical analysis of the computational complexity of worst case scenarios. The results that we present in Chapters 3 and 6 are based on the probabilistic model specified in *Definition 2*; whereas the results that we present in Chapters 4 and 5 were obtained in the deterministic model, stated in *Definition 1*.

Chapter 3

Scalability of Connectivity in Ad Hoc Obstructed Wireless Networks

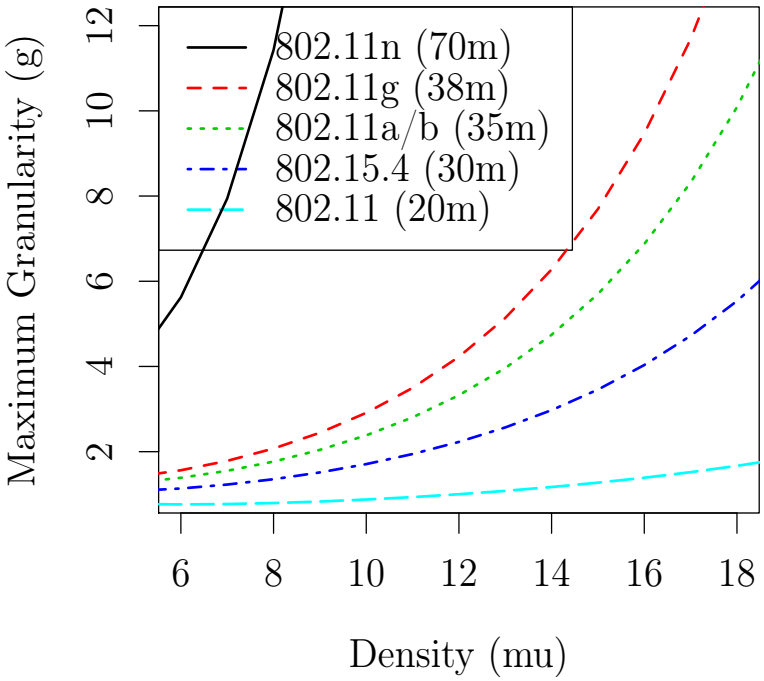
In Chapter 2, we presented a model that can be easily adjusted for *real* city proportions. We will analyze the scalability of connectivity in ad hoc Random OWN, considering cities dimensions. For such, expression (2.1) computes the CTR for connectivity, for a given grid size g and device density μ in a ad hoc Random OWN. Alternatively, we can think about the transmission range as a constant $\tilde{r}$ and determine, afterward, the scalability of the network in terms of size.

Note that when $g \to \infty$, an infinite amount of connected components emerges, independently of the value of $\tilde{r}$. As a consequence, for a fixed value of $\tilde{r}$, we have a restriction on values for g and μ. More specifically, big values of g require higher values of μ. This relationship is given by manipulating the expression (2.1) as follows:

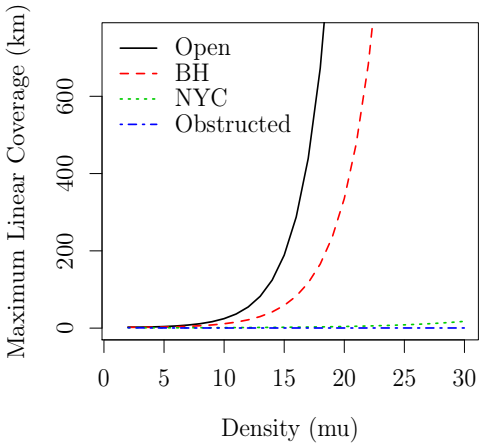
$$g = \left(\frac{e^{\mu\tilde{r}}}{\mu - 1} \right)^{\frac{2}{3}}, \tag{3.1}$$

whenever we use the convergence factor $a = 1$.

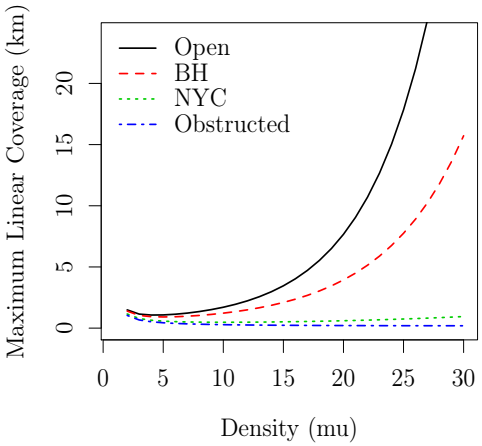
In order to demonstrate the practical implications of Expression (3.1), let us take as an example typical values of transmission range for the IEEE 802.15.4 and 802.11 legacy, a, b, g and n standards in a "Standard City" with street blocks of size 100m. Figure 3.1a illustrates the upper bound for the grid size as a function of device density, for each communication technology. For instance, a network with IEEE 802.15.4 devices configured to transmit up to 30m at 250kbps requires a dense deployment, about 30 devices per segment, to be able to scale to a grid of 42×42, while maintaining connectivity. On the other hand, if communication technology can be updated such that the transmission range is 40m, with the same device density,



(a) Relation between density μ and grid size g in a Standard City.



(b) IEEE 802.11n devices with 70m of range transmitting at 72Mbps.



(c) IEEE 802.15.4 devices with 30m of range transmitting at 250kbps.

Figure 3.1: Scalability of connectivity in different urban environments and communication technologies.

connectivity is achieved w.h.p. in much larger urban areas, with grid sizes up to 315. It can be seen that small variations in the configuration of transmission range have a high impact on the scalability of connectivity in ad hoc OWN.

Next, we analyze how the geometry of specific cities, such as Standard City (SC), Belo Horizonte (BH), and New York City (NYC), can influence the scalability of a ad hoc network topology. These three cities present similar street widths (22m, 20m, and 18m, respectively) but quite different block lengths (100m, 120m, and 274m, respectively). Figures 3.1b and 3.1c show the maximum *linear coverage*, using IEEE 802.11n and IEEE 802.15.4, respectively. Linear coverage is defined as the total street length covered by the maximum grid size connected w.h.p., normalized by the segment length of each urban scenario, i.e., $y = 2g(g-1)$, where g is given by expression (3.1), using the normalized radii values $\tilde{r}[SC, BH, NYC] = \left\{1, \frac{100}{120}, \frac{100}{274}\right\} \times r$.

We can see that, for the same radio technology, the NYC topology presented the worst scaling properties, due to its long and narrow street blocks. In fact, we can see that the IEEE 802.15.4 standard is not suitable to ensure connectivity in a ad hoc grid topology like that of NYC. BH and Standard City, on the other hand, present significantly better scaling properties than NYC, due to shorter and proportionally wider street blocks, which, for a given linear coverage, provide more options (paths) to establish connectivity.

To sum up, we showed that the Random OWN model is suitable to represent real urban obstructed scenarios and, provided there is enough flexibility in the model parameters, connectivity scales relatively well with the size of the grid. If some parameters, such as the radio transmission range, are restricted, however, the necessary device density explodes to meet the connectivity requirement. This behavior points to a necessity of other network topologies, such as connected backbone structures, in order to obtain connectivity in OWN.

Chapter 4

Obstructed Wireless Network Backbone Cover Problem

In this chapter, we turn our attention to *sub-critical deterministic* scenarios of an OWN with a combination of parameters D, g, ϵ and r that results in a disconnected network. A network comprised of communication devices with arbitrary coordinates on the grid, partitioned into $|C| > 0$ connected components, where each component $C_i \in C$ is comprised of one or more interconnected devices.

In such a scenario, we are interested in the following question: given a global connectivity infrastructure, e.g. the Internet, how many access points, or base stations, do we need to add to the OWN so that every component has access to the backbone. Firstly, we elaborate on how to enumerate all possible locations for backbone access points, given an instance of OWN. Then, we formulate the problem of how to select the minimum number of access points and guarantee that all devices will have connectivity to the global backbone. We assume that we interchangeably use the terms *base station* and *access point*, as well as *connecting* and *covering* (a component). Moreover, the term *backbone* is used to denote both a set of access points to the global connectivity infrastructure (e.g. the Internet) and the infrastructure itself.

Candidate locations for access points: Consider an instance of OWN, comprised of a set C of components. An access point can be positioned at any candidate location on the grid, as long as it has a communication link to at least one device within at least one component $C_i \in C$. The existence of a communication link is determined by the distance ($\leq r$) and visibility criteria, defined in Chapter 2. Note that the communication range of a base station is equal to that of a regular device. If a candidate location does not cover any combination of components already covered by some other candidate location, then it is considered redundant. All non-redundant candidate loca-

tions for potential access points to the backbone, denoted by $\mathcal{B}$, can be enumerated as outlined in Algorithm 1. Note that in Algorithm 1 there is one base station $\mathcal{B}_j = \{C_i\}$, of size one, for each component (lines 1-2), one base station $\mathcal{B}_k = \{C_i, C_j\}$, of size two, for every pair of components at distance $\leq 2r$ on the same street (lines 3-6), and one for every pair, triple and quadruple of components that could be connected to an access point located at (close to) each street intersection of the grid (lines 7-19). Note that no base station can connect more than 4 components simultaneously. The running time of Algorithm 1 is $O(|D|) + O(g^2)$.

Definition 4. *OWN Backbone Cover Problem (OWN-BC): Consider an OWN, partitioned into a set of components C. Consider a set $\mathcal{B}$ of candidate locations for access points to the backbone, and denote by $\mathcal{B}(C_j) \subseteq \mathcal{B}, C_j \in C$, the subset of candidate locations that cover each component of the OWN. The objective of the OWN-BC is to connect every component to the backbone by activating at least one access point in $\mathcal{B}_i \in \mathcal{B}(C_j) \subseteq \mathcal{B}, \forall C_j \in C$, while minimizing the total number of active access points:*

$$\begin{array}{lll} min & \sum_{\mathcal{B}_i \in \mathcal{B}} b_i & \\ s.t. & \sum_{\mathcal{B}_i \in \mathcal{B}(C_j)} b_i \geq 1, & \forall C_j \in C \\ & b_i \in \{0,1\}, & \forall \mathcal{B}_i \in \mathcal{B}. \end{array} \tag{4.1}$$

Note that a trivial upper bound for OWN-BC is to activate one access point in every single component: this will connect every device to the backbone at cost $|C|$. A trivial lower bound is $|C|/4$, since no access point can connect more than 4 components. Therefore, the optimum solution has cost in the range $[|C/4|, |C|]$.

The problem of finding a minimum-size backbone cover is an optimization problem. Its decision version asks for the existence of a backbone cover with at most k access points for an instance of a OWN with parameters g, ϵ and r. We next prove that OWN-BC is NP-complete.

4.1 NP-Completeness

We prove that OWN-BC is NP-hard by reduction from the Exact Cover by 3-Sets (X3C) Problem, which is a classical NP-complete problem, formulated by Garey and Johnson [2002] as follows.

Definition 5. *Exact Cover by 3-Sets Problem (X3C): Given a set $X = \{x_1, x_2, \ldots, x_{3q}\}$ and a collection $Y = \{Y_1, Y_2, \ldots, Y_{|Y|}\}$, such that Y_i is a 3-element subset of X, X3C consists in deciding whether there is an exact cover of X by sets of*

Algorithm 1: Candidate locations for base stations

Input: Instance OWN(D, g, ϵ, r)
Output: $\mathcal{B}$: set of candidate base stations

Let C be the set of components in OWN; $\mathcal{B} = \emptyset$;
foreach $C_i \in C$ **do** $\mathcal{B} = \mathcal{B} \cup \{\{C_i\}\}$;
Let S be the set of streets and D_i the set of devices located on street S_i;
foreach $S_i \in S$ **do**
    Let A be an array with the positions (in increase order) of D_i set at S_i;
    $d_1 = 0$;
    while $d_1 < |A| - 1$ **do**
        $d_1 = d_1 + 1$; $d_2 = d_1 + 1$;
        if $d_1 \in C_1 \wedge d_2 \in C_2 \wedge C_1 \neq C_2 \wedge r < A[d_2] - A[d_1] \leq 2r$
        then $\mathcal{B} = \mathcal{B} \cup \{\{C_i, C_j\}\}$;
    end
end
Let I be the set of g^2 street intersections in OWN;
foreach $I_i \in I$ **do**
    Let $\mathcal{F} = \{d_N \in C_N, d_E \in C_E, d_W \in C_W, d_S \in C_S\}$ be the 4 closest devices to I_i that belong to components $C_N, C_E, C_W, C_S \in C$, respectively;
    foreach $(d_1, d_2) \in \mathcal{F}^2 \mid C_1 \neq C_2 \wedge d_1 \in C_1 \wedge d_2 \in C_2$
     do if $\exists$ location $L \mid \exists$ links $(L, d_1) \wedge (L, d_2)$
     then $\mathcal{B} = \mathcal{B} \cup \{\{C_1, C_2\}\}$;
    foreach $(d_1, d_2, d_3) \in \mathcal{F}^3 \mid C_1 \neq C_2 \neq C_3$ **do**
     if $\exists$ location $L \mid \exists$ links $(L, d_1) \wedge (L, d_2) \wedge (L, d_3)$
     then $\mathcal{B} = \mathcal{B} \cup \{\{C_1, C_2, C_3\}\}$;
    foreach $(d_1, d_2, d_3, d_4) \in \mathcal{F}^4 \mid C_1 \neq C_2 \neq C_3 \neq C_4$
     do if $\exists L \mid \exists\ (L, d_1) \wedge (L, d_2) \wedge (L, d_3) \wedge (L, d_4)$
     then $\mathcal{B} = \mathcal{B} \cup \{\{C_1, C_2, C_3, C_4\}\}$;
end
return $\mathcal{B}$;

Y (i.e., if it possible to select mutually disjoint sets from Y such that their union is exactly X).

In the following lemma, an instance of X3C is reduced, in polynomial time, to an instance of OWN-BC, the decision version of OWN-BC, such that the existence of solution on the latter implies the existence of a solution on the former, and conversely.

