Università degli Studi di Padova
Facoltà di Ingegneria
Dipartimento di Ingegneria dell'Informazione
Corso di Laurea Specialistica in Ingegneria Informatica

On the Diameter of Bluetooth-Based Ad Hoc Networks

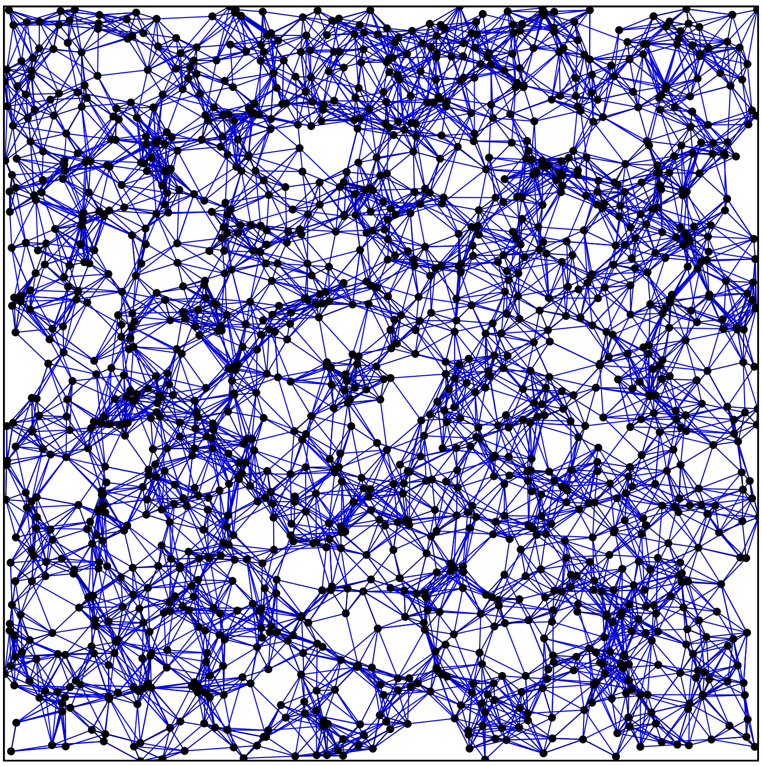
Relatori
Ch.mo Prof. Andrea Pietracaprina
Ch.mo Prof. Geppino Pucci

Laureando
Alberto Pettarin

Anno Accademico 2007/2008
Padova, 7 Luglio 2008

On the Diameter of Bluetooth-Based Ad Hoc Networks

Alberto Pettarin



Alberto Pettarin

On the Diameter of Bluetooth-Based Ad Hoc Networks

WEB PAGE

http://www.dei.unipd.it/~pettarin/

E-MAIL

pettarin@dei.unipd.it

pettarin@computer.org

In the previous page: an instance of the Bluetooth Topology with $n = 1500$ nodes, visibility range $r(n) = 0.075$ and $c(n) = 5$ choices per node.

A Clelia e Guido

ABSTRACT

In recent years, *ad hoc networks* have attracted growing interest. We briefly discuss their characteristics and design challenges as well as some theoretical models of *random graphs* which are nowadays used to study their topological properties, namely, Erdős-Rényi $G_{n,p}$, Random Geometric $G_{n,\lambda}$ and Bluetooth Topology $BT(r(n), c(n))$ graphs.

The latter is a good model for the *device discovery phase* in the Bluetooth protocol, since it accounts for the limited number of links that a node can mantain active towards other devices. A $BT(r(n), c(n))$ graph consists of n nodes randomly placed in the unit square $[0, 1]^2$ and each node establishes a link (i.e. an edge) to $c(n)$ neighbours chosen uniformly at random among those within distance $r(n)$.

Our study focuses on the *diameter* of $BT(r(n), c(n))$, for all the possible values of $r(n) \geq r_{MIN} = \gamma_0 \sqrt{\frac{\log n}{n}}$ and $c(n)$ that guarantee connectivity.

Setting $c(n) = \log \frac{1}{r(n)}$, we demonstrate that with high probability the diameter is $O\left(\frac{1}{r(n)} + \log n\right)$. We also prove a "geometric lower bound" of $\Omega\left(\frac{1}{r(n)}\right)$ valid for $r(n) \geq r_{MIN}$ and for any $c(n)$; furthermore, this result can be strengthen to $\Omega\left(\frac{\log n}{\log \log n}\right)$

for constant radii. Note that, for almost all the parameter values, we asymptotically match the aforementioned upper bound.

SOMMARIO

Negli ultimi anni le *reti di comunicazione ad hoc* hanno riscosso crescente interesse. Qui ne discuteremo brevemente le caratteristiche e le sfide progettuali per poi passare a presentare alcuni modelli teorici di *grafi random* che sono tutt'oggi utilizzati per lo studio delle loro proprietà topologiche; in particolare, il modello di Erdős-Rényi $G_{n,p}$, i Random Geometric Graphs $G_{n,\lambda}$ e la Bluetooth Topology $BT(r(n), c(n))$.

A differenza dei primi due, l'ultimo pare un paradigma particolarmente appropriato per catturare le caratteristiche della *fase di scoperta* di dispositivi vicini in quelle reti che adottano il protocollo Bluetooth, perché esso tiene conto del numero limitato di canali che un nodo può mantenere simultaneamente attivi verso altri dispositivi.

Un grafo $BT(r(n), c(n))$ è formato da n nodi distribuiti casualmente nel quadrato unitario $[0, 1]^2$. Ciascuno di essi stablisce un link (ossia un lato del grafo) con $c(n)$ nodi vicini, scegliendoli uniformemente a caso tra quelli distanti al più $r(n)$.

Il nostro lavoro si concentra sul *diametro* del $BT(r(n), c(n))$, per tutti quei valori dei parametri $r(n) \geq r_{MIN} = \gamma_0 \sqrt{\frac{\log n}{n}}$ e $c(n)$ che garantiscono la connessione del grafo risultante.

Fissando $c(n) = \log \frac{1}{r(n)}$, siamo riusciti a dimostrare che, con alta probabilità, il diametro è $O\left(\frac{1}{r(n)} + \log n\right)$.

Abbiamo inoltre ottenuto un lower bound "geometrico" di $\Omega\left(\frac{1}{r(n)}\right)$ valido per ogni $c(n)$ non appena $r(n) \geq r_{MIN}$. Inol-

tre, quest'ultimo risultato può essere rafforzato a $\Omega\left(\frac{\log n}{\log\log n}\right)$ se il raggio di visibilità $r(n)$ diviene costante. Si noti che, per quasi tutti i valori assumibili dai parametri del modello, questi lower bound corrispondono all'upper bound descritto in precedenza.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ACL	ACL Logical Transport
ASN	Ad Hoc Sensor Network
$BT(r(n), c(n))$	Bluetooth Topology
C4ISRT	Command, Control, Communications, Computing, Intelligence, Surveillance, Reconnaisance and Targeting
$G_{n,p}$	Erdős-Rényi Random Graph
ETSI	European Telecommunications Standards Institute
HSN	Hybrid Sensor Network
GPS	Global Positioning System
IC	Integrated Circuit
IR	Infrared
ISM	Industrial, Scientific and Medical
IT	Information Technology
L2CAP	Logical Link Control and Adaptation Protocol
LMP	Link Manager Protocol
MAC	Medium Access Control
MANET	Mobile Ad Hoc Network
MEMS	Micro Electromechanical System
NBC	Nuclear, Biological and Chemical

P2P	Peer To Peer
PAN	Personal Area Network
PDA	Personal Digital Assistant
QoS	Quality of Service
RAM	Random Access Memory
$G_{n,\lambda}$	Random Geometric Graph
ROM	Read Only Memory
SAN	Sensor Area Network
UWB	Ultra Wide Band
VLSI	Very Large Scale Integration
WSN	Wireless Sensor Network

1 INTRODUCTION

1.1 AD HOC NETWORKS

SINCE the late 1990s, rapid advances in several fields of Information Technology (IT) have permitted the large-scale development of an innovative type of distributed network, commonly called *ad hoc network*.

Informal description of ad hoc networks

Basically, an ad hoc network consists of various autonomous devices performing a specialized task (e.g., sensing a physical phenomenon in a given site) which are able to communicate among themselves and possibly with "base stations" (or "data sinks") located inside or at the border of the covered area. Moreover, little computation and/or data gathering can be performed *in situ*. An introduction to this subject can be found in [39].

Wireless communication

In ad hoc networks, information is transmitted wirelessly among nodes, and this characteristic makes them suitable to establish communication channels in devastated or belligerent regions where regular infrastructures have been destroyed or to observe and measure a physical process from the inside.

Typical applications of ad hoc networks

The most typical applications include environmental control and surveillance, health, business and military monitoring. A survey focussing on general issues of Wireless Sensor Networks (WSNs) but also discussing Sensor Area Networks (SANs) can be found in [2].

Due to technological and applicative constraints, The designer of an ad hoc network has to cope with many challenges. Among the others, BLUETOOTH-based networks seem more suitable for sensing tasks since they require extremely low energy over short communication ranges.

1.2 ANALYTICAL MODELLING OF AD HOC NETWORKS

While there are many experimental data...

Although a plethora of combinations of different devices, protocols, routing algorithms and management policies has been proposed in the literature, there is a stringent need for solid theoretically grounded studies of the properties of these networks. In fact, state-of-the-art results are almost always empirical and thus dramatically depend on the settings and parameters chosen for those particular applications.

... we need more theoretic results...

More specifically, classical network attributes have to be investigated, (connectivity, bandwidth, delay, robustness against device failures, etc.) as well as specific properties like energy consumption, robustness against topology changes due to mobility, data/policy management, just to cite a few (cf. [51]).

... using probabilistic analysis

Since the exact position of the nodes cannot usually be determined, the resulting topology has to be thought as a "random placement". In other words, we are dealing with intrinsically aleatory system configurations and therefore we have to perform a probabilistic analysis of the problem.

With reference to well-established theoretical random structures which can model ad hoc network, namely the Erdős-Rényi Random Graph ($G_{n,p}$) [5], the Random Geometric Graph ($G_{n,\lambda}$) [37] and the Bluetooth Topology ($BT(r(n), c(n))$) [34], in this work we will concentrate on the diameter of $BT(r(n), c(n))$, that is, the maximum length of a shortest path between two nodes, which is a good approximation of the maximum delay to be expected when communicating within the network.

1.3 THE BLUETOOTH TOPOLOGY AND OUR CONTRIBUTION

Informal definition of $BT(r(n), c(n))$

An informal description of the $BT(r(n), c(n))$ model is the following. Select n points uniformly at random in the unit square $[0, 1]^2$, to model the device placement. Each node has a "visibility radius" $r(n)$ and chooses $c(n)$ nodes as its neighbours among all the visible ones (i.e. among the nodes that are within distance $r(n)$). The resulting graph is called $BT(r(n), c(n))$.

Results in literature on $BT(r(n), c(n))$

Experimental studies [18] have demonstrated that this random graph approximates quite well the BLUETOOTH device discovery phase. In two seminal papers [12, 10], the connectivity of $BT(r(n), c(n))$ was investigated, in the first paper fixing parameters $r(n)$ and $c(n)$ to be constant, in the latter article expressing those parameters in function of the number of nodes n.

Other properties of this topology are of interest for practical purposes, among the others the diameter (representing maximum delay) and the expansion (ensuring bandwidth), just to cite the main two.

Our contribution

Our work focuses on the diameter of $BT(r(n), c(n))$, starting from the work on connectivity done in [10]. Specifically, we will prove asymptotic lower and upper bounds to the diameter, as function of the number of nodes, the visibility range and the number of neighbours that each device is allowed to choose.

We are able to prove that the diameter is (within a constant) the inverse of the visibility range when the latter is minimum, formally:

$$\mathrm{diam}(BT(r(n), c(n))) = \Theta\left(\frac{1}{r(n)}\right). \quad (1.1)$$

With "minimum radius" we intend that it is the "shortest" still able to guarantee the connectivity of the whole graph, which turns out to be

$$r(n) = \Omega\left(\sqrt{\frac{\log n}{n}}\right).$$

To obtain (1.1), we allow neighbourhoods to be large enough; roughly speaking, of logarithmic size in the visibility radius:

$$c(n) = \Omega\left(\log \frac{1}{r(n)}\right).$$

This result holds also for longer transmission radii, namely when $r(n) \leq n^{-\delta}$ for a certain constant $\delta \approx \frac{1}{8}$. Clearly these results have practical appeal and are provably optimum, in the sense that a simple "geometric" argument yields the matching lower bound

$$\mathrm{diam}(BT(r(n), c(n))) = \Omega\left(\frac{1}{r(n)}\right).$$

Moreover, we prove that for even longer (actually, constant) transmission radii, the diameter is $O(\log n)$, thus very small for any resonable n of interest. When $r(n) = \Theta(1)$ we achieve a better lower bound of $\Omega(\log n / \log\log n)$, which is just slightly weaker than the upper bound $O(\log n)$.

Our results are summarized in Table 1 and Table 2.

Proof techniques

The main techniques used in the proofs are common in probabilistic analysis and include stochastic majorizations of random variables, properties of branching processes and probability amplification. Furthermore, concentration bounds like Chernoff bounds are extensively applied.

λ	$\text{diam}(G_{n,\lambda})$	Ref.
$\Omega\left(\sqrt{\frac{\log n}{n}}\right)$	$\Omega(\frac{1}{\lambda})$	4.1
$\Omega\left(\sqrt{\frac{\log n}{n}}\right)$	$\leq \frac{2\sqrt{5}}{\lambda} + 2 \approx \frac{4.472}{\lambda} + 2$	B.1

Table 1: A concise view of our results concerning the $\text{diam}(G_{n,\lambda})$. They are asymptotic estimates in high probability, i.e. they hold with probability $\rightarrow 1$ as $n \rightarrow \infty$. The third column contains the reference to the Theorem or Lemma where the corresponding bound is proved.

Proof idea

The results on the diameter are proved for three cases, corrisponding to three disjoint radii intervals — here referred to as "short", "medium" and "long" radii. The main idea behind the proof is the following. Divide $[0,1]^2$ in square cells as in a chessboard. The edge of a cell has length $1/k$, where $k = \left\lceil \sqrt{5}/r(n) \right\rceil$. With this choice, two nodes residing in adjacent cells are within distance $r(n)$.

For short ($r(n) \leq n^{-1/3}$) and medium ($n^{-1/3} < r(n) \leq n^{-1/8}$) radii, we can prove that, for every cell, the subgraph induced by nodes in that cell is connected and has "small" diameter. In the "short" case, approximating the diameter of a cell by the length of an hamiltonian path touching all the $m \approx n/k^2$ nodes inside the cell suffices, since each cell contains "few" nodes. In the "medium" case, we resort to a stochastic majorization involving $G_{n,p}$ to demonstrate that the diameter of each cell is approximately $\log m / \log\log m$.

$r(n)$	$c(n)$	$\mathrm{diam}(BT)$	Ref.
$\leq n^{-\frac{1}{3}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)})$	4.5
$n^{-\frac{1}{3}} < r(n) \leq n^{-\frac{1}{8}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)})$	4.7
$> n^{-\frac{1}{8}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)} + \log n)$	4.11

(a) Upper bounds for the $\mathrm{diam}(BT(r(n), c(n)))$.

$r(n)$	$c(n)$	$\mathrm{diam}(BT)$	Ref.
$\Omega\left(\sqrt{\frac{\log n}{n}}\right)$	Any	$\Omega(\frac{1}{r(n)})$	4.1
$\sqrt{2}$	$\Theta(1)$	$\Omega(\log \log n)$	4.12
$\sqrt{2}$	$\Theta(1)$	$\Omega(\frac{\log n}{\log \log n})$	4.14

(b) Lower bounds for the $\mathrm{diam}(BT(r(n), c(n)))$.

Table 2: A concise view of our results concerning the $\mathrm{diam}(BT(r(n), c(n)))$. These are asymptotic estimates in high probability, i.e. they hold with probability $\to 1$ as $n \to \infty$. The fourth column contains the reference to the Theorem or Lemma where the corresponding bound is proved.

Then, showing that there exists an edge between adjacent cells with high probability, we prove the connectivity and contemporarly bound the diameter of $BT(r(n), c(n))$, achieving

$$\mathrm{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)}\right).$$

For long radii (i.e. $r(n) > n^{-1/8}$), internal connectivity of the cells is not guaranteed. Instead, we can demonstrate that with high probability from every node we can reach with a sufficiently small number of hops a giant connected component which contains at least $\Theta(nr^2)$ nodes. The upper bound on the diameter is obtained limiting the diameter of this giant connected component, yielding

$$\operatorname{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)} + \log n\right).$$

1.4 STRUCTURE OF THIS THESIS

This dissertation is organized as follows. *Thesis organization*

- Chapter 2 summarizes the peculiar characteristics and challenges of ad hoc networks, with special attention to BLUETOOTH-based networks.
- The aforementioned theoretic models, suitable to capture the topology properties of an ad hoc network, are discussed in Chapter 3 and known results in literature are stated.
- The main contribution of our work is presented in Chapter 4, where we give asymptotic lower and upper bounds for the diameter of $BT(r(n), c(n))$.
- Finally, Chapter 5 contains some observations about the work done and the possible research directions to further extend the results of this thesis.

The reader mainly interested in graph-theoretical matters can skip Chapter 2.

2

AD HOC NETWORKS

In this chapter we briefly present the main characteristics, application fields and design challenges of ad hoc networks. In particular, Bluetooth-based networks are discussed since they are notably efficient and economic for localized sensing tasks.

2.1 AD HOC NETWORKS

2.1.1 *Historical development*

The first military projects about ad hoc networking date back to the 1970s and 1980s. The original intent was dealing with the lack of communication infrastructures on battlefields [27, 52].

However, it was just since the 1990s that ad hoc networks became a very active field of research. In fact, many circumstances contributed to the flourishing of new paradigms, models, protocols and eventually commercial products.

In 1990s three factors fostered ad hoc network development:

We can identify three main causes. At first, VLSI production processes favoured the introduction of low-cost, portable devices capable of generating and processing data, such as laptops, PDAs, mobile phones, etc. Then, new emerging wireless technologies, such as Bluetooth IEEE 802.11 and HyperLAN 2 allowed the interconnection of such different units to exchange data and fostered the research in ultra-low power communications. Finally,

Very Large Scale Integration (VLSI) production processes, emerging wireless technologies, *MEMS and new energy supplies*

MEMS-based devices and more efficient and manageable energy supplies made possible to develop more durable systems, with months or even years of operating autonomy.

2.1.2 *Main features*

Ad hoc networks are distributed systems whose name derives from the fact that they are tailor-made for a particular applicative goal.

Nodes placement is random...

The exact placement of nodes cannot usually be pre-determined (since very often random placement occurs, as in a smart dust, in intelligent fabrics or when sensor are literally thrown onto geographic area) and can even vary during time (e.g. when devices are mobile, as in a Mobile Ad Hoc Network (MANET) or WSN).

