Budapest University of Technology and Economics
Institute of Mathematics
Department of Stochastics

Networks and fractals

BSc Thesis

Roland Molontay

Supervisors:

Dr. Júlia Komjáthy
Postdoctoral Research Fellow, Eindhoven University of Technology

Prof. Károly Simon
Professor of Mathematics, Head of Department of Stochastics, Budapest University of Technology and Economics

2013

Contents

Preface

My thesis was written based on the following thesis topic:

Fraktálok és hálózatok

Az elmúlt években egyre nagyobb figyelem irányul a hálzatok elméletére. A téma dr. Simon Károly, dr. Komjáthy Júlia és Vágó Lajos jelenleg is zajló kutatásaihoz kapcsolódik. A hallgató megismerkedik a hálózatelmélet legfontosabb fogalmaival, áttekintést ad néhány fontos hálózatmodellról. Körüljárja a fraktálok elméletét, illetve hálózatokkal való kapcsolatát. Tanulmányozza a különböző fraktáldimenziók fogalmát, törekszik összefüggést találni a gráfok fizikusok által definiált fraktáldimenziója és a matematikusok által ismert fogalom között. Vizsgálja, hogy lehet-e kapcsolat az euklideszi térbe ágyazott fraktál dimenziója és a hálózat gráftávolsággal definiált fraktáldimenziója között.

Acknowledgements

First and foremost I would like to express my greatest gratitude to both of my supervisors, dr. Júlia Komjáthy and Prof. Károly Simon, without their kind help and invaluable advices I would not have been able to write my thesis. I have benefited greatly from their helpful comments and suggestions during the preparation of my thesis.

I am thankful for the enthusiastic supervision of dr. Júlia Komjáthy within the scope of Individual Research Project last year and for introducing me to the world of network theory. I would like to say a special thank for the guidance and support received from her this year in spite of the fact that she moved to Eindhoven.

I am greatly indebted to Prof. Károly Simon for his continuous assistance and tremendous support. I feel motivated and encouraged every time I attend his meeting. I also wish to thank him for the interesting course on Fractals and geometric measure theory.

I am grateful to all of my teachers who helped me during my academic progress. A special thank of mine goes to my secondary school math teacher, Helga Kertai whose way of teaching inspired me to study mathematics.

Finally, I owe a great debt of gratitude to my family for their understanding and support all over the years.

Chapter 1

Introduction

The study of complex networks has received tremendous amount of attention recently, mainly because they are used in several disciplines of science, such as in Information Technology (World Wide Web, Internet), Sociology (social relations), Biology (cellular networks) etc. Understanding the structure of such networks has become essential since the structure affects their performance, for example the topology of social networks influences the spread of information and disease.

1.1 Historical overview

In most cases real networks are too large to describe them explicitly. In order to surmount this problem, network models must be considered. Mathematical modeling of networks dates back to the late 1950s, when thanks to Erdős and Rényi the field of random graphs came into the focus of research interest. The Erdős-Rényi random graph arises by taking a finite set of vertices and placing an edge between any pair of distinct vertices with some fixed probability p [11].

Later, the advent of computer age has opened up new approaches to the investigation of real networks. It has become possible to study real networks which lead to the striking conclusion that many real networks - irrespective of their function share fascinating features, for instance the fact that they are small worlds and scale-free (this means that they obey a power-law degree distribution). The Erdős-Rényi random graph and its generalizations fail to match some of these very important properties, and therefore, new graph models were introduced. In fact already in [11] Erdős and Rényi remark that:

> *"Of course, if one aims at describing such a real situation, one*

should replace the hypothesis of equiprobability of all connection by some more realistic hypothesis."

The question naturally came up: what more realistic procedure could give rise to networks with the above features? Barabási and Albert found an answer to this question in their ground-breaking paper [2] in 1999, which has attracted an enormous amount of attention within the research community. Their answer was the heuristically introduced Preferential Attachment Model (also known as Barabási-Albert model). The principle behind the model is that our world is not perfectly egalitarian, the rich get richer: if a newcomer arrives to a community, they are more likely to get acquainted with socially active people, a new webpage is more likely to link to a popular site rather than a randomly chosen one. In other words, new elements are more likely to attach to elements with high degree compared to elements with small degree. The mathematically rigorous construction of this model is linked with the names of Bollobás, Riordan, Spencer and Tudnády [6].

At the same time another pioneering paper turned the attention to complex networks: Watts and Strogatz results on modeling small-world networks [28]. Their model is given as follows: One starts with a ring of n nodes in which each node is connected to its k nearest neighbours, for a given k. Then, each link is rewired with probability p by choosing randomly a new extremity. A completely different observation - that real networks often obey some hierarchical structure - motivated Barabási, Ravasz and Vicsek [3] to introduce deterministic network models generated by a method which is common in constructing fractals. A similar, fractal based model was introduced by Zhang, Comellas, Fertin and Rong [30].

1.2 Real networks and their fascinating features

In this section we describe a number of real networks and observe some interesting phenomenon. First of all, we present some of the most researched complex networks [14] with the pretension of spanning the variety of contexts in which complex networks appear.

- **Internet:** The topology of the Internet can be modeled by graphs where the nodes are routers (or autonomous systems such as computers at a company/university) linked with physical links.
- **World Wide Web:** The elements of the WWW are web pages, and

there is a directed connection between two pages when the first links to the second.

- **Co-occurrence:** Considering any specified unit of text (a book, the queries to a search engine, or a chat on an interactive system), co-occurrence networks are generated by connecting two words if they appear in the same sentence or query. An example is shown in Figure 1.1.

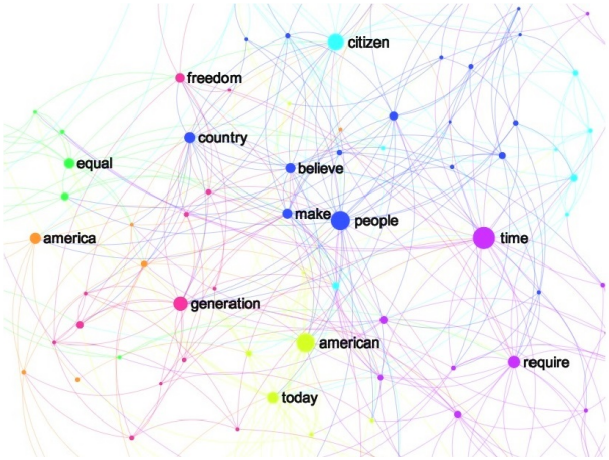

Figure 1.1: The co-occurence network of U.S. President Barack Obama's 2013 inaugural speech. Text network visualization - in comparison with "tag clouds" - emphasizes both the most frequently mentioned words, as well as the relationships between them, making it much easier to understand what the text is about. The figure is from Nodus Labs [2].

- **Social contacts:** We define the vertices of the social network to be the inhabitants of the world, an edge is drawn between two people if they know each other. More precisely "know each other" could mean that the two people involved have shaken hands, or that they know each other on a first name basis. The main difficulties of such networks is that they are notoriously hard to measure.

- **Actors:** In this social network two actors are connected if they have played together in at least one movie. The data of this evolving network is easily available through the Internet Movie Database[3].

[2]http://noduslabs.com/cases/presidents-inaugural-speeches-text-network-analysis/
[3]http://www.imdb.com/

- **Co-authoring:** In this example of complex networks the vertices are mathematicians and two of them share an edge if they have co-authored a paper. This is popularized under the name "Erdős number project" and has drawn substantial attention [12],[4].

- **Proteins:** This cellular network is composed of proteins of a given biological system (Escherichia coli, Homo sapiens, etc.) that linked together if they influence each other. More specifically, the edges are physical, biochemical or functional interactions [27], [15].

The first observation that we can make concerning the real complex networks is that most of them are large. For example the Indexed Web contains at least 14.52 billion pages[4] according to WorldWideWebSize.com[5], and there are over 7.1 billion (and counting) people on the Earth according to Worldometers[6].

Another conspicuous feature of real-world networks is being sparse, in the sense that the nodes have a relatively low degree compared to the size of the graph. To put in other words, according to Dunbar's number [10] a human being can maintain approximately 150 social relationships [9], i.e. approx. 150 edges start from an arbitrary node of the graph of social contacts, which is negligible small in comparison with the all 7 billion nodes of the network.

In addition, many real networks display small world effect, in the sense that typical distances between vertices are small. In the context of social networks this phenomenon is popularized as the six degrees of separation. The idea that any number of people can be reached by a chain of at most 6 intermediaries was originally set out by Frigyes Karinthy. He writes in the short story called "Chain-links", published in the volume titled "Everything is different" in 1929 [16]:

> *"[...] A fascinating game grew out of this discussion. One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth - anyone, anywhere at all. He bet us that, using no more than five individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances. [...]"*

[4]Retrieved Tuesday, 12 February, 2013

[5]http://www.worldwidewebsize.com/

[6]http://www.worldometers.info/world-population/

In 1967 Stanley Milgram confirmed Karinthy's intuition with the help of an interesting experiment based on observing the path of letters [19]. The idea was refreshed by Duncan Watts in 2001, and the new experiment lead to the same result: the average number of intermediaries was six [29]. It makes the common phrase "What a small world!" scientifically based.