Lemma 1. *OWN-BC is* NP*-hard.*

Proof. Consider an arbitrary X3C instance $\{X, Y\}$, $|X| = 3q$. Consider a set D of communication devices, to be placed at specified locations on the grid. Construct an instance $\{D, g, \epsilon, r, \mathsf{k}\}$ of the decision version of OWN-BC as follows:

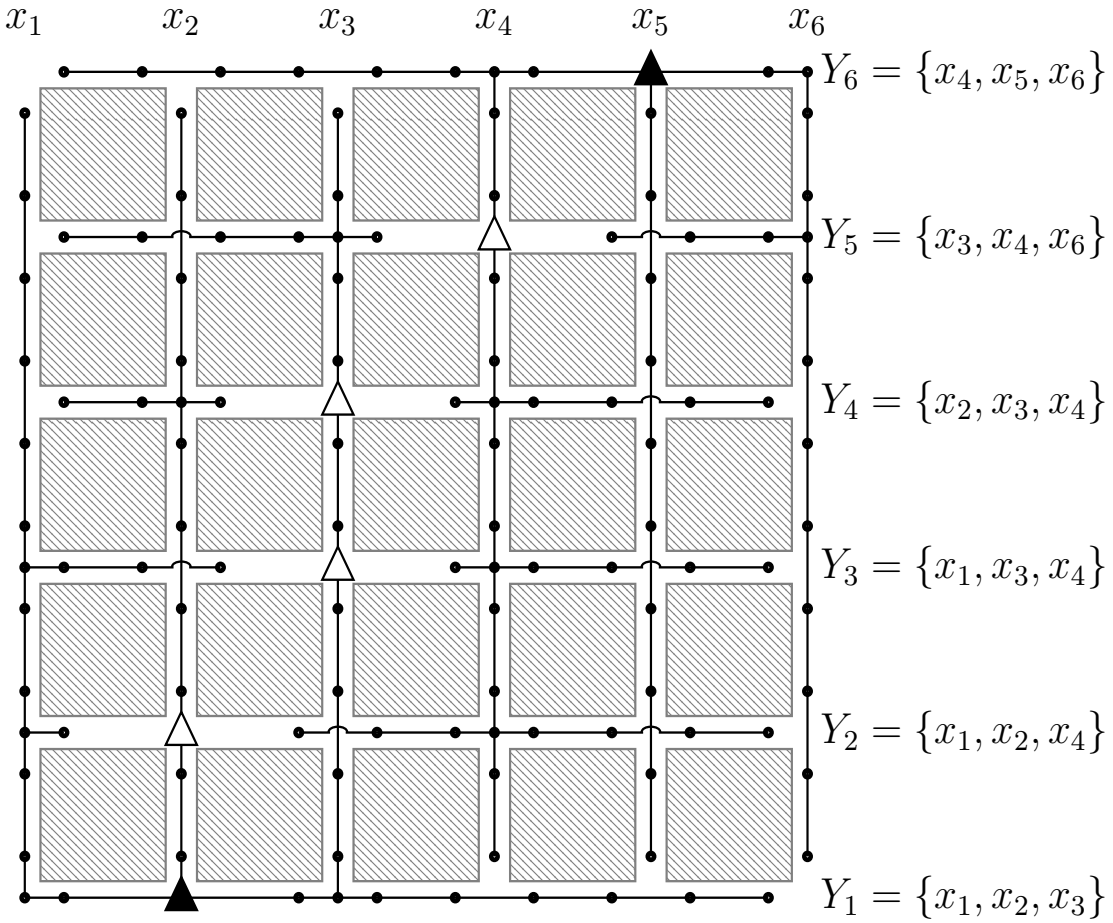


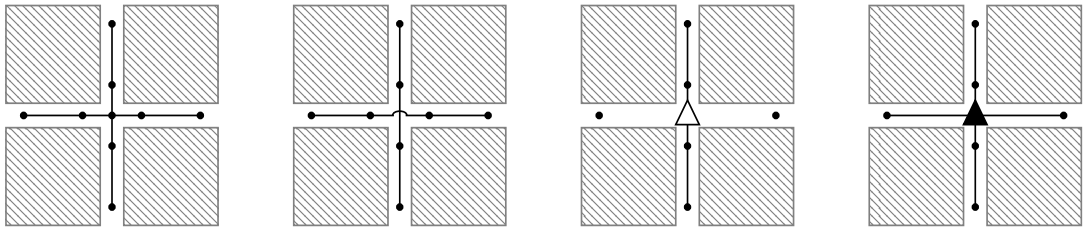
Figure 4.1: X3C to OWN-BC reduction example: $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, $Y = \{Y_1 = \{x_1, x_2, x_3\}, Y_2 = \{x_1, x_2, x_4\}, Y_3 = \{x_1, x_3, x_4\}, Y_4 = \{x_2, x_3, x_4\}, Y_5 = \{x_3, x_4, x_6\}, Y_6 = \{x_4, x_5, x_6\}\}$

1. If $|X| > |Y|$ then do step (1.a): add $|X| - |Y|$ copies of an arbitrary set $Y_k \in Y$ to the X3C instance. Else if $|Y| > |X|$ then do step (1.b): add variables in groups of 3 (a total of $3(|Y| - |X|)$) and add one set containing each group of 3 new variables to the X3C instance. Next, since now $|X| > |Y|$, repeat step (1.a) until $|X| = |Y|$.

2. Set parameters of the OWN-BC instance as follows: $g = |X|$, $0 < \epsilon < \frac{1}{8}$, $r = \frac{3}{4}$, $\mathsf{k} = |X|/3 = q$, and $|D| = 4g(g-1)$;

3. Place two devices per grid segment at positions $\frac{1}{4}$ and $\frac{3}{4}$ of each street segment;

4. Note that the proximity and visibility criteria induce a network with a total of $2g$ components: one on each column and one on each row of the grid. This occurs because all street intersections are *obstructed* (Figure 4.2b) due to lack of visibility, i.e., no column can communicate to any row of the grid. On the other



(a) connected (b) obstructed (c) candidate $\mathcal{B}_i$ (d) active $\mathcal{B}_i$

Figure 4.2: Intersection types in the OWN-BC instance.

hand, the distance criterion on street segments is satisfied, so all devices on the same line (vertical or horizontal) are connected.

5. Consider a one-to-one correspondence between each column (vertical street) of the OWN-BC instance to an element $x_i \in X$ in the X3C instance. Moreover, consider a one-to-one relation between each row (horizontal street) of the OWN-BC instance and a 3-set $Y_j \in Y$ in the X3C instance, as explained below (Figure 4.1).

6. For each row Y_j of the grid, consider the ordered sequence $Y_j = \{x_l^j, x_m^j, x_r^j\} \in Y$, such that $l < m < r$, i.e., x_l^j is the leftmost element, x_m^j is in the middle, and x_r^j is the rightmost element in row Y_j (Figure 4.3). Consider the intersection (Y_j, x_m^j) on the grid and move the closest device to the left of this intersection to the center of intersection (Y_j, x_l^j), which becomes *connected* (Figure 4.2a). In a similar way, move the closest device to the right of intersection (Y_j, x_m^j) to the center of intersection (Y_j, x_r^j), which also becomes *connected* (Figure 4.2a). Note that, after this modification, intersection (Y_j, x_m^j) becomes a *candidate* location for a base station (Figure 4.2c), and row Y_j no longer contains a unique component, but rather two components, that could be considered as extensions of the (vertical) components on columns x_l^j and x_r^j, respectively, while all vertical components remain connected, resulting in a total of $3\mathsf{k}$ disconnected components.

7. Note that now there is a candidate location for an access point at every *horizontally disconnected intersection* (Y_j, x_m^j), $\forall Y_j \in Y, x_m^j \in X$ (Figures 4.1, 4.2c, 4.3). In fact, these are the only candidate locations for access points connecting more than two components, namely the vertical component on column x_m^j, and the two horizontal components on row Y_j (see Lemma 2).

It is easy to see that the reduction process has cost polynomial in $|X|$ and $|Y|$. Moreover, it can be seen that $3\mathsf{k} = 3g = 3q = |X|$ components are formed in the process. Next, we show that there is a backbone cover of size k iff there is a X3C of the same size.

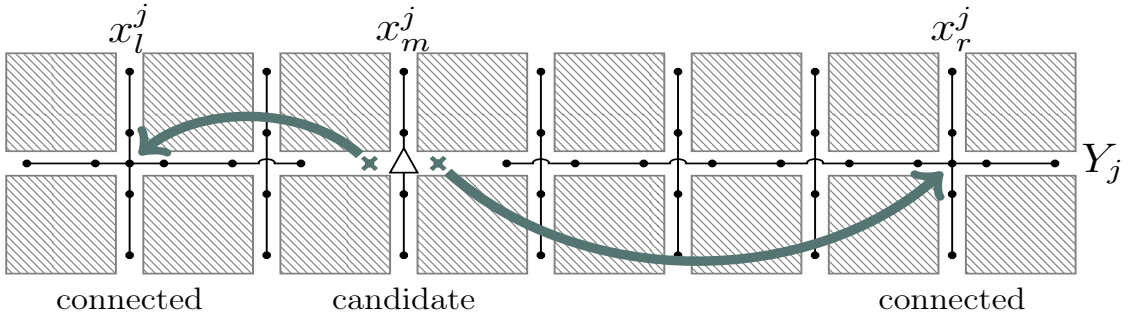


Figure 4.3: Step (6) of Lemma 1

X3C $\Rightarrow$ *OWN-BC:* Suppose there exists a solution S_{X3C} containing exactly k disjoint 3-sets $Y_j \in Y$ that cover all ($3\mathsf{k}$) elements in X. We construct a backbone cover by activating an access point at the candidate location created in step (6) at the grid intersection (Y_j, x_m^j), $\forall Y_j \in S_{\text{X3C}}, x_m^j \in Y_j$ ((Y_j, x_m^j) is now *active*, Figure 4.2d). Since S_{X3C} covers all elements in X, by construction of the OWN-BC instance, k access points will be activated at the horizontally disconnected grid intersections $(Y_j, x_i), \forall x_i \in X$, each access point connecting two horizontal components on row Y_j to one vertical component on column x_i, thus connecting all $3\mathsf{k}$ components of the OWN.

OWN-BC $\Rightarrow$ *X3C:* Suppose there is a solution S_{BC} containing exactly k access points that connect all the devices, originally partitioned into $3\mathsf{k}$ components. By Lemma 2, the only candidate locations in the OWN-BC instance that connect ≥ 3 components are those at horizontally disconnected grid intersections (denoted by the triangles in Figure 4.1). Therefore, there can be no backbone cover with less than k access points, and any backbone cover with k elements must contain only access points that cover exactly 3 components. By construction, there is one such access point at every horizontally disconnected intersection $(Y_j, x_i), x_i \in X, Y_j \in Y$. Thus, the corresponding k 3-sets $Y_j \in Y$ do cover all $3\mathsf{k}$ elements $x_i \in X$ in the X3C instance. This concludes the Lemma. □

Lemma 2. *Given an instance of OWN-BC constructed in Lemma 1, the grid intersections (Y_j, x_m^j), horizontally disconnected in step (6), are the only candidate locations on the grid, where one access point could connect more than 2 components simultaneously to the backbone.*

Proof. Consider all candidate locations for an access point in the instance of OWN-BC, constructed in Lemma 1. Since devices located on vertical lines are all connected, any base station on a vertical street segment and not close enough ($d > 1/4$) to a street

intersection, would connect at most one component to the backbone. Therefore, in order to connect two components (simultaneously) to the backbone, an access point has to be positioned close to a street intersection. Furthermore, in order to connect 3 components simultaneously, an access point has to be placed at the center of one of the horizontally disconnected intersections (Y_j, x_m^j), which will connect the two horizontal components on row Y_j, since their distance to the center is exactly $r = 3/4$, plus the vertical component on column x_m^j, since its distance to the closest device in that column is $1/4 < r$ and there is no line-of-sight obstruction. Finally, there is no possibility to connect more than three components simultaneously by one base station. We conclude that the only candidate locations for access points connecting at least (and exactly) three components are those at the center of the horizontally disconnected intersections. □

Lemma 3. *OWN-BC is in* NP.

Proof. Using a deterministic Turing Machine, it is possible to check a given candidate solution $\mathcal{S}$ in time $O(|D|^2)$, $|D| = 4g(g-1)$, as follows. Fist, check if $|\mathcal{S}| \leq \mathsf{k}$. Then, check if the set of access points in $\mathcal{S}$ induces a single connected component in the OWN network. For that, compute the set C of all components, using, for example, the union-find algorithm Sedgewick and Wayne [2011]. Then, for each component $\forall C_i \in C$, check if there is an active access point $\exists \mathcal{B}_j \in \mathcal{S}$ with a communication link (distance and visibility criteria met) to at least one device $d \in C_i$. Clearly, all steps can be done in polynomial time in $|D|$. □

Theorem 2. *OWN-BC is* NP*-complete.*

Proof. By Lemma 1, OWN-BC is NP-hard, By Lemma 3 it is in NP. □

Theorem 2 implies that, unless $P = NP$, there is no polynomial-time algorithm to solve arbitrary instances of OWN-BC. Therefore, in Chapter 5, we discuss approximation algorithms for the problem. In particular, we propose a factor-2 approximation algorithm, which returns optimal solutions for certain instances. In Chapter 6, we investigate problem parameters that lead to optimal solutions, in other words, we show how to ensure necessary conditions for obtaining instances of OWN-BC that can be solved optimally in polynomial time.

Chapter 5

Approximation algorithm

In this chapter, we turn our attention to finding close to optimum solutions for OWN in polynomial time. Note that we are considering the *deterministic* version of the problem, in which all communication devices of an instance have arbitrary coordinates on the grid.

Note that the OWN-BC problem can be modeled as the Set Cover problem, where network components correspond to the elements and base stations to the sets. Given that an access point to the backbone can cover at most 4 disconnected network components, as described in Algorithm 1, OWN-BC can be viewed as a special case of the k-Set Cover (k-SC) problem, where $k = 4$. We first discuss the applicability of algorithms that already exist in the literature for the k-Set Cover problem and then propose a new approximation algorithm, based on properties specific to OWN. The k-Set Cover can be defined as follows.