... and thus algorithms are self-organizing

This feature forces algorithms and management policies to be "self-organizing", i.e. they should be able to modify duty-cicles, data routes, energy management, etc. without a centralized supervisor.

Location/topology obliviousness is desiderable

A remarkable property which is highly desiderable is the capability of working without having any information about the network topology and, more generally, absolute device location. In fact, putting a Global Positioning System (GPS) receiver or other localization unit on a device could be neither practical nor affordable.

Devices are autonomous...

... and can perform data processing locally

Each element is supposed to operate autonomously, i.e. it performs its tasks quite "independently" from the presence of other devices in his neighbourhood. Devices are usually equipped with some onboard processor that can perform local computation on the raw data, thus reducing the amount of network traffic by transmitting partially processed information. The latter is a crucial ability since, with state-of-the-art Integrated Circuits (ICs) and over the entire system life, the energy required to receive and transmit data can be three or four orders of magnitude higher than the energy spent to perform local computation (cf. [40]).

Communication phase

Data created at each point has to be transmitted to gathering points, usually one or more *base stations* (also called *data sinks*) where it will be further processed or used. These endpoints can be inside the area where nodes are placed or in proximity of its borders and usually are directly connected to a wired backbone.

Wireless communication...

Inter-node communication is performed wirelessly, but wired shortcuts can be issued to increase the overall system performances, giving birth to a Hybrid Sensor Network (HSN) (see [41]). Radio tecnologies adopted in ad hoc networks include low-energy protocols such as Bluetooth[1], IEEE 802.11 (WiFi)[2], HyperLAN 2[3] and more recently other IEEE 802.15 protocols (including Ultra Wide Band (UWB) and ZigBee) and IEEE 802.16 (WiMax).

... based on flooding or gossiping...

Typically, communication takes place with broadcast algorithms, such as *flooding* or *gossiping*, rather than adopting *address-based* routing policies [39, 2].

... with multi-hop stategies

Additionally, multi-hop strategies are preferred upon classical single-hop communication since they are expected to reduce the interference among nodes and with the environment (which is auspicable in covert scenarios) and to consume less power. In fact, the transmitting power required by state-of-the-art anten-

1 Specifications are available at `http://www.bluetooth.com/Bluetooth/Technology/Building/Specifications/Default.htm`

2 Specifications of IEEE 802 standards are available at `http://standards.ieee.org/getieee802/`

3 Standardized by European Telecommunications Standards Institute (ETSI) as ETS 300 652 and ETS 300 893 and available at `http://pda.etsi.org/`.

nas is proportional to the covered distance elevated to an exponent between two and four, depending on the directionality of the antenna, reflectivity of the terrain and the fading model for carrying medium ([3]).

Energy efficiency is critical

Keeping the amount of traffic low and adopting multi-hop transmission are not always sufficient to preserve battery life. Usually the power source is irreplaceable and, once it becomes depleted, the single device becomes permanently unavailable. In order to prolong system life, management policies are power-aware and it is not uncommon that, unlike classical networks, Quality of Service (QoS) policies are not issued for the sake of energy efficiency.

Other design concerns

Power consumption issues are not the unique concern in designing such systems. A careful identification of the task to be performed, usually the same for all the components, has to be pursued to produce a tailor-made project of final devices. Many constraints have to be satisfied, just to cite a few: minimize the form factor, reduce the manufacturing budget, reduce environmental interferences, enforce system robustness or security, guarantee the reconfigurability of the system.

2.1.3 *Typical applications*

Strictly speaking, we should classify as "ad hoc network" every distributed system wherein autonomous devices can communicate via radio equipment [39].

Communication-centric goals

The main goal of an ad hoc network consists of coping with the lack of regular infrastructures, for example due to war, natural disasters or electrical blackouts. During rescue or emergency operations in a devastated area, nodes cannot be placed according to a regular *a priori* scheme, since the orography of the area or the presence of obstacles does not permit it.

A little taxonomy

Although our definition is quite general, we can identify a sort of taxonomy depending on the features of the elements of the system. When each node is equipped with sensors (temperature, humidity, fire, IR, chemical, etc.), we label the system as an *Ad Hoc Sensor Network*. Moreover, if the nodes are mobile or the topology is subject to changes, we can speak of a *Mobile Ad Hoc Network*. Mixing these two characteristics, we obtain a so-called *Wireless Sensor Network*.

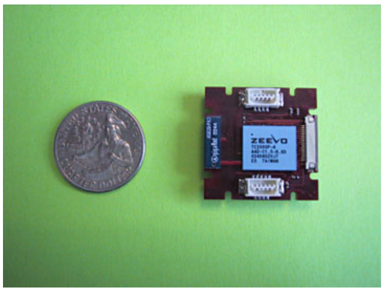


Figure 1: Intel Mote prototype [32] (original size: 30×30 mm). Image from [22].

Sensing applications

The latter kind is of particular interest, since sensing applications range over a wide spectrum, including continuous sensing, event detection/identification and local control of actuators [2, 11].

Some examples of applications of WSNs follow.

- *Home/business*: home automation, smart control of rooms and buildings, industrial production processes [22] (Figure 1 illustrates a prototype of Intel Mote node for industrial applications), object/vehicle tracking [33];
- *Environmental*: fire/flood/pollution detection [43] (in Figures 2 and 3 is depicted a Firebug, a fire-sensor deployed

by UC Berkeley), ambiental protection [6, 1], precision agricolture [42];

- *Health*: telemonitoring, patients/doctors tracking inside a hospital [35], drug administration;
- *Entertainment*: governance of visitors in a site [26], interactive museums, 3D reconstruction [50];
- *Engineering*: infrastructure health monitoring [29], pipe inspection [45], plant safety control;
- *Military*: accounting equipment/ammunition, battlefield surveillance, NBC attacks detection, damage assessment, intelligent clothing (cumulatively called C4ISRT).

Figure 2: A close-up of FIREBUG, a fire sensor designed by UC Berkeley. Image from [48].

2.2 BLUETOOTH-BASED NETWORKS

Among lots of proprietary solutions, BLUETOOTH seems viable for wide-spread applicability, since it is sufficiently mature and economic while guaranteeing good performaces in terms of power

Figure 3: A FIREBUG mote deployed onto a tree. Image from [48].

consumption, quality of service, and easiness of design and integration.

Target applications

As any innovation, BLUETOOTH was developed with a specific set of target applications in mind: specifically, the set up of *Personal Area Networks (PANs)*, i.e., the interconnection of a few devices operated by a single user. It is well suited for voice and data streams over short-ranges (in fact, BLUETOOTH was originally dubbed as a "cable replacement" technology) exchanged among a small number of near devices in a hierarchical fashion. As an example, think to a mobile phone connected to a headset or to a Personal Digital Assistant (PDA) connected to a peripheral like a keyboard or printer.

Moreover, BLUETOOTH has been considered as a possible radio technology to interconnect also larger networks, especially those where multi-hop transmission is carried out to connect distant nodes, thus limiting the area covered by a single device and thus the power required to mantain the coverage.

BLUETOOTH might not be the best choice for sensor networs consisting of a very large number of devices which have to communicate in a full-distributed fashion, like a Peer To Peer (P2P) network [25]. However, this kind of networks generally requires longer transmission ranges which cannot be provided by any BLUETOOTH-enabled equipment.

For the details of the protocol, we invite the reader to read the specifications[4].

2.2.1 *Physical Operation Overview*

Radio features

BLUETOOTH is a radio technology operating in the unlicensed Industrial, Scientific and Medical (ISM) band at 2.4 GHz with frequency hopping and time division features. Transceivers have low complexity since frequency modulation is shaped-binary; the symbol rate is 1 Megasymbol per second and thus the bit rate is approximately 1 Mb/s in the Basic Rate mode (722 kb/s netting the overhead off). Higher capacity (about 2–3 Mb/s) can be achieved in Enhanced Data Rate mode. BLUETOOTH does not require line-of-sight between connected devices, can penetrate solid objects and is omni-directional.

Piconets, master/slave devices

A physical radio channel is shared by a group of devices that are synchronized to a common clock and follow the same frequency hopping pattern; this group is called "piconet". Synchronization is provided by the "master" device, while all others are known as "slaves".

Frequency hopping with time division

The basic hopping pattern is a pseudo-random ordering of the 79 frequencies in the ISM band but some of those can be skipped to prevent interferences with other systems operating in ISM band. This adaptative scheme is particularly efficient

4 Available at `http://www.bluetooth.com/Bluetooth/Technology/Building/Specifications/Default.htm`

against static (i.e., non-hopping) interference sources. Frequency hopping takes place between the transmission or reception of packets. A channel is divided into time slots of 0.625 ms, which are the basic units for physical occupancy of the channel during transmission. Data is encapsulated into packets which are fit into consecutive slots as much as possible, to achieve a time-division full duplex scheme.

2.2.2 *Logical Operation Overview*

Device address

When a Bluetooth device is manufactured, it receives a unique 48-bit long address. Additionally, since actual implementations of the standard consider up to eight components in each piconet, a local 3-bit address is leased when a piconet is formed or entered by a device and it is recovered when it leaves.

Layered stack

A layered stack of links, channels and associated protocols governs the logical operation of a Bluetooth system; from bottom to upper levels:

1. physical channel,
2. physical link,
3. logical transport,
4. logical link,
5. L2CAP channel.

Physical link

In a piconet, there is a physical link between each slave and the master but not directly between the slaves, so the master acts as a router for its piconet.

Logical link

Traffic on a physical link is transported over logical links both in synchronous/asynchronous unicast or broadcast and the actual data is multiplexed in the time slots decided by the scheduling function residing at resource manager of each device.

Link Manager Protocol

Additionally, a Link Manager Protocol (LMP) controls the actual utilization of baseband and physical layer, setting up and controlling logical transports and logical links, and managing physical links. Active devices in a piconet establish a default asynchronous connection-oriented logical transport used by the LMP signals. This transport is called ACL Logical Transport (ACL) and is created whenever a device joins a piconet.

High-level layers

Above all there is the Logical Link Control and Adaptation Protocol (L2CAP) layer which provides a channel-based high-level abstraction to applications and services. This layer performs segmentation and reassembly of packets and multiplexing/demultiplexing of multiple channels over a shared logical link. Also, it conveys QoS messages.

2.2.3 *Performances*

Transmission and rate ranges

Typical transmission ranges for a Bluetooth wireless device range from 10 meters to 100 meters depending on the device class and peak data rate can touch 3 Mb/s with input signal sensivity of about 82 dBm.

Energy efficiency

The current absorbed varies from 20 mA to 50 mA during normal operation for Bluetooth Class 2 cores. The consumed power in a final product is nearly 10 mW in stand-by, 100 mW when listening the channel and about 200 mW while sending/receiving data. Note, however, that only 5–10% of this power is needed for actual transmission while the remainder power is for feeding processors, memories, sensors, transducers, etc.

Figure 4: KC22 BLUETOOTH OEM MICRO MODULE. The chip measures 10 mm × 13 mm. Image from [24].

Cost and integrability

Nowadays (2008) a single chip costs less than \$5 and occupies less than 100 mm² of area, making it very easy to integrate in any electronic device.

For example, TEXAS INSTRUMENTS BRF6300 BLUELINK 5.0, whose block diagram is shown in Figure 5, supports BLUETOOTH Specification v2 and integrates the BLUETOOTH baseband, RF transceiver, an ARM processor, memory (ROM and RAM) and power management in a single chip of area 45 mm². The current consumption is about 100 μA during page/inquiry phases [47].

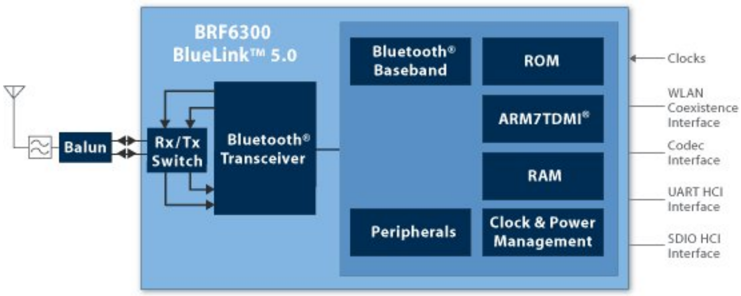


Figure 5: A logic block diagram of BRF6300. Image from [47].

2.2.4 *Communication topology*

Besides technical physical/procedural aspects, we are interested in the communication topology that arises when BLUETOOTH-enabled devices interact. Two detailed surveys on the argument are [49, 46].

Piconet formation

As said before, a piconet is formed by two or more devices that share the same physical channel, synchronized to a common clock and hopping sequence. More in detail, the common (piconet) clock is the one of the piconet master. The frequency hopping sequence is derived from the master's clock and device address, too. In current implementations, it is common to find a maximum of eight nodes per piconet, althought core specifications do not state an explicit maximum.

Many piconets may coexist

Within a common location a number of independent piconets may exist, each with its own physical channel, master device and thus piconet clock and hopping sequence.

A device can partecipate in many piconets...

A BLUETOOTH-enabled device may participate in two or more piconets at the same moment, on a time-division multiplexing basis. However, it can never be a master of more than one piconet, so it may be a slave in many independent piconets. If so, that device is said to be involved in a "scatternet".

... but routing functions are high-level!

Moreover, involvement in a scatternet does not necessarily imply any network routing capability or function, since BLUETOOTH core protocols do not offer such functionality, which is responsibility of higher level protocols.

Discovering procedure

Devices use an inquiry procedure to discover nearby devices, or to be discovered by counterparts in their proximity. This phase is asymmetrical: a device that tries to find other nearby devices ("inquiring device") actively sends inquiry requests. On the other hand, devices that are available to be found ("discoverable devices") listen for these inquiry requests and possibly send responses. The inquiry handshake takes place on a reserved physical channel and does not require the intervention of high-level layers. Both inquiring and discoverable devices may already be part of another piconet.

Connected Mode

Upon successful connection, the two parties enter the so-called "Connected Mode" within a piconet. Additional logical channels can be established and released on demand through LMP, and also new nodes can join the piconet.

Parked State

When a slave device is actively connected (i.e. if produces/-consumes data traffic), the default ACL connects it to the master. This transport ceases only when the slave is detached from the piconet or when it enters a "Parked State", which is similar to a suspension: the master suspends traffic from or to the slave, until it wakes up again. The parking mode is a power-saving feature particularly useful when battery life is a strict constraint.

Additional operation modes exist, like "Hold Mode" (a severe hibernation) or "Role Switch Mode" (to exchange master/slave role within a piconet). For the details, see the BLUETOOTH core specifications.

2.3 MOTIVATION OF OUR WORK

A search for "ad hoc network" or "wireless sensor network" in scientific databases provides evidence that thousands of combinations of different types of devices, protocols, routing algorithms and system control policies have been proposed in the literature. Most of then have been already implemented and tested in real-life tasks, while some have become commercial products.

The very largest part of these works comply with an experimental approach which can be depicted as a sort of "guess-try-validate" process: guess suitable values for system parameters, build up a simulation or test-scale prototype, and check if the system behaves as expected. If it does, deploy it; if it does not, go back tuning the choice of parameters.

Even if such an approach could highlight some properties and also lead to the formulation of useful practical rules-of-thumb, a more "foundational" analysis of ad hoc networks is needed to be guaranteed about the final performances.

In particular, we want to be able to predict or estimate some characteristics of the network that we are planning in a simple, straightforward manner. To accomplish this, in the next chapter we will examine various mathematical models for an ad hoc network which hopefully preserve its most meaningful features.

3

RANDOM GRAPH MODELS

THREE main models of random graphs are presented in this chapter, namely, $G_{n,p}$, $G_{n,\lambda}$ and $BT(r(n), c(n))$. We discuss how they are related to the topologies arising in real-life ad hoc networks and report on the known results about their topological properties.

3.1 PRELIMINARY DEFINITIONS

We will introduce some notations and definitions in order to be consistent during the subsequent exposition. For basic concepts and terms not explicitly defined here, the reader can refer to any standard text in graph theory, like [4].

Definition 3.1 (Distance between two nodes)**.** *Given an undirected graph $G = (V, E)$ and two nodes $u, v \in V$, their* distance, *denoted by* $\text{dist}(u, v)$, *is the number of edges in a shortest path starting at u and ending at v, if such a path exists, and $+\infty$ otherwise.*

$\text{dist}(u, v)$

Observe that every node is at distance zero from itself and that if $(u, v) \in E$ then $\text{dist}(u, v) = 1$.

Two notions that will be useful later are the following:

Definition 3.2 (Crown and Neighbourhood of a node)**.** *Given a graph $G = (V, E)$ and a vertex $x \in V$, we define the* crown of x at distance $i \geq 0$ *to be the set*

$\Gamma_i(x), N_i(x)$

$$\Gamma_i(x) = \{y \in V : \text{dist}(x, y) = i\}.$$

Additionally, we define the neighbourhood of x at distance $i \geq 0$ *to be the set*

$$N_i(x) = \bigcup_{j=0}^{i} \Gamma_j(x) = \{y \in V : \text{dist}(x,y) \leq i\}.$$

Note that:

- $\Gamma_0(x) = \{x\}$;
- $\Gamma_1(x)$ is the set of vertices adjacent to x;
- $\Gamma_i(x) \cap \Gamma_j(x) = \varnothing$ for $i \neq j$;
- if $\Gamma_i(x) \cap \Gamma_j(y) \neq \varnothing$ then $\text{dist}(x,y) \leq i + j$.

If a graph is (strongly) connected, i.e. there exists a path joining every pair of nodes, it makes sense to consider the maximum distance between any two nodes. Formally:

diam(G)

Definition 3.3 (Diameter of a connected graph)**.** *Given a connected graph $G = (V, E)$, the* diameter *of G, denoted by* diam(G), *is the maximum among the lengths of shortest paths between pairs of nodes in V, i.e.*

$$\text{diam}(G) = \max_{u,v \in V} \text{dist}(u,v).$$

Intuitive meaning of diam(G)

Informally, the diameter measures the "maximum distance" between any two nodes in a (strongly) connected graph. The lower the diameter, the nearer to one another are all the nodes. For example, the complete graph on n vertices $\mathbb{K}_n$ has $\text{diam}(\mathbb{K}_n) = 1$, a linear array (i.e., a chain) or a ring have diameter $\Theta(n)$, while a complete tree with ariety d has diameter $\Theta(\log_d n)$. Other popular graphs, like the hypercube or the shuffle exchange (cf. [9]) have logarithmic diameter with respect to the number of vertices.

With refence to our intent to model ad hoc networks, it is resonable to assume the cost of transmitting a packet from one node to another be the number of traversed edges in a shortest path between them. The unitary cost of an edge represents the delay[1] needed to pass through that link, including physical transmission time, queueing and processing at the endpoints, etc. Under this cost model, the diameter upperbounds the maximum delay in the whole network, provided that an end-to-end communication takes place over an available shortest path.