We have seen that in most real networks typical vertices can be reached from every other by a small number of steps. It may be of interest to examine a node's degree of separation from a particular vertex. With respect to the collaboration graph of mathematicians this particular vertex is mostly considered to be Pál Erdős, since he published more papers than any other mathematician in history [20], most of them with co-authors. The Erdős number of a mathematician is how many papers they are away from the legendary mathematician, thus Pál Erdős has an Erdős number of zero. Anybody else's Erdős number is $k+1$ where k is the lowest Erdős number of any coauthor.[7] See Table 1.1 for the number of mathematicians with a certain Erdős number.

A second similar example that has attracted some attention is the Kevin Bacon number regarding the movie actor network. The computation of a Bacon number for an actor uses the same algorithm that we have introduced in the previous paragraph. 1902 actors have played in a movie starring Kevin Bacon, hence their Bacon number is one, and 160463 actors have played in a movie in which another movie star played who had played in a movie starring Kevin, but did not work directly with him, which implies these actors have a Bacon number of two.[8] For more details see Table 1.1. We must remark that it turned out that Kevin Bacon is not the most central vertex in the graph, Sean Connery is a more central actor in terms of the average Connery number is 2.731 in comparison with the average Bacon number of 2.954 [26].

A further important structural feature of small-world networks is the high clustering coefficient, which means it is likely that two neighbours of a vertex are adjacent to each other. From a social network perspective it means that the friend of one's friend is also likely to be their friend. This property implies that real networks tend to contain cliques, and near-cliques, meaning sub-networks which have connections between almost any two nodes within them, as we can see in Figure 1.2.

[7] One can see what Erdős number their own professors have or what the distance is between them on the following website operated by the American Mathematical Society: http://www.ams.org/msnmain/cgd/index.html

[8] One can see what Bacon number one's favourite actors have or what the distance is between them at the Oracle of Bacon maintained by Virginia University on http://www.cs.virginia.edu/oracle/

Erdős number	# of mathematicians	Bacon number	# of actors
0	1	0	1
1	504	1	1902
2	6593	2	160463
3	33605	3	457231
4	83642	4	11310
5	87760	5	8168
6	40014	6	810
7	11591	7	81
8	3146	8	14
9	819		
10	244		
11	68		
12	23		
13	5		

Table 1.1: Erdős and Bacon Numbers

Another, maybe more surprising, fundamental property of many real networks is that they are scale-free, meaning the degree distribution obeys a power law: the number of nodes with a degree k decays polynomially with k. It implies that although most nodes have a low degree, there exists few hubs, nodes with very large degree. An example of this is Pál Erdős who had 509 collaborators in comparison with the average number of collaborators per person of 3.36 [26]; popular websites, hyperlinked by several other pages are also hubs. Power laws deeply connected to the Pareto principle, best known under the name "80/20" rule: 80% of the land is owned by 20% of the population, 80% of the total income is earned by 20% of the people, or even maybe 80% of the papers is written by 20% of the scientists. Note that there is nothing special about the 80% mathematically, but it is in strong connection with the typical value of exponent k of real networks [21].

The above features make real networks signicantly different from the classical Erdős-Rényi random graph and its generalizations. This recognition lead to the introduction of many network models.

1.3 Structure of the thesis

My thesis is organized as follows:

- This is the first chapter. We have already given a brief summary of the

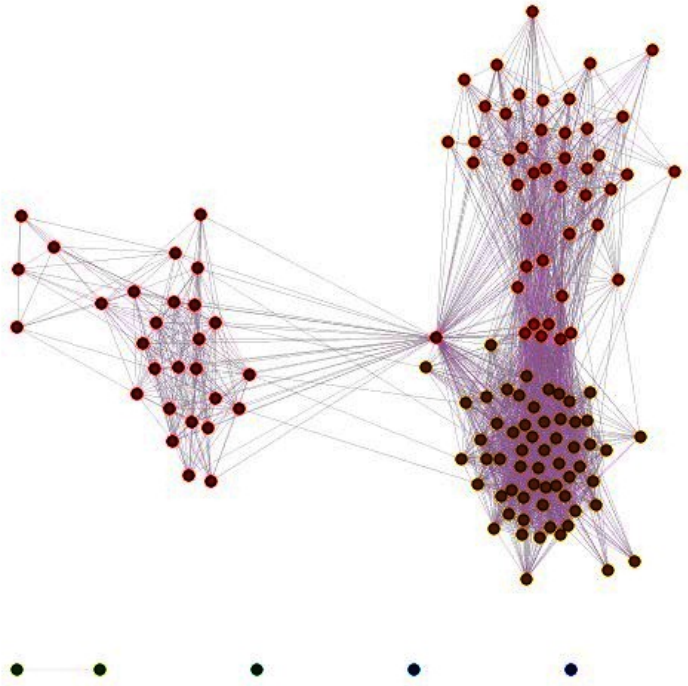

Figure 1.2: This is my social network as seen through my Facebook. The graph is made with the help of Wolfram Alpha's Facebook report [10]. Basically, my social network consists of three near-cliques: my primary school mates, my friends from secondary school and my fellow students at university. We can also observe that there are some "social outsiders" (the isolated vertices) and some "social connectors" or "bridges" (friends of mine with whom I attended the same schools) in the graph.

ever-growing importance of network theory. After a short historical overview of the topic we have made mention of some real complex networks. In particular, their common properties have been illustrated heuristically. Mathematical modeling and analysis of such networks provides the motivation of this thesis.

In the remainder of this chapter we offer a brief survey on what the reader can find in each chapter.

- In Chapter 2 we study how one can model real complex networks. We cover the background material related to network theory, including the

[10] http://www.wolframalpha.com/facebook/

terminology of graph theory. We observe some fundamental properties that many real-world networks share. A few different approaches of modeling real networks are shown in this chapter. We give the description of the Barabási-Albert model, more versions of Apollonian networks and the hierarchical graph sequence model.

- Chapter 3 is devoted to the self-similarity and fractality of complex networks. We show some methods to investigate the self-similar nature and we give the definitions of the most important concepts of network dimension.

- In Chapter 4 we investigate the hierarchical graph sequence model. Using box-counting method we prove the existence of a limit that can serve as an analogue box dimension of the hierarchical graph sequence model. The proof consists of two main parts: giving a lower and an upper bound on the dimension.

- We give a short summary in chapter 5 of this thesis. In addition, we give some plans, and discuss future research directions in this area.

Chapter 2

Mathematical modeling of complex networks

Modeling a real-world object means producing an artificial object that is similar to the real one. In addition, the purpose is to create a mathematically tractable model that allows for rigorous analysis. The first step is getting some information on the properties of the network using a measurement procedure and an analysis of the result. After collecting information there are basically two ways to propose a model. One may focus on static snapshots of networks and obtain a model with the general properties. The other modeling approach attempts to reproduce real world construction processes with the aim of generating networks with such structures. It is essential to create more accurate network models that can produce more realistic models for the future, for instance: How well will a given protocol run on the Internet five years from now? How can search engines explore the WWW more efficiently? How to reduce the spread of the next influenza pandemic?

2.1 Notation

Considering that when we think of a network we mean a system that can be modeled by a graph, to start with, we need to introduce the most important definitions of graph theory and fix the notation used throughout this paper. Here we follow [7] and the corresponding Wikipedia articles[1].

- A *graph* G is an ordered pair $G = (V, E)$, where V is the set of vertices or nodes together with a set E of edges, which are two-element subsets

[1] http://en.wikipedia.org/wiki/List_of_graph_theory_topics, http://en.wikipedia.org/wiki/Glossary_of_graph_theory

of V. To be more precise, this is the definition of undirected and simple graph, since it does not allow neither loops (self-edges) nor multiple edges between elements of V.

- An *undirected graph* is one in which edges have no orientation. The edge (i, j) is identical to the edge (j, i). A *directed graph* is an ordered pair D = (V, A) with V as before, and A a set of ordered pairs of vertices, called directed edges, or arrows.

- The *degree* of a vertex v is the number of edges that connect to it, where an edge that connects to the vertex at both ends (a loop) is counted twice, usually denoted by $\deg(v)$. By simple consideration $\sum_{v \in V} \deg(v) = 2\,|E|$.

- The *neighbourhood* Γ_u of a vertex u is the set of vertices v which are connected to u by an edge. Clearly $\deg(u) = |\Gamma_u|$.

- An *isolated vertex* is a vertex with degree zero.

- In an undirected graph G, two vertices u and v are called *connected* if G contains a path from u to v. Otherwise, they are called *disconnected*. A graph is said to be connected if every pair of vertices in the graph is connected. A *connected component* is a subgraph in which any two vertices are connected to each other, and which is connected to no additional vertices in the graph.

- A *bridge* is an edge whose deletion increases the number of connected components.

- The *maximal subgraph* of $V_1 \subset V$ is $G_1 = (V_1, E_1)$, where $(i, j) \in E_1$ if and only if $i, j \in V_1$ and $(i, j) \in E$.

- *Dyads* and *triads* are the maximal subgraphs of two or three vertices respectively.

- A *clique* in an undirected graph is a subset of its vertices such that every two vertices in the subset are connected by an edge.

- A *complete graph* is a simple undirected graph in which every pair of distinct vertices is connected by a unique edge. A complete graph on n vertices is denoted by K_n.

- A *planar graph* is a graph that can be embedded in the plane, i.e., it can be drawn on the plane in such a way that no edges cross each other.

A simple graph is called *maximal planar* if it is planar but adding any edge (on the given vertex set) would destroy that property.