Definition 6. *k-Set Cover Problem (k-SC): Given a set of elements* $X = \{x_1, x_2, \ldots, x_n\}$*, and a set* $C = \{C_1, C_2, ..., C_m\}$*,* $C_i \subseteq X, |C_i| \leq k$ *of subsets of* X *of size at most* k*. The objective of k-SC is to find a subset* $\mathcal{S} \subseteq C$ *of minimum cardinality (or cost) that covers all elements of* X*, i.e.,* $\cup_{S_i \in \mathcal{S}} S_i = X$*.*

k-SC is a well-studied problem, and many approximation algorithms can be found in the literature. It is known that the greedy algorithm gives a H_k-approximation, where $H_k = \sum_{i=1}^{k} 1/i$ is the k-th harmonic number (Garey and Johnson [2002]). Hence, this gives an easy $\frac{25}{12}$-approximation for OWN-BC, since $k = 4$. More involved approximation algorithms have been developed and the current state-of-the-art for k-SC is an $(H_k - \frac{196}{390})$-approximation proposed in Levin [2008], resulting in a 1.58-approximation for OWN-BC.

We are particularly interested in OWN. During our simulations, we observed that frequently only a small fraction of the connected components can be covered by 3- or 4-sets. Motivated by this observation, we propose a simpler approximation algorithm, Algorithm 2, that guarantees a factor-2 approximation in the worst case, but performs better in practice, yielding optimal solutions in certain instances. More simulations aspects and the quality of the solutions obtained by approximation algorithms are discussed in Chapter 7.

Algorithm 2 works as follows. Given an instance of the OWN-BC problem, it builds an instance of the Edge Cover problem [1]. The minimum edge cover can be found in polynomial time by finding a maximum matching and extending it greedily so that all vertices (or network components) are covered Garey and Johnson [2002]. We show that the optimum solution to the Edge Cover problem provides a 2-approximation for the OWN-BC problem.

Consider the instance of OWN-BC, in which the communication devices are partitioned into a set C of components. Moreover, consider the set $\mathcal{B}$ of candidate locations for access points to the backbone, returned by Algorithm 1. Firstly, note that each component $C_i \in C$ that is covered by only one candidate location $\mathcal{B}_i \in \mathcal{B}, |\mathcal{B}_i| = 1$ of cardinality 1, i.e., C_i can not share an access point with any other component $C_j \in C, j \neq i$, must have that access point activated exclusively for itself. We call them *isolated* components. Before solving the edge covering problem, isolated components, as well as the corresponding 1-sets in $\mathcal{B}$, can be safely removed from the input (Algorithm 2, lines 1-4).

Algorithm 2 builds a graph $G = \{V, E\}$ from the OWN-BC instance as follows (lines 5-10). A vertex set $V(G)$ includes a vertex for every component $C_i \in C$ covered by a candidate location $\mathcal{B}_i \in \mathcal{B}, |\mathcal{B}_i| \geq 2$ of cardinality at least two. The edge set $E(G)$ includes an edge for every 2-set in $\mathcal{B}, |\mathcal{B}_i| = 2$. Furthermore, all 3- and 4-sets $\mathcal{B}_i \in \mathcal{B}, |\mathcal{B}_i| \geq 3$ are broken into all possible pairs of their elements. Then, the minimum edge cover E' is computed on graph G (line 11). Note that every edge in the minimum edge cover E' comes from one (or more) sets $\mathcal{B}_i \in \mathcal{B}'$, hence from an edge cover we can recover a backbone cover. The resulting solution $\mathcal{S}$, together with the 1-sets covering the isolated components in the OWN instance corresponds to a solution to the OWN-BC (line 12).

It remains to show that Algorithm 2 is actually a factor-2 approximation that runs in polynomial time, which is proved in Theorem 3.

[1] An Edge Cover of a graph is a set of edges, such that every vertex of the graph is incident to at least one edge of the set.

Algorithm 2: 2-approximation algorithm for OWN-BC.

Input: OWN-BC=$\{C, \mathcal{B}\}$: network components and candidate base stations (sets);
Output: $\mathcal{S} \subset \mathcal{B}$: backbone covering all components.

$\mathcal{B}' = \{\mathcal{B}_i \mid \mathcal{B}_i \in \mathcal{B} \wedge 2 \leq |\mathcal{B}_i| \leq 4\}$;
$C' = \{C_i \mid C_i \in C \wedge \exists \mathcal{B}_i \in \mathcal{B}' \wedge C_i \in \mathcal{B}_i\}$; (non-isolated components)
$C'' = \{C_i \mid C_i \in C \wedge C_i \notin C'\}$; (isolated components)
$\mathcal{B}'' = \{\mathcal{B}_i \in \mathcal{B} \mid |\mathcal{B}_i| = 1 \wedge \exists C_i \in C'' \mid C_i \in \mathcal{B}_i\}$; (isolated base stations)
$G = (V = C', E = \emptyset)$;
for *each* $\mathcal{B}_i \in \mathcal{B}'$ **do**
    for *each* $u, v \in \mathcal{B}_i$ **do**
        Add $e = (u, v)$ to E with label $l(e) = \mathcal{B}_i$;
    end
end
Compute an Edge Cover E' of G;
return $\mathcal{S} = \mathcal{B}'' \cup \{\mathcal{B}_i \in \mathcal{B}' \mid \exists e \in E' \wedge l(e) = \mathcal{B}_i\}$;

Theorem 3. *Algorithm 2 gives a* 2*-approximation for OWN-BC in polynomial time.*

Proof. Since a minimum edge cover can be found in polynomial time (Garey and Johnson [2002]), the algorithm clearly runs in polynomial time.

For the proof of the approximation factor, note that an optimal solution for 4-Set Cover C^* can be converted to an edge cover E^* of the constructed graph as follows: For each 1-set S, take any edge containing S as a subset; For each 2-set S, take the edge corresponding to it; For each 3-set S, take two distinct edges with endpoints in S; For each 4-set S, take two disjoint edges with endpoints in S. Note that E^* is an edge cover with at most twice the number of elements of C^*. Since the edge cover E' is minimum, hence we have $|E'| \leq |E^*| \leq 2|C^*|$, which completes the proof. □

We hypothesize that the quality of the approximation tends to be better if the number of 3- and 4-sets is small relatively to the total number of candidate locations, since, as the above proof suggests, these are the kind of sets that can increase the gap between $|C^*|$ and $|E^*|$. We investigate deeper into this issue in the following Chapters 6 and 7.

Chapter 6

Characterizing Polynomial Complexity Instances

In this chapter, we turn our attention back to the Random OWN model, defined in Chapter 2. We assume that μ communication devices are distributed uniformly at random along each segment of the grid. So, given parameters g, μ, ϵ, and r, only the distribution function, and not the exact coordinates of the devices, are known.

As we discussed in Chapter 5, OWN-BC can be viewed as a special case of the k-SC problem, where network components correspond to the elements, access points correspond to the sets in the k-SC instance, and $k = 4$. In cases where $k = 2$, the problem reduces to the Edge Cover problem and can be solved optimally in polynomial time by Algorithm 2. Therefore, we characterize OWN-BC instances in the Random OWN model, in which all candidate locations for access points to the backbone connect at most two network components. Next, we define the probability of this event.

Definition 7. P_{poly} ***(probability of polynomial connectivity with a backbone):*** *Consider an instance I of the OWN-BC problem in a Random OWN. We say that the event of polynomial connectivity with a backbone occurs when all candidate locations for access points in I connect at most two network components. When this event occurs, Algorithm 2 solves the OWN-BC problem instance optimally in polynomial time. We refer to the probability of this event as P_{poly}.*

Consider the set $\mathcal{B}$ of all (non-redundant) candidate locations for access points to the backbone, returned by Algorithm 1. Note that candidate locations of cardinality $|\mathcal{B}_i| > 2$ can occur exclusively at (or *close to*[1]) street intersections, since street segments

[1]The exact distance from the center of the street intersection depends on the visibility criterion, computed for specific values of ϵ and r.

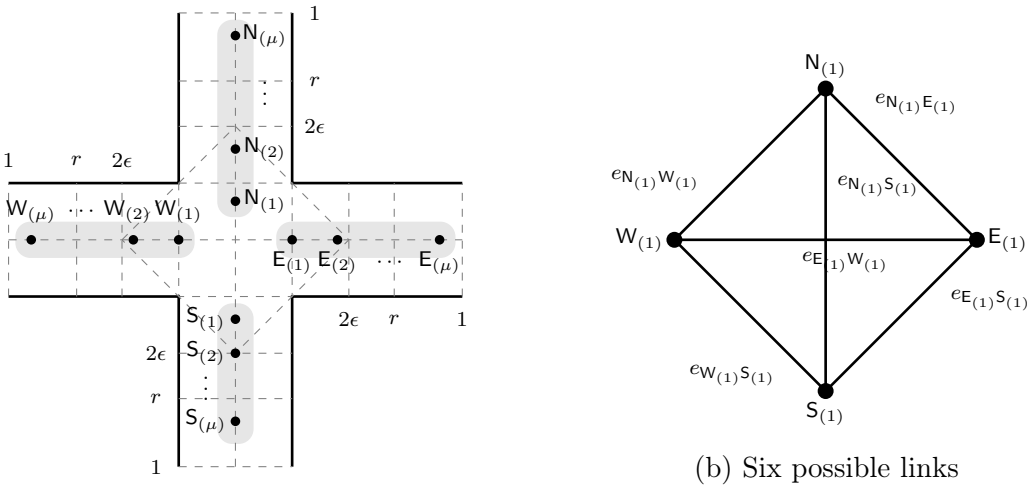
are uni-dimensional and, therefore, any access point placed on a street segment can connect either an already connected component of devices to the backbone (in which case $|\mathcal{B}_i|= 1$) or connect two endpoint devices that belong to two different components (in which case $|\mathcal{B}_i|= 2$). From now on, we will turn our attention to candidate locations at street intersections.

Consider an arbitrary street intersection, such as the one illustrated in Figure 6.1a, and the four closest devices to the center of the intersection, in each direction (North (N), East (E), West (W), and South (S)). A sufficient condition for an access point located at this street intersection to connect > 2 components is that there are < 2 communication links connecting these four devices. This is due to the fact that, the presence of less than two communication links induces the existence of at least three mutually disconnected components at (adjacent to) that street intersection.

Therefore, a sufficient condition for the event *polynomial connectivity with a backbone* is that, for all (g^2) street intersections in the (Random) OWN instance, there exist ≥ 2 communication links among the four closest devices to the center of each intersection. Note that this condition is not necessary since the event can also occur if some of the street intersections have no candidate access point or if all candidate access points in the problem instance are located on street segments.

Probability $P_{\texttt{poly}}$ depends on parameters g, μ, ϵ, and r. An analytical formulation of this probability is quite complex and involves a careful analysis of geometric correlations between combinations of links at a street intersection. In Theorem 4 we compute a lower bound for $P_{\texttt{poly}}$. In Section 7.2, we present a comparison of the obtained analytical bound to the simulated empirical probability, corroborating that the approximation is of good quality.

Firstly, let us introduce some definitions. For each of the four segments adjacent to a street intersection, we instantiate a random variable sequence $X = \{X_1, X_2, ..., X_\mu\}$ (Figure 6.1a) with the positions of μ devices deployed at each (adjacent) street segment (since there are at most four adjacent street segments to each intersection, we use the notation $X \in \{\mathsf{N}, \mathsf{E}, \mathsf{W}, \mathsf{S}\}$). For a random variable sequence $X = \{X_1, X_2, ..., X_\mu\}$, let the first order statistics (f.o.s.) be denoted and obtained by $X_{(1)} := \min\{X_1, X_2, ..., X_\mu\}$. We are interested in measuring the probability that $X_{(1)}$ (of each segment) belongs to a specific interval $[\alpha, \omega]$, such that $P_{\texttt{poly}} > 0$. Let $f_{X_{(1)}}$ denote the probability density function (p.d.f.) for the f.o.s. $X_{(1)}$, then the integral $\int_\alpha^\omega f_{X_{(1)}} dX$ represents the probability of $X_{(1)}$ to belong to the interval $[\alpha, \omega]$. We are assuming in the Random OWN model that X has a uniform distribution. It is well known that the k^{th} order statistics of the uniform distribution is a Beta random variable. In particular, the p.d.f. for f.o.s. of a uniform distribution can be expressed by $\mu(1 - X)^{\mu-1}$.



(a) Closest points to an intersection

(b) Six possible links

Figure 6.1: Intersection of a Random OWN with density μ.

Definition 8. ***Probability of*** $X_{(1)}$ ***to belong to the interval*** $[\alpha, \omega]$***:*** *(X has uniform distribution, and the dependence on μ is implicit in the notation):*

$$\begin{aligned} \int_{\alpha}^{\omega} &= \int_{\alpha}^{\omega} f_{X_{(1)}} dX = \int_{\alpha}^{\omega} \mu(1-X)^{\mu-1} dX \\ &= [-(1-X)^{\mu}]_{\alpha}^{\omega} = (1-\alpha)^{\mu} - (1-\omega)^{\mu} \end{aligned} \tag{6.1}$$

Definition 9. ***Intersection graph*** $\mathcal{G}(\mathbb{V}, \mathbb{E})$***:*** *Consider an arbitrary street intersection, we define the graph $\mathcal{G}(\mathbb{V}, \mathbb{E})$ as follows (Figure 6.1b). The vertex set is built with the f.o.s. of each segment, $\mathbb{V} = \{\mathsf{N}_{(1)}, \mathsf{E}_{(1)}, \mathsf{W}_{(1)}, \mathsf{S}_{(1)}\}$. Let $X_{(1)}$ and $Y_{(1)}$ be f.o.s. of segments sharing an intersection, we denote the communication link between them by $e_{X_{(1)}Y_{(1)}}$. The edge set $\mathbb{E}$ is determined by the Random OWN model criteria of distance and visibility, i.e., $\mathbb{E} := \{e_{X_{(1)}Y_{(1)}} | X_{(1)}$ and $Y_{(1)}$ satisfy Definition's 1 criteria$\}$.*

Definition 10. ***Sample space partition*** $\tilde{P}$***:*** *Let $\tilde{P}_e = \{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}, e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}, \overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}, \overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}\}$ be a partition of the sample space Ω based on the existence/absence of the edges $e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ and $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ in the intersection graph $\mathcal{G}$. Note that the probability of these two edges to occur is independent since they do not share an endpoint. We refer to partitions in $\tilde{P}_e$ as $\tilde{P}$. The partitions $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ and $\overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ are equivalent due to symmetry. Thus, we will use, for convenience, the symbols "+" for $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$, "|" for $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}} \equiv \overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ and "$\diamond$" to $\overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$. For example, the partition $\overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ can also be denoted by $\tilde{P}_{\diamond}$.*

Definition 11. ***Sample classes*** $\tilde{C}_{ji}$***:*** *Let $\tilde{P}_{\epsilon,r} = \{\mathsf{N}_{(1)} \in \tilde{I}_j\} \times \{\mathsf{E}_{(1)} \in \tilde{I}_i\}$ where*

$j \in \{1,2,3,4\}$, $i \in \{1,2,3,4\}$, $\tilde{I}_1 := [0,\epsilon)$, $\tilde{I}_2 := [\epsilon, 2\epsilon)$, $\tilde{I}_3 := [2\epsilon, r)$ *and* $\tilde{I}_4 := [r,1]$. *We will refer to partitions in* $\tilde{P}_{\epsilon,r}$ *as classes* $\tilde{C}_{ji}$. *These ranges are purposely related with the* r *and* ϵ *parameters so that each class allows to separate which scenarios contribute positively to* $P_{\texttt{poly}}$.