3.2 ERDŐS-RÉNYI RANDOM GRAPH

The classical Erdős-Rényi model is perhaps the simplest and most studied random graph model. This model has only one simple assumption: each edge exists with a fixed probability, independently of the existence of other edges.

The definition dates back to the first two seminal papers of Erdős and Rényi [16, 17] in 1959 and 1960.

Definition 3.4 ($G_{n,p}$). *Given an integer n and a real* $0 \leq p \leq 1$, *the* Erdős-Rényi Random Graph, *denoted by* $G_{n,p}$, *is the undirected random graph on n vertices where each potential edge is chosen with probability p, independently of other edges.* $G_{n,p}$

A detailed discussion of $G_{n,p}$ and its properties can be found in the classical book [5]. A shorter introduction with references to the literature can be found in [8]. If the reader is more interested in algorithmics, the exhaustive survey paper by Frieze and McDiarmid [20] is a milestone in the area.

The reader should be warned that when we say "$G_{n,p}$ has property X" we actually mean "$\Pr\left[G_{n,p} \text{ has property } X\right] \to 1$ as $n \to \infty$".

1 Assumed equal for all the possible links.

3.2.1 *The evolution of $G_{n,p}$*

Connectivity of $G_{n,p}$

One obvious question about this kind of graph is whether there is connectivity when the edge probability is expressed as function of the number of nodes.

Surprisingly, six different phases are clearly distinguishable in the evolution of $G_{n,p}$.

Range I $p = o(1/n)$

$G_{n,p}$ is the union of disjoint trees, a tree on k vertices appearing when p is in the order of $n^{-k(k-1)}$. For $p = cn^{-k(k-1)}$ with $c > 0$, the distribution of the number of components which are trees of k vertices tends to a Poisson distribution with expected value of $\lambda = (2c)^{k-1}k^{k-2}/k!$.

Range II $p \sim c/n$ for $0 < c < 1$

All connected components of $G_{n,p}$ are trees or unicyclic (a tree plus an additional edge) and almost all vertices are in components which are trees. The largest connected component has about $\frac{1}{c}\log n$ nodes and the expected number of components is $n - p\binom{n}{2} + O(1)$. The distribution of the number of cycles on k vertices is approximately Poisson with mean value $\lambda = c^k/(2k)$.

Range III $p \sim 1/n$

The behaviour of $G_{n,p}$ is dramatically different when $p < 1/n$ or $p > 1/n$ (this is the reason why this phase is called "the double jump"). In the former case, the largest connected component has size $O(\log n)$ and all components are trees or unicyclic. In the latter, most of the trees have merged to form a giant connected component of size $O(n)$

and the remaining components are still trees or unicyclic of logarithmic size. For $p = 1/n$ the largest component has size of about $n^{2/3}$.

Range IV $p \sim c/n$ for $c > 1$

There is a giant component and all others are relatively small, most of them being trees, which consist of

$$n - f(c)n + o(n)$$

vertices, where

$$f(c) = 1 - \frac{1}{c} \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} (ce^{-c})^k$$

is the fraction of nodes residing in the giant component. Clearly, $f(1) = 0$ and $f(c) \to 1$ as $c \to \infty$.

Range V $p = c \log n/n$ for $c \geq 1$

In this range, $G_{n,p}$ becomes connected with high probability as soon as $c > 1$.

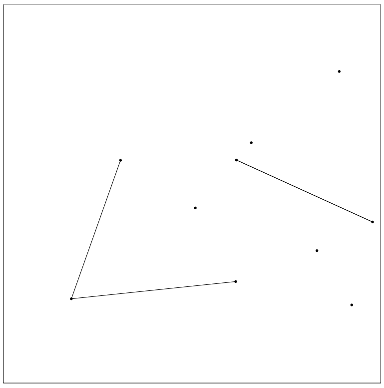
If $p = \frac{\log n}{n} + \frac{y}{n} + o(1/n)$, $\Pr\left[G_{n,p} \text{ is connected}\right] \to e^{e^{-y}}$ as $n \to \infty$ and thus it approaches 1 as $y \to \infty$.

Range VI $p = \omega(n) \log n/n$ for $\omega(n) \to \infty$

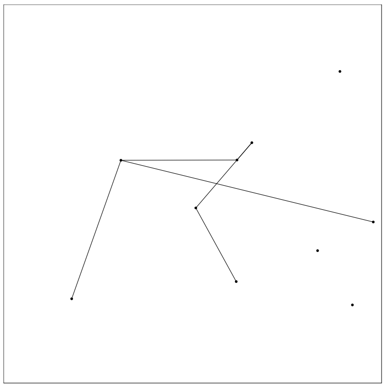
$G_{n,p}$ is connected with high probability and asimptotically the degrees of almost all vertices become equal.

Figure 6 portrays indicatively the evolution of $G_{n,p}$ when the edge probability p increases. The choice of a small $n = 10$ (which makes questionable any probabilistic approach) was done for the sake of clarity when viewing the resulting graph.

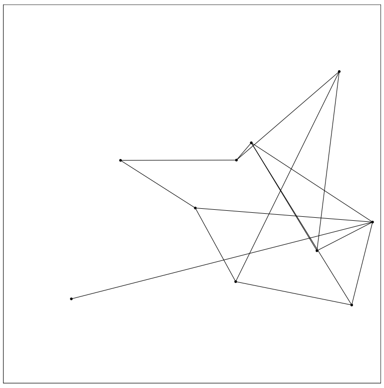
A sufficient condition for the connectivity of $G_{n,p}$ is stated in the following theorem:



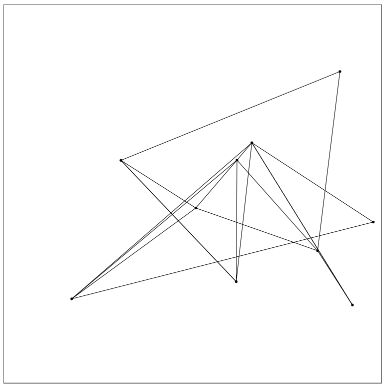
(a) $G_{n,p}$ with $p = 0.05$.



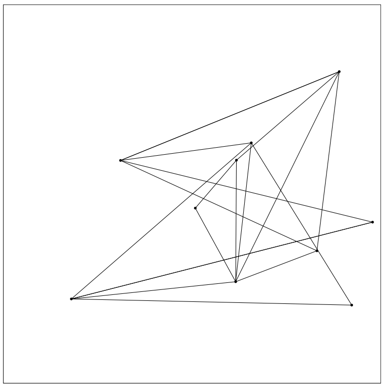
(b) $G_{n,p}$ with $p = 0.10$. This instance is disconnected.



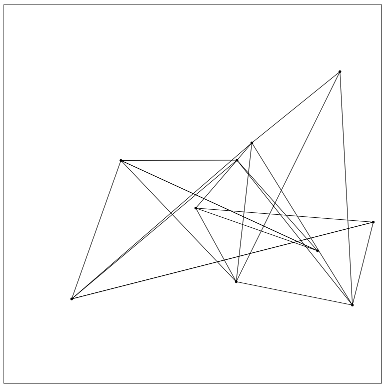
(c) $G_{n,p}$ with $p = 0.10$. This instance is connected.



(d) $G_{n,p}$ with $p = 0.15$.



(e) $G_{n,p}$ with $p = 0.20$.



(f) $G_{n,p}$ with $p = 0.25$.

Figure 6: The evolution of $G_{n,p}$ with $n = 10$ when the edge probability increases. The threshold for connectivity is $p = \frac{1}{n} = 0.10$.

Theorem 3.1 ([16, 17]). *If $p = \frac{\log n + c_n}{n}$, for some $c_n \to \infty$, then $G_{n,p}$ is connected with high probability.*

Sufficient condition for connectivity

If we allow $G_{n,p}$ to be disconnected, we still are interested in knowing when a giant connected component arises in the graph. "When" actually means "which is the minimum p for which one can determine with strictly positive probability the presence of a connected component with cardinality linear in n".

The rising of giant connected component

The following theorem precises what Erdős and Rényi observed in Ranges II–IV. When p is greater than $\frac{1}{n}$, there is a giant connected component in the graph with strictly positive probability, depending on how much p is beyond that critical threshold.

Theorem 3.2 ([23, Theorem 5.4]). *Let $L_1(G)$ be the number of vertices in the greatest connected component of G, and let $L_2(G)$ represent the size of the second greatest connected component of G. If $c > 0$, as $n \to \infty$,*

Giant connected component threshold

$$\frac{1}{n} L_1 \left(G_{n,\frac{c}{n}} \right) \to \phi(c)$$

and

$$L_2 \left(G_{n,\frac{c}{n}} \right) \to 0$$

where $\phi(\cdot)$ satisfies $\phi(c) = 0$ for $c \leq 1$ and $\phi(c) > 0$ for $c > 1$.

Another important property related to the $G_{n,p}$ model, which we will encounter later, is the following.

Property 3.3 ([5]). *If the edges of a complete graph on n vertices are added in an order chosen uniformly at random from all $\binom{n}{2}!$ possibilities, then with high probability, the resulting graph becomes t-connected roughly at the instant when it achieves a minimum degree t.*

t-connectivity hitting time

3.2.2 *The diameter of $G_{n,p}$*

Diameter of $G_{n,p}$

As noted before, if $np < \log n$, $G_{n,p}$ is almost surely disconnected. However, let us convene that, if the whole graph is disconnected, the diameter of $G_{n,p}$ is equal to the maximum diameter of its connected components. When $\frac{np}{\log n}$ is greater than a constant quantity it was shown that the diameter is concentrated on at most a constant number of values (depending on the parameter values) around $\frac{\log n}{\log np}$. Also when the graph is disconnected but $np \to \infty$ the diameter is, within a constant factor, $\frac{\log n}{\log np}$. More interestingly, in [28, 7], Chung and Lu proved that even if np is a constant $t > 1$ the diameter is at most $k\frac{\log n}{\log np}$ with the factor k ranging between 1 and a function of t alone.

For our purposes, we just summarize the known results of [8, p. 96] in Table 3, for some intervals of interest of p. Detailed proofs can be found in [7]. Note that some cases remain open; among all, $p = \frac{1}{n}$.

Parameters	Diameter of $G_{n,p}$
$\frac{np}{\log n} \to \infty$	Concentrated on ≤ 2 values
$\frac{np}{\log n} = t > 8$	Concentrated on ≤ 2 values
$8 \geq \frac{np}{\log n} = t > 2$	Concentrated on ≤ 3 values
$2 \geq \frac{np}{\log n} = t > 1$	Concentrated on ≤ 4 values
$1 \geq \frac{np}{\log n} = t > t_0$	Concentrated on $\leq 2\lfloor \frac{1}{t_0} \rfloor + 4$ values
$\log n > np \to \infty$	$(1 + o(1))\frac{\log n}{\log np}$
$np \geq t > 1$	The ratio $\mathrm{diam}(G_{n,p})/\frac{\log n}{\log np}$ is finite and between 1 and a function of t alone

Table 3: The diameter of $G_{n,p}$, depending on the relative values of n and p. "Concentration" is meant around $\frac{\log n}{\log np}$.

3.2.3 *The usefulness of $G_{n,p}$ in modelling ad hoc networks*

$G_{n,p}$ is (too) simple...

Although some proofs are very technical, the flourishing of results about the properties of $G_{n,p}$ graphs is essentially due to the extremly compact and straightforward formulation of the Erdős-Rényi model.

... to model ad hoc networks

On the other hand, the assumption that an edge exists independently from the existence of others is quite unjustified in practice. Not surprisingly, some characteristics of $G_{n,p}$ make it inappropriate while modelling ad hoc networks.

A spatial problem

Above others, $G_{n,p}$ is a *global* model, where one node can potentially "connect" to every other node. There is no *spatial* notion here — the "actual" placement of nodes in geometric space being not included in the model. In real networks, it is more likely that a node close to another (under some metric) will establish a link with the latter rather than with another device which is very distant. Most importantly, if the transmission range is limited, a node could see only a "small" fraction of the nodes in the system, and thus $G_{n,p}$ is a poor abstraction for this scenario since it allows all the possible edges, i.e. also those whose existence is prohibited by the limited transmission range.

A degree problem

A more subtle aspect of $G_{n,p}$ is the fact that the degrees of the nodes are not controllable without modifying the model. In real ad hoc networks, if there are some devices in the visibility range of a specific node, we might force the node to connect to them, but only until the maximum number of affordable connections (say, c) is reached[2]. This limitation is due to process-

2 If the possible neighbours are more than that limit, we simply select among them using some policy (proximity, signal strength, randomization,...).

ing/memory limits of the single devices or can be caused by energy-preserving mechanisms[3].

In dense scenarios, i.e. when almost always there are enough nodes to choose from, this mechanism ensures that each node mantains a bounded number of connections, say $2c$ (which is the expected degree of a node, under those hypoteses) or just a factor more than that.

Unfortunately, $G_{n,p}$ cannot guarantee such a desiderable property. In fact, using a simple "balls-and-bins" argument [31, 30], one can show that the maximum degree is as high as $\Omega(\log n)$ for those values of p that ensure the global connectivity of $G_{n,p}$.

Starting from these observations, other models have been proposed in literature. The Random Geometric Graph is a well-known probabilistic structure which has been extensively studied and it is the subject of the next section.

3.3 RANDOM GEOMETRIC GRAPH

$G_{n,\lambda}$ accounts for spacial proximity

A mathematical model that considers the "spatial placement" of the nodes is the Random Geometric Graph $G_{n,\lambda}$. In this model, n vertices are chosen uniformly at random from a metric space and are connected by an edge if and only if they are within a given distance.

The formal definition of this random graph is the following and applies to a number of dimensions $d \geq 1$.

$G_{n,\lambda}$

Definition 3.5 ($G_{n,\lambda}$)**.** *Given an integer n and a real $\lambda > 0$, let V_n be a set of n points chosen uniformly and independently at random in $\mathcal{S} = [0,1]^d$. The* Random Geometric Graph $G_{n,\lambda}$ *is the undirected*

3 Notice that a device in a deep sleep state can consume up to 10^4 times less than an inquiring/transmitting device.

graph with vertices V_n in which there exists an edge between two nodes if and only if their Euclidean distance is at most λ.

For our purposes, we will consider only the case $d = 2$, which is of interest for ad hoc network applications.

Since in finite-dimensional spaces every norm is within a constant from any other norm, the choice of the Euclidean distance is not fundamental for the deployment of the results concerning $G_{n,\lambda}$ and any other distance function will work with small changes to the associated constants.

Analogously, in the definition, the unit hypercube can be replaced by any convex subset of $\mathbb{R}^d$ with changes only to the constants appearing in the results. For example, in [15] the unit disk $\mathcal{D} = \{x \in \mathbb{R}^2 : ||x||_2 \leq 1\}$ is used instead of $\mathcal{S} = [0,1]^2$ and the authors call the resulting graph "unit disk random graph".

Historical development

While the first practical studies on $G_{n,\lambda}$ involved optimization topics like the construction of minimum spanning trees, k-nearest neighbour selection (cf. [44]), or layout approximation (as in [13]), the recent developments in the field of ad hoc networks put emphasis on connectivity and routing properties.

Numerical simulations were carried out by engineers and physicists to determine the critical transmission power, i.e. the minimum λ that guarantees the connectivity of the resuting graph. One of the first attempts to study this problem analytically was done by Gupta and Kumar in [21]. Subsequently, the effect of faulty nodes/links has also been investigated, for example in [14].

The reference for all the results mentioned here is the monograph [37] written by Penrose. There, he presents a number of results about various properties of $G_{n,\lambda}$, from component counts to vertex degree sequence, clique number and colourings, from percolative giant component analysis to orderings and partitioning problems.

3.3.1 *Properties of $G_{n,\lambda}$*

$G_{n,\lambda}$ shares some features with $G_{n,p}$

An interesting characteristic of $G_{n,\lambda}$, first proved by Penrose in [36] is the fact that the hitting time of t-connectivity is asymptotically the same as the time needed to achieve minimum degree t. This is quite similar to Property 3.3 of $G_{n,p}$.

Characteristic thresholds

There are two characteristic thresholds that influence the behaviour of $G_{n,\lambda}$. The *thermodynamic limit* occurring at $\lambda \approx n^{-1/d}$ and the *connectivity limit* which is near $\lambda \approx \left(\frac{\log n}{n}\right)^{1/d}$.

Thermodynamic limit

When λ is greater than the thermodynamic threshold, $G_{n,\lambda}$ is said to be in the *supercritical phase*. The expected vertex degree tends to a constant as $n \to \infty$ and a giant connected component appears[4]. On the other hand, when λ is smaller than the thermodynamic threshold, $G_{n,p}$ is said to be in *subcritical phase*.

Connectivity limit

The connectivity threshold discriminates between the *superconnectivity phase* ($\frac{n\lambda^d}{\log n} \to \infty$) where the graph is connected with high probability, and the *subconnectivity phase* ($\frac{n\lambda^d}{\log n} \to 0$) where the graph is disconnected with high probability.

For $G_{n,\lambda}$, it is common to use a different nomenclature from the standard graph-theory. We say that $G_{n,\lambda}$ is *sparse* if $n\lambda^d \to 0$ and that it is *dense* if $n\lambda^d \to \infty$. This convention is justified by the behaviour nearby the connectivity limit.

The reader should be warned that from now on we will discuss only the case $d = 2$, however general results for higher dimensions hold and can be found in [37]. Furthermore, we re-

4 This is similar to the property of $G_{n,p}$ established in Theorem 3.2.

view only some results that are of interest for our subsequent discussion.

A fundamental negative result on the connectivity of $G_{n,\lambda}$ is the following theorem, which establishes a necessary condition on the visibility radius in order to obtain a strongly connected graph.

Connectivity of $G_{n,\lambda}$

Theorem 3.4 (See [36, 37, 15]). *If $\lambda \leq \delta\sqrt{\log n/n}$, for some constant $0 < \delta < 1$, $G_{n,\lambda}$ is disconnected with high probability.*

In contrast, the following theorem states a sufficient condition for connectivity:

Theorem 3.5 (See [38, 21]). *If $\pi\lambda^2 = \frac{\log n + c_n}{n}$, for some $c_n \to \infty$, then $G_{n,\lambda}$ is connected with high probability.*

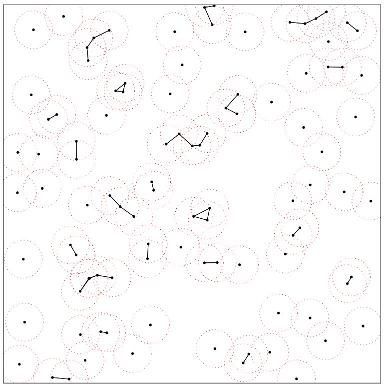
Note that the latter theorem is strikingly similar to Theorem 3.1 derived for $G_{n,p}$.

In Figure 7 the evolution of an instance of $G_{n,\lambda}$ is evident; the subfigures are obtained by increasing λ from a small value until the connectivity threshold is reached while mantaining fixed the location of the vertices.