- A *path* is a sequence of edges such that the target of the previous edge is the source of the next edge. The length of a path is the number of its edges.

- A path is *geodesic* if its end points cannot be connected by shorter path.

- The length of a geodesic between to vertices u and v is the *distance* $d(u, v)$ of these vertices.

- We write $\mathrm{Diam}(G)$ (*diameter*) for the maximal graph-distance in the graph G within components of G. It is not hard to see that $Diam(G) \leq |V(G)| - 1$.

- *Local clustering coefficient* at $v \in V$:

$$C_v(G) := \frac{\text{number of edges between neighbours of } v}{\binom{\deg(v)}{2}}.$$

- The degree of clustering of a whole network can be captured by the *average clustering coefficient*. The most often stated, "$C(G)$ is the average of $C_v(G)$", i.e.:

$$C(G) :=< C(G) >= \sum_{v \in V} \frac{C_v(G)}{|V|}.$$

- Another way to define the clustering coefficient of a whole network is *global clustering coefficient* :

$$C(G) := C_g(G) = \left(\sum_{v \in V} \binom{\deg(v)}{2} C_v(G) \right) / \sum_{v \in V} \binom{\deg(v)}{2}.$$

An equivalent form of this:

$$C_g(G) = \frac{\#\{((a,b),(a,c)) \in E \times E : (b,c) \in E\}}{\#\{((a,b),(a,c)) \in E \times E\}}.$$

We can re-write the equation above as

$$\begin{aligned} C_g(G) &= \frac{3 \times \text{number of triangles}}{\text{number of pairs of adjacent edges}} \\ &= \frac{\text{number of closed triplets}}{\text{number of connected triples of vertices}}. \end{aligned}$$

The second definition has the advantage that it makes sense when some degrees are less than 2 and there is no problem applying the definition to graph that are not simple, thus we use the second one. It is very important to point out that these definitions are not equivalent.

- The *degree distribution* $P(k)$ is the probability that the degree of a randomly (uniformly) chosen vertex is equal to k.

- Given two variables x and y, y is directly *proportional* to x if there is a non-zero constant C such that $y = Cx$. In this paper sometimes we denote this relation by $x \propto y$ or by $x \approx y$.

2.2 Properties of real networks

In this section, we describe what it means for a model of a real network to satisfy the properties that we have introduced in Section 1.2 empirically.

- **Large:** The number of nodes are on the scale of ten of thousands to billions.

- **Sparse:** The number of edges is a constant multiple of vertices, i.e. a node has a bounded number of neighbours.

- **Small worlds:** The characteristic path length ℓ grows proportionally to the logarithm of the number of nodes in the network, i.e. $\ell \propto \log |V|$. The characteristic path length ℓ is defined as the number of edges in the shortest path between two vertices, averaged over all pairs of vertices. It can be shown analytically that scale-free networks (for the definition see below) are ultra-small worlds, in the sense that: $\ell \propto \log \log |V|$[8], [7].

- **Highly clustered:** A network is said to be highly clustered, if it has a much higher clustering coefficient than expected for an Erdős-Rényi random graph of similar number of nodes and edges. In a random graph clustering coefficient C depends on the system size as $\frac{1}{N}$. In contrast, measurements indicate that for real networks C decreases with the degrees of vertices and is largely independent of the system size [26].

- **Scale free:** The degree distribution of the graph follow a power law, i.e. $P(k)$ is proportional to a power of k, for some number $\gamma \geq 1$: $P(k) \propto k^{-\gamma}$. (This definition is widely spread among network scientists, however more sophisticated definitions - two weaker formulation - can

be found in [26] on page 5.) For real networks the degree exponent γ usually satisfies $2 < \gamma < 3$. This property is visualized in Figure 2.1, where the degree distribution is plotted on log-log scale. Thus, we see a plot of $\log(k) \mapsto \log P(k)$. Therefore, $\log P(k) = \log C - \gamma \log k$, so that the slope is a straight line and the negative of the slope of the line is the power law exponent. One can estimate the exponentγ of a real network by substituting $P(k)$ with N_k, the number of vertices with degree k, since - counter to the probability - it can be measured empirically.

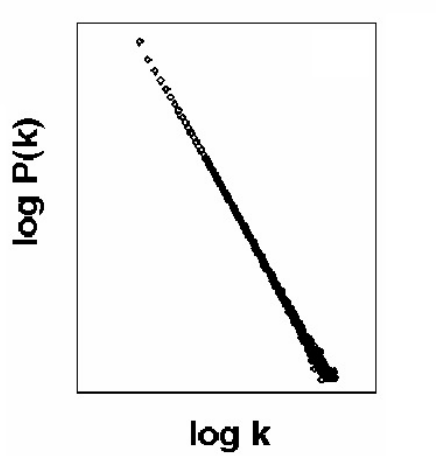

Figure 2.1: The visualization of the scale free property. The figure is from [7].

2.3 Preferential Attachment Model

A possible and convincing explanation for the scale-free nature of real-world networks is offered by the preferential attachment model introduced by Barabási and Albert in [2]. The principle behind the model is the "popularity is attractive" or the "rich get richer" paradigm. In the preferential attachment model, vertices arrive one by one with a number of edges connected to them. The newly added vertices are connected to the already existing vertices with a probability proportional to the degree of the receiving vertex at that time, thus favoring vertices with large degrees.

In [2] the authors describe the preferential attachment graph informally by the following ingredients:

- Initial condition: To start with the network consists of a small number (m_0) of isolated vertices.

- Growth: One vertex v with m ($\leq m_0$) edges is added every time step. The edges link the new vertex to m different vertices already present in the system.

- Preferential attachment: The probability P_u for a new vertex v to be attached to u is:
$$P_u = \frac{\deg(u)}{\sum_{w \in V} \deg(w)}.$$

The first to investigate the model rigorously, were Bollobás et al.[6], they criticize [2] due to lack of a formal definition. In [7] Bollobás writes: "*From a mathematical point of view, however, the description above, repeated in many papers, does not make sense.*"

The description of the model does not explain how to get started. Initially there are no edges, thus we can not take probabilities proportional to the degrees, since they are all zero. (Provided that, we start instead from a small graph G_0 with no isolated vertices, then the choice of G_0 can make notable difference.) The second problem is with the preferential attachment rule itself, and arises only for $m \geq 2$; it is not clear whether these edges are independent, whether we allow for self-loops, whether we should update the degrees after each attachment of a single edge, etc. (One might hope that the exact choice does not make much difference, but in [7] it is shown that there is a wide range of models fitting the Barabási-Albert description with very different properties.)

Now we define precisely a random graph model introduced in [5] satisfying the obscure description above. Here we consider graphs of labelled graphs, meaning that if the graph has n vertices, the vertex set is $[V_n] = \{v_1, v_2, \ldots, v_n\}$. Let $m \geq 1$ be a natural number. We construct a sequence of graphs $\{G_m^t\}_{t=1}^{\infty}$ as follows: We define inductively a random graph process $\{G_1^t\}_{t=1}^{\infty}$, starting with $\{G_1^1\}$ a graph with one vertex and one loop. Given $\{G_1^{t-1}\}$ with $V(\{G_1^{t-1}\}) := [V_{t-1}] = \{v_1, v_2, \ldots, v_{t-1}\}$. Then we form $\{G_1^t\}$ by adding the vertex v_t and one more additional edge between v_t and a vertex $v_i \in [V_t]$ which is chosen randomly proportionally to the degree of v_i at the time counting the new edge as already contributing one to the degree of v_t, i.e.:
$$P(i = s) = \begin{cases} \frac{\deg_{G_1^{t-1}}(v_s)}{(2t-1)}, & \text{if } 1 \leq s \leq t-1; \\ \frac{1}{(2t-1)}, & \text{if } s = t. \end{cases}$$

Finally, we can define $\{G_m^t\}_{t=1}^{\infty}$ by identifying the consecutive m tuples of vertices in G_1^{mt}. Bollobás and Riordan proved rigorously the most important features of the model: The power law exponent turned to be $\gamma = 3$ and the diameter is $\log n$ for $m = 1$ and $\frac{\log n}{\log(\log n)}$ for $m > 1$, where $|V| = n$.

2.4 Apollonian networks

The class of Apollonian networks was introduced by Andrade et al. in [1]. These networks have some advantageous properties making them interesting for a wide range of applications. Apollonian networks - that can be either deterministic or random - are scale free, small world and in the two-dimensional case they are maximal planar graphs.

2.4.1 Deterministic Apollonian Network

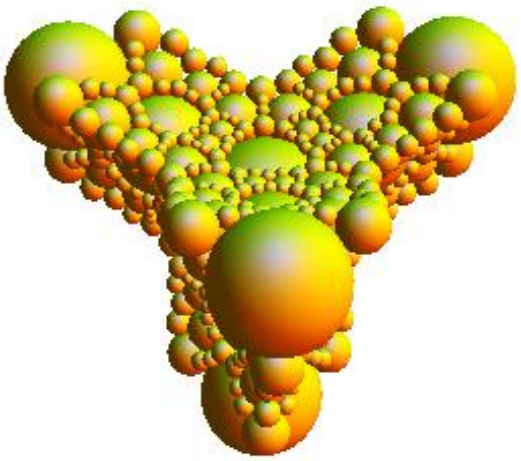

Figure 2.2: The three-dimensional Apollonian gasket. The figure is from Wolfram MathWorld [3].