Definition 12. ***Joint probability function*** $\mathcal{C}_{\epsilon,r,\mu}(\tilde{P}, \tilde{C}_{ji})$*: probability of* $X_{(1)}$ *to simultaneously belong to intervals* $[\alpha_X, \omega_X]$ *for each random variable* $X \in \{\mathsf{N}, \mathsf{E}, \mathsf{W}, \mathsf{S}\}$, *i.e.,* $\int_{\alpha_\mathsf{N}}^{\omega_\mathsf{N}} f_{\mathsf{N}_{(1)}} d\mathsf{N} \int_{\alpha_\mathsf{E}}^{\omega_\mathsf{E}} f_{\mathsf{E}_{(1)}} d\mathsf{E} \int_{\alpha_\mathsf{W}}^{\omega_\mathsf{W}} f_{\mathsf{W}_{(1)}} d\mathsf{W} \int_{\alpha_\mathsf{S}}^{\omega_\mathsf{S}} f_{\mathsf{S}_{(1)}} d\mathsf{S}$.
If X *has uniform distribution then, by Definition 8, the product reduces to:* $\int_{\alpha_\mathsf{N}}^{\omega_\mathsf{N}} \int_{\alpha_\mathsf{E}}^{\omega_\mathsf{E}} \int_{\alpha_\mathsf{W}}^{\omega_\mathsf{W}} \int_{\alpha_\mathsf{S}}^{\omega_\mathsf{S}}$ *and function* $\mathcal{C}$ *is defined as:*

$$\mathcal{C}_{\epsilon,r,\mu}(\tilde{P}, \tilde{C}_{ji}) = \underbrace{\int_{\alpha_j}^{\omega_j}}_{\mathsf{N}} \underbrace{\int_{\alpha_i}^{\omega_i}}_{\mathsf{E}} \underbrace{\int_{\alpha(\tilde{P},\tilde{C}_{ji})}^{\omega(\tilde{P},\tilde{C}_{ji})}}_{\mathsf{W}} \underbrace{\int_{\alpha(\tilde{P},\tilde{C}_{ji})}^{\omega(\tilde{P},\tilde{C}_{ji})}}_{\mathsf{S}}, \tag{6.2}$$

where the implicit dependence on the model parameters ϵ, r, μ *are reinforced with the subscripts and the explicit dependence on partition and class has direct effect on the intervals* $[\alpha_X, \omega_X]$ *of each random variable* X. *Note that* $\mathcal{C}$ *dependence on* μ *is due to* $\int$ *(as stated in Definition 8).*

We are now ready to formulate a lower bound on the probability $P_{\texttt{poly}}$ of the event *polynomial connectivity with a backbone.*

Theorem 4. *Given an instance OWN-BC(*g, ϵ, r, μ*) in the Random OWN model, the probability of polynomial connectivity with a backbone can be lower bounded as follows:*

$$P_{\texttt{poly}} \geq p(\epsilon, r, \mu)^{g^2}, \tag{6.3}$$

where $p(\epsilon, r, \mu)$ *is defined in (6.4);*

Proof. As discussed earlier, we need that candidate access points at all g^2 grid intersections connect at most 2 network components. We denote as p the probability that this condition is satisfied at a given intersection. Since these events at different intersections are independent, we have that $P_{\texttt{poly}} = p^{g^2}$.

Using the *law of total probability*, the probability p can be expressed as the sum over the partitions $\tilde{P} \in \tilde{P}_e$ (Definition 10):

$$p(\epsilon, r, \mu) = p_+(\epsilon, r, \mu) + 2p_|(\epsilon, r, \mu) + p_\diamond(\epsilon, r, \mu), \tag{6.4}$$

where $p_+, p_|, p_\diamond$ are defined in (6.5), (6.6), and (6.7), respectively. The term $p_|$ appears twice because the partitions $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ and $\overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ are equivalent due

Table 6.1: Integration limits α's and ω's for the $\tilde{P}_+$ partition.

N	E	W	S
$[0, \epsilon]$	$[0, \epsilon]$	$[0, r-\epsilon]$	$[0, r-\epsilon]$
$[0, \epsilon]$	$[\epsilon, 2\epsilon]$	$[0, r-2\epsilon]$	$[0, r-\epsilon]$
$[0, \epsilon]$	$[2\epsilon, \frac{r}{2}]$	$[0, \frac{r}{2}]$	$[0, r-\epsilon]$
$[\epsilon, 2\epsilon]$	$[0, \epsilon]$	$[0, r-\epsilon]$	$[0, r-2\epsilon]$
$[\epsilon, 2\epsilon]$	$[\epsilon, 2\epsilon]$	$[0, r-2\epsilon]$	$[0, r-2\epsilon]$
$[\epsilon, 2\epsilon]$	$[2\epsilon, \frac{r}{2}]$	$[0, \frac{r}{2}]$	$[0, r-2\epsilon]$
$[2\epsilon, \frac{r}{2}]$	$[0, \epsilon]$	$[0, r-\epsilon]$	$[0, \frac{r}{2}]$
$[2\epsilon, \frac{r}{2}]$	$[\epsilon, 2\epsilon]$	$[0, r-2\epsilon]$	$[0, \frac{r}{2}]$
$[2\epsilon, \frac{r}{2}]$	$[2\epsilon, \frac{r}{2}]$	$[0, \frac{r}{2}]$	$[0, \frac{r}{2}]$
$[0, \frac{r}{2}]$	$[\frac{r}{2}, 1]$	$\varnothing$	$\varnothing$
$[\frac{r}{2}, 1]$	$[0, 1]$	$\varnothing$	$\varnothing$

to symmetry. Then, for each probability (p_+, $p_|$ and $p_\diamond$), we obtain a lower bound as a sum over the classes $\tilde{C}_{ji} \in \tilde{P}_{\epsilon,r}$ (Definitions 10 and 11):

$$p_+(\epsilon, r, \mu) \geq \sum_{\tilde{C}_{ji} \in \tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\underbrace{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}_{\tilde{P}_+}, \tilde{C}_{ji}), \tag{6.5}$$

$$p_|(\epsilon, r, \mu) \geq \sum_{\tilde{C}_{ji} \in \tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_|, \tilde{C}_{ji}), \tag{6.6}$$

$$p_\diamond(\epsilon, r, \mu) \geq \sum_{\tilde{C}_{ji} \in \tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{ji}), \tag{6.7}$$

in which is employed the $\mathcal{C}$ function from Definition 12.

Function $\mathcal{C}$ is constructed case by case for each pair of partition $\tilde{P}$ and class $\tilde{C}_{ji}$. This process consists in the instantiation of the α's and ω's integration limits at each segment adjacent to the intersection, where the event probability is positive ($p > 0$). The values of the integration limits used in the analysis are summarized in Tables 6.1, 6.2 and 6.3. We prove that these integration limits lead to a lower bound on probability p in Lemma 4. □

Lemma 4. *The values of the integration limits for function $\mathcal{C}$, as stated in Definition 12, listed in Tables 6.1, 6.2 and 6.3, provide lower bounds on the probabilities in expressions (6.5), (6.6), and (6.7).*

Table 6.2: Integration limits α's and ω's for the $\tilde{P}_|$ partition.

N	E	W	S
$[0, \epsilon]$	$[0, \epsilon]$	$[r, 1]$	$[0, r - \epsilon]$
$[0, \epsilon]$	$[\epsilon, 2\epsilon]$	$[r - \epsilon, 1]$	$[0, r - \epsilon]$
$[0, \epsilon]$	$[2\epsilon, \sqrt{r^2 - \epsilon^2}]$	$[r - 2\epsilon, 1]$	$[0, r - \epsilon]$
$[0, \epsilon]$	$[r, 1]$	$[0, \sqrt{r^2 - \epsilon^2}]$	$[0, r - \epsilon]$
$[\epsilon, 2\epsilon]$	$[0, \epsilon]$	$[r, 1]$	$[0, r - 2\epsilon]$
$[\epsilon, 2\epsilon]$	$[\epsilon, 2\epsilon]$	$[r - \epsilon, 1]$	$[0, r - 2\epsilon]$
$[\epsilon, 2\epsilon]$	$[2\epsilon, \sqrt{r^2 - \epsilon^2}]$	$[r - 2\epsilon, 1]$	$[0, \epsilon]$
$[\epsilon, 2\epsilon]$	$[r, 1]$	$[0, 2\epsilon]$	$[0, r - 2\epsilon]$
$[2\epsilon, r - 2\epsilon]$	$[\epsilon, 2\epsilon]$	$[r - \epsilon, 1]$	$[0, 2\epsilon]$
$[2\epsilon, r - 2\epsilon]$	$[r, 1]$	$[0, 2\epsilon]$	$[0, 2\epsilon]$
$[2\epsilon, r - \epsilon]$	$[2\epsilon, \sqrt{r^2 - \epsilon^2}]$	$[r - 2\epsilon, 1]$	$[0, \epsilon]$
$[2\epsilon, \sqrt{r^2 - \epsilon^2}]$	$[0, \epsilon]$	$[r, 1]$	$[0, r - \sqrt{r^2 - \epsilon^2}]$
$[0, r - \epsilon]$	$[\sqrt{r^2 - \epsilon^2}, r]$	$\varnothing$	$\varnothing$
$[r - 2\epsilon, r]$	$[\epsilon, 2\epsilon]$	$\varnothing$	$\varnothing$
$[r - 2\epsilon, r]$	$[r, 1]$	$\varnothing$	$\varnothing$
$[r - \epsilon, r]$	$[2\epsilon, r]$	$\varnothing$	$\varnothing$
$[\sqrt{r^2 - \epsilon^2}, r]$	$[0, \epsilon]$	$\varnothing$	$\varnothing$
$[r, 1]$	$[0, 1]$	$\varnothing$	$\varnothing$

Proof. We planned a general strategy to instantiate all intervals $[\alpha_X, \omega_X]$ which contribute positively to the probability p, defined in (6.4). Observe that a class $\tilde{C} \in \tilde{C}_{\epsilon,r}$ delimits $\mathsf{N}_{(1)}$ and $\mathsf{E}_{(1)}$ ranges. Then we just have to derive the $\mathsf{S}_{(1)}$ and $\mathsf{W}_{(1)}$ ranges according to the partition $\tilde{P} \in \tilde{P}_e$. For instance, as shown in Figure 6.2, if the edge $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ exists, then $\mathsf{S}_{(1)}$ range will be $[0, \omega_{\mathsf{S}}]$. However, if the edge $\overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ does not exist, the range will be $[\alpha_{\mathsf{W}}, 1]$. When segments are orthogonal, the integration intervals will be different, depending on the partition/class being analyzed (as explained below). Many partition/classes cases will not contribute to the event and will have probability zero (denoted by $\varnothing$ in Tables 6.1, 6.2 and 6.3).

We begin considering the partition $\tilde{P}_+ = e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$. This partition contains seven classes where $p = 0$, namely: $\{\tilde{C}_{14}, \tilde{C}_{24}, \tilde{C}_{34}, \tilde{C}_{41}, \tilde{C}_{42}, \tilde{C}_{43}, \tilde{C}_{44}\}$. These classes make the partition unfeasible, i.e., it is impossible that both edges occur. The remaining classes, $\{\tilde{C}_{11}, \tilde{C}_{12}, \tilde{C}_{13}, \tilde{C}_{21}, \tilde{C}_{22}, \tilde{C}_{23}, \tilde{C}_{31}, \tilde{C}_{32}, \tilde{C}_{33}\}$, contribute, respectively, with:

Table 6.3: Integration limits α's and ω's for the $\tilde{P}_\diamond$ partition.