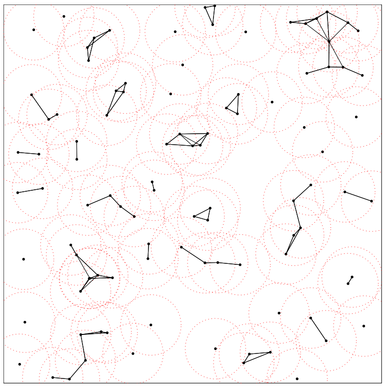
Obviusly, one is also interested in the "quality" of the connectedness attained by $G_{n,\lambda}$. In particular, in the field of ad hoc networks one might know the diameter of the graph.

Diameter of $G_{n,\lambda}$

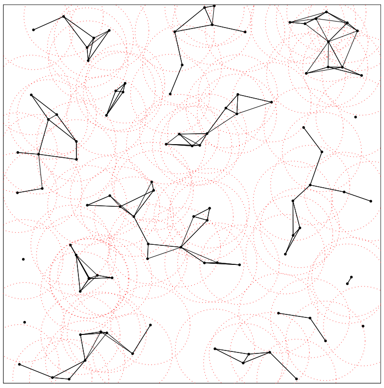
The results collected in [37] — mainly derived using *continuum percolation theory* techniques in $\mathbb{R}^d$ — deal with the "metric diameter" (cf. [37, Section 10.3]) of $G_{n,\lambda}$. Informally, the metric diameter is the actual distance (measured accordingly to a chosen norm) between the two most distant nodes in a realization of $G_{n,\lambda}$. This concept is related to but differs from the diameter as stated in Definition 3.3.



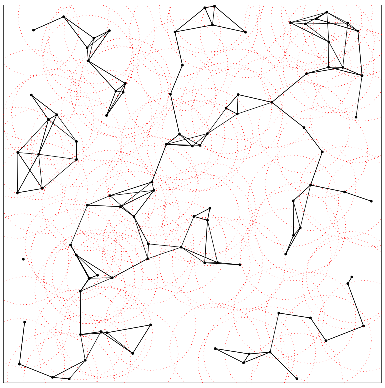
(a) $\lambda = 0.050$ Only few nodes are not isolated.



(b) $\lambda = 0.080$ A non-negligible number of isolated nodes remains.

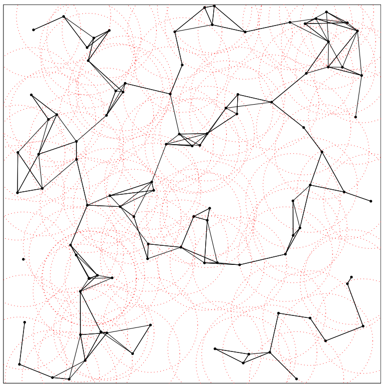


(c) $\lambda = 0.105$ Many different connected components, few isolated nodes.

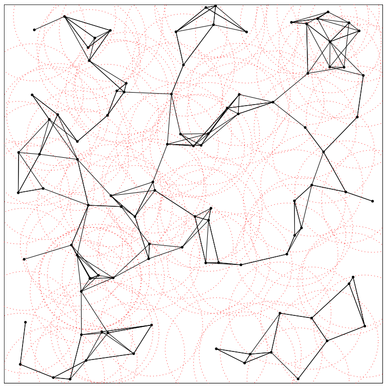


(d) $\lambda = 0.120$ Most of the components have merged.

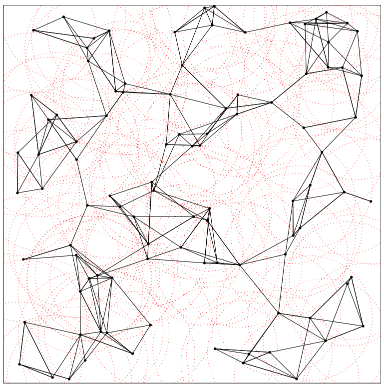
Figure 7: The evolution of an instance of $G_{n,\lambda}$ on $n = 100$ vertices in function of the visibility range λ. The connectivity threshold is around $\sqrt{\log n / n} \approx 0.215$.



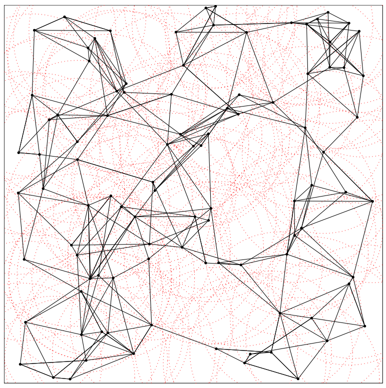
(e) $\lambda = 0.125$ There are only two big connected components. There is still an isolated node.



(f) $\lambda = 0.135$ There are only two big connected components, no isolated point.



(g) $\lambda = 0.165$ The graph is now connected. Note that some nodes have low degree.



(h) $\lambda = 0.215$ The graph exhibits a large degree for all nodes.

Figure 7: The evolution of an instance of $G_{n,\lambda}$ on $n = 100$ vertices in function of the visibility range λ. The connectivity threshold is around $\sqrt{\log n/n} \approx 0.215$.

However some recent articles in the literature present new results on the "graph-theoretic" diameter. Here we briefly present the principal ones contained in [15] where the unitary disk $\mathcal{D}$ is used instead of the unitary square $\mathcal{S}$.

A first theorem states that as soon as $G_{n,\lambda}$ gets connected, its diameter can be upper bounded by $K \cdot 2/\lambda$ where K is a positive constant (not depending from n nor λ). For the reader's ease, we report it here:

$\mathrm{diam}(G_{n,\lambda}) < K \cdot 2/\lambda$

Theorem 3.6 ([15, Theorem 4]). *Let $\phi(n) \to \infty$ be nonnegative. If $\lambda \geq \sqrt{(\log n + \phi(n))/n}$ then there exists an absolute constant $K > 0$ such that almost always, the unit disk random graph is connected with diameter $< K \cdot 2/\lambda$.*

With a little stronger hypotheses, a fine refinement could be achieved, namely:

A better upper bound for appropriate λ

Theorem 3.7 ([15, Theorem 7]). *Let $\lambda = c\sqrt{\log n/n}$. If $c > \frac{\sqrt{12\pi}}{\sqrt{4\pi - 3\sqrt{3}}} \approx 2.26164$, then almost always, the unit disk random graph is connected with diameter $\leq (4 + o(1))/\lambda$.*

Our results on $\mathrm{diam}(G_{n,\lambda})$

It is worth indicating that, as a by-product of our studies on $BT(r(n), c(n))$, we obtained a nice, simple proof that

$$\mathrm{diam}(G_{n,\lambda}) = \Theta\left(\frac{1}{\lambda}\right).$$

Specifically, we prove in Theorem 4.1 that

$$\mathrm{diam}(G_{n,\lambda}) = \Omega\left(\frac{1}{\lambda}\right)$$

and then we match this lower bound by proving in Theorem B.1 that, w.h.p.

$$\mathrm{diam}(G_{n,\lambda}) \leq \frac{2\sqrt{5}}{\lambda} + 2 \approx \frac{4.472}{\lambda} + 2.$$

3.3.2 *The usefulness of $G_{n,\lambda}$ in modelling ad hoc networks*

As said before, $G_{n,\lambda}$ solves the main drawback of $G_{n,p}$ since it accounts for the intrinsic geometric nature of modelling the topology of an ad hoc network.

$G_{n,\lambda}$ is better than $G_{n,p}$...

Additionally, even if the model is relatively recent, it has been studied extensively and the demonstration techniques applied to it are now mature and agreed upon. They combine combinatorial arguments to probabilistic analysis and other theories — like *continuum percolation* — derived from physics and natural science fields.

However, the assumption that a node is connected to every other node in its visibility range is too stringent in practice. Remember that in a real scenario, thousands of nodes can be within transmission distance but each device can interact with only a few of them (say, some dozens), because of memory limits or energy consumption, as explained in Section 3.2.3.

... but not enough!

It is crucial, therefore, to introduce a parameter in the mathematical model to have an adequate tool to describe the neighbour selection operated by a device in a real-life ad hoc network. This observation gives birth to the $BT(r(n), c(n))$ graph which will be discussed in the next pages.

Neighbour selection has to be considered

3.4 BLUETOOTH TOPOLOGY

The Bluetooth Topology $BT(r(n), c(n))$ is an undirected random graph very similar to $G_{n,\lambda}$ with the additional parameter c controlling the degree of the nodes therein.

Informally, $BT(r(n), c(n))$ is a subgraph of $G_{n,r(n)}$ obtained by letting each node choose only $c(n)$ neighbours among the visible

Informal definition

nodes, i.e. the nodes at distance at most $r(n)$. Its name derives from the fact that it approximates very well the graph arising during the piconet formation, the so-called BLUETOOTH device discovery phase [18].

$BT(r(n), c(n))$

Definition 3.6 (Bluetooth Topology)**.** *Given an integer n and two functions $r(n)$ ("visibility range") and $c(n)$ ("selection function"), let V_n a set of n nodes chosen uniformly and independently at random in $\mathcal{S} = [0, 1]^2$. Each node selects $c(n)$ neighbours independently, uniformly at random among the set of nodes within distance $\leq r(n)$ (and all the visible ones, if they are less than $c(n)$). An edge $\{u, v\} \in E_n$ exists if and only if u has selected v or viceversa. The resulting undirected graph $G = (V_n, E_n)$ is called* Bluetooth Topology *and is denoted by $BT(r(n), c(n))$.*

Dubahashi et al. [12] considered $r(n)$ and $c(n)$ as fixed constants, i.e. $r(n) = r$ and $c(n) = c$ with $r, c = O(1)$, and called the resulting $BT(r, c)$ graph the "irrigation graph" $G^n_{r,c}$.

Expansion properties

In 2004, Panconesi and Radhakrishnan [34] showed that with $r(n), c(n)$ fixed constants, $G^n_{r,c}$ is a linear expander, although for a quite huge $c = 10^7$.

Defining parameters as functions of n

Rather than fixing r, c arbitrarily, a better way to characterize the properties of $BT(r(n), c(n))$ is trying to express the visibility range and the number of neighbours a node can choose as a function of the total number of nodes in the system. Succeding in that, we could predict the asymptotic behaviour of any BLUETOOTH-based network by just scaling the results obtained for the unitary square to the actual covered area. In fact, working in the unitary square actually means that n is the *areal density* of the nodes in the region.

3.4.1 *The connectivity of* $BT(r(n), c(n))$

A lower bound on $r(n)$ for the connectivity

By viewing $BT(r(n), c(n))$ as a particular geometric random graph in $d = 2$ dimensions and a more restrictive edge selection policy, we can impose a lower bound on the visibility range required to the full connectivity of the graph. In fact, Theorem 3.4 states that even if we allow a node to select all the possible neighbours, if $\lambda < \sqrt{\log n / n}$ then the $G_{n,\lambda}$ is disconnected w.h.p.. Since selecting (removing) edges cannot improve the connectivity, it is clear that there is no hope for a $BT(r(n), c(n))$ to be connected unless

$$r(n) = \Omega\left(\sqrt{\frac{\log n}{n}}\right).$$

Figure 8 provides a graphical comparison between $G_{n,r(n)}$ and $BT(r(n), 3)$ evolution as $r(n)$ increases. The placement of points is the same for both graphs. It is clear that as the $r(n)$ grows, $G_{n,r(n)}$ contains a lot of additional edges over those established by $BT(r(n), 3)$. In particular, it is evident how the "backbone" of the network arises as soon as the radius is sufficiently large to cover "enough" area.

Connectivity with fixed constant r, c

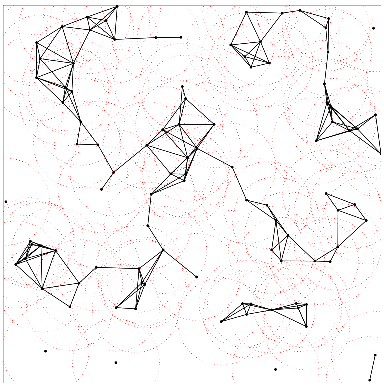
In a seminal paper, Dubahashi et al. [12] considered a first version of $BT(r(n), c(n))$, fixing $r(n), c(n)$ to be constants. They proved that the obtained graph (dubbed "irrigation graph" $G^n_{r,c}$) is connected with high probability as soon as $c \geq 2$. Precisely, they stated the following

Theorem 3.8 (Theorem 4 of [12])**.** *Fix $r > 0$ and $c \geq 2$. Then*

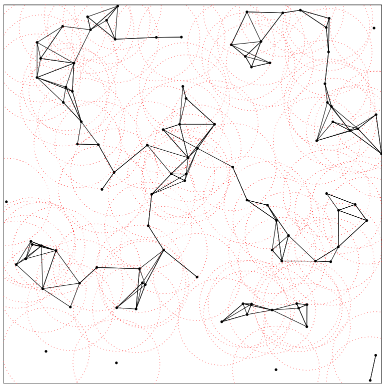
$$\lim_{n\to\infty} \Pr\left[G^n_{r,c} \text{ is connected}\right] = 1.$$

Connectivity for $c(n) = \Omega\left(\log \frac{1}{r(n)}\right)$

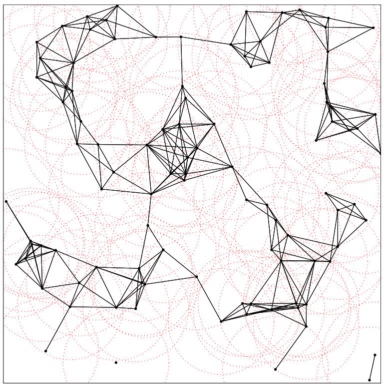
An effective improvement upon this result for the entire range of possible visibility radii was proposed by Crescenzi et al. in [10]. They proved formally that for $r(n) = \Omega\left(\sqrt{\frac{\log n}{n}}\right)$, allowing a



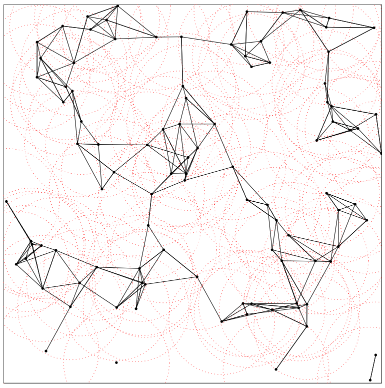
(a) $G_{n,\lambda}$ with $\lambda = 0.115$.



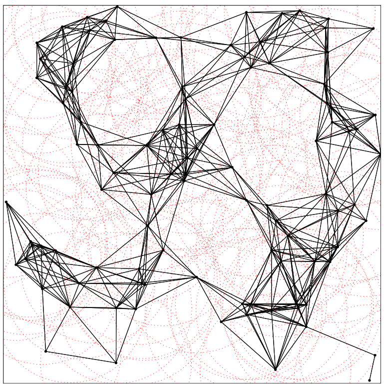
(b) $BT(r, 3)$ with $r = 0.115$.



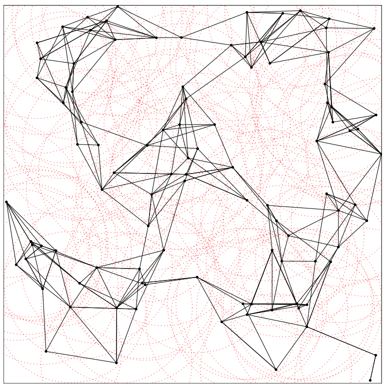
(c) $G_{n,\lambda}$ with $\lambda = 0.140$.



(d) $BT(r, 3)$ with $r = 0.140$.



(e) $G_{n,\lambda}$ with $\lambda = 0.200$.



(f) $BT(r, 3)$ with $r = 0.200$.

Figure 8: A comparison between $G_{n,r(n)}$ and $BT(r(n), 3)$ as function of the visibility range $r(n)$. Both graphs have the same placement of $n = 100$ vertices. The connectivity threshold is around $\sqrt{\log n/n} \approx 0.215$.

node to select $c(n) = \Omega\left(\log \frac{1}{r(n)}\right)$ neighbours suffices w.h.p. to guarantee the connectivity of the resulting $BT(r(n), c(n))$.

Connectivity for the double-choice protocol

Moreover, this lower bound on $c(n)$ seems to be amenable to improvement when the visibility radius[5] is $\Omega(n^{-1/8})$. In fact, under a slightly modified selection protocol, just three neighbours seem sufficient to achieve full connectivity w.h.p.. However, the proof of this latter theorem assumes a "double-choice protocol": one neighbour has to be chosen sufficiently close and the other two randomly among the visible ones [10, Section 3].

Some experimental results

Previous experimental results [12] showed that $c = 4$ is enough to guarantee the connectivity of $BT(r(n), c(n))$, when $r = \Theta(1)$ and n is sufficiently large. More refined simulations reported in [10] seem to confirm that with as few choices as three, $BT(r(n), 3)$ is connected for various ranges of $r(n)$ and n.

3.4.2 *Properties of* $BT(r(n), c(n))$

We remark upon some characteristics of the $BT(r(n), c(n))$ graph model, in order to draw attention to its weaknesses and strengths. Clearly, it is a simplified abstract model of the actual discovery phase in Bluetooth protocol, and thus some aspects are neglected to make the mathematics manageable.

Uniform distribution of nodes

The assumption on the uniform distribution of nodes may be questionable, althought it is a resonable first zero-knowledge approximation. In a real application, another distribution might be interesting or a deterministic *a priori* placement of nodes could be imposed. In the former case, we believe that the results holding for the "native" $BT(r(n), c(n))$ model can be ported to ac-

5 The exponent $-1/8$ has no intrinsic meaning and emerges from the technical details of the proof.

complish other (non-pathological) random distributions of the nodes. In the latter case, systems with reasonably many nodes (say, in the order of the thousands or less) can be engineered with a combinatorial optimization approach.

Location obliviousness

A point of strength of the $BT(r(n), c(n))$ model is its obliviousness about the actual absolute placement of nodes. In fact, all the results presented here do not assume that nodes are able to know either their position in space nor their relative distances. This is consistent with the observation stated in Section 2.1.2.

Note that the "geographic" data could significantly improve the performances of the system if no full connectivity is required, which is the typical case of Ad Hoc Sensor Network (ASN). In [19], Flaxman, Frieze and Upfal proposed a limited depth-first search algorithm that can achieve the goal of routing a message from any point of $\mathcal{S} = [0,1]^2$ to a boundary edge with a visibility range $r(n) = \Theta(\frac{1}{\sqrt{n}})$ which is under the minimum range needed for the global connectivity of the resulting graph.

Clearly, if we are modelling a network of vehicles or soldiers equipped with a GPS receiver or any other localization device, we can assume that each node knows its own position. If we are dealing with distributed, low-power sensors, simply we cannot. In short, the information derived by these means could or could not be incorporated in the mathematical model depending on its correspondence with real available data.