Apollonian networks are named after the ancient Greek mathematician Apollonius of Perga (ca. 262 BC - ca. 190 BC)[25]. These networks are derived from the Apollonian packing, hence first we introduce the problem of Apollonian packing. Apollonian gasket is constructed as follows: Initially there are three circles C_1, C_2 and C_3, each one of which is tangent to the other two (in the general construction, these three circles can be any size, as

[3]http://mathworld.wolfram.com/ApollonianGasket.html

long as they have common tangents) inside a circular space which is to be filled. Apollonius discovered that there are two other non-intersecting circles, C_4 and C_5, which have the property that they are tangent to all three of the original circles these are called Apollonian circles or Soddy circles (this is a consequence of Descartes' theorem [18]). Take one of the two Apollonian circles, without loss of generality let it be C_4. It is tangent to C_1 and C_2, so the triplet of circles C_4, C_1 and C_2 has its own two Apollonian circles: one of them is C_3, but the other is a new circle C_6. In a similar way we get C_7 that is tangent to C_4, C_2 and C_3 and so on. Continuing the construction stage by stage in this way, we add $2 \cdot 3^n$ new circles at stage n, giving a total of $3^{n+1} + 2$ circles after n stages. In the limit of infinite generations, we obtain the two-dimensional Apollonian gasket.

The two-dimensional Apollonian gasket can be generalized to higher dimensions (d-dimensional, $d \geq 2$). Describing the construction we follow [30]. We start with $d + 1$ mutually tangent d-dimensional hyperspheres (d-hyperspheres). (The 3-dimensional case is visualized in Figure 2.2) The hyperspheres are in a d-dimensional simplex which is enclosed and tangent to a larger hypersphere. The interstices correspond to curvilinear d-dimensional simplices (d-simplices). In the first iteration, and inside each of the interstices, we add d-hyperspheres tangent to each of the $d + 1$ d-hyperspheres bounding the curvilinear d-simplex. The added hyperspheres will not fill the interstice and each produces $d + 1$ smaller interstices. In the second iteration, $d + 1$ d-hyperspheres are added inside all of the $d + 1$ new interstices, being again tangent to the enclosing d-hyperspheres. The process is repeated iteratively obtaining the high-dimensional Apollonian packing.

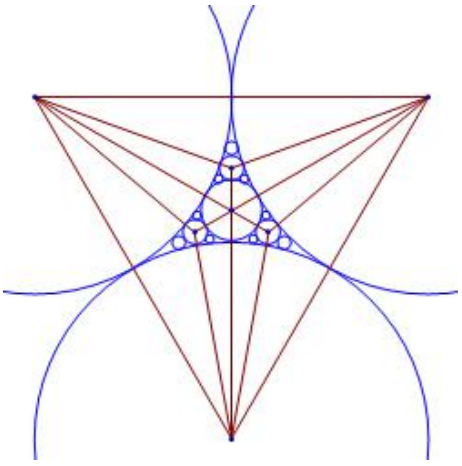

Figure 2.3: This is how to derive the two-dimensional Apollonian network from the packing. The figure is from Wolfram MathWorld [5].

This Apollonian packing can be used to design the Deterministic Apollonian Network (DAN), where the vertices of the network correspond to the circles (d-hypersperhes in higher dimension) and vertices are connected by an edge if the corresponding circles (d-hypersperhes) are tangent. Figure 2.3 shows a network based on the two-dimensional packing.

2.4.2 Random Apollonian Network

The Random Apolloninan Network (abbreviated as RAN or HDRAN if we would like to emphasize the high-dimensional nature of these networks) can be derived from DAN. In the d-dimensional Apollonian packing, if we add only one d-hypersphere inside a randomly selected interstice, then we get a high dimensional random Apollonian packing. Analogously, if each d-hypersphere corresponds to a vertex and vertices are connected by an edge if the corresponding d-hyperspheres are tangent, then we obtain a d-dimensional random Apollonian network.

Here we delineate the iterative algorithm of RAN based on [30], then we implement the algorithm in Mathematica with purpose of further investigation and simulation. The code is shown in Figure 2.4. In the iterative process for the construction of RAN, for each new vertex added, $d+1$ new d-simplices are created in the network, into which vertices may be inserted in one of the following iterations. We denote the d-dimensional random Apollonian network after n iterations by RAN(n, d), $d \geq 2, n \geq 0$. Initially ($n = 0$), RAN(0,d) is the complete graph K_{d+2}. At each step, we choose an existing subgraph isomorphic to a $(d+1)$-clique that has never been selected before (we can call them active cliques), then we add a new vertex and join it to all the vertices of the selected subgraph. The growing process is repeated until the network reaches the desired size.

Despite the stochastic nature of this network it has some deterministic properties. It is easy to see that after n steps the network consists of $V = n+d+2$ vertices. The total degree equals $(d+1)(2n+d+2)$. So, when n is large the average vertex degree at step n is equal approximately to a constant value $2(d+1)$, which shows that this network is sparse like many real-life networks. In [30] the authors showed that RANs obey power-law degree distributions and the average path length of the networks grows logarithmically with the number of nodes.

```
<< Combinatorica`
Needs["GraphUtilities`"];

RAN[n_Integer, d_Integer] := Module[(*n:number of iterations,d:dimension*){g, i, j, rand, clique, newedges},
   g = AdjacencyGraph[ToAdjacencyMatrix[CompleteGraph[d + 2]]];
   (*we start with the complete graph on d+2 vertices*)
   clique = KSubsets[VertexList[CompleteGraph[d + 2]], d + 1]
      (*initially the active cliques are the d+1-cliques of the complete graph*) newedges = {};
   For[i = 1, i < n + 1, i++, g = VertexAdd[g, d + 2 + i];
    (*in each step we add a vertex to the graph*) rand = RandomInteger[{1, Length[clique]}];
    (*we choose a clique from the active cliques randomly*) newedges = Map[{i + d + 2, #} &, clique[[rand]]];
    (*we connect the newly added vertex to all the vertices of the randomly chosen clique*)
    g = EdgeAdd[g, Apply[UndirectedEdge, newedges, 2]];
    clique = Delete[Join[clique, ReplacePart[Table[clique[[rand]], {d + 1}], {j_, j_} → i + d + 2]], rand];
    (*we delete the currently chosen clique from the active cliques,
    and add the newly created d+1 cliques to the list of the active cliques*)];
   g];
```

Figure 2.4: The iterative algorithm of the RAN implemented in Mathematica by me.

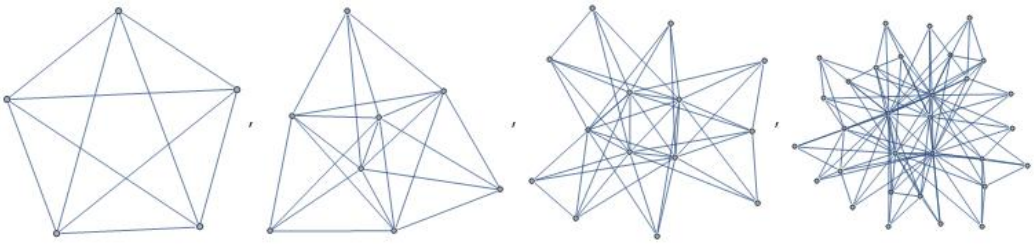

Figure 2.5: The visulaization of the EAN[3, 3, 0.5] in each step: $n = 0, n = 1, n = 2$ and $n = 3$ respectively. The graphs are made with the help of the programme code shown in Figure 2.6.

2.4.3 Evolutionary Apollonian Network

The evolving Apollonian networks (EAN) were introduced by Zhang et al. in [31]. The initial configuration is the same as for the DAN or RAN. Then in each subsequent generation, each d-simplex is filled with probability q. In a special case $q = 1$, it is reduced to the DAN. If q approaches but is not equal to 0, it coincides with the RAN.

Here we also give the iterative algorithm of EAN based on [31], using which we can write a computer program. The Mathematica implementation is shown in Figure 2.6. The d-dimensional EAN with filling probability q after n generations are denoted by EAN(n, d, q), $d \geq 2, n \geq 0, 0 \leq q \leq 1$. Then at step n, the d-dimensional EAN is constructed as follows: For n = 0,