N	E	W	S
$[0,\epsilon]$	$[\epsilon,2\epsilon]$	$[r-\epsilon,\sqrt{r^2-\epsilon^2}]$	$[r,1]$
$[0,\epsilon]$	$[2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[r-2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[r,1]$
$[0,\epsilon]$	$[r,1]$	$[r,1]$	$[r,1]$
$[\epsilon,2\epsilon]$	$[0,\epsilon]$	$[r,1]$	$[r-\epsilon,\sqrt{r^2-\epsilon^2}]$
$[\epsilon,2\epsilon]$	$[2\epsilon,r]$	$[r,1]$	$[r,1]$
$[\epsilon,2\epsilon]$	$[r,1]$	$[0,\epsilon]$	$[r-\epsilon,\sqrt{r^2-\epsilon^2}]$
$[2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[0,\epsilon]$	$[r,1]$	$[r-2\epsilon,\sqrt{r^2-\epsilon^2}]$
$[2\epsilon,r]$	$[\epsilon,2\epsilon]$	$[r,1]$	$[r,1]$
$[2\epsilon,r]$	$[2\epsilon,r]$	$[r,1]$	$[r,1]$
$[2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[r,1]$	$[0,\epsilon]$	$[r-2\epsilon,\sqrt{r^2-\epsilon^2}]$
$[r,1]$	$[0,\epsilon]$	$[r,1]$	$[r,1]$
$[r,1]$	$[\epsilon,2\epsilon]$	$[r-\epsilon,\sqrt{r^2-\epsilon^2}]$	$[0,\epsilon]$
$[r,1]$	$[2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[r-2\epsilon,\sqrt{r^2-\epsilon^2}]$	$[0,\epsilon]$
$[r,1]$	$[r,1]$	$[r,1]$	$[r,1]$
$[0,\epsilon]$	$[0,\epsilon]$	$\varnothing$	$\varnothing$
$[0,\epsilon]$	$[\sqrt{r^2-\epsilon^2},r]$	$\varnothing$	$\varnothing$
$[\epsilon,2\epsilon]$	$[\epsilon,2\epsilon]$	$\varnothing$	$\varnothing$
$[\sqrt{r^2-\epsilon^2},r]$	$[0,\epsilon]$	$\varnothing$	$\varnothing$
$[\sqrt{r^2-\epsilon^2},r]$	$[r,1]$	$\varnothing$	$\varnothing$
$[r,1]$	$[\sqrt{r^2-\epsilon^2},r]$	$\varnothing$	$\varnothing$

$$
\begin{aligned}
\sum_{\tilde{C}_{ji}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_+,\tilde{C}_{ji}) \;\geq\;& \int_0^\epsilon\int_0^\epsilon\int_0^{r-\epsilon}\int_0^{r-\epsilon} \\
&+ \int_0^\epsilon\int_\epsilon^{2\epsilon}\int_0^{r-2\epsilon}\int_0^{r-\epsilon} + \int_0^\epsilon\int_\epsilon^{\frac{r}{2}}\int_0^{\frac{r}{2}}\int_0^{r-\epsilon} + \int_\epsilon^{2\epsilon}\int_0^\epsilon\int_0^{r-\epsilon}\int_0^{r-2\epsilon} \\
&+ \int_\epsilon^{2\epsilon}\int_\epsilon^{2\epsilon}\int_0^{r-2\epsilon}\int_0^{r-2\epsilon} + \int_\epsilon^{2\epsilon}\int_\epsilon^{\frac{r}{2}}\int_0^{\frac{r}{2}}\int_0^{r-2\epsilon} + \int_{2\epsilon}^{\frac{r}{2}}\int_0^\epsilon\int_0^{r-\epsilon}\int_0^{\frac{r}{2}} \\
&+ \int_{2\epsilon}^{\frac{r}{2}}\int_\epsilon^{2\epsilon}\int_0^{r-2\epsilon}\int_0^{\frac{r}{2}} + \int_{2\epsilon}^{\frac{r}{2}}\int_{2\epsilon}^{\frac{r}{2}}\int_0^{r-2\epsilon}\int_0^{\frac{r}{2}}
\end{aligned}
\tag{6.8}
$$

where we use the fact that for the edge to exist, the S (W) intervals are $[0,\omega_\mathsf{S}]$ ($[0,\omega_\mathsf{W}]$). ω_S is determined by the worst case scenario for this edge to happen, when $\mathsf{N}_{(1)}$ is located farthest away from the center of the intersection, i.e, for the ω_N value. In this scenario, the edge will be possible only if the $\mathsf{S}_{(1)}$ coordinates are restricted to $[0, r-\omega_\mathsf{N}]$.

We proceed with the analysis of partitions $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}\cap\overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ and $\overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}}\cap e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$. Due to symmetry, we just calculate one, $\tilde{P}_|$, and multiply by two. The contribution of

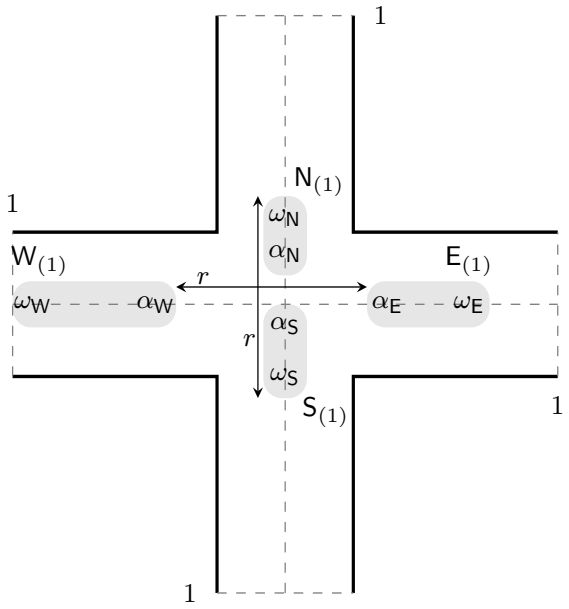


Figure 6.2: Integration interval example: For partition $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ the edge $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ implies for a given $\mathsf{N}_{(1)} \in [\alpha_{\mathsf{N}}, \omega_{\mathsf{N}}]$ that ω_{S} cannot be greater than $r - \omega_{\mathsf{N}}$, i.e., that $\mathsf{S}_{(1)} \in [0, r - \omega_{\mathsf{N}}]$. On the other hand, the edge's $\overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ absence implies that, for a given $\mathsf{E}_{(1)} \in [\alpha_{\mathsf{E}}, \omega_{\mathsf{E}}]$, α_{W} cannot be smaller than $r - \alpha_{\mathsf{E}}$, i.e., $\mathsf{W}_{(1)} \in [r - \omega_{\mathsf{E}}, 1]$.

this partition is given by:

$$
\begin{aligned}
\sum_{\tilde{C}_{ji} \in \tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_|, \tilde{C}_{ji}) \;\geq\; & \int_0^{\epsilon}\int_0^{\epsilon}\int_r^1\int_0^{r-\epsilon} + \int_0^{\epsilon}\int_{\epsilon}^{2\epsilon}\int_{r-\epsilon}^1\int_0^{r-\epsilon} \\
+ \; & \int_0^{\epsilon}\int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}}\int_{r-2\epsilon}^1\int_0^{r-\epsilon} + \int_0^{\epsilon}\int_r^1\int_0^{\sqrt{r^2-\epsilon^2}}\int_0^{r-\epsilon} \\
+ \; & \int_{\epsilon}^{2\epsilon}\int_0^{\epsilon}\int_r^1\int_0^{r-2\epsilon} + \int_{\epsilon}^{2\epsilon}\int_{\epsilon}^{2\epsilon}\int_{r-\epsilon}^1\int_0^{r-2\epsilon} \\
+ \; & \int_{\epsilon}^{2\epsilon}\int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}}\int_{r-2\epsilon}^1\int_0^{\epsilon} + \int_{\epsilon}^{2\epsilon}\int_r^1\int_0^{2\epsilon}\int_0^{r-2\epsilon} \\
+ \; & \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}}\int_0^{\epsilon}\int_r^1\int_0^{r-\sqrt{r^2-\epsilon^2}} + \int_{2\epsilon}^{r-2\epsilon}\int_{\epsilon}^{2\epsilon}\int_{r-\epsilon}^1\int_0^{2\epsilon} \\
+ \; & \int_{2\epsilon}^{r-\epsilon}\int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}}\int_{r-2\epsilon}^1\int_0^{\epsilon} + \int_{2\epsilon}^{r-2\epsilon}\int_r^1\int_0^{2\epsilon}\int_0^{2\epsilon}
\end{aligned}
\tag{6.9}
$$

Four classes $\{\tilde{C}_{41}, \tilde{C}_{42}, \tilde{C}_{43}, \tilde{C}_{44}\}$ will contribute with zero to the summation, because we are not able to build the partition, i.e., there is no way to produce the $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ edge. Determining S range is similar to the first partition analysis since we have the $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ edge in both partitions. The difference is due to $\overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$ edge absence. In contrast, the $\mathsf{E}_{(1)}$ worst case scenario now is the position most probable to produce the edge, which is nearest to the origin (center of the intersection), i.e., α_{E}. Thus, the $\mathsf{W}_{(1)}$ device must be positioned in interval $[r - \alpha_{\mathsf{E}}, 1]$. After guaranteeing $e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}$ and prohibit $e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ edges, one more edge, from the set $\{e_{\mathsf{N}_{(1)}\mathsf{E}_{(1)}}, e_{\mathsf{N}_{(1)}\mathsf{W}_{(1)}}, e_{\mathsf{E}_{(1)}\mathsf{S}_{(1)}}, e_{\mathsf{W}_{(1)}\mathsf{S}_{(1)}}\}$, must

be present. Let's consider $\mathsf{N}_{(1)} \in \tilde{I}_1$ classes, the second edge more probable to happen is $e_{\mathsf{N}_{(1)}\mathsf{E}_{(1)}}$ for $\{\tilde{C}_{11}, \tilde{C}_{12}, \tilde{C}_{13}\}$ and $e_{\mathsf{N}_{(1)}\mathsf{W}_{(1)}}$ for $\tilde{C}_{14}$. For the first two classes nothing should be made, because the $e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}$ existence is assured by default. Nevertheless, for the $\tilde{C}_{13}$ and $\tilde{C}_{14}$ classes, we have to shorten the E and W intervals, respectively, to assure the second edge. The contributions are:

$$\begin{aligned}\sum_{\tilde{C}_{1i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_{|}, \tilde{C}_{1i}) \;&\geq\; \int_0^\epsilon \int_0^\epsilon \int_r^1 \int_0^{r-\epsilon} + \int_0^\epsilon \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^1 \int_0^{r-\epsilon} \\ &+\; \int_0^\epsilon \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^1 \int_0^{r-\epsilon} + \int_0^\epsilon \int_r^1 \int_0^{\sqrt{r^2-\epsilon^2}} \int_0^{r-\epsilon} \end{aligned} \tag{6.10}$$

The $\mathsf{N}_{(1)} \in \tilde{I}_2$ classes are very similar, the contributions are:

$$\begin{aligned}\sum_{\tilde{C}_{2i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_{|}, \tilde{C}_{2i}) \;&\geq\; \int_\epsilon^{2\epsilon} \int_0^\epsilon \int_r^1 \int_0^{r-2\epsilon} + \int_\epsilon^{2\epsilon} \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^1 \int_0^{r-2\epsilon} \\ &+\; \int_\epsilon^{2\epsilon} \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^1 \int_0^\epsilon + \int_\epsilon^{2\epsilon} \int_r^1 \int_0^{2\epsilon} \int_0^{r-2\epsilon} \end{aligned} \tag{6.11}$$

Lastly, for the $\mathsf{N}_{(1)} \in \tilde{I}_3$ classes, we perceive that it is necessary to shorten the N range, which in turn, reduces the S range. The contributions are:

$$\begin{aligned}\sum_{\tilde{C}_{3i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_{|}, \tilde{C}_{3i}) \;&\geq\; \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon \int_r^1 \int_0^{r-\sqrt{r^2-\epsilon^2}} + \int_{2\epsilon}^{r-2\epsilon} \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^1 \int_0^{2\epsilon} \\ &+\; \int_{2\epsilon}^{r-\epsilon} \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^1 \int_0^\epsilon + \int_{2\epsilon}^{r-2\epsilon} \int_r^1 \int_0^{2\epsilon} \int_0^{2\epsilon} \end{aligned} \tag{6.12}$$

Where we used $e_{\mathsf{W}_{(1)}\mathsf{S}_{(1)}}$ as second edge in classes $\tilde{C}_{33}$ e $\tilde{C}_{34}$. We get (6.9) by the sum of terms (6.10), (6.11) and (6.12).

Finally, we evaluate the partition $\tilde{P}_\diamond = \overline{e_{\mathsf{N}_{(1)}\mathsf{S}_{(1)}}} \cap \overline{e_{\mathsf{E}_{(1)}\mathsf{W}_{(1)}}}$, in which the two edges of the event must be orthogonal. Since the classes $\tilde{C}_{11}$ and $\tilde{C}_{22}$ do not permit this, their contribution is zero. The partition will contribute with:

$$\begin{aligned}\sum_{\tilde{C}_{ji}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{ji}) \;&\geq\; \int_0^\epsilon \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 + \int_0^\epsilon \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 \\ &+\; \int_0^\epsilon \int_r^1 \int_r^1 \int_r^1 + \int_\epsilon^{2\epsilon} \int_0^\epsilon \int_r^1 \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} + \int_\epsilon^{2\epsilon} \int_{2\epsilon}^r \int_r^1 \int_r^1 \\ &+\; \int_\epsilon^{2\epsilon} \int_r^1 \int_0^\epsilon \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} + \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon \int_r^1 \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \\ &+\; \int_{2\epsilon}^{2\epsilon} \int_\epsilon^{2\epsilon} \int_r^1 \int_r^1 + \int_{r-\epsilon}^{2\epsilon} \int_{r-\epsilon}^r \int_r^1 \int_r^1 \\ &+\; \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 \int_0^\epsilon \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} + \int_r^1 \int_0^\epsilon \int_r^1 \int_r^1 \\ &+\; \int_r^1 \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon + \int_r^1 \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon \\ &+\; \int_r^1 \int_r^1 \int_r^1 \int_r^1 \end{aligned} \tag{6.13}$$

These classes' contributions are evaluated in a similar way to the partitions analyzed above and the details of these calculations will be omitted. Note that for all cases where W and S are in the $[r, 1]$ interval we are computing the occurrence of the graph $\mathcal{G}$ with no edges ($|\mathbb{E}|= 0$) which contribute positively to probability p.