Neighbour selection protocol

A more controversial point is the neighbour selection protocol. Again, the fact that a node selects its neighbours uniformly at random among all the visible nodes is a first zero-knowledge approximation.

A natural improvement could be developed considering other features like the signal strength (approximated by a suitable function of the Euclidean distance between nodes) to weight the probability of a certain choice.

Another potential refinement is modelling the discover procedure as a discrete dynamic edge-selection process taking place in a discrete (or better continuous) time domain. In such a case, degree balancing policies could be issued aiming to create a more regular topology.

3.4.3 *The diameter of* $BT(r(n), c(n))$

A naturally arising question is:

> Can we characterize the maximum delay of a BLUETOOTH-based network?

If we adopt the $BT(r(n), c(n))$ model, we can restate this question as:

What's the diameter of $BT(r(n), c(n))$*?*

> Can we prove lower and upper bounds to the diameter of the $BT(r(n), c(n))$ graph?

To the best of our knowledge, this question is still open in the scientific literature. In the next chapter, we will present our contribution toward an answer.

4

DIAMETER OF BLUETOOTH TOPOLOGY

This chapter presents our contribution in establishing analytical bounds for the diameter of $BT(r(n), c(n))$. Section 4.1 summarizes our results, while the subsequent sections contain the complete proofs.

4.1 OUTLINE OF OUR RESULTS

A first, quite obvious result is the "geometric" lower bound of Theorem 4.1 on the diameter of $BT(r(n), c(n))$. The proof is given in Section 4.6.

Theorem 4.1 (Geometric Lower Bound)**.** *There exists a positive real constant γ_1 such that, if*

$$r(n) \geq \gamma_1 \sqrt{\frac{\log n}{n}}$$

then w.h.p.

$$\mathrm{diam}(BT(r(n), c(n))) = \Omega\left(\frac{1}{r(n)}\right).$$

We called it "geometric" since it holds for *any* $c(n)$ and thus even if we allow to choose all the nodes in the visibility range, i.e. it holds even for $G_{n,\lambda}$ whenever $\lambda = r(n)$. This is not surprising, since we already know (cf. Section 3.3.1) that

$$\mathrm{diam}(G_{n,\lambda}) = \Theta\left(\frac{1}{\lambda}\right)$$

with high probability. Incidentally, we also proved with a elementary proof technique a matching asymptotic upper bound for $G_{n,\lambda}$: with high probability

$$\operatorname{diam}(G_{n,\lambda}) \leq \frac{2\sqrt{5}}{\lambda} + 2 \approx \frac{4.472}{\lambda} + 2$$

as Theorem B.1 in Appendix B.

A very interesting result is a good upper bound for the diameter of $BT(r(n), c(n))$, obtained setting $c(n) = \Theta\left(\log \frac{1}{r(n)}\right)$. As the radius increases, we allow each node to select a decreasing number of neighbours and still obtain connectivity and a good worst-case distance between pairs of nodes.

Theorem 4.2 (Upper Bound)**.** *There exist two positive real constants* γ_1, γ_2 *such that, if*

$$r(n) \geq \gamma_1 \sqrt{\frac{\log n}{n}}$$

and

$$c(n) = \gamma_2 \log \frac{1}{r(n)}$$

then both the following events occur w.h.p.:

1. $BT(r(n), c(n))$ *is connected;*
2. $\operatorname{diam}(BT(r(n), c(n))) = \begin{cases} O\left(\frac{1}{r(n)}\right) & \text{if } r(n) \leq n^{-\epsilon} \\ O\left(\frac{1}{r(n)} + \log n\right) & \text{if } r(n) > n^{-\epsilon} \end{cases}$

with $\epsilon = 1/8$*.*

Note that when $r(n) = O\left(\frac{1}{\log n}\right)$, we match the geometric lower bound of Theorem 4.1. Indeed, if $r(n)$ becomes constant, the logarithmic term dominates over $1/r(n)$. Nevertheless, for any actual n of interest, $\log n$ is still a reasonable quantity.

An improved asymptotic lower bound when $c(n)$ is constant and $r(n)$ is maximum (i.e. each node can select its $c(n)$ neighbours out of all others) is given in Section 4.6.

There we show two different results. With a simple counting argument, we demonstrate that the diameter cannot be constant (it is at least $\Omega(\log \log n)$). In addition, we ameliorate this estimate upper bounding the node degree, a technique that gives us a lower bound of $\Omega\left(\frac{\log n}{\log \log n}\right)$, which is very close to the upper bound $O(\log n)$.

Our results are summarized in Table 2, reproduced here for convenience.

$r(n)$	$c(n)$	$\text{diam}(BT)$	Ref.
$\leq n^{-\frac{1}{3}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)})$	4.5
$n^{-\frac{1}{3}} < r(n) \leq n^{-\frac{1}{8}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)})$	4.7
$> n^{-\frac{1}{8}}$	$\Theta(\log \frac{1}{r(n)})$	$O(\frac{1}{r(n)} + \log n)$	4.11

(a) Upper bounds to $\text{diam}(BT(r(n), c(n)))$.

$r(n)$	$c(n)$	$\text{diam}(BT)$	Ref.
$\Omega\left(\sqrt{\frac{\log n}{n}}\right)$	Any	$\Omega(\frac{1}{r(n)})$	4.1
$\sqrt{2}$	$\Theta(1)$	$\Omega(\log \log n)$	4.12
$\sqrt{2}$	$\Theta(1)$	$\Omega(\frac{\log n}{\log \log n})$	4.14

(b) Lower bounds to $\text{diam}(BT(r(n), c(n)))$.

Table 4: A concise view of our bounds to concerning the $\text{diam}(BT(r(n), c(n)))$. These are asymptotic estimates in high probability, i.e. they hold with probability $\rightarrow 1$ as $n \rightarrow \infty$. The fourth column contains the reference to the Theorem or Lemma where the corresponding bound is proved.

4.2 THE FRAMEWORK

We consider the $BT(r(n), c(n))$ model as described in Definition 3.6. Let $BT(r(n), c(n)) = (V_n, E_n)$ the undirected graph over the vertex set V_n made by n points, chosen uniformly at random from $\mathcal{S} = [0, 1]^2$. Each node selects $c(n)$ neighbours choosing among all the other nodes in its visibility range, that is to say among all the other nodes which are at distance at most $r(n)$. If there are less than $c(n)$ possible neighbours in the visibility range of a node, the latter selects all of them. An edge $\{u, v\} \in E_n$ if and only if u has selected v as its neighbour or viceversa (or both).

4.2.1 *Standard tessellation*

For our arguments, we imagine to tessellate $\mathcal{S} = [0, 1]^2$ with k^2 non-overlapping square cells of side $1/k$ (see Figure 9a) where

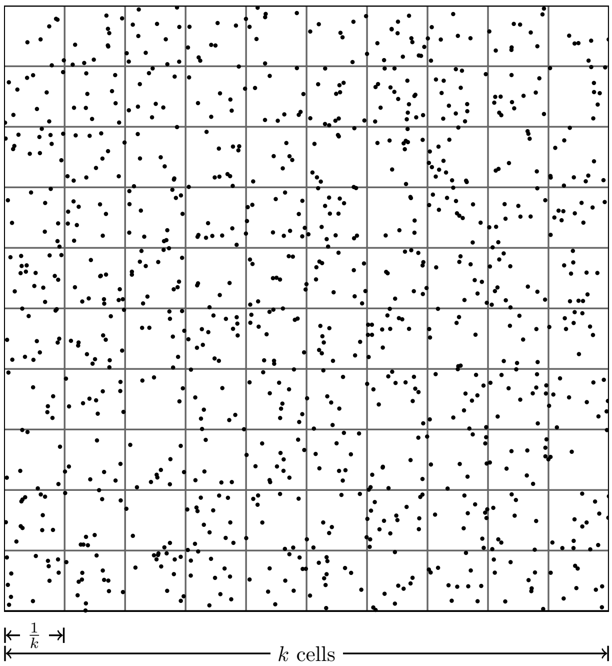
$$k = \lceil \sqrt{5}/r \rceil.$$

This choice is justified by the fact that if two nodes reside in adjacent cells, then their Euclidean distance is not greater than $r(n)$, as shown in Figure 9b.

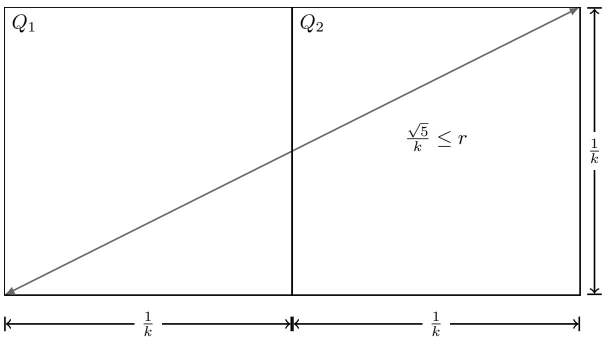
4.2.2 *Sequential discovery procedure*

A *sequential discovery procedure* (or breadth-first search) of $BT(r(n), c(n))$ is an exploration of the graph performed as follows:

- pick a node v_0 in $BT(r(n), c(n))$ — possibly at random if the search is not rooted at any specific v_0;



(a) The tessellation of $\mathcal{S} = [0,1]^2$ into k^2 cells of side $\frac{1}{k}$. A sample uniform distribution of $n = 1000$ nodes is shown; grid spacing is $\frac{1}{k} = 0.1$.



(b) Defining $k = \lceil \frac{\sqrt{5}}{r} \rceil$ allows two nodes residing in two adjacent cells to be within distance r.

Figure 9: The natural tessellation of $\mathcal{S} = [0,1]^2$.

- consider the neighbours

$$v_1, \ldots, v_t \in \Gamma_1(v_0)$$

(with $t \leq c(n)$) chosen by v_0 and continue exploring *in order* the nodes belonging to

$$\Gamma_1(v_1), \ldots, \Gamma_1(v_t)$$

chosen by these latters and so on in a breadth-first manner (see also [9]).

At any given node being considered two things can happen:

- the node has not been seen before: we say it is a *new node*;
- the node has been previously visited: it is a *failure*.

We can stop the search at any time, perhaps because no new nodes can be discovered or a prescribed number of nodes is reached. The nodes which are the leaves of this branching process constitute the *terminal set* of the discovery procedure.

4.2.3 *Some useful facts*

A fundamental observation, which will be exploited many times, states that the uniform distribution of points and a sufficiently large $r(n)$ ensure that each cell has approximately n/k^2 nodes therein and every node can choose its $c(n)$ neighbours out of $\Theta(nr^2)$ visible nodes.

Proposition 4.3 ([10, Proposition 1])**.** *Let* $\alpha = 9/10$ *and* $\beta = 11/10$*. There exists a constant* $\gamma_1 > 0$ *such that for every* $r(n) \geq \gamma_1\sqrt{\log n/n}$ *the following events occur w.h.p.:*

1. *every cell contains at least* $\alpha n/k^2$ *and at most* $\beta n/k^2$ *nodes;*

2. *every node has at least* $(\alpha/4)\pi n r^2(n)$ *and at most* $\beta\pi n r^2(n)$ *nodes in its visibility range.*

Naturally, Proposition 4.3 holds until $r(n) \le 1$. Note that Point 1 is a strengthening of Lemma 9 in [12] which states that w.h.p. each of the k^2 cells contains at least $n/(2k^2)$ nodes.

We already know that the resulting graph is connected w.h.p. as soon as

- the visibility radius is $\approx \sqrt{\log n/n}$;
- we allow $c(n) \approx \log(1/r(n))$ choices per node.

Formally:

Theorem 4.4 ([10, Theorem 1])**.** *There exist two positive real constants* γ_1, γ_2 *such that, if*

$$r(n) \ge \gamma_1 \sqrt{\log n/n}$$

and

$$c(n) = \gamma_2 \log(1/r(n))$$

then $BT(r(n), c(n))$ *is connected w.h.p.*

We know that for small and medium radii, i.e. $\gamma_1\sqrt{\log n/n} \le r(n) \le n^{-1/8}$, with $c(n) = \gamma_2 \log(1/r(n)) = \Theta(\log n)$ choices for each node, the following events hold w.h.p.:

1. for each cell Q, G_Q the subgraph of $BT(r(n), c(n))$ formed by nodes and edges internal to that cell is connected [10, Lemma 1] and
2. for every pair of adjacent cells Q_1, Q_2 there is an edge connecting a node residing in Q_1 and a node residing in Q_2 [10, Lemma 2].

For long radii, i.e. $r(n) > n^{-1/8}$, we can prove that w.h.p.:

1. for $c(n) \geq 2$, $BT(r(n), c(n))$ contains a giant connected component C of size $n/(8k^2)$ [10, Lemma 3] and
2. for $c(n) = \gamma_2 \log(1/r(n)) = \Theta(\log k)$, there exists a path from each node $u \in BT(r(n), c(n))$ to some node in $V(Q, C)$ [10, Lemma 4].

$V(Q, C)$ is set of nodes of C residing in a cell Q' containg at least $n/(8k^4)$ nodes of C. Q' exists by Point 1 and the pigeonhole principle.

To prove Theorem 4.2, we distinguish between three cases, corresponding to three disjoint intervals for $r(n)$, precisely:

- "short" radii: $\gamma_1 \sqrt{\log n/n} \leq r(n) \leq n^{-1/3}$ (Section 4.3);
- "medium" radii: $n^{-1/3} < r(n) \leq n^{-1/8}$ (Section 4.4);
- "long" radii: $r(n) > n^{-1/8}$ (Section 4.5).

The next three sections illustrate the proofs of these three cases of the upper bound stated in Theorem 4.2.

4.3 UPPER BOUND CASE 1: $r(n) \leq n^{-1/3}$

To prove Theorem 4.2 in the case of short radii, we state and prove the following.

Lemma 4.5. *Let*

$$\gamma_1 \sqrt{\frac{\log n}{n}} \leq r(n) \leq n^{-1/3}$$

and

$$c(n) = \gamma_2 \log \frac{1}{r(n)}$$

for a suitable constant $\gamma_2 > 0$. *Then, w.h.p.*

$$\operatorname{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)}\right).$$

Proof. We want to upperbound the number of hops needed to move from a generic node u to another node v. The argument proceeds as follows: we first show that we can start a sequential discovery procedure to reach m (to be determined later) nodes from which we can begin a second phase whose aim is to reach the cell Q_v containing v, as depicted in Figure 10. Then, since Q_v is internally connected, we can upperbound the number of hops needed to actually reach v with the number of nodes contained in the cell.

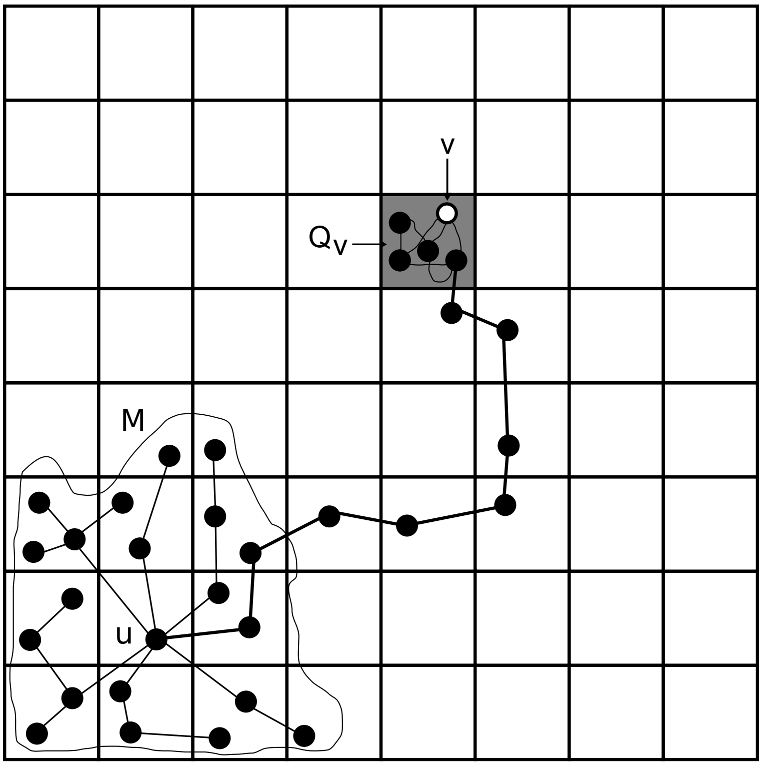


Figure 10: A sketch of the main idea underlying the proof of the upper bound for short and medium radii. M is the set of m nodes reached in the first phase; a succeeding path from u to Q_v is highlighted.

If we can prove that m is sufficiently small and the lengths of the path from u to the m "starting" nodes and then from them to Q_v are "short" we can obtain an upper bound to the diameter of the whole graph using the union bound.

So, run a sequential breadth-first exploration of the directed version[1] of $BT(r(n), c(n))$ and terminate it as soon as m different nodes have been discovered but not yet explored.

Let $w_1, w_2, \ldots, w_m$ be the m nodes reached by this first phase. We calculate the probability that there exists a path from some w_i to the Q_v. Note that the number of vertices of Q_v is $\Theta(n/k^2) = \Theta(nr^2)$ by Proposition 4.3.

We constrain the paths to be disjoint and consisting of exactly one node per cell. Since there is a path of at most $2k$ cells from every w_i to Q_v, we can say that the probability of the existence of a path is at least $p^{2k}q$ where the probability of prolonging a given path by a single cell is

$$p \geq \left(1 - \left(1 - \frac{\alpha n/k^2 - 3m}{\beta\pi n r^2(n)}\right)^{c(n)}\right)$$

and the probability of ending in Q_v is

$$q \geq \frac{\alpha n/k^2}{\beta\pi n r^2(n)} = \sigma$$

with σ being a positive constant less than one.

Now, letting $m = o(n/k^2)$ and γ_2 large enough, we have that $p^{2k} \geq \tau$ for some constant $0 < \tau < 1$.

So, the probability that u is not connected to G_v is at most

$$(1 - \tau\sigma)^m$$

By choosing $m = \gamma_3 \log n$ for a suitable constant $0 < \gamma_3 < \alpha/15$ and a suitable (high) γ_2 we ensure that the above probability be less than $1/n^3$ when $r(n) < n^{-1/3}$. The connectivity is

1 In the directed version, an arc goes from w to y if node w has chosen y as its neighbour.

obtained applying the union bound over $\Theta(k^2) = O(n^2)$ pairs of cells.

Now we calculate the maximum length of the path found in this manner. We need at most $m = \gamma_3 \log n$ hops to reach the "actual" w_i from which we can start the path leading to Q_v.

Afterward, we need at most $2k$ cell-to-cell hops to jump into Q_v and then we can suppose the worst-case situation of being forced to visit all the nodes therein. So, we can say that for each pair of nodes (u, v),

$$\text{dist}(u, v) = O(\log n + 1/r(n) + n/k^2) = O(1/r(n))$$

since the $1/r(n)$ term dominates over $\log n$ and $n/k^2 \approx nr^2(n)$ in the given range of visibility radii.