[5] http://mathworld.wolfram.com/ApollonianGasket.html

EAN(0,d,q) is a complete graph K_{d+2}. For $n \geq 1$, EAN(n, d, q) is obtained from EAN($n-1$, d, q). For each of the existing subgraphs of EAN($n-1$, d, q) that is isomorphic to a $(d+1)$-clique and has never generated a node before (i.e. active $(d+1)$-cliques), with probability q, a new node is created and connected to all the nodes of this subgraph. The growing process is repeated until the network reaches a desired order. The illustration of the evolving can be seen in Figure 2.5.

```
<< Combinatorica`
Needs["GraphUtilities`"];

EAN[n_Integer, d_Integer, q_Real] := Module[(*n: number of iterations, d: dimension,
   q: the probability that an active (d+1)-clique generates a new vertex *)
   {g, i, j, k, rand, position, position2, clique, add, newedges, order, newvertices},
   g = AdjacencyGraph[ToAdjacencyMatrix[CompleteGraph[d + 2]]];
   (* we start with the complete graph on d+2 vertices*)
   clique = Subsets[VertexList[CompleteGraph[d + 2]], {d + 1}];
   (*initially the active cliques are the d+1 cliques of the complete graph*)
   For[i = 1, i < n + 1, i++,
    rand = RandomVariate[BinomialDistribution[Length[clique], q]];
    (* we give rand new vertices to the graph in the current step*)
    If[rand == 0, g,
     position = RandomSample[Range[Length[clique]], rand];
     (*we choose the cliques from the active cliques that we add the new vertex to*)
     order = VertexCount[g];
     newvertices = Table[j, {j, order + 1, order + rand}];
     g = VertexAdd[g, newvertices];
     newedges = {};
     For[k = 1, k < rand + 1, k++,
      newedges = Join[newedges, Map[{order + k, #} &, clique[[position[[k]]]]]];
      (*we connect the rand new vertices to the nodes of the chosen their cliques*)
     ];
     g = EdgeAdd[g, Apply[UndirectedEdge, newedges, 2]];
     add = {}; (*we add the newly created ran * (d+1) cliques to the list of the active cliques*)
     For[k = 1, k < rand + 1, k++,
      add = Join[add, ReplacePart[Table[clique[[position[[k]]]], {d + 1}], {j_, j_} → order + k]];
     ];
     position2 = Partition[position, 1];
     clique = Join[Delete[clique, position2], add];
     (* we delete the currently chosen cliques from the active cliques*)
    ];
   ];
   g];
```

Figure 2.6: The iterative algorithm of the EAN implemented in Mathematica by me.

2.5 Hierarchical graph sequence model

In this section we follow Komjáthy and Simon [17] who introduced a general hierarchical graph sequence model derived from a graph directed self-similar fractal, motivated by the deterministic scale-free network model of Barabási, Ravasz and Vicsek [3].

Starting from an arbitrary initial bipartite graph G on N vertices, we construct a hierarchical sequence of deterministic graphs G_n. Namely, $V(G_n)$, the set of vertices of G_n is $\{0, 1, \dots, N-1\}^n$. To construct G_n from G_{n-1}, we take N identical copies of G_{n-1}, each of them identified with a vertex of G. Then we connect these components in a complicated way described in 2.1. In this way, G_n contains N^{n-1} copies of G_1, which are connected in a hierarchical manner, see 2.3 and 2.4 for two examples.

Let G, our *base graph*, be any labeled bipartite graph on the vertex set $\Sigma_1 = \{0, \dots, N-1\}$. We partition Σ_1 into the non-empty sets V_1, V_2 and one of the end points of any edge is in V_1, and the other is in V_2. We write $n_i := |V_i|$, $i = 1, 2$ for the cardinality of V_i. The edge set of G is denoted by $E(G)$. If the pair $x, y \in \Sigma_1$ is connected by an edge, then this edge is denoted by $\binom{x}{y}$.

Now we define our graph sequence $\{G_n\}_{n\in\mathbb{N}}$ generated by the base graph G.

The vertex set is $\Sigma_n = \{(x_1 x_2 \dots x_n) : x_i \in \Sigma_1\}$, all words of length n above the alphabet Σ_1. To be able to define the edge set, we need some further definitions.

Definition 2.1.

1. *We assign a type to each element of* Σ_1. *Namely,*

$$typ(x) = \begin{cases} 1, & \text{if } x \in V_1; \\ 2, & \text{if } x \in V_2. \end{cases}$$

2. *We define the* ***type*** *of a word* $\underline{z} = (z_1 z_2 \dots z_n) \in \Sigma_n$ *as follows: if all the elements* $z_j, j = 1, \dots, n$ *of* $\underline{z}$ *fall in the same* V_i, $i = 1, 2$ *then* $typ(\underline{z})$ *the type of* $\underline{z}$ *is* i. *Otherwise* $typ(\underline{z}) := 0$.

3. *For* $\underline{x} = (x_1 \dots x_n), \underline{y} = (y_1 \dots y_n) \in \Sigma_n$ *we denote the* ***common prefix*** *by*

$$\underline{x} \wedge \underline{y} = (z_1 \dots z_k) \text{ s.t. } x_i = y_i = z_i, \forall i = 0, \dots, k \text{ and } x_{k+1} \neq y_{k+1}.$$

4. *Given* $\underline{x} = (x_1 \dots x_n), \underline{y} = (y_1 \dots y_n) \in \Sigma_n$, *the* ***postfixes*** $\tilde{\underline{x}}, \tilde{\underline{y}} \in \Sigma_{n-|\underline{x} \wedge \underline{y}|}$ *are determined by*

$$\underline{x} = (\underline{x} \wedge \underline{y})\tilde{\underline{x}}, \ \underline{y} = (\underline{x} \wedge \underline{y})\tilde{\underline{y}},$$

where the concatenation of the words $\underline{a}, \underline{b}$ *is denoted by* $\underline{ab}$.

Now we can define the edge set $E(G_n)$. Two vertices $\underline{x}$ and $\underline{y}$ in G_n are connected by an edge if and only if the following assumptions hold:

(a) One of the postfixes $\tilde{\underline{x}}, \tilde{\underline{y}}$ is of type 1, the other is of type 2,

(b) for each $i > |x \wedge y|$, the coordinate pair $\binom{x_i}{y_i}$ forms an edge in G.

That is, $E(G_n) \subset \Sigma_n \times \Sigma_n$:

$$E(G_n) = \left\{ \binom{\underline{x}}{\underline{y}} \quad \middle| \quad \{\mathrm{typ}(\tilde{\underline{x}}), \mathrm{typ}(\tilde{\underline{y}})\} = \{1, 2\} \text{ and } \forall |\underline{x} \wedge \underline{y}| < i \leq n, \binom{x_i}{y_i} \in E(G) \right\}$$

Remark 2.2 (Hierarchical structure of G_n)**.** *For every initial digit* $x \in \{0, 1, \dots, N-1\}$, *consider the set* W_x *of vertices* $(x_1 \dots x_n)$ *of* G_n *with* $x_1 = x$. *Then the induced subgraph on* W_x *is identical to* G_{n-1}.

The following two examples satisfy the requirements of our general model.

Example 2.3 (Cherry)**.** *The "cherry" model was introduced in [3], and is presented in Figure 2.7: Let* $V_1 = \{1\}$ *and* $V_2 = \{0, 2\}$, $E(G) = \{(1, 0), (1, 2)\}$.

Example 2.4 (Fan)**.** *Our second example is called "fan", and is defined in Figure 2.8. Note that here* $|V_1| > 1$.

2.5.1 Properties of the graph sequence

In this subsection we see that this model shares three very important features of many real networks. Namely

(1) the hierarchical structure (as we have seen in 2.2),

(2) the power law decay and

(3) the small world property.

Komjáthy and Simon computed [17]the degree distribution of the graph sequence G_n under some regularity assumption on the base graph G and found that it has an extreme-end power law decay.

They also calculated the average length of shortest path between two nodes in G_n, for arbitrary bipartite graph G; and proved that the magnitude of the average length of a shortest path between two uniformly chosen vertices in G_n is the logarithm of the size of G_n.

In Chapter 5 we make a further investigation on the hierarchical graph sequence model.

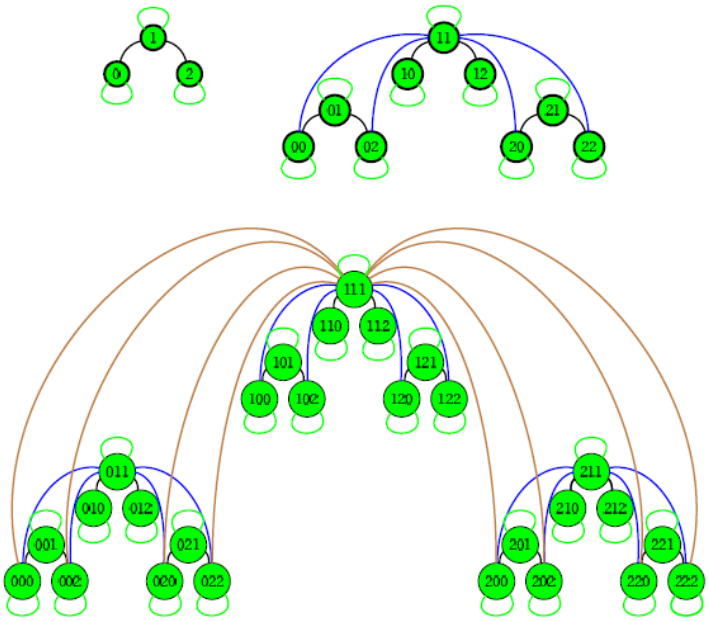

Figure 2.7: The first three elements of the "cherry" model: G_1, G_2 and G_3. Note that in this figure the graphs contain all of the loops additionally. The figure is from [17].