Considering $\mathsf{N}_{(1)} \in \tilde{I}_j, \forall j \in \{1, 2, 3, 4\}$ classes, in that order, we have the following probability contributions:

$$\begin{aligned}\sum_{\tilde{C}_{1i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{1i}) &\geq \int_0^\epsilon \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 + \int_0^\epsilon \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 \\ &+ \int_0^\epsilon \int_r^1 \int_r^1 \int_r^1 \end{aligned} \tag{6.14}$$

$$\begin{aligned}\sum_{\tilde{C}_{2i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{2i}) &\geq \int_\epsilon^{2\epsilon} \int_0^\epsilon \int_r^1 \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} + \int_\epsilon^{2\epsilon} \int_{2\epsilon}^{r} \int_r^1 \int_r^1 \\ &+ \int_\epsilon^{2\epsilon} \int_r^1 \int_0^\epsilon \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} \end{aligned} \tag{6.15}$$

$$\begin{aligned}\sum_{\tilde{C}_{3i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{3i}) &\geq \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon \int_r^1 \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} + \int_{2\epsilon}^{2\epsilon} \int_\epsilon^{2\epsilon} \int_r^1 \int_r^1 \\ &+ \int_{r-\epsilon}^{2\epsilon} \int_{r-\epsilon}^{r} \int_r^1 \int_r^1 + \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_r^1 \int_0^\epsilon \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \end{aligned} \tag{6.16}$$

$$\begin{aligned}\sum_{\tilde{C}_{4i}\in\tilde{P}_{\epsilon,r}} \mathcal{C}_{\epsilon,r,\mu}(\tilde{P}_\diamond, \tilde{C}_{4i}) &\geq \int_r^1 \int_0^\epsilon \int_r^1 \int_r^1 + \int_r^1 \int_\epsilon^{2\epsilon} \int_{r-\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon \\ &+ \int_r^1 \int_{2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_{r-2\epsilon}^{\sqrt{r^2-\epsilon^2}} \int_0^\epsilon + \int_r^1 \int_0^1 \int_r^1 \int_r^1 \end{aligned} \tag{6.17}$$

We get (6.13) adding up terms (6.14), (6.15), (6.16) and (6.17). Expressions (6.8), (6.9) and (6.13) has the values of Tables 6.1, 6.2 and 6.3, which completes the proof. □

In the next chapter, we show how well the expression (6.3) derived in Theorem 4 adjusts to simulated data, as a validation of the quality of the analytical lower bound. Moreover, we analyze the variation of the relative contributions of terms (6.5), (6.6), and (6.7) in the expression (6.4), for different values of parameters ϵ, r and μ.

Chapter 7

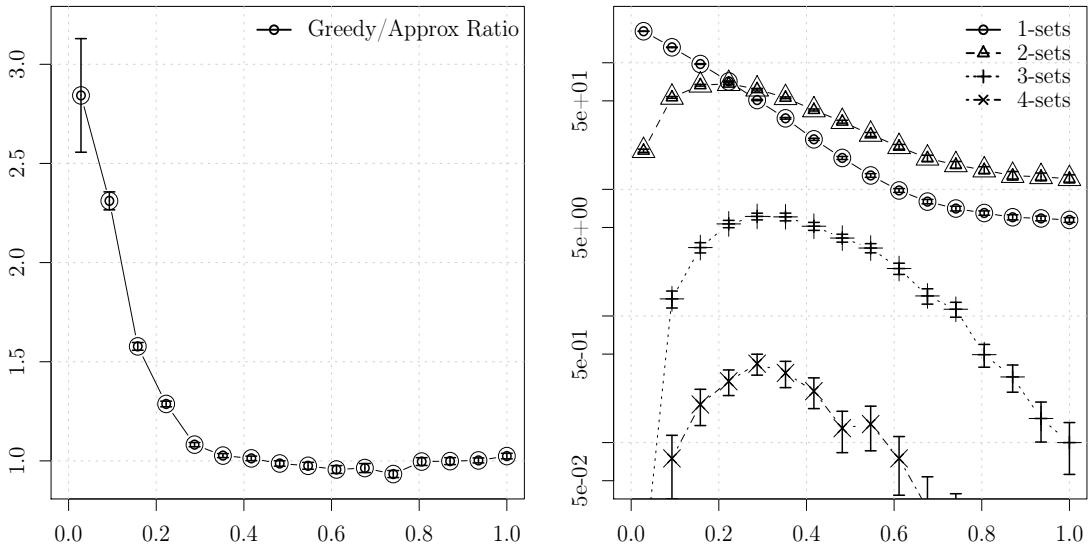
Experimental Results

In this chapter, we validate our analytical results through simulations. We simulated several instances of the Random OWN model, defined in Chapter 2, by using different combinations of parameters g, μ, ϵ and r. Due to limited space, we focus our analysis on the following parameter values: $g = 5$, $\mu \in \{5, 10\}$, $\epsilon \in \{0.01, 0.033\}$, and $0 < r < 1$. Note that the $\epsilon = 0.033$ scenario illustrates a geometry similar to that of NYC, as discussed in Chapter 3, while lower values of ϵ induce more obstructed scenarios. The x-axis represents the increasing transmission range r in all plots. Each value in the plot represents the mean value over 200 samples, with the respective 95% confidence interval.

7.1 Computed backbone structures

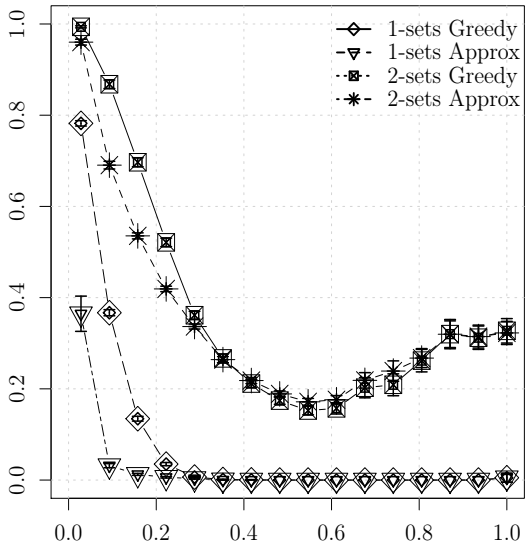
We start by analyzing the size of the backbones computed by Algorithm 2 (referred as Approx), proposed in Chapter 5. As a baseline, we also implemented the greedy H_k-approximation algorithm (referred as Greedy) for the k-SC problem (Chvatal [1979]). Figures 7.1 and 7.2 illustrate (a) the ratio of the backbone sizes computed by the two algorithms, and (b) the number of available k-sets in the simulated instances, and (c-d) the percentage of k-sets in each solution, relative to the respective availability of each kind of set in the instance, $k \in \{1, 2, 3, 4\}$, for $\epsilon = 0.01$ (more obstructed scenario) and $\epsilon = 0.033$ (NYC), respectively.

Looking at plots 7.1a and 7.2a, we can see that Approx finds smaller backbones in nearly all of the simulated scenarios. The gain of the algorithm, relative to Greedy, is particularly high for smaller transmission range values ($r \leq 0.4$), when the network is partitioned into more components and, consequently, the problem instance size is larger.

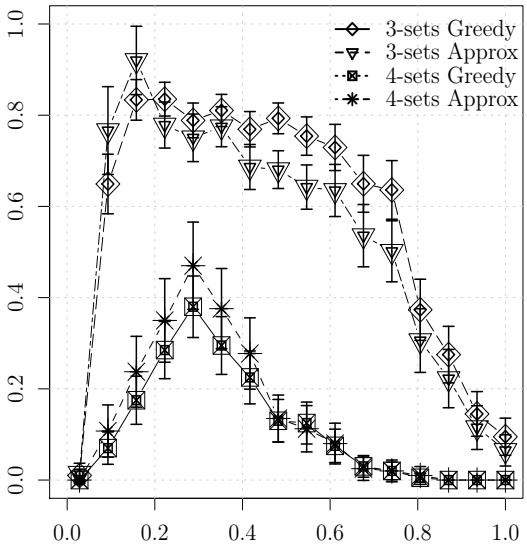


(a) Backbone size ratio

(b) k-set availability

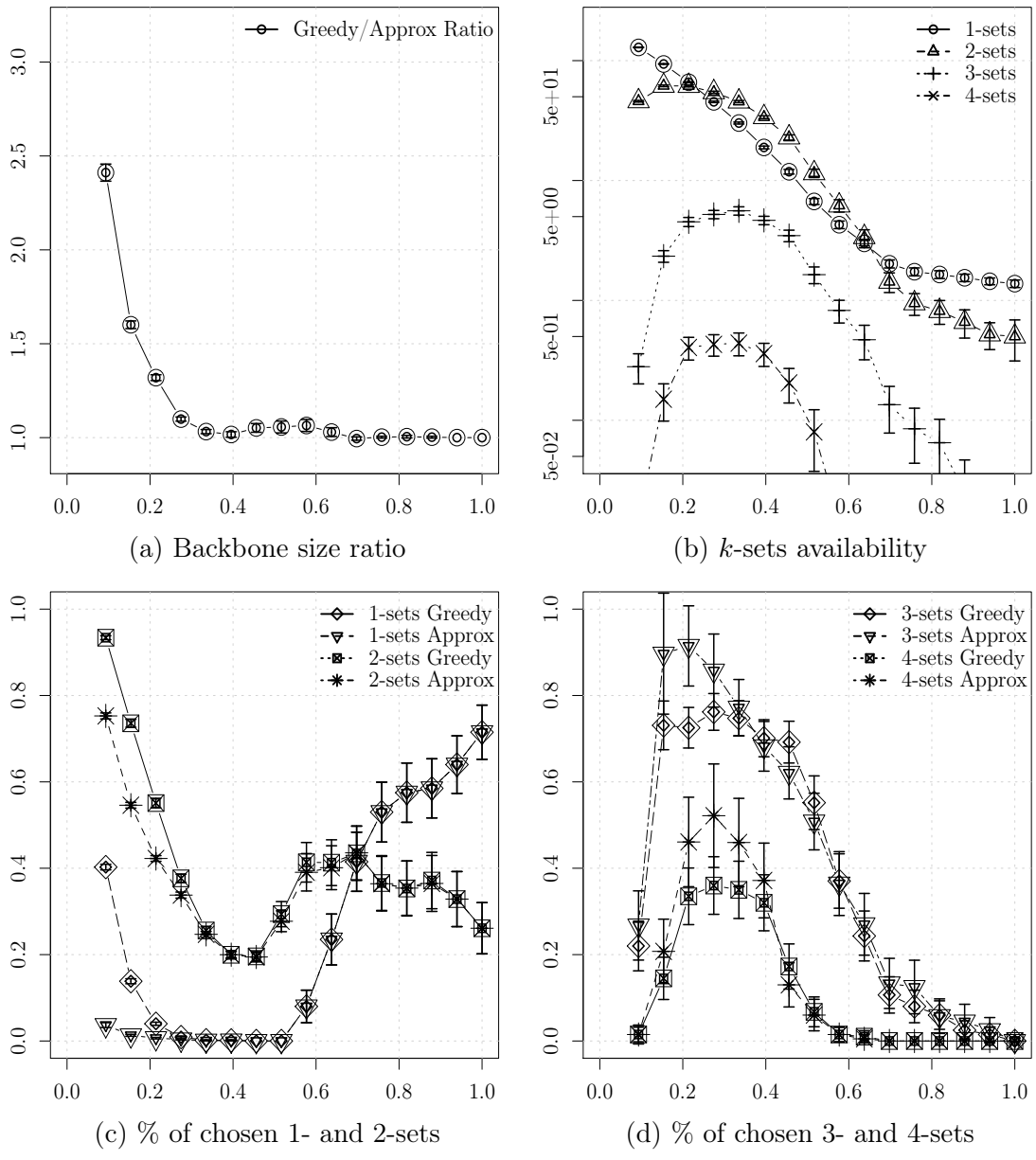


(c) % of chosen 1- and 2-sets



(d) % of chosen 3- and 4-sets

Figure 7.1: Computed backbones (obstructed scenario): OWN parameters: $g = 5$, $\mu = 5$, $\epsilon = 0.01$



(a) Backbone size ratio

(b) k-sets availability

(c) % of chosen 1- and 2-sets

(d) % of chosen 3- and 4-sets

Figure 7.2: Computed backbones (NYC): OWN parameters: $g = 5$, $\mu = 5$, $\epsilon = 0.033$

In plots 7.1b and 7.2b, we can see the distribution of sets of different sizes in the problem instances. As expected, the total number of 1- and 2-sets decreases with increasing r, and the total number of 3- and 4-sets peaks at around $r = 0.3$ and quickly goes to zero in both scenarios. Note that, in the more obstructed scenario, the number of 3- and 4-sets decays more slowly with increasing transmission range r, than in the NYC scenario. This happens because connectivity at intersections is harder to achieve with less visibility, which makes it possible to connect 3 or more components with a single base station.

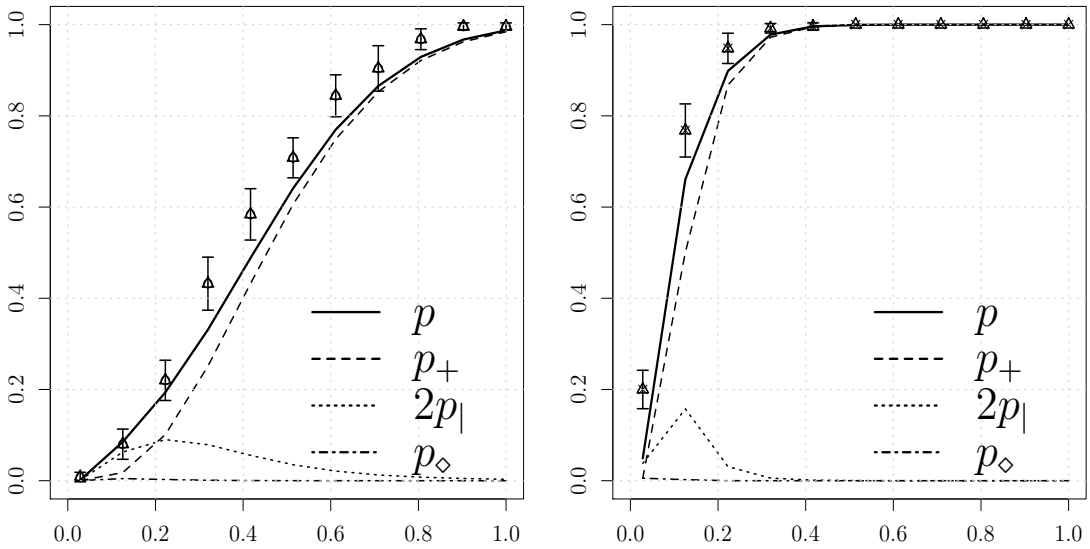
In plots 7.1c and 7.2c we can see how many of the available 1- and 2-sets both algorithms selected into the backbone cover. Both algorithms pick proportionally more 1-sets in the $\epsilon = 0.01$ scenario, since it has harder visibility conditions at street intersections, which makes the network more disconnected. For small values of r, Approx selects significantly fewer 1-sets and slightly fewer 2-sets than Greedy. For larger values of r, results differ significantly between the two scenarios, but both algorithms behave similarly: when $\epsilon = 0.01$, almost no 1-sets are picked, and when $\epsilon = 0.033$, the larger the r the more 1-sets are picked by both algorithms, peaking at 80%. This can be explained by the limited availability of larger sets in these instances, as can be seen in Figure 7.2b.