Concluding, we have that

$$\text{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)}\right)$$

with high probability. □

Please note that for very short radii — which incidentally are the most interesting for practical purposes — we match the Geometric Lower Bound (Theorem 4.1), thus obtaining a tight (asymptotical) result.

4.4 UPPER BOUND CASE 2: $n^{-1/3} \leq r(n) \leq n^{-\epsilon}$

The proof given for the case of short radii is a good starting point to prove the upper bound for the diameter of $BT(r(n), c(n))$ when $n^{-1/3} \leq r(n) \leq n^{-\epsilon}$. In fact, we can mimic that proof (two phases of breadth-first search) until we arrive in the destination cell.

Then, we only have to cope with the fact that the number of nodes residing in a cell — which is $\approx \frac{n}{k^2} \approx nr^2(n)$ — dominates over $1/r(n)$ as soon as $r(n) > n^{-1/3}$ and thus we need an upper bound to the diameter of the subgraph G_Q induced by the nodes residing in a cell Q. Given that, we will procede in establishing the main result (Lemma 4.7).

Lemma 4.6. *Let*

$$n^{-1/3} < r(n) \leq n^{-\epsilon}$$

and

$$c(n) = \gamma_2 \log \frac{1}{r(n)}$$

for a suitable constant $\gamma_2 > 0$. Then, for each cell Q, w.h.p.

$$\text{diam}(G_Q) = O(\log n).$$

Proof. We will prove that from each node u of a specific cell Q, we are able to reach at least $\frac{m}{2} + 1$ nodes of Q with high probability with only $O(\log m)$ steps, where m denotes the number of nodes redsiding in Q. Hence, w.h.p. $\text{diam}(G_Q) = O(\log m)$; the final result will be obtained through the union bound over $k^2 = O(m^2)$ cells.

Consider a specific cell Q and let $m = \Theta(n/k^2)$ be the number of nodes in Q and let $t = \Theta(nr^2(n))$ be the number of nodes in the visibility range of each node by Proposition 4.3. Each node is allowed to choose

$$c(n) = \gamma_2 \log \frac{1}{r(n)} = \gamma \log m$$

neighbours out of the t visible ones, with constant γ proportional to γ_2.

Start a breadth-first search from u. We say that a *failure* occurs at node v if that node selects less than two *new nodes* belonging to Q (i.e., nodes that have not been previously selected). A single choice of v is *good* if it selects a new node, otherwise it is said to be *bad*. Note that in the latter case, the node selected by v is either outside Q or it has been previously discovered.

At each time before $\frac{m}{2}+1$ nodes are discovered starting from u, the probability that a failure occurs is

$$
\begin{aligned}
&\Pr[\text{failure at } v] \leq \\
&\leq \Pr[\text{all } c(n) \text{ choices of } v \text{ are bad}] + \\
&\quad \Pr[\text{1 choice of } v \text{ is good}] \Pr[c(n)-1 \text{ choices of } v \text{ are bad}] \\
&\leq \left(1-\frac{m-(\frac{m}{2}+1)}{t}\right)^{c(n)} + \frac{m-(\frac{m}{2}+1)}{t}\left(1-\frac{m-(\frac{m}{2}+1)}{t}\right)^{c(n)-1} \\
&\leq 2\left(1-\frac{m-(\frac{m}{2}+1)}{t}\right)^{c(n)-1} = 2\left(1-\frac{m-(\frac{m}{2}+1)}{t}\right)^{\gamma \log m - 1} \\
&\leq e^{-\gamma' \log m} = o\left(\frac{1}{m^5}\right)
\end{aligned}
$$

since $0 < \frac{m-(\frac{m}{2}+1)}{t} < 1$ is constant and γ' is an appropriate (high) positive constant proportional to γ_2.

Now, the probability that there is *at least* one failure among the first $M = \frac{m}{2}+1$ nodes is

$$\leq M \quad o\left(\frac{1}{m^5}\right) = o\left(\frac{1}{m^4}\right).$$

Then, we can say that we reach any of these M nodes from u with at most $\Theta(\log_2 M) = \Theta(\log m)$ hops.

Applying the union bound over m nodes, we obtain that from every node $u \in Q$ we can reach M nodes with probability at least $1 - o\left(\frac{1}{m^3}\right)$. Since two nodes $u, v \in Q$ share at least one node out of M, we can conclude that

$$\forall u, v \in Q \quad \text{dist}(u,v) \leq 2log_2 M + 1 = O(\log m)$$

and hence

$$\text{diam}(G_Q) = O(\log m) = O(\log n)$$

with probability $\geq 1 - o\left(\frac{1}{m^3}\right)$.

Considering the union bound over $k^2 = O(m^2)$ cells, we can state that

$$\forall \text{cell } Q \quad \text{diam}(G_Q) = O(\log n)$$

with probability $\geq 1 - o\left(\frac{1}{m}\right) = 1 - o\left(\frac{1}{n^{1/3}}\right)$.

□

Finally, we can state the main result of this section:

Lemma 4.7. *Let*

$$n^{-1/3} < r(n) \leq n^{-\epsilon}$$

and

$$c(n) = \gamma_2 \log \frac{1}{r(n)}$$

for a suitable constant $\gamma_2 > 0$*. Then, w.h.p.*

$$\text{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)}\right).$$

Proof. Consider the proof of Lemma 4.5 which still holds for the current visibility radius, until the cell Q_v is reached.

Then, instead of upper bounding the diameter of Q_v with its cardinality, we exploit the result of Lemma 4.6, which states that every cell has diameter not greater than $O(\log n)$.

So, we can say that for each pair of nodes (u, v),

$$\text{dist}(u, v) = O\left(\log n + 1/r + O(\log n)\right) = O(1/r(n))$$

since the $1/r(n)$ term dominates over the others.

Concluding, we have that

$$\text{diam}(BT(r(n), c(n))) = O(1/r(n))$$

with high probability even for medium radii. □

4.5 UPPER BOUND CASE 3: $r(n) > n^{-\epsilon}$

In this section we will study the diameter of $BT(r(n), c(n))$ when the visibility radius is long, i.e. $r(n) > n^{-\epsilon}$.

Remember that under this assumption, we can no longer claim that each induced subgraph G_Q is connected, hence we cannot exploit the "local" connectedness as done for the previous cases.

The main idea is building a (short) path from both the starting node and the ending node towards $V(Q, C)$, and then upper bounding the diameter of the giant connected component, which automatically implies upper bounding the distance between the nodes residing in $V(Q, C)$ belonging to the aforementioned paths starting in the two points. Then, the shortest path between two nodes has length not greater than the sum of the diameter of $V(Q, C)$ plus two times the number of hops needed to reach $V(Q, C)$ from the two end-points, as depicted in Figure 11.

4.5.1 *A $O(\log n)$ bound for the diameter of the giant connected component*

The following theorem shows that the giant connected component has (at most) logarithmic diameter.

Theorem 4.8. *Let $r(n) > n^{-\epsilon}$. For every $c(n) \geq 2$ the giant connected component C (as in 4.2.3) of size $n/(8k^2)$ has diameter $O(\log n)$ w.h.p.*

To prove Theorem 4.8 we need to establish two technical lemmas.

Lemma 4.9. *Let $r(n) > n^{-\epsilon}$ and $c(n) \geq 2$. Run a sequential discovery procedure until $m = O(n^{1/3})$ nodes are explored. Let M to be the*

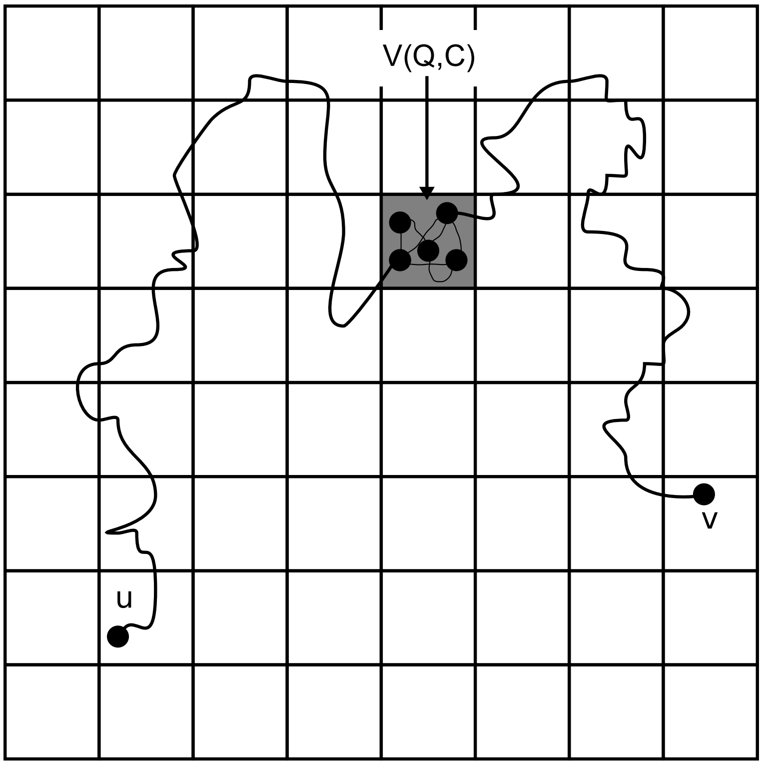


Figure 11: A sketch of the main idea underlying the proof of the upper bound for long radii.

terminal set of the discovery procedure starting at v_0 *and denote the maximum distance of a node in* M *from* v_0 *with* $\text{dist}(v_0, M)$*. Then w.h.p.*

$$\text{dist}(v_0, M) = O(\log n).$$

Proof. We conduct the proof for $c = 2$ since adding further choices can only improve the convergence of the search and shorten the length of the paths.

When selecting a neighbour (i.e., establishing an edge), at most $2m$ nodes have already been seen but each node can select from at least $n/(2k^2)$ nodes w.h.p. by our choice of k. Thus, the probability that a single edge is not valid is at most $\frac{4m}{n/k^2}$. It follows immediately that the probability that at least one edge is a failure is at most $2m\frac{4m}{n/k^2} = \frac{8k^2m^2}{n}$.

With $m = O(n^{1/3})$ and $r(n) > n^{-\epsilon}$, the fraction tends to zero as n approaches infinity, that is the (binary) search tree is complete w.h.p. and thus the maximum distance of the root from a leaf is $\log_2 m = O(\log n)$. □

Given the terminal set M of the first phase of the discovery procedure as described in Lemma 4.9, we denote with T_i the search tree rooted at $w_i \in M$, consisting of all the nodes discovered from w_i and its descendants. Let P the terminal set of the whole discovery procedure, that is the set of nodes discovered but not explored when $n/(8k^2)$ nodes in total are eventually found.

Define the *depth* of M as

$$\text{depth}(M) = \max_{w_i \in M} \max_{y_j \in T_i \cap P} \text{dist}(w_i, y_j).$$

Lemma 4.10. *Let $r(n) > n^{-\epsilon}$ and $c(n) \geq 2$ as in Lemma 4.9. Let M the terminal set of the first phase of the discovery procedure, consisting of $m = \Theta(n^{1/3})$ nodes. Then w.h.p.*

$$\text{depth}(M) = O(\log n).$$

Proof. Again, we might reduce to consider only the case $c = 2$. Define a node in the second phase to be *good* if both of its choices are successful, *bad* otherwise. Each edge has probability at most $\frac{n/(8k^2)}{n/(2k^2)} = \frac{1}{4}$ of not being successful, so $\Pr[\text{a node is good}] \geq 9/16$.

We stochastically upper bound this branching process allowing $\Pr[\text{a node is good}] = 9/16$ and $\Pr[\text{a node is bad}] = 7/16$.

Now we will bound the depth of the search starting from a given $w_i \in M$. Let l_i be the number of leaves in the branching tree T_i rooted at w_i, that is $l_i = |T_i \cap P|$. A graphical sketch is shown in Figure 12.

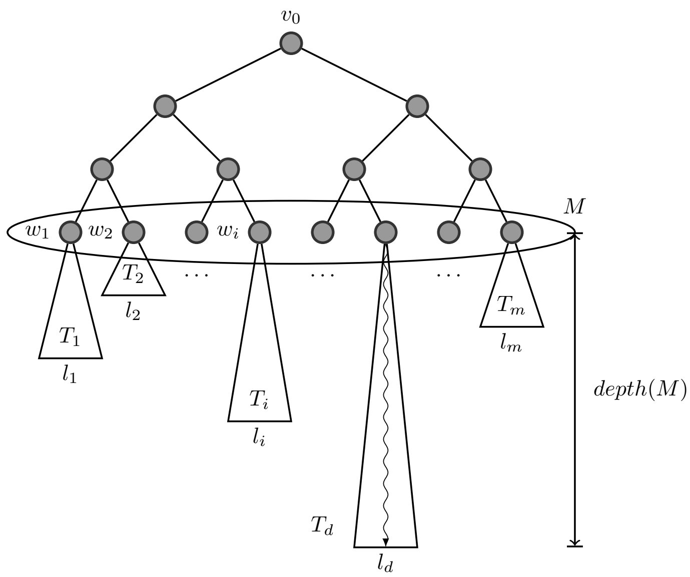


Figure 12: A sketch of the branching tree rooted at v_0 which we use to bound w.h.p. the diameter of the giant connected component.

Consider a particular path Π from w_i to a leaf $y \in T_i \cap P$. We want to compute the probability

$$\Pr\left[|\Pi| > s\right]$$

where $s = a \log_2 l_i$ for a positive constant a to be determined in the analysis.

For every node in Π, define an indicator variable

$$X_j = \begin{cases} 1 & \text{the } j\text{-th node is good} \\ 0 & \text{otherwise} \end{cases}$$

and let $X = \sum_{j=1}^{s} X_j$ be the number of good choices in the first s nodes and define $\mu = \mathrm{E}\left[X\right] = \frac{9}{16}s$.

We compute the probability that

$$\Pr[X < \log_2 l_i] = \Pr[X < (1-\delta)\mu]$$

allowing $(1-\delta)\mu = \log_2 l_i$.

With a simple application of the Chernoff bound (Theorem A.1) for the sum of Bernoulli trials, we get

$$\Pr[X < (1-\delta)\mu] \le \exp\left(-\frac{1}{2}\mu\delta^2\right) < 1/l_i^3$$

where the last inequality holds for a suitable high a ($a = 11$ suffices).

Since there are l_i root-to-leaf paths in the search tree T_i, the probability that there exists one with length greater than $a \log_2 l_i$ is at most $1/l_i^2$.

We know that $\sum_{i=1}^{m} l_i = l = n/(8k^2)$ and thus the application of the union bound over the m starting nodes gives

$$\Pr[A] \le m/l^2$$

where A denotes the existence of a leaf with depth greater than $a \log_2 l$ out of the w_i's belonging to M. For $m = \Theta(n^{1/3})$ and $r(n) > n^{-\epsilon}$, $l = n/(8k^2) = \Omega(n^{1-2\epsilon})$ and therefore

$$\Pr[A] < 1/n.$$

Then, we can conclude that w.h.p.

$$\text{depth}(M) = O(\log n).$$

□

Proof of Theorem 4.8. We know from [10, Lemma 3] that such a giant connected component C exists with high probability.

For every pair of nodes $c_1, c_2 \in C$, we have that

$$\text{dist}(c_1, c_2) = O(\log n)$$

applying Lemma 4.9 and Lemma 4.10.

In fact, if c_1 and c_2 are found in the starting phase, then Lemma 4.9 suffices; otherwise, we could follow the discovery process from c_1 back to v_0 and then re-explore down to c_2 but each of these steps costs at most $O(\log n)$ as a result of Lemma 4.10.

□

4.5.2 *The diameter of $BT(r(n), c(n))$*

The part of Theorem 4.2 relative to the case to long radii, $r(n) > n^{-\epsilon}$, is the following Lemma 4.11.

Lemma 4.11. *Let*

$$r(n) > n^{-\epsilon}$$

and

$$c(n) = \gamma_2 \log \frac{1}{r(n)}$$

for a suitable constant $\gamma_2 > 0$. *Then, w.h.p.*

$$\mathrm{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)} + \log n\right).$$

Proof. The existence in high probability of a path that eventually leads from any node $u \in BT(r(n), c(n))$ to $V(Q, C)$ is proved in [10, Lemma 3, Lemma 4] and, by construction, it requires no more than $O(1/r(n) + \log n)$ hops.

As proved in Theorem 4.8, also the diameter of the giant connected component C is $O(\log n)$ with high probability.

Therefore, we can conclude that, in the case of long radii, the diameter of $BT(r(n), c(n))$ is

$$\mathrm{diam}(BT(r(n), c(n))) = O\left(\frac{1}{r(n)} + \log n\right)$$

with high probability. □

4.6 LOWER BOUNDS

In Section 4.6.1 we prove the Geometric Lower Bound (Theorem 4.1) to the diameter of $BT(r(n), c(n))$.

The upper bounds illustrated in the previous sections match the Geometric Lower Bound when $r(n) = O\left(\frac{1}{\log n}\right)$.

However, when the visibility radius becomes constant, the $\log n$ term in Lemma 4.11 dominates over $1/r(n)$ and thus there is a logarithmic gap between the lower and the upper bound.

In Section 4.6.2 we present a simple counting argument which shows that $\text{diam}(BT(r(n), c(n))) = \Omega(\log \log n)$ when we allow a node to select only a constant number of neighbours out of all other $n - 1$ nodes.

An asymptotic improvement upon this result is demonstrated in Section 4.6.3 with a more sophisticated technique, yielding $\text{diam}(BT(r(n), c(n))) = \Omega(\frac{\log n}{\log \log n})$.

4.6.1 *Geometric Lower Bound*

Proof of Theorem 4.1. Consider the natural tessellation introduced in Section 4.2.1.

By Proposition 4.3, the two cells Q_1 and Q_2 in Figure 13 contain at least one node each, with high probability.

Since the Euclidean distance between Q_1 and Q_2 is $\sqrt{2}\,(1 - \frac{2}{k})$, we need at least

$$\left\lceil \frac{\sqrt{2}\,(1 - \frac{2}{k})}{r(n)} \right\rceil$$

hops in a path from a $u \in Q_1$ and a $v \in Q_2$.