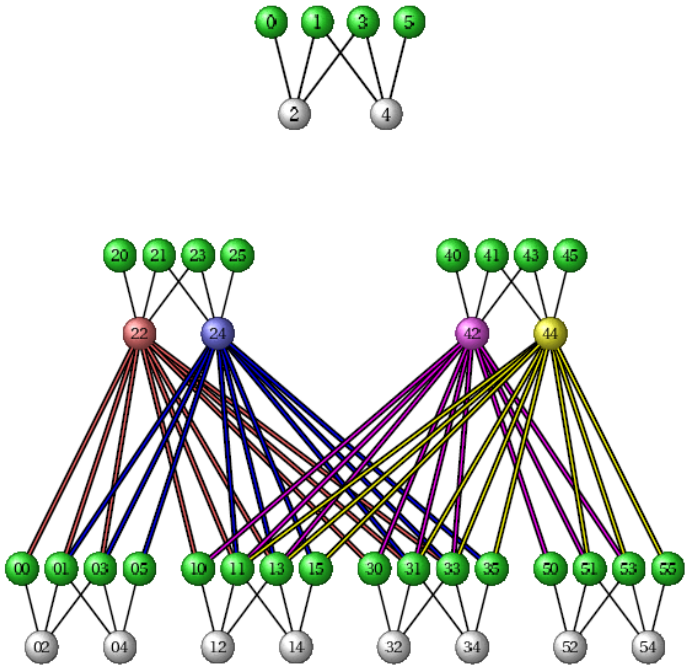

Figure 2.8: The first two elements of the "fan". Here $V_1 = \{2, 4\}$ and V_2={0,1,3,5}. The figure is from [17].

Chapter 3

Self-similarity and fractality of complex networks

In addition to the fascinating features of real networks presented in Section 1.2 we mention two more essential properties of some complex networks: self-similarity and fractality, introduced by Song, Havlin and Makse in[23]. An object is called self-similar, if it is exactly or approximately similar to a part of itself, i.e it looks the same on all length scales. Having regard to the fact that self-similarity is a typical property of fractals, to unravel the self-similar property of some networks we use methods that were introduced to investigate fractals.

3.1 Box-counting and renormalization

If one would like to unfold the self-similar property of real-world networks, box-counting method turns to be practical [23]. The method works as follows: We partition the vertices into boxes of size l_B. The maximal distance between nodes within a box is at most $l_b - 1$. The resulting number of boxes needed to tile the network is denoted by $N_B(l_b)$.

Definition 3.1. *The box dimension d_B is defined by*

$$\frac{N_B(l_B)}{|V(G)|} \approx l_B^{d_B} \tag{3.1}$$

in the sense that

$$d_B :\approx \frac{\log \frac{N_B(l_B)}{|V(G)|}}{\log l_B} \tag{3.2}$$

exists.

In [23] it was pointed out based on numerical evidence that such limit exists for a number of real networks including WWW and actor network.

After tiling the system with l_B-boxes we can apply a renormalization procedure to it. We replace each box by a single node. Two vertices of the renormalized graphs are connected if there is at least one link between the corresponding boxes in the original graph. The renormalization procedure is applied again and repeated until we get a single node. The procedure is visualized in Figure 3.1.

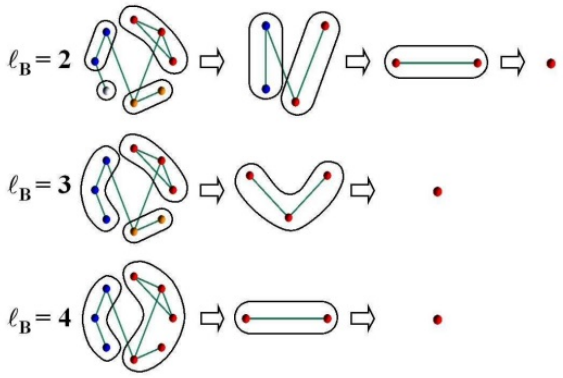

Figure 3.1: Demonstration of the box-counting and renormalization method for different l_B in a sample network. The figure is from [23].

It appears to be useful to investigate this process since the probability distribution of some real-world networks is invariant under the renormalization

$$P'(k) \propto k^{-\gamma}, \tag{3.3}$$

where $P'(k)$ is the probability that a node chosen randomly in the renormalized graph has degree k. (In [23] the authors considered the first three elements of the sequence of graphs obtained by successive renormalizations for $l_B = 3$ and get the result in 3.3 by numerical evidence.)

Now we can define the two properties appearing in the title of this chapter.

Definition 3.2. *The **self-similarity** of a network means that the degree distribution is invariant under the renormalization, i.e. a network is self-similar if satisfies 3.3 for an appropriate renormalization procedure.*

Definition 3.3. *The **fractality** of a network (also called fractal scaling or topological fractality) stands for the power-law relation between the minimum number of boxes needed to cover the entire network and the size of the boxes, i.e a network is fractal if the box dimension d_B exists, in the sense of 3.2.*

Regarding that this topic is very fresh and researched mainly empirically from a more practical point of view, the terminology is not always consequent in the papers, here we followed [13].

Topological fractality have been observed in some complex networks, such as the WWW, actor network or protein interaction networks [24]. In addition in [24] the authors find that the degree distribution of these networks is invariant under renormalization, thus they are also self-similar. On the other hand, some networks are self-similar, yet are not fractal. A typical example of such networks is the Internet and hierarchical graph sequence model (introduced in Section 2.5) also belongs here. Thus the self-similarity and the fractality do not always imply each other with respect to complex networks [13].

3.2 Other definitions for network dimension

Another method for calculation of fractal dimension on networks is the cluster-growing method [23]. The method works as follows: We pick a vertex v at random and consider the ℓ-neighourhood Γ_v^ℓ which is the set of vertices having distance not greater than ℓ from v, Γ_u^ℓ can be called a cluster. For the cluster $c = \Gamma_u^\ell$ let

$$M_c := |c| .$$

We write

$$< M_c >:= \mathbb{E}(M_c).$$

Definition 3.4. *The fractal cluster dimension d_f is defined by*

$$\mathbb{E}(M_c) \approx \ell^{d_f}. \tag{3.4}$$

in the sense that

$$d_f :\approx \frac{\log \mathbb{E}(M_c)}{\log \ell} \tag{3.5}$$

exists.

For a complex network of small world property $d_f = \infty$ holds. Since for sufficiently big ℓ and for an average v, we have $\left|\Gamma_v^\ell\right| \approx e^\ell$. Hence

$$\frac{\log < Mc >}{\log \ell} \approx \frac{\ell}{\log \ell} \to \infty.$$

On the other hand for a homogenous network characterized by a narrow degree distribution the cluster-growing and box-counting methods yield the

same exponent, i.e. $d_b = d_f$, since in this case every node typically has the same number of neighbours [23].

We can get an other interesting result after introducing the scaling transformation of the degree distribution. Referring to the box-counting method, for a given l_b we define the scaling factor

$$s(l_B) := \frac{\text{the degree of the greatest hub in the renormalized network}}{\text{the degree of the greatest hub in the unrenormalized network}}.$$

It was found in [23] that the scaling factor $s < 1$ scales with l_B with an exponent d_k:

$$s(l_B) \approx l_B^{d_k}. \tag{3.6}$$

The following assertion was made in [23]:

$$\gamma = 1 + \frac{d_B}{d_k}, \tag{3.7}$$

where γ is the power law exponent of the complex network. Other definitions for network dimension also appear in the literature (for instance in [22]). In Chapter 4 we also introduce a new concept of dimension that turned to be useful to investigate the hierarchical graph sequence model.

Chapter 4

Modified box dimension of the hierarchical graph sequence model

In this chapter we investigate the hierarchical graph sequence model introduced in Section 2.5. Using the box counting method (see Section 3.1) we prove the existence of the following limit.

Definition 4.1.

$$\widetilde{dim}\left(\{G_n\}_{n\in\mathbb{N}}\right) := \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log\frac{B_k^n}{|V(G_n)|}}{-l_k}, \tag{4.1}$$

where $l_k := diam(G_k) + 1$ and B_k^n denotes the minimal number of boxes of size l_k that we need to cover G_n.
This limit can serve as a certain box dimension of the hierarchical graph sequence model.

The introduction of this modified definition of box dimension was motivated by the fact that in the case of the hierarchical graph sequence model the original definition of box dimension (see Definition 3.1) is infinite. On the other hand our new concept of dimension does exist and is finite for this model as the next result shows.

Theorem 4.2. *For the hierarchical graph sequence model the modified box dimension:*

$$\widetilde{dim}\left(\{G_n\}_{n\in\mathbb{N}}\right) = \frac{\log|V(G)|}{2} \tag{4.2}$$

In the rest of this chapter we prove the theorem above. The proof consists of two main parts: In the first part we give an upper bound on B_k^n that gives a lower bound on the dimension. In the second part of the proof we use a similar procedure with the aim of constructing an upper bound on the dimension.

4.1 Proof of the lower bound

The following lemma is a useful tool to examine the modified box dimension of the graph sequence.