Finally, in plots 7.1d and 7.2d, the proportion of selected 3- and 4-sets is shown. We can see that both scenarios and algorithms present similar behavior: the more 3- and 4-sets there are available, the more are selected into the backbone. Interestingly, even though Greedy gives priority to larger sets while Approx is oriented to 2-set maximization, the end solutions contain similar numbers of 3- and 4-sets.

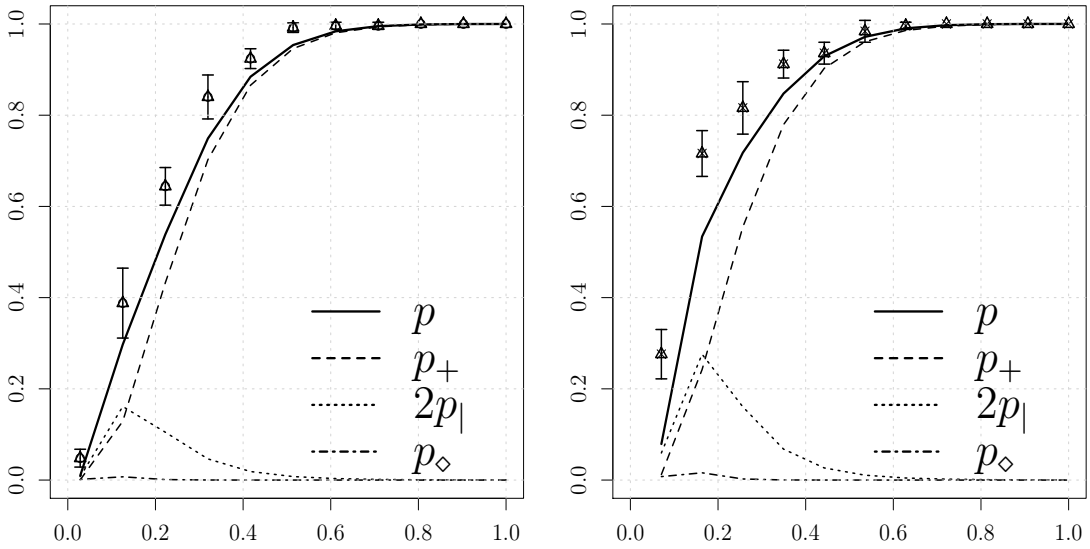
Overall, the relatively worse performance of Greedy in low-connectivity instances ($r \leq 0.2$) is due to more frequent selection of 1-sets. Another reason why Approx performs better in these instances is due to the low number of available 3- and 4-sets, which makes the maximum matching core of this algorithm highly efficient. Moreover, these low-connectivity instances seem to be the hardest to solve, not only because of their larger size, due to a larger number of sets, but also due to a lower probability of "polynomial complexity", as illustrated in the next part of our simulation study.

7.2 Polynomial complexity instances

We turn our attention to the probability of a Random OWN instance to be solved optimally in polynomial time, as described in Chapter 6. In Figure 7.3, we plot the empirical and the analytical expression 6.4 for probability p, which is a sum of the

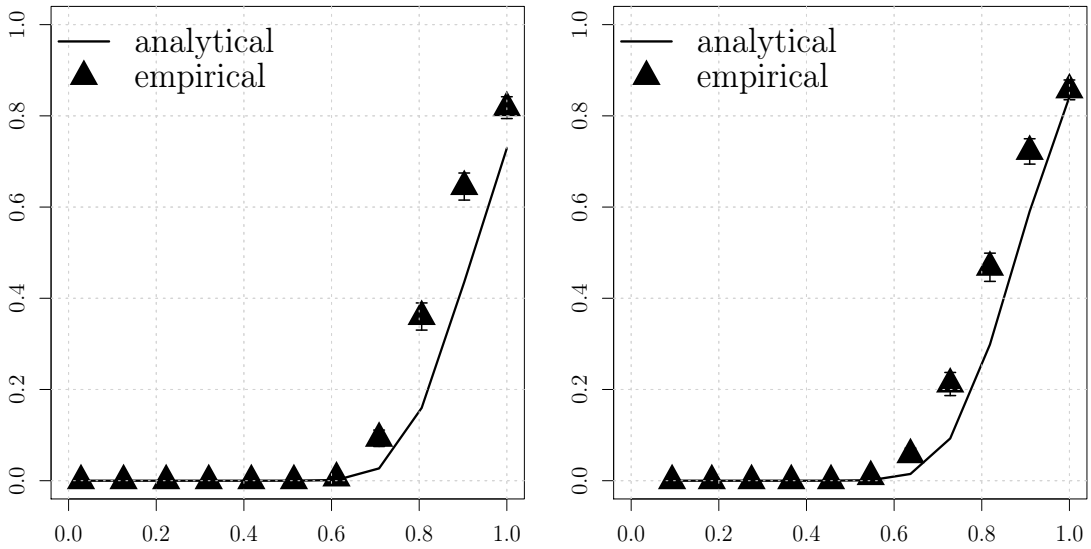


(a) $g = 5$, $\mu = 5$, $\epsilon = 0.01$ (b) $g = 5$, $\mu = 20$, $\epsilon = 0.01$



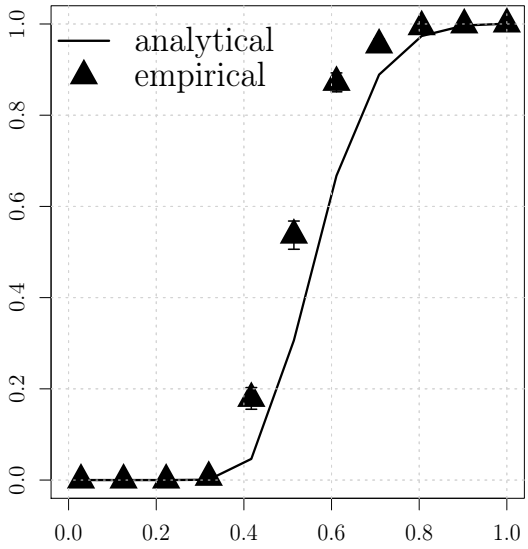
(c) $g = 5$, $\mu = 10$, $\epsilon = 0.01$ (d) $g = 5$, $\mu = 10$, $\epsilon = 0.025$

Figure 7.3: Probability of "polynomial connectivity w/ backbone" (at a street intersection) $p = p_+ + 2p_| + p_\diamond$: analytical (lines) $\times$ empirical (triangles)

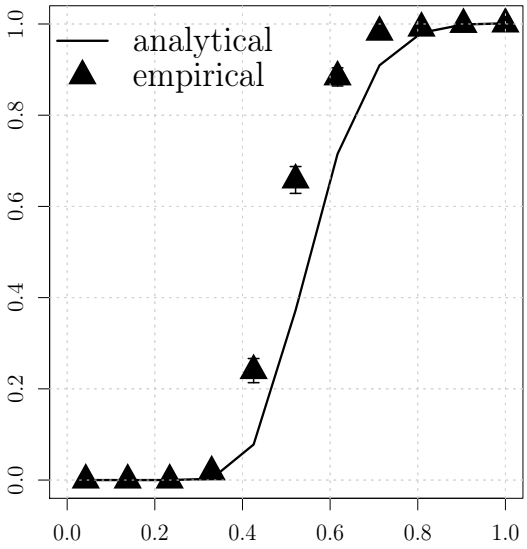


(a) $g = 5$, $\mu = 5$, $\epsilon = 0.01$

(b) $g = 5$, $\mu = 5$, $\epsilon = 0.033$



(c) $g = 5$, $\mu = 10$, $\epsilon = 0.01$



(d) $g = 5$, $\mu = 10$, $\epsilon = 0.015$

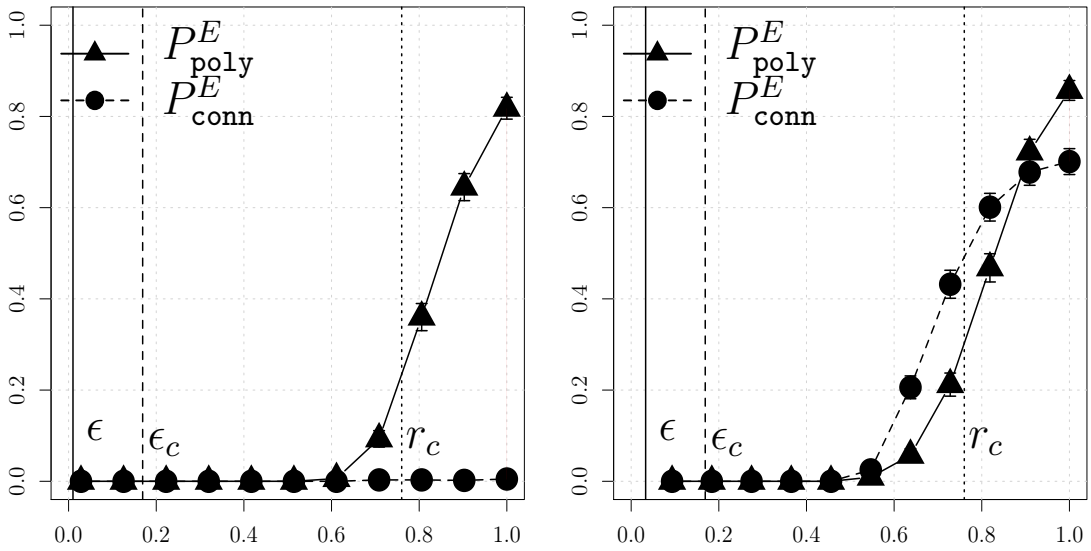
Figure 7.4: Probability of "polynomial connectivity w/ backbone" $P_{\texttt{poly}}$: analytical (solid line) $\times$ empirical (triangles).

terms 6.8, 6.9 and 6.13, denoted by p_+, $p_|$ and $p_\diamond$, , respectively, also plotted separately. In Figure 7.3a, we can see the analytical lower-bound is well-adjusted. Moreover, in the simulated scenarios, partition $\tilde{P}_\diamond$ does not significantly contribute to "polynomial complexity", for all values of r. The most important contribution comes from partition $\tilde{P}_+$. The contribution of $\tilde{P}_|$ becomes more relevant for small transmission ranges, with a peak around $r = 0.15$. The most relevant aspect of this experiment is to show that the analytical results provide a good approximation for the probability of the event for different combinations of parameters ϵ, μ and r. In Figure 7.4, we will validate "polynomial complexity" on the whole grid, expression 6.3, which involves the dependence on g.

Recall that an OWN-BC instance can be solved optimally in polynomial time, if it contains no 3- or 4-sets. For this event to happen, a sufficient condition is that all (g^2) street intersections of the grid are adjacent to no more than two components, which is true whenever there exist at least two edges between the 4 closest nodes to the center of the intersection. Note that this condition is not necessary since the event can also occur in highly fragmented networks when most street intersections have no candidate base stations and most candidate base stations are located on street segments. Such scenarios can be seen in Figures 7.1b and 7.2b, for small values of r. In Figure 7.4, we can see that the adjust of the expression 6.3 follows the data. It should be noted that we can clearly distinguish regions of zero probability as opposed to those of probability one and that the curve captures very well these different regions. Next, we evaluate how the gap between the event of "polynomial connectivity with a backbone" and the event of the empirical probability of connectivity without a backbone infrastructure.

In Figure 7.5, we plot the empirical probability $P^E_{\texttt{poly}}$ of the event of "polynomial connectivity with a backbone", as well as $P^E_{\texttt{conn}}$, the empirical probability of connectivity without a backbone infrastructure, i.e., the probability of connectivity of a *ad hoc* Random OWN. The vertical lines in Figure 7.5 represent the simulated value of ϵ, the critical value ϵ_c (below which the CTR bound is not valid), and the value of CTR r_c. Note that all the simulated scenarios are sub-critical, so $P^E_{\texttt{conn}}$ is significantly below 1, even when $r > r_c$. This is due to the fact that the street width ϵ is below its critical value ϵ_c as shown in Almiron et al. [2013]. As r grows, $P^E_{\texttt{poly}}$ increases from 0 to 1 in different rates in each scenario.

Comparing $P^E_{\texttt{poly}}$ to $P^E_{\texttt{conn}}$, different cases are possible: (1) the former can grow while the latter remains close to zero (7.5a), (2) the latter can grow faster than the former (7.5b), or (3) the former can grow faster than the latter (7.5c and 7.5d). These different scenarios illustrate different practical implications for our results. In situation (1), a backbone is necessary to ensure connectivity, no mater how high the transmission



(a) $g = 5$, $\mu = 5$, $\epsilon = 0.01$

(b) $g = 5$, $\mu = 5$, $\epsilon = 0.033$

(c) $g = 5$, $\mu = 10$, $\epsilon = 0.01$

(d) $g = 5$, $\mu = 10$, $\epsilon = 0.015$

Figure 7.5: Empirical Probability of "polynomial connectivity w/ backbone" ($P^E_{\texttt{poly}}$) and "ad hoc" ($P^E_{\texttt{conn}}$) connectivity. Shaded region: $\geq 95\%$ optimal backbone.

range, close to optimum (and optimum, with probability $P^E_{\texttt{poly}}$) backbone structures can be computed by Algorithm 2. In case (2), $P^E_{\texttt{conn}}$ is slightly higher than $P^E_{\texttt{poly}}$, so some instances might not require a backbone and some may, in which case, Algorithm 2 computes near optimal solutions (optimal with probability $P^E_{\texttt{poly}}$). In scenario (3), $P^E_{\texttt{poly}}$ grows much faster than $P^E_{\texttt{conn}}$, so in the shaded areas of Figures 7.5c and 7.5d, an optimal backbone can be computed by Algorithm 2 in polynomial time with probability $P^E_{\texttt{poly}} \geq 95\%$.