The above quantity can be rewritten as

$$\left\lceil \frac{\sqrt{2}}{r(n)} - \Theta(1) \right\rceil$$

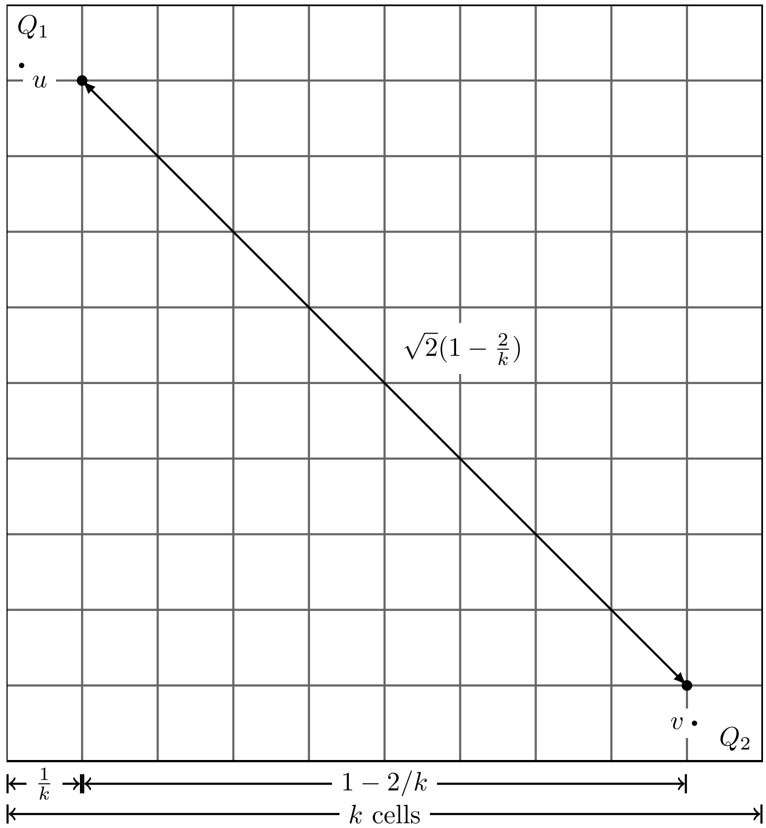


Figure 13: A graphical sketch for proving the geometric lower bound to $\text{diam}(BT(r(n), c(n)))$.

an thus we have that

$$\text{diam}(BT(r(n), c(n))) = \Omega\left(\frac{1}{r(n)}\right).$$

□

As explained in Section 4.1, this lower bound is independent from the choice protocol, and thus we called it "geometric" since it holds even when we allow the nodes to select all the other vertices in its visibility range. In other words, it holds also for $G_{n,\lambda}$ with the convention $\lambda = r(n)$. A matching asymptotic upper bound for $G_{n,\lambda}$, specifically

$$\text{diam}(G_{n,\lambda}) \leq \frac{2\sqrt{5}}{\lambda} + 2 \approx \frac{4.472}{\lambda} + 2$$

can be found in Appendix B.

4.6.2 *A* $\Omega(\log\log n)$ *lower bound for the extreme case* $r(n) = \sqrt{2}$

Theorem 4.12. *Let* $r(n) = \sqrt{2}$ *and* $c(n) = \Theta(1)$. *Then, w.h.p.*

$$\operatorname{diam}(BT(r(n), c(n))) = \Omega(\log\log n).$$

Proof. When $r = \sqrt{2}$ each device has every other node into its visibility range and therefore it selects $c(n)$ neighbours out of the remaining $n-1$ nodes.

Consider two arbitrary nodes $u, v \in BT(r(n), c(n))$ and define

$$p_1 = \Pr\left[\operatorname{dist}(u,v) = 1\right] = \frac{2c(n)}{n-1}$$

to be the probability that they are directly connected, i.e. at least one of the two selects the other as its neighbour.

We can state that the probability p_i that two nodes are exactly at distance $i \geq 1$ is at most the probability that there exists a sequence of $i-1$ intermediate distinct nodes $w_1, w_2, \ldots, w_{i-1}$, with adjacent nodes directly connected and w_1, w_{i-1} directly connected to u, v respectively. So

$$\begin{aligned} p_i &= \Pr\left[\operatorname{dist}(u,v) = i\right] \\ &\leq \Pr\left[\exists w_1, w_2, \ldots, w_{i-1} : \right. \\ &\qquad \left. \operatorname{dist}(u, w_1) = \operatorname{dist}(w_1, w_2) = \cdots = \operatorname{dist}(w_{i-1}, v) = 1\right] \\ &\leq p_1^i \binom{n-2}{i-1}(i-1)! \end{aligned}$$

Now, exploiting the fact that

$$\binom{t}{i} i! \leq t^i$$

we compute

$$
\begin{aligned}
\Pr\left[\operatorname{dist}(u,v) \le \log\log n\right] &\le \sum_{i=1}^{\log\log n} p_i \\
&\le \sum_{i=1}^{\log\log n} p_1^i \binom{n-2}{i-1} (i-1)! \\
&\le \sum_{i=0}^{\log\log n - 1} p_1^{i+1} \binom{n-2}{i} i! \\
&\le \frac{2c}{n-1} \sum_{i=0}^{\log\log n-1} \left(2c\frac{n-2}{n-1}\right)^i \\
&\le \frac{2c}{n-1} \sum_{i=0}^{\log\log n-1} (2c)^i \\
&\le \frac{2c}{n-1} \frac{(2c)^{\log\log n} - 1}{2c-1} \\
&\le \frac{2c}{2c-1} \frac{1}{n-1} \left((\log n)^{\log(2c)} - 1\right)
\end{aligned}
$$

Therefore, assuming $c(n) = \Theta(1)$, we have that

$$
\Pr\left[\operatorname{dist}(u,v) \le \log\log n\right] \to 0
$$

as n goes to infinity and thus w.h.p.

$$
\operatorname{diam}(BT(r(n), c(n))) = \Omega(\log\log n).
$$

□

4.6.3 *A* $\Omega(\frac{\log n}{\log\log n})$ *lower bound*

In oreder to improve the previous result, we show in Lemma 4.13 that the degree of each node is at most logarithmic with probability $> 1 - \frac{1}{n^2}$ as $n \to \infty$.

Then, we can exploit this observation to state that the diameter of the $BT(r(n), c(n))$ graph is at least equal to the diameter of a tree with logarithmic ariety, i.e. $\Omega(\frac{\log n}{\log\log n})$.

Lemma 4.13. *Let $r(n) = \sqrt{2}$ and $c = c(n) = \Theta(1)$. Then each node has degree $\leq c + 3\log n$ with probability $> 1 - \frac{1}{n^2}$ when $n > e^{2.464c}$.*

Proof. Let

$$p = \frac{\binom{n-2}{c-1}}{\binom{n-1}{c}} = \frac{c}{n-1}$$

denote the probability that a node v selects u as its neighbour.

Define the random variables

$$X_u^v = \begin{cases} 1 & \text{if } v \text{ selects } u \text{ as its neighbour} \\ 0 & \text{otherwise} \end{cases}$$

and

$$X_u = \sum_{v \neq u} X_u^v.$$

Note that the latter represents the "in-degree" of node u. Consequently $\Pr[X_u^v = 1] = p$ and $\mu = \mathrm{E}[X_u] = c$.

Let Y_u be the degree of node u: from the above definitions, we have

$$Y_u \leq c + X_u$$

and thus,

$$\Pr[Y_u \geq c + t] \leq \Pr[X_u \geq t]$$

for any integer $t \geq 0$. By the linearity of expectation,

$$\mathrm{E}[Y_u] \leq c + \mathrm{E}[X_u] = 2c.$$

To upper bound the probability of the total degree of node u we upper bound the probability of its "in-degree" via the Chernoff bound for the sum X_u of i.i.d. variables X_u^v.

Let $\delta = \frac{t \log n}{c} - 1$:

$$\Pr[X_u \geq t \log n] \leq \left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\mu}$$
$$\leq \left[\left(\frac{ce}{t \log n}\right)^{\frac{t \log n}{c}}\right]^{c}$$
$$\leq e^{-t \log n} = \frac{1}{n^t}$$

as soon as $\frac{ce}{t \log n} \leq e^{-1}$, i.e. $n \geq e^{\frac{ce^2}{t}}$.

Now, setting $t = 3$, we have that

$$\Pr[X_u \geq 3 \log n] \leq \frac{1}{n^3}$$

for a sufficiently high $n > e^{2.464c}$.

Using the union bound over the n nodes, we can conclude that

$$\Pr[\forall u : \quad Y_u \leq c + 3 \log n] > 1 - \frac{1}{n^2}.$$

□

Recall that in $G_{n,p}$, when $p \approx \frac{c}{n}$, the limit distribution of node degree tends to a Poisson distribution with expectation c as n goes to infinity [8].

An experimental confirmation is given by Figure 14 where the frequency of node in-degrees is shown. We generated a $BT(r(n), c(n))$ graph with $n = 10^6$ vertices, each of them being able to express $c = 32$ choices out of the remaining $n - 1$ nodes. We run 100 simulations and took the average for each value of in-degree.

Theorem 4.14. *Let $r(n) = \sqrt{2}$ and $c(n) = \Theta(1)$. Then, as $n \to \infty$,*

$$\Pr\left[\operatorname{diam}(BT(r(n), c(n))) = \Omega\left(\frac{\log n}{\log \log n}\right)\right] \to 1.$$

Proof. With the same notation of Lemma 4.13, we have that

$$\Pr[\forall u : \quad Y_u \leq c + 3 \log n] > 1 - \frac{1}{n^2}.$$

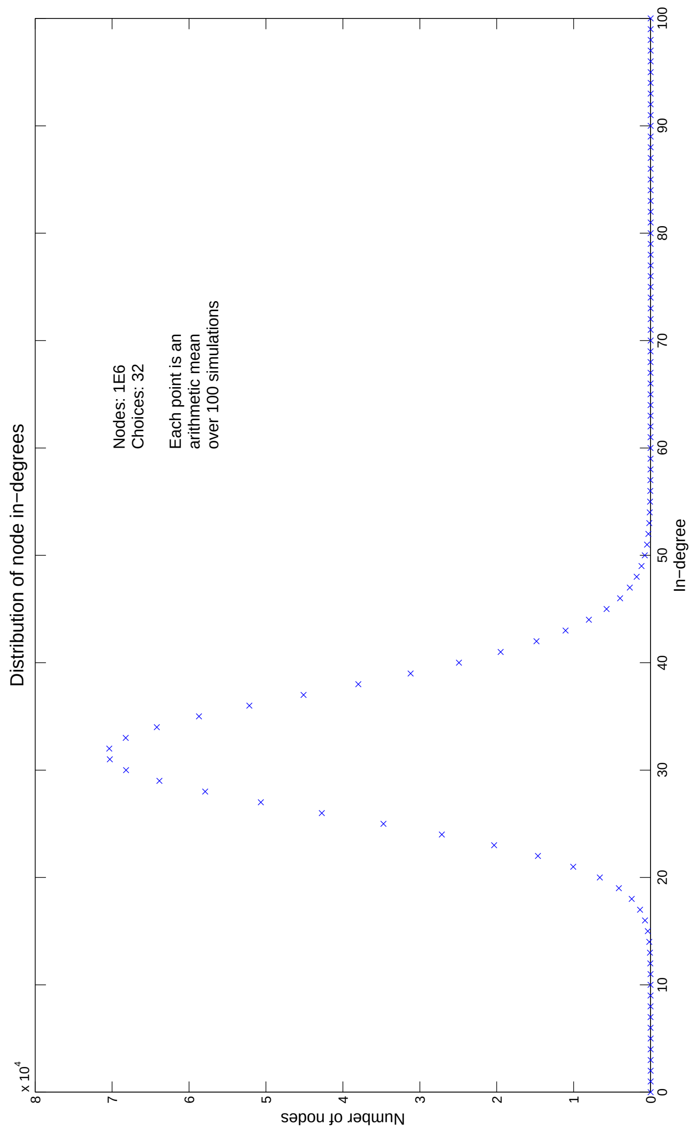


Figure 14: The frequency of node in-degrees of $BT(r(n), c(n))$ with $n = 10^6$ nodes, and $c = 32$ choices for $r = \sqrt{2}$. Each point is an average over 100 simulations.

Then, the diameter of $BT(r(n), c(n))$ cannot be smaller than the diameter of a complete tree of n nodes with ariety $a = (c + 3\log n) - 1$.

The latter is

$$\Theta(\log_a n) = \Theta\left(\frac{\log n}{\log a}\right)$$

and since $c = \Theta(1)$, we have that the diameter of $BT(r(n), c(n))$, when $r(n) = \sqrt{2}$ and $c(n) = \Theta(1)$, is

$$\text{diam}(BT(r(n), c(n))) = \Omega\left(\frac{\log n}{\log \log n}\right)$$

with probability $\rightarrow 1$ as $n \rightarrow \infty$. $\square$

5

CONCLUSIONS

AD HOC NETWORKS are one of the most promising technologies among those mature enough to be deployed on large scale in a near feature, especially for sensing and controlling tasks. Therefore, it is quite easy to predict that their impact on our lives will be significant.

However, a lot of work has to be done on the front of the analysis of their performances and, from an engineering point of view, their design.

The quest for analytical guarantees

In the present work, we reported the properties of some random graphs which are nowadays used to model some ad hoc networks, trying to highlight the aspects corresponding to real behaviours but also spotting out the discrepancies. We focused on Bluetooth Topology, because it seems suitable to model the device discovery phase since takes into account the selection of a small number of neighbours that a device is forced to operate due to its limited resources.

The spirit of our thesis

In literature there were only analytical results on the connectivity of this topology and then we concentrated on proving lower and upper bounds to the diameter of $BT(r(n), c(n))$. We succeeded in proving that the diameter is "small" for almost all the values that the visibility radius $r(n)$ can assume.

Main results

Specifically, if we allow

$$c(n) = \Omega\left(\log \frac{1}{r(n)}\right)$$

choices per node, we can prove our main Theorem 4.2: the diameter satisfies

$$\mathrm{diam}(BT(r(n), c(n))) = \begin{cases} O\left(\frac{1}{r(n)}\right) & \text{if } r(n) \leq n^{-\frac{1}{8}} \\ O\left(\frac{1}{r(n)} + \log n\right) & \text{if } r(n) > n^{-\frac{1}{8}} \end{cases}$$

We also proved a matching "geometric" lower bound $\Omega\left(\frac{1}{r(n)}\right)$ (Theorem 4.1), and for $r(n) = \sqrt{2}$, we are able to increase the bound to $\Omega\left(\frac{\log n}{\log \log n}\right)$, as shown in Theorem 4.14.

$c(n) = \Theta(1)$ suffices?

Experimental results indicate that even a small constant number of choices suffices in practice to guarantee that the resulting graph is connected. A possible improvement of our results goes in this direction: proving that $c(n) = \Theta(1)$ neighbours suffice for the connectedness of the graph or, at least, for the emergence of a giant connected component, as some similar works in the literature indicate.

Selection protocol

Another controversial aspect of the $BT(r(n), c(n))$ model is the fact that neighbours are chosen uniformly at random among all the nodes in the visibility range. This assumption is reasonable but it is not clear whether another selection procedure could lead to more regular topologies or not. Perhaps, we can obtain better overall performances considering other features, like the Euclidean spatial separation between the two parties establishing a communication channel.

Time matters

Moreover, our model is static in the sense that the formation of the graph is done in a single step with all nodes selecting their neighbours at the same time. Clearly this constraint is very strong: ad hoc networks are characterized by the fact that a node can join and leave the network at any time, due to several causes, including clock skews, duty cicle with sleep phases, mobility. Then, a far-reaching improvement over the current $BT(r(n), c(n))$

model would be embedding the selection phase in a temporal structure (maybe discrete) with a different selection procedure. For example, policies aimed to control (bound) the degree of each node could lead to better network performances.

Mobility: problem or opportunity?

Some real networks like MANETs have no fixed structure because nodes can wander in a given area. Mobility poses novel problems, beginning with modelling the paths followed by devices. The "immutable placement" approach used in our work is no more appropriate. The connection between two nodes can be lost — temporarily or permanently — while node density changes and so forth. Some authors have tried to model the evolution of dynamic geometric random graphs but none (at the best of our knowledge) has introduced both the mobility and the neighbour selection. Note that mobility is not necessarly a problem: for example, allowing nodes to change their position, pathological events like a very biased distribution of nodes in the surveilled area are more improbable to occur than with a *a priori* fixed dissemination.

Location/topology awareness of routing

Obviously other properties beyond the diameter are of interest, like expansion and min-cut cardinality, modelling the minimum bandwidth available to a node. A special mention goes to the routing algorithms since they should be network-oblivious because location or topology awareness cannot be assured for, say, sensor networks.

We hope that the reader has found our work worthwhile and has enjoyed the reading as we have enjoyed writing this thesis and, first of all, working on this subject.

A

CHERNOFF BOUNDS

CHERNOFF bounds are powerful tools often used in probabilistic analysis. They estimate the deviation of a random variable from its expected value by providing a strict bound to the probability of that event. They can be derived for any random variable using Markov's inequality and the moment generating function.

The material of this appendix is adapted from [30]; the reader can find there the proofs and an exhaustive set of examples of application of this inequalities.

Here we reproduce only the material needed to understand our proofs, specifically the Chernoff bounds for the tail distribution of a sum of independent Poisson trials.

Poisson trial

Definition A.1 (Poisson trial)**.** *A* Poisson trial *(or* indicator variable*) is a discrete random variable with alphabet* $\{0,1\}$.

Theorem A.1 (Chernoff bounds for the sum of Poisson trials)**.** *Let* $X_1, \ldots, X_n$ *be independent Poisson trials such that* $\Pr[X_i] = p_i$*. Let* $X = \sum_{i=1}^{n} X_i$ *and* $\mu = \mathrm{E}[X] = \sum_{i=1}^{n} p_i$*.*

Chernoff bounds above the mean

Then the following Chernoff bounds *for the deviation* above the mean *hold:*

1. *for any* $\delta > 0$*,*

$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu};$$

2. *for* $0 < \delta \leq 1$*,*

$$\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3};$$

3. *for* $R \geq 6\mu$,

$$\Pr[X \geq R] \leq 2^{-R}.$$

Chernoff bounds below the mean

Moreover, the following Chernoff bounds *for the deviation* below the mean *hold for* $0 < \delta < 1$*:*

1.

$$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu};$$

2.

$$\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}.$$

Bernoulli trials

When considering identically distributed Poisson trials, we often use the term "Bernoulli trials".

Often we cannot derive a precise probability distribution for a random variable of interest, but only establish a domination relation, that is we can find another random variable with known distribution such that the probability that the latter assumes the same value a of the former is smaller for all values a. Formally:

Stochastic domination

Definition A.2 (Stochastic domination)**.** *Given two random variables* X, Y *we say that* Y is stochastically bounded by X *(*$X \gtrsim Y$*) if, for all* a*,*

$$\Pr[X \geq a] \geq \Pr[Y \geq a].$$

Theorem A.2. *Let* Y *be a binomial random variable* $B(n, p)$*.*

1. *If* $X \gtrsim Y$*, for* $a > 0$*,*

$$\Pr[X \leq np - a] \leq e^{-\frac{a^2}{2np}};$$

2. *if* $Y \gtrsim X$*, for* $a > 0$*,*

$$\Pr[X \geq np + a] \leq e^{-\frac{a^2}{2np} + \frac{a^3}{(np)^3}}.$$

B

DIAMETER OF $G_{n,\lambda}$

The tessellation introduced in Section 4.2.1 and the consequent Proposition 4.3 let us state that the diameter of $G_{n,\lambda}$ is $\Theta\left(\frac{1}{\lambda}\right)$.