Lemma 4.3. *The diameter of the graph G_n is $2(n-1) + Diam(G)$.*

Proof of Lemma 4.3. (The proof follows the line of [17, Lemma 2.3.8].) For two arbitrary vertices $\underline{x}, \underline{y} \in \Sigma_n$ we denote the length of their common prefix by $k = k(\underline{x}, \underline{y}) := |\underline{x} \wedge \underline{y}|$. Furthermore, let us decompose the postfixes $\tilde{\underline{x}}, \tilde{\underline{y}}$ into blocks of digits of the same type:

$$\tilde{\underline{x}} = \underline{b}_1\underline{b}_2 \dots \underline{b}_r, \ \tilde{\underline{y}} = \underline{c}_1\underline{c}_2 \dots \underline{c}_q, \tag{4.3}$$

such that all of the blocks have a nonzero type and the consecutive blocks are of different types. That is, for $i = 1, \dots, r-1, \ j = 1, \dots q-1$ we have

$$\mathrm{typ}(b_i) \neq \mathrm{typ}(b_{i+1}) \in \{1, 2\}, \text{ and } \mathrm{typ}(c_j) \neq \mathrm{typ}(c_{j+1}) \in \{1, 2\}.$$

Note, that we denoted the number of blocks in $\tilde{\underline{x}}, \tilde{\underline{y}}$ by r and q, respectively. Recall that it follows from Definition 2.1 that for any path $Q(\underline{x}, \underline{y}) = (\underline{x} = \underline{q}^0, \dots, \underline{q}^\ell = \underline{y})$, the consecutive elements of the path only differ in their postfixes, which have different types. That is,

$$\forall i, \underline{q}^i = \underline{w}^i\underline{z}^i, \ \underline{q}^{i+1} = \underline{w}^i\tilde{\underline{z}}^i, \text{ with } \mathrm{typ}(\underline{z}^i) \neq \mathrm{typ}(\tilde{\underline{z}}^i) \in \{1, 2\}.$$

Now we fix an arbitrary self-map p of Σ_1 such that

$$(x, p(x)) \in E(G) \ \forall x \in G.$$

Most commonly, $p(p(x)) \neq x$. Note that x and $p(x)$ have different types since G is bipartite. For a word $\underline{z} = (z_1 \dots z_m)$ with $\mathrm{typ}(\underline{z}) \in \{1, 2\}$ we define $p(\underline{z}) := (p(z_1) \dots p(z_m))$. Then,

$$(\underline{t}\underline{z}, \underline{t}p(\underline{z})) \text{ is an edge in } G_{\ell+m}, \forall \underline{t} = (t_1 \dots t_\ell), \tag{4.4}$$

follows from Definition 2.1. Mind that in each step on the path, the number of blocks in 4.3 changes by at most one. Recall that $|\underline{x} \wedge \underline{y}| = k$, so $x_{k+1} \neq y_{k+1}$.

Since the digit on the $k+1$-th position changes on the path, we have to reach a point where all the digits to the right from the k-th position are of the same type. Starting from $\tilde{p}^0 = \underline{x}$, to reach the first vertex $\underline{a}$ of this property, we need at least $r-1$ steps on any path $\tilde{P}$, where r was defined in formula 4.3. Similarly, starting from $\underline{y}$, we need at least $q-1$ steps to reach the first vertex $\underline{b}$ where all the digits after the k-th position are of the same type. Because $x_{k+1} \neq y_{k+1}$, we need at least one more edge and at most $\mathrm{Diam}(G)$ edges.

Considering the worst case scenario, i.e. choosing $\underline{x}$ and $\underline{y}$ such that

- $|\underline{x} \wedge \underline{y}| = k$,
- $q = r = n$ and
- $d(x_1, y_1) = \mathrm{Diam}(G)$ (using Definition 2.1 it is not hard to see that this can be done), yields that

$$\mathrm{Length}(P(\underline{x}, \underline{y})) \geq n + n + \mathrm{Diam}(G) - 2 = 2(n-1) + \mathrm{Diam}(G) \qquad (4.5)$$

On the other hand, now we construct a path $P(\underline{x}, \underline{y})$ between two arbitrary vertices $\underline{x}$ and $\underline{y}$ that is no longer than $2(n-1) + \mathrm{Diam}(G)$. Starting from $\underline{x}$ the first half of the path $P(\underline{x}, \underline{y})$ is as follows:

$$\begin{aligned}
\hat{\underline{x}}^0 &= \underline{x} = (\underline{x} \wedge \underline{y})\underline{b}_1 \dots \underline{b}_{r-1}\underline{b}_r \\
\hat{\underline{x}}^1 &= (\underline{x} \wedge \underline{y})\underline{b}_1 \dots \underline{b}_{r-1} p(\underline{b}_r) \\
&\dots \\
\hat{\underline{x}}^{r-1} &= (\underline{x} \wedge \underline{y})\underline{b}_1 p(\underline{b}_2 \dots p(\underline{b}_{r-1} p(\underline{b}_r))),
\end{aligned}$$

Starting from $\underline{y}$ the first half of the path $P(\underline{x}, \underline{y})$ is as follows:

$$\begin{aligned}
\hat{\underline{y}}^0 &= \underline{y} = (\underline{x} \wedge \underline{y})\underline{c}_1\underline{c}_2 \dots \underline{c}_r \\
\hat{\underline{y}}^1 &= (\underline{x} \wedge \underline{y})\underline{c}_1 \dots \underline{c}_{r-1} p(\underline{c}_r) \\
&\dots \\
\hat{\underline{y}}^{q-1} &= (\underline{x} \wedge \underline{y})\underline{c}_1 p(\underline{c}_2 \dots p(\underline{c}_{r-1} p(\underline{c}_q))).
\end{aligned}$$

It follows from 4.4 that

$$\begin{aligned}
P_x &:= (\hat{\underline{x}}^0, \hat{\underline{x}}^1, \dots, \hat{\underline{x}}^{r-1}) \\
P_y &:= (\hat{\underline{y}}^{q-1}, \cdots \hat{\underline{y}}^1, \hat{\underline{y}}^0)
\end{aligned}$$

are two paths in G_n. To construct $P(\underline{x}, \underline{y})$ the only thing remained is to connect $\hat{\underline{x}}^{r-1}$ and $\hat{\underline{y}}^{q-1}$. Using 4.4 it is easy to see that this can be done with a path P_c of length at most $\mathrm{Diam}(G)$. In this way,

$$P(\underline{x}, \underline{y}) := P_x P_c P_y.$$

Clearly,

$$\text{Length}(P(\underline{x},\underline{y})) \leq r + q + \text{Diam}(G) - 2 \leq 2(n-1) + \text{Diam}(G)$$

This completes the proof. □

The following two lemmas give us an upper bound on B_k^n. For a better understanding these lemmas are visualized in Table 4.1: primarily we show that the first row can serve as an upper bound for the other rows, then we give an upper bound for the first row.

$N_B(l_b)$	G_1	G_2	G_3	G_4	G_5	...	G_n	...
⋮	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱
l_1	$B_1^1 = 1$	$B_1^2 \leq N$	$B_1^3 \leq N^2$	$B_1^4 \leq N^3$	$B_1^5 \leq N^4$	...	$B_1^n \leq N^{n-1}$	...
	↘	↘	↘	↘	↘		↘	↘
l_2	...	$B_2^2 = 1$	$B_2^3 \leq B_1^2$	$B_2^4 \leq B_1^3$	$B_2^5 \leq B_1^4$	...	$B_2^n \leq B_1^{n-1}$	...
		↘	↘	↘	↘		↘	↘
l_3	...	...	$B_3^3 = 1$	$B_3^4 \leq B_1^2$	$B_3^5 \leq B_1^3$	...	$B_3^n \leq B_1^{n-2}$	...
			↘	↘	↘		↘	↘
l_4	...	...	...	$B_4^4 = 1$	$B_4^5 \leq B_1^2$	...	$B_4^n \leq B_1^{n-3}$	...
⋮	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱
l_k	...	...	...	...	...	...	$B_k^n \leq B_1^{n-k+1}$	...
⋮	⋱	⋱	⋱	⋱	⋱	⋱	⋱	⋱

Table 4.1: $N_b(l_b)$: the number of l_b boxes that we need to cover the graph sequence

Lemma 4.4 (Upper bound on B_k^n)**.** *The following inequality holds for* $\forall n > k$ *:*

$$B_k^n \leq B_1^{n-k+1};$$

and if $n \leq k$ *holds then* $B_k^n = 1$*.*

Proof of Lemma 4.4. The second part of the lemma follows immediately from the fact that $k \geq n \Rightarrow l_k = \text{Diam}(G_k) + 1 \geq \text{Diam}(G_n) + 1$, i.e. all of the vertices of G_n can be covered by one box. For the first part we consider the meaning of the inequality in the lemma. It means that if we can cover G_{n-k+1} with some number of l_1-boxes, then the same number of l_k-boxes is sufficient to cover G_n with. The proof works with an "anti-projection" procedure, we "blow up" all of the vertex of G_{n-k+1} with a G_{k-1} in order to create G_n. We show that the same l_1- boxing that we have used in G_{n-k+1} is an appropriate