Note that from the practical point of view, this is a valuable result, since it allows to determine when optimal solutions can be found to ensure connectivity w.h.p. in ad hoc and "with a backbone" Random OWN, for any network size, communication technology, or geometry of a particular city or another kind of grid structure. An advantage of our 2-approximation algorithm is that it returns the optimum when the problem instance falls within the "polynomial complexity area", described above.

Chapter 8

Related Work

Relatively few attempts have been made to analyze obstructed wireless networks, many of which are quite complex and not easily generalizeable, e.g., Nekoui and Pishro-Nik [2009]. One interesting model for obstructed wireless networks is the so-called "line-of-sight networks" (Frieze et al. [2009]), which represents the network by a grid and places devices on the grid vertices with some constant probability. Among other results, the authors derive asymptotic bounds for k-connectivity of such networks.

The Random OWN model was proposed in Almiron et al. [2013], with the objective to characterize the CTR for connectivity. In Giles et al. [2016], Giles et al. extended the soft connection model (Penrose [2016]) into non-convex spaces based on circular or spherical obstacles (as opposed to fractal boundaries in Dettmann et al. [2014], internal walls in Georgiou et al. [2013] or fixed obstacles on a grid in Almiron et al. [2013]). The authors characterized situations, where obstacles are (and are not) important influences on connectivity. For instance, they show that small obstacles have little impact on connectivity, large obstacles have a similar impact on connectivity as the enclosing perimeter, and multiple obstacles can have the dominant effect on the connection in certain density regimes, particularly ad hoc communication networks deployed in urban environments, such as 5G wireless networks.

The impact of a connected backbone on the capacity of random wireless networks *without obstacles* has been investigated before (Dousse et al. [2002]; Liu et al. [2007]). In Dousse et al. [2002], for instance, one-dimensional, two-dimensional and strip models were considered, leading to different capacity scaling laws. In the one-dimensional network, the gain in capacity was shown to increase linearly with the number of base stations. In the two-dimensional case, in order to improve capacity that significantly, the number of base stations was shown to be $\Omega(\sqrt{n})$. For the strip network, if the width of the strip is at least on the order of the logarithmic of its length, it was

shown that capacity scales as in the two-dimensional case; otherwise, it scales as in the one-dimensional case.

Another problem related to our work is the k-SC. Many approximation algorithms can be found in the literature in non-geometric domains, the most recent being an $(H_k - \frac{196}{390})$-approximation developed in Levin [2008]. Finally, several related problems, such as domination and connected domination, were shown to be NP-complete in unit disk graphs, in particular, the domination problem was shown to remain NP-complete even for grid graphs, a subclass of unit disk graphs (Clark et al. [1990]).

A literature review of "Backbone Structure and Wireless Networks" return results in the so-called *virtual* backbones. For instance, Das and Bharghavan [1997] shows how to impose such virtual backbone on ad-hoc networks. Min et al. [2004] also studies the virtual infrastructures in wireless ad-hoc networks in the hope of reducing the communication overhead, presenting a work on how to construct and maintain reliable and efficient structures. Jurdzinski and Kowalski [2012] demonstrated how to obtain a backbone structure in a deterministic distributed way on the top of a given wireless network modeled with Signal-to-Interference-and-Noise-Ratio (SINR) physical model. It is important to stress that this virtual backbone differs from the wired backbone of access points (used in the present work).

The work of Zhang [2012] is about the renewed focus on the need to develop more energy-efficient underlying network infrastructures. They recall that once the base station deployment is done, the next problem is how to efficiently operate the base stations for energy conservation during the off-peak period. The problem of activate-deactivate base stations has great potential in being modeled in a very similar way to ours OWN-BC problem.

The literature review of "Connectivity and Obstructed Wireless Networks" returns the work of Bian et al. [2015a] in vehicular ad hoc networks (VANET) which studies how to quantify (together with theoretical analysis) the affecting factors on caching mechanisms because of the complex urban environment and high mobility of vehicles. Their model parameters include vehicle density, transmission range, and the ratio of caching vehicles. In Nekoui [2013], the authors argue that the road and obstacle geometry are two important factors that should be appropriately addressed when studying the communications throughput of VANETs. Our work points out possibilities of modeling in this sense.

The works of Almiron et al. [2013] and Almiron et al. [2014] are the foundations on which we construct our results. At best of our knowledge, they are the best effort in modeling OWN with results, at the same time, practical and mathematically simple. We planned to repeat the analysis for the CTR taking into account the backbone.

However, this was partially possible. Instead of a critical range, we obtained a probability of obtaining instances with "polynomial complexity". We have recorded here that the works of Ghasemi and Nader-Esfahani [2006] and Franceschetti and Meester [2008] are great sources of understanding and inspiration to perform this first intent.

Ghasemi and Nader-Esfahani [2006] demonstrates rigorously how to obtain the connectivity along a finite one-dimensional segment, taking into account the presence of a given number of devices with a given transmission range. The work of Franceschetti and Meester [2008] has many results on scalability and proofs on physical limits of wireless communications. Limits that could be determinant for design protocols and algorithms on scenarios that reassemble the proposed model. This reference is an example of how Information Theory can help to push beyond and prove limits for algorithms.

To conclude, we remember that Du and Hu [2008] illustrates several applications of Steiner tree problems in communication wireless networks. Which can constitute an alternative approach to modeling the effects of base stations introduction in wireless networks. We could treat devices and access points, respectively, as terminals and Steiner points, which should form a tree (acyclic communication network) with some cost function. In the bottleneck Steiner tree problem, for instance, the objective is to minimize the length of the longest edge in the tree [1]. Wang and Li [2002] give a ratio-1.866 approximation algorithm for this problem.

[1] We can easily interpret the minimum longest edge as a critical length above which any communication link can be established. This Steiner tree approach, of course, lacks the obstacles' modeling.

Chapter 9

Conclusions

In this work we analyzed how connectivity scales in large ad hoc obstructed wireless networks and focused on the problem of establishing connectivity using a connected backbone infrastructure in sub-critical scenarios.

We present a problem formulation, study its complexity, prove NP-completeness and develop an approximation algorithm for it. The proposed algorithm is based on a maximum matching core, and finds optimal solutions in polynomial time for certain network instances w.h.p.. We study the probability of obtaining optimal solutions under random device deployments. This is an original work that provides insights into theoretical aspects of a potentially important practical problem of obtaining connectivity in ad hoc wireless networks with restrictions on transmission power (e.g. 6LoWPAN) or device density that are deployed in obstructed environments, such as dense urban grids, tunnels or indoor networks.

These results have relevance in practice, since we provide an efficient way to ensure connectivity in OWN with no restriction on the network's size, geometry, transmission range or device density. Furthermore, we make it possible, for this specific problem, to explore the barriers between hard and easy instances. This may be seen as further indication of how intricate can be the relationship between *verified* and *solved*.

9.1 Future Directions

Academically, we plan to formalize the notion of "polynomial complexity" regimes expressed by the analytical probability $P_{\texttt{poly}}$. One possible direction is the phase transition and criticality *modus operandi* of Percolation Theory (Stauffer and Aharony [1994]). With regard to percolation, geometry/locality play crucial roles in global be-

havior (Oliveira and Braga [2002]). Hence, a primary question emerge: how take into account the fact that backbone structure *dribbles* geometry/locality?

Personally, two interesting directions (possibly combined) would be: format OWN-BC as an International Collegiate Programming Contest problem and apply the analytical framework developed on *real* data. One possible idea would be to study the fleet of personal cars of UBER as an OWN, obtaining from ride data, the g, ϵ, r and μ values for study the "polynomial complexity" regions of this case of study. Assuming, e.g., the company's interest in investing in a vehicular ad hoc network with low-power wireless technologies, "UBERnet", we are able to respond to design issues such as: When would it be necessary to insert backbone infrastructures? When would it be possible to obtain optimum minimum backbone? When the solution should be approximated since optimum search cost is prohibitive?

Bibliography

Almiron, M. G., Goussevskaia, O., Frery, A. C., and Loureiro, A. A. (2014). Connectivity at crossroads. In *Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014 IEEE 25th Annual International Symposium on*, pages 1437--1441. IEEE.

Almiron, M. G., Goussevskaia, O., Loureiro, A. A., and Rolim, J. (2013). Connectivity in obstructed wireless networks: From geometry to percolation. In *Proc.of ACM*, MobiHoc'13, pages 157--166, New York, NY, USA. ACM.

Bian, C., Zhao, T., Li, X., Du, X., and Yan, W. (2015a). Theoretical analysis on caching effects in urban vehicular ad hoc networks. *Wireless Communications and Mobile Computing*.

Bian, C., Zhao, T., Li, X., and Yan, W. (2015b). Boosting named data networking for efficient packet forwarding in urban vanet scenarios. In *LANMAN*, pages 1–6. IEEE.

Chvatal, V. (1979). A greedy heuristic for the set-covering problem. *Mathematics of operations research*, 4(3):233--235.

Clark, B. N., Colbourn, C. J., and Johnson, D. S. (1990). Unit disk graphs. *Discrete Mathematics*, 86(1-3):165--177.

Das, B. and Bharghavan, V. (1997). Routing in ad-hoc networks using minimum connected dominating sets. In *Communications, 1997. ICC'97 Montreal, Towards the Knowledge Millennium. 1997 IEEE International Conference on*, volume 1, pages 376--380. IEEE.

Dettmann, C. P., Georgiou, O., and Coon, J. P. (2014). More is less: Connectivity in fractal regions. *CoRR*, abs/1409.7520.

Dousse, O., Thiran, P., and Hasler, M. (2002). Connectivity in ad-hoc and hybrid networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE*

Computer and Communications Societies. Proceedings. IEEE, volume 2, pages 1079--1088. IEEE.

Du, D. and Hu, X. (2008). *Steiner tree problems in computer communication networks*. World Scientific.

Florian, M., Andreev, S., and Baumgart, I. (2013). Overdrive: An overlay-based geocast service for smart traffic applications. In *Annual International Conference on Mobile Computing & Networking (MobiCom)*, pages 147--150.

Forooshani, A., Bashir, S., Michelson, D., and Noghanian, S. (2013). A survey of wireless communications and propagation modeling in underground mines. *IEEE Communications Surveys & Tutorials*, 15(4):1524–1545.

Franceschetti, M. and Meester, R. (2008). *Random networks for communication: from statistical physics to information systems*, volume 24. Cambridge University Press.

Frieze, A. M., Kleinberg, J. M., Ravi, R., and Debany, W. (2009). Line-of-sight networks. *Combinatorics, Probability & Computing*, 18(1–2):145--163.

Garey, M. R. and Johnson, D. S. (2002). *Computers and intractability*, volume 29. wh freeman New York.

Georgiou, O., Dettmann, C. P., and Coon, J. P. (2013). Network connectivity through small openings. *CoRR*, abs/1304.3646.

Ghasemi, A. and Nader-Esfahani, S. (2006). Exact probability of connectivity one-dimensional ad hoc wireless networks. *IEEE Communications Letters*, 10(4):251--253.

Giles, A. P., Georgiou, O., and Dettmann, C. P. (2016). Connectivity of soft random geometric graphs over annuli. *Journal of Statistical Physics*, 162(4):1068--1083.

Goussevskaia, O., Halldórsson, M. M., and Wattenhofer, R. (2013). Algorithms for wireless capacity. *IEEE/ACM Transactions on Networking*, page 11.

Gupta, P. and Kumar, P. R. (2000). Capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388--404.

Jurdzinski, T. and Kowalski, D. R. (2012). Distributed backbone structure for algorithms in the sinr model of wireless networks. In *International Symposium on Distributed Computing*, pages 106--120. Springer.

Kawsar, F. and Brush, A. B. (2013). Home computing unplugged: Why, where and when people use different connected devices at home. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, pages 627--636.

Levin, A. (2008). Approximating the unweighted k-set cover problem: greedy meets local search. *SIAM Journal on Discrete Mathematics*, 23(1):251--264.

Liu, B., Thiran, P., and Towsley, D. (2007). Capacity of a wireless ad hoc network with infrastructure. In *Proc. of the ACM*, MobiHoc, pages 239--246, New York, NY, USA. ACM.

Min, M., Wang, F., Du, D.-Z., and Pardalos, P. M. (2004). A reliable virtual backbone scheme in mobile ad-hoc networks. In *Mobile Ad-hoc and Sensor Systems, 2004 IEEE International Conference on*, pages 60--69. IEEE.

Nekoui, M. (2013). *Vehicular ad hoc networks: Interplay of geometry, communications, and traffic*. PhD thesis.

Nekoui, M. and Pishro-Nik, H. (2009). Geometrical analysis of obstructed wireless networks. In *IEEE Information Theory Workshop*, pages 589--593.

Oliveira, M. M. d. and Braga, G. A. (2002). O fenômeno de transição de fase no modelo de percolação de elos* em d dimensões. *Revista Brasileira de Ensino de Física*, 24(4):448--454.

Penrose, M. D. (2016). Connectivity of soft random geometric graphs. *The Annals of Applied Probability*, 26(2):986--1028.

Sedgewick, R. and Wayne, K. (2011). *Algorithms Fourth Edition*. Addison-Wesley Professional.

Stauffer, D. and Aharony, A. (1994). *Introduction to percolation theory*. CRC press.

Wang, L. and Li, Z. (2002). An approximation algorithm for a bottleneck k-steiner tree problem in the euclidean plane. *Information Processing Letters*, 81(3):151--156.

Zhang, H. (2012). *Green communications: theoretical fundamentals, algorithms and applications*. CRC Press.