This result was already known (cf. Section 3.3.1). Although the constant factor is slightly worse, our proof is much simpler than those present in the literature.

The lower bound $\mathrm{diam}(G_{n,\lambda}) = \Omega\left(\frac{1}{\lambda}\right)$ follows from Theorem 4.1.

The upper bound is proved as Theorem B.1. We assume that the visibility radius λ is long enough to satisfy the hypoteses of Proposition 4.3 where we identify $\lambda = r(n)$.

Theorem B.1. *With high probability the diameter of $G_{n,\lambda}$ satisfies*

$$\mathrm{diam}(G_{n,\lambda}) \leq \frac{2\sqrt{5}}{\lambda} + 2 \approx \frac{4.472}{\lambda} + 2.$$

Proof. Consider two arbitrary nodes $u, v \in G_{n,\lambda}$. Since their distance is at most $\sqrt{2}$, the result follows immediately from Lemma B.2. □

Lemma B.2. *Consider two arbitrary nodes $u, v \in G_{n,\lambda}$ and let $\mathrm{d}(u, v)$ their Euclidean distance. Then w.h.p.*

$$\mathrm{dist}(u, v) \leq \sqrt{10}\frac{\mathrm{d}(u, v)}{\lambda} + \sqrt{2}\,\mathrm{d}(u, v).$$

Proof. Given two arbitrary nodes $u, v \in G_{n,\lambda}$, with high probability we can construct a "L-shaped" path from u to v by first moving along x-coordinate and then along y-coordinate, since *every* cell is not empty w.h.p. by Proposition 4.3. An example of such a path is shown in Figure 15.

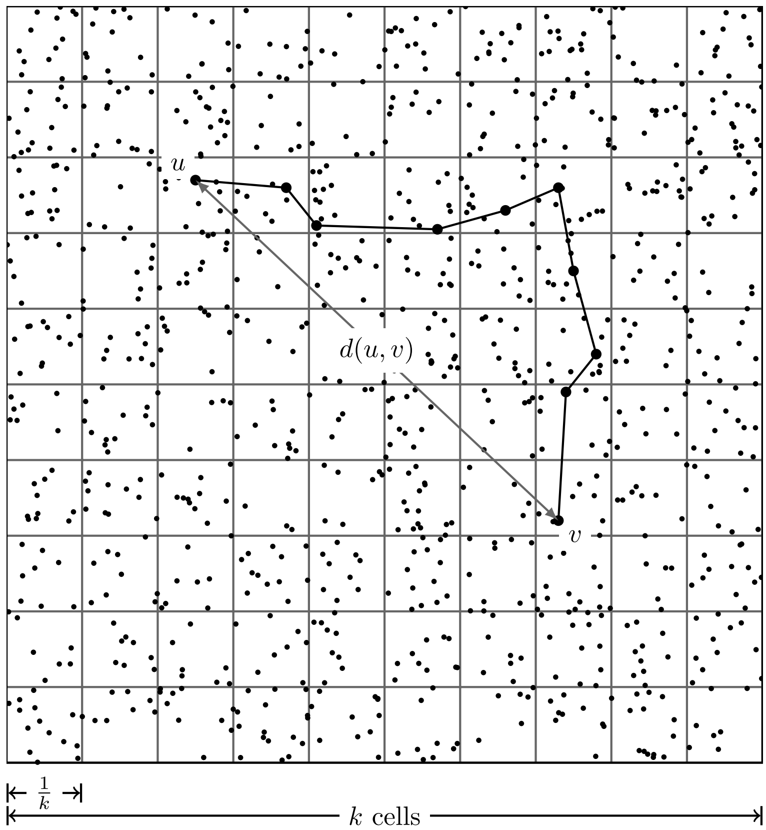


Figure 15: An example of "L-shaped" path.

We can jump with a single hop from a node to the following since they belong to adjacent cells and thus they are adjacent in $G_{n,\lambda}$. The number of cells in the path (and then its length) is

$$\leq \sqrt{2}\frac{\mathrm{d}(u,v)}{1/k}.$$

Therefore, by our choice of

$$k = \left\lceil \frac{\sqrt{5}}{\lambda} \right\rceil \leq \frac{\sqrt{5}}{\lambda} + 1,$$

we have that:

$$\mathrm{dist}(u,v) \leq \sqrt{10}\frac{\mathrm{d}(u,v)}{\lambda} + \sqrt{2}\,\mathrm{d}(u,v).$$

□

BIBLIOGRAPHY

[1] Akyildiz, I. F., Pompili, D., and Melodia, T. State of the art in protocol research for underwater acoustic sensor networks. *SIGMOBILE Mob. Comput. Commun. Rev. 11*, 4 (2007), 11–22. (Cited on page 14.)

[2] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirci, E. Wireless sensor networks: a survey. *Computer Networks (Amsterdam, Netherlands: 1999) 38*, 4 (2002), 393–422. (Cited on pages 1, 11, and 13.)

[3] Benvenuto, N., and Cherubini, G. *Algorithms for Communication Systems and their Applications*. Wiley, 2005. (Cited on page 12.)

[4] Bollobás, B. *Modern Graph Theory*. Graduate Texts in Mathematics. Springer-Verlag, 1998. (Cited on page 23.)

[5] Bollobás, B. *Random Graphs*, second ed. Cambridge University Press, 2001. (Cited on pages 2, 25, and 29.)

[6] Chen, C.-E., Ali, A. M., and Wang, H. Design and testing of robust acoustic arrays for localization and enhancement of several bird sources. In *IPSN '06: Proceedings of the fifth international conference on Information processing in sensor networks* (New York, NY, USA, 2006), ACM, pp. 268–275. (Cited on page 14.)

[7] Chung, F., and Lu, L. The diameter of sparse random graphs. *Advances in Applied Mathematics 26* (2001), 257–279. (Cited on page 30.)

[8] Chung, F., and Lu, L. *Complex Graphs and Networks*. No. 107 in Regional Conference Series in Mathematics. American Mathematical Society, 2004. (Cited on pages 25, 30, and 72.)

[9] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. *Introduction to Algorithms*, second ed. MIT Press, 2001. (Cited on pages 24 and 52.)

[10] Crescenzi, P., Nocentini, C., Pietracaprina, A., Pucci, G., and Sandri, C. On the connectivity of Bluetooth-based ad hoc networks. In *Proc. EURO-PAR'07 Parallel Processing* (Berlin / Heidelberg, 2007), Springer LNCS, pp. 960–969. (Cited on pages 3, 41, 43, 52, 53, 54, 65, and 66.)

[11] Culler, D., Estrin, D., and Srivastava, M. Guest editors' introduction: Overview of sensor networks. *Computer 37*, 8 (2004), 41–49. (Cited on page 13.)

[12] Dubhashi, D., Johansson, C., Häggström, O., Panconesi, A., and Sozio, M. Irrigating ad hoc networks in constant time. In *SPAA '05: Proceedings of the seventeenth annual ACM symposium on Parallelism in algorithms and architectures* (New York, NY, USA, 2005), ACM, pp. 106–115. (Cited on pages 3, 40, 41, 43, and 53.)

[13] Díaz, J., Penrose, M. D., Petit, J., and Serna, M. Approximating layout problems on random geometric graphs. *Journal of Algorithms 39*, 1 (Apr 2001), 1–7. (Cited on page 33.)

[14] Díaz, J., Petit, J., and Serna, M. J. Faulty random geometric networks. *Parallel Processing Letters 10*, 4 (2000), 343–358. (Cited on page 33.)

[15] Ellis, R., Jia, X., and Yan, C. On random points in the unit disk. *Random Structures and Algorithms 29*, 1 (2005), 14–25. (Cited on pages 33, 35, and 38.)

[16] Erdős, P., and Rényi, A. On random graphs. *Publ. Math. Debrecen 6* (1959), 290–297. (Cited on pages 25 and 29.)

[17] Erdős, P., and Rényi, A. On the evolution of random graphs. *Mayyar Tud. Akad. Mat. Kutató Int. Közl. 5* (1960), 17–61. (Cited on pages 25 and 29.)

[18] Ferraguto, F., Mambrini, G., Panconesi, A., and Petrioli, C. A new approach to device discovery and scatternet formation in Bluetooth networks. In *IPDPS '04: Proceedings of the 18th International Parallel and Distributed Processing Symposium* (2004). (Cited on pages 3 and 40.)

[19] Flaxman, A., Frieze, A., and Upfal, E. Efficient communication in an ad-hoc network. *Journal of Algorithms 52*, 1 (Jul 2004), 1–7. (Cited on page 44.)

[20] Frieze, A. M., and McDiarmid, C. Algorithmic theory of random graphs. *Random Structures and Algorithms 10*, 1-2 (1997), 5–42. (Cited on page 25.)

[21] Gupta, P., and Kumar, P. The capacity of wireless networks. *IEEE Transactions on Information Theory 46*, 2 (2000), 388–404. (Cited on pages 33 and 35.)

[22] Intel Corp. *Intel Mote research project (http://www.intel.com/research/exploratory/motes.htm)*, 2007. (Cited on page 13.)

[23] Janson, S., Łuczak, T., and Ruciński, A. *Random Graphs*. Wiley, 2000. (Cited on page 29.)

[24] KC Wirefree LLC. *KC-22 Bluetooth OEM Micro Module Datasheet*, 2007. (Cited on page 19.)

[25] Kettimuthu, R., and Muthukrishnan, S. Is Bluetooth suitable for large-scale sensor networks? In *Proceedings of the 2005 International Conference on Wireless Networks* (2005), pp. 448–454. (Cited on page 16.)

[26] Klingbeil, L., and Wark, T. A wireless sensor network for real-time indoor localisation and motion monitoring. In *IPSN '08: Proceedings of the 2008 International Conference on Information Processing in Sensor Networks* (Washington, DC, USA, 2008), IEEE Computer Society, pp. 39–50. (Cited on page 14.)

[27] Leiner, M., Nielson, D., and Tobagi, F., Eds. *Special Issue on Packet Radio Networks* (Jan 1987), vol. 75, IEEE. (Cited on page 9.)

[28] Lu, L. The diameter of random massive graphs. In *SODA '01: Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms* (Philadelphia, PA, USA, 2001), Society for Industrial and Applied Mathematics, pp. 912–921. (Cited on page 30.)

[29] Mehta, V., and Zarki, M. E. A bluetooth based sensor network for civil infrastructure health monitoring. *Wirel. Netw. 10*, 4 (2004), 401–412. (Cited on page 14.)

[30] Mitzenmacher, M., and Upfal, E. *Probability and Computing*. Cambridge University Press, 2005. (Cited on pages 32 and 79.)

[31] Motwani, R., and Raghavan, P. *Randomized algorithms*. Cambridge University Press, New York, NY, USA, 1995. (Cited on page 32.)

[32] Nachman, L., Kling, R., Adler, R., Huang, J., and Hummel, V. The Intel Mote platform: a Bluetooth-based sensor network for industrial monitoring. In *IPSN '05: Proceedings of the 4th international symposium on Information processing*

in sensor networks (Piscataway, NJ, USA, 2005), IEEE Press, p. 61. (Cited on page 13.)

[33] Oh, S., and Sastry, S. Tracking on a graph. In *IPSN '05: Proceedings of the 4th international symposium on Information processing in sensor networks* (Piscataway, NJ, USA, 2005), IEEE Press, p. 26. (Cited on page 13.)

[34] Panconesi, A., and Radhakrishnan, J. Expansion properties of (secure) wireless networks. In *SPAA '04: Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures* (New York, NY, USA, 2004), ACM, pp. 281–285. (Cited on pages 2 and 40.)

[35] Park, C., Liu, J., and Chou, P. H. Eco: an ultra-compact low-power wireless sensor node for real-time motion monitoring. In *IPSN '05: Proceedings of the 4th international symposium on Information processing in sensor networks* (Piscataway, NJ, USA, 2005), IEEE Press, p. 54. (Cited on page 14.)

[36] Penrose, M. On k-connectivity for a geometric random graph. *Random Structures and Algorithms 15*, 2 (1999), 145–164. (Cited on pages 34 and 35.)

[37] Penrose, M. *Random Geometric Graphs*. Oxford University Press, 2003. (Cited on pages 2, 33, 34, and 35.)

[38] Penrose, M. D. The longest edge of the random minimal spanning tree. *The Annals of Applied Probability 7* (1997), 340—-361. (Cited on page 35.)

[39] Perkins, C. E. *Ad hoc networking: an introduction*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001, ch. 1, pp. 1–28. (Cited on pages 1, 11, and 12.)

[40] Pottie, G. J., and Kaiser, W. J. Wireless integrated network sensors. *Commun. ACM 43*, 5 (2000), 51–58. (Cited on page 11.)

[41] Sharma, G., and Mazumdar, R. Hybrid sensor networks: a small world. In *MobiHoc '05: Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing* (New York, NY, USA, 2005), ACM, pp. 366–377. (Cited on page 11.)

[42] Sikka, P., Corke, P., Valencia, P., Crossman, C., Swain, D., and Bishop-Hurley, G. Wireless adhoc sensor and actuator networks on the farm. In *IPSN '06: Proceedings of the*

fifth international conference on Information processing in sensor networks (New York, NY, USA, 2006), ACM, pp. 492–499. (Cited on page 14.)

[43] Steed, A., and Milton, R. Using tracked mobile sensors to make maps of environmental effects. *Personal Ubiquitous Comput. 12*, 4 (2008), 331–342. (Cited on page 13.)

[44] Steele, J. M. Probabilistic and worst case analyses of classical problems of combinatorial optimization in euclidean space. *Math. Oper. Res. 15*, 4 (1990), 749–770. (Cited on page 33.)

[45] Stoianov, I., Nachman, L., Madden, S., and Tokmouline, T. PIPENET: a wireless sensor network for pipeline monitoring. In *IPSN '07: Proceedings of the 6th international conference on Information processing in sensor networks* (New York, NY, USA, 2007), ACM, pp. 264–273. (Cited on page 14.)

[46] Stojmenovic, I., and Zaguia, N. *Bluetooth scatternet formation in ad hoc wireless networks*. Auerbach Publications, 2006, pp. 147–171. (Cited on page 20.)

[47] Texas Instruments Inc. *TI Bluelink 5.0 Platform Product Bulletin (http://www.ti.com/bluelink_5)*, 2005. (Cited on page 19.)

[48] UC Berkeley. *UCB Firebug Project (http://firebug.sourceforge.net/)*. (Cited on pages 14 and 15.)

[49] Whitaker, R. M., Hodge, L., and Chlamtac, I. Bluetooth scatternet formation: A survey. *Ad Hoc Networks 3* (2005), 403–450. (Cited on page 20.)

[50] Wu, C., Aghajan, H., and Kleihorst, R. Real-time human posture reconstruction in wireless smart camera networks. In *IPSN '08: Proceedings of the 2008 International Conference on Information Processing in Sensor Networks* (Washington, DC, USA, 2008), IEEE Computer Society, pp. 321–331. (Cited on page 14.)

[51] Wu, J. *Handbook on theoretical and algorithmic aspects of sensor, ad hoc wireless, and peer-to-peer networks*, first ed. Auerbach, Boca Raton, 2006. (Cited on page 2.)

[52] Xu, B., Hischke, S., and Walke, B. The role of ad hoc networking in future wireless communications. *ICCT '03: Proceedings of the International Conference on Communication Technology* 2 (April 2003), 1353–1358. (Cited on page 9.)

ACKNOWLEDGMENTS

Writing acknowledgments is never easy since one undergoes the risk of forgetting somebody. So, if you do not see your name here and feel deluded, I will inform you it is just a lack of memory, not a lack of gratefulness.

I want to thank my parents and, more generally, my family because they gave me the chance of studying, helped me in the hardest times, tolerated my restless nature. Your little but numerous favours have contributed to my academic success.

I remember also my friends, for their moral support and the pleasant (always interesting) conversations of these last five years. I will not compile a list of names: it would have been too long!

A special appreciation goes to my advisors, professors Andrea Pietracaprina and Geppino Pucci, who have shown extraordinary patience and generosity, combined with a warm courtesy, without mentioning their professional experience.

I want to thank professors Gianfranco Bilardi, Enoch Peserico and Eli Upfal for their kindness toward me. I acknowledge also Ph.D. students Marco Bressan, Mattia Tonetto and Fabio Vandin for their esteem and cordiality.

To conclude, I want to express my recognition to who has believed in me and also to who has not. The formers for their regard and committment. The latters because they encouraged me to improve myself.

Padua, July 2008 A. P.

RINGRAZIAMENTI

Scrivere dei ringraziamenti non è mai semplice perché si corre sempre il rischio di dimenticare qualcuno. Perciò chi non si vedesse citato e ne fosse giustamente deluso, sappia che ciò è dovuto non a ingratitudine ma a semplice dimenticanza.

Innanzitutto voglio ringraziare i miei genitori e, più in generale, tutta la mia famiglia per avermi dato la possibilità di studiare, avermi supportato nei periodi più impegnativi e faticosi, aver sopportato la mia turbolenta irrequietezza. I vostri piccoli ma numerosi aiuti hanno contribuito al successo della mia carriera accademica.

In secondo luogo, voglio ricordare in questa sede tutti i miei amici, per il sostegno morale e per le piacevoli (e sempre brillanti) discussioni di questi ultimi cinque anni. Non scrivo una lista di nomi: sarebbe troppo lunga!

Uno speciale attestato di stima e riconoscenza va ai miei due relatori, professori Andrea Pietracaprina e Geppino Pucci, che hanno dimostrato una pazienza e una magnanimità eccezionali unite a una squisita cortesia, per non citare la loro esemplare professionalità.

Desidero menzionare anche i professori Gianfranco Bilardi, Enoch Peserico e Eli Upfal per la disponibilità dimostrata nei miei confronti. Ringrazio inoltre i dottorandi Marco Bressan, Mattia Tonetto e Fabio Vandin per la loro stima e cordialità.

Per concludere, voglio formulare la mia riconoscenza a chi ha creduto in me e anche a chi non l'ha fatto. I primi per la fiducia e l'interessamento. I secondi per avermi spronato a migliorare.

Padova, Luglio 2008 A. P.

COLOPHON

This thesis has been typeset using LaTeX 2_ε on `Fedora 8`.
I'd like to thank André Miede for his stylesheet `classicthesis`, based on the beautiful essay *The Elements of Typographic Style* written by Robert Bringhurst.
The skeleton of the document is due to Lorenzo Pantieri but the original template of his guide to `classicthesis` style has been heavily adapted to fit comfortably a graduation thesis.
Text body adopts `Palatino` font designed by Herman Zapf.
Fixed-space font is `Bera Mono` created by Jim Lyles for Bitstream Inc.

Last revision: July, 7^{th} 2008.