l_k-boxing for G_n (in the sense that G_{k-1} is substituted for each vertex) since the substitution increases the maximum distance within the box with $2(k-1)$. In other words, let $\underline{x} = (x_1, \ldots, x_{n-k+1})$ and $\underline{y} = (y_1, \ldots, y_{n-k+1})$ two arbitrary vertices in G_{n-k+1}, contained by the same l_1-box, i.e. the distance between $\underline{x}$ and $\underline{y}$ is not greater than $\mathrm{Diam}(G)$. If we blow them up, we get two sets of vertices: $X = \{(\breve{x}_1, \ldots, \breve{x}_n) | (\breve{x}_1, \ldots, \breve{x}_{n-k+1}) = \underline{x}\}$ and $Y = \{(\breve{y}_1, \ldots, \breve{y}_n) | (\breve{y}_1, \ldots, \breve{y}_{n-k+1}) = \underline{y}\}$. It is not hard to see that within X or Y the maximal distance is $\mathrm{Diam}(G_{k-1})$, since they are isomorphic to G_{k-1}. To end the proof, it remains to calculate the maximal distance between the element of X and Y. Let $\underline{\breve{x}} \in X$ and $\underline{\breve{y}} \in Y$ and we compute the distance between them. Considering the worst case scenario, namely that $\underline{\breve{x}} = \{(\breve{x}_1, \ldots, \breve{x}_n) | (\breve{x}_1, \ldots, \breve{x}_{n-k+1}) = \underline{x}$ and $\mathrm{typ}(\breve{x}_j) \neq \mathrm{typ}(\breve{x}_{j+1}), \forall j \geq n-k+1\}$ and similarly $\underline{\breve{y}} = \{(\breve{y}_1, \ldots, \breve{y}_n | (\breve{y}_1, \ldots, \breve{y}_{n-k+1}) = \underline{y}$ and $\mathrm{typ}(\breve{y}_j) \neq \mathrm{typ}(\breve{y}_{j+1}), \forall j \geq n-k+1\}$. Starting from $\underline{\breve{x}}$ it takes $k-1$ steps to reach the first vertex $\underline{a}$ where all the digits after the $n-k$-th position are of the same type. Similarly, starting from $\underline{\breve{y}}$ we need $k-1$ steps to reach the first vertex $\underline{b}$ of the same property. To connect $\underline{a}$ and $\underline{b}$ we need at most $\mathrm{Diam}(G)$, since their "preimages" $\underline{x}$ and $\underline{y}$ are in the same l_1-box in G_{n-k+1}. Thus the distance between $\underline{\breve{x}}$ and $\underline{\breve{y}}$ is not greater than $2(k-1)+\mathrm{Diam}(G) = \mathrm{Diam}(G_k)$. This is what we wanted to show. □

Lemma 4.5 (Upper bound on B_1^n)**.** *The following inequality holds for* $\forall i \geq 1$:

$$B_1^i \leq N^{i-1}.$$

Proof of Lemma 4.5. It is easy to see that we need one l_1-box to cover G_1. It follows from the hierarchical structure of G_i (Remark 2.2) that G_i contains N^{i-1} copies of G_1, since for every prefix $\underline{z} = (z_1 \ldots z_{i-1}) \in \Sigma_{i-1}$, consider the set $W_{\underline{z}}$ of vertices $(x_1 \ldots x_i)$ of G_i with $(x_1 \ldots x_{i-1}) = \underline{z}$. Then the induced subgraph on $W_{\underline{z}}$ is identical to G_1 and $|\Sigma_{i-1}| = N^{i-1}$. This implies that we can cover G_i with N^{i-1} l_1-boxes. □

Corollary 4.6. *Putting together Lemma 4.4 and Lemma 4.5 we get*

$$\forall n > k : B_k^n \leq B_1^{n-k+1} \leq N^{n-k}. \tag{4.6}$$

Proof of the lower bound in Theorem 4.2. Using Corollary 4.6 we get a lower bound on 4.1:

$$
\begin{aligned}
\widetilde{\dim}\left(\{G_n\}_{n\in\mathbb{N}}\right) &= \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log\frac{B_k^n}{|V(G_n)|}}{-l_k} = \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log\frac{|V(G_n)|}{B_k^n}}{l_k} \\
&= \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log|V(G_n)| - \log B_k^n}{\mathrm{Diam}(G_k)+1} \\
&= \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log N^n - \log B_k^n}{\mathrm{Diam}(G)+2k-1} \\
&\geq \lim_{k\to\infty}\lim_{n\to\infty}\frac{\log N^n - \log N^{n-k}}{\mathrm{Diam}(G)+2k-1} \\
&= \lim_{k\to\infty}\lim_{n\to\infty}\frac{n\log N - (n-k)\log N}{\mathrm{Diam}(G)+2k-1} \\
&= \lim_{k\to\infty}\frac{k\log N}{\mathrm{Diam}(G)+2k-1} = \lim_{k\to\infty}\frac{k\log N}{k\left(2+\frac{(\mathrm{Diam}(G)-1)}{k}\right)} \\
&= \frac{\log N}{2} = \frac{\log|V(G)|}{2}
\end{aligned}
$$

□

In the second row we use the definition of l_k (see 4.1) and the bound comes from Corollary 4.6.

4.2 Proof of the upper bound

Lemma 4.7 (Lower bound on B_1^n)**.** *The following inequality holds for* $\forall i \geq n_1 + 1$:

$$B_1^i \geq N^{i-n_1},$$

where $n_i := |V_i|$, $i = 1, 2$ *and we assume that* $n_1 \leq n_2$ *without loss of generality.*

Proof of Lemma 4.7. Notice that $\mathrm{Diam}(G) \leq 2n_1$ (since the connectivity of the base graph). It is enough to show that we can find N^{i-n_1} vertices in G_i ($\forall i \geq n_1 + 1$) such that the pairwise distances between the vertices are greater than $2n_1$ (hence greater than $\mathrm{Diam}(G)$) so all of these vertices must be in distinct boxes, i.e. we need at least N^{i-n_1} boxes to cover G_i.

First let us consider the $i = n_1 + 1$ case: we need $N^{n_1+1-n_1} = N$ witness vertices such that all the pairwise distances are large enough. The construction of these $\underline{z}_0^w, \ldots, \underline{z}_{N-1}^w$ vertices are the following: For every initial digit $x \in \{0, 1, \ldots, N-1\}$, we define an appropriate $\underline{z}_x^w$ witness as follows:

$\underline{z}_x^w \in \underline{Z}_x^w$, and the set $\underline{Z}_x^w$ consists of words $(z_1 \dots z_i)$ with $z_1 = x$ and $\text{typ}(z_j) \neq \text{typ}(z_{j+1})$, $j \in \{1, \dots, i-1\}$. Let $\underline{z}_x^w$ be an arbitrary element of $\underline{Z}_x^w$. (Note that choosing $\underline{z}_x^w$ is not unique, all of the vertices that satisfies the above condition are appropriate witnesses.)

Now we give a lower bound on the shortest path between $\underline{z}_{x_j}^w$ and $\underline{z}_{x_k}^w$, where $x_j, x_k \in \{0, 1, \dots, N-1\}, x_j \neq x_k$. Recall the method and notation that we have used in Lemma 4.3. Notice that $|\underline{z}_{x_j}^w \wedge \underline{z}_{x_k}^w| = 0$ and $q = r = i = n_1 + 1$, thus we need at least $r - 1 + q - 1 + 1 = 2n_1 + 1$ steps on any path between $\underline{z}_{x_j}^w$ and $\underline{z}_{x_k}^w$. Using $\text{Diam}(G) \leq 2n_1$ these witnesses must be in distinct $l_1$boxes, so we need at least N l_1-boxes to cover G_{n_1+1}.

We have proved the lemma for $i = n_1 + 1$, but this procedure works for arbitrary $i : i \geq n_1 + 1$. Let be $i = n_1 + 1 + j, j > 0$. The extensibility of the procedure follows from the hierarchical structure of G_n, i.e. G_{n_1+1+j} consists of N^j copies of G_{n_1+1}. In other words in G_{n_1+1+j} the corresponding witnesses are the following: For every $\underline{y}$ word of length j over alphabet $\{0, 1, \dots, N-1\}$ we construct words $\underline{y}\underline{z}_x^w$ with all $x \in \{0, 1, \dots, N-1\}$, where $\underline{y}\underline{z}_x^w$ means the concatenation of $\underline{y}$ and $\underline{z}_x^w$. Hence we have created $N \cdot N^j = N^{j+1} = N^{i-n_1}$ vertices whom pairwise distance is greater than $\text{Diam}(G)$ (it is easy to see in the same way as we did in $i = n_1 + 1$ case). Thus we need at least N^{i-n_1} boxes to cover G_i: $B_1^i \geq N^{i-n_1}$. □

Lemma 4.8 (Lower bound on B_k^n)**.** *Using the notation of the previous lemma, the following inequality holds if* $n - k \geq n_1$:

$$B_k^n \geq N^{n-k-n_1+1} = N^{n-k} \cdot C.$$

Proof of Lemma 4.8. We have seen in Lemma 4.7 that $B_1^i \geq N^{i-n_1} \forall i \geq n_1 + 1$. We have constructed N^{i-n_1} vertices in G_i whose pairwise distance is greater than $\text{Diam}(G)$. It is enough to show that we can find the same number of vertices (i.e. N^{i-n_1}) in $G_{j+i-1} \forall j \geq 1$, such that the pairwise distances between them are greater than $\text{Diam}(G_j)$, this implies $B_j^{j+i-1} \geq N^{i-n_1}$. We can find them with the method of "lifting up" or "anti-projection", meaning that we "lift up" the witnesses from G_i to G_{j+i-1}: we continue the code of a witness in a way that the type alternates (i.e. it is changed in every character), otherwise arbitrarily. To put in mathematically precise terms the "anti-projection" works as follows:

$$\underline{z}_x^w \in \Sigma_i \longmapsto \underline{\acute{z}}_x^w \in \Sigma_{i+j-1}; \underline{\acute{z}}_x^w = (\acute{z}_1, \acute{z}_2, \dots, \acute{z}_{i+j-1}),$$

where $(\acute{z}_1, \acute{z}_2, \dots, \acute{z}_i) := \underline{z}_x^w$ and $\text{typ}(\acute{z}_k) \neq \text{typ}(\acute{z}_{k+1}), k \in i, \dots, i+j-1$. Notice that due to having chosen $\underline{z}_x^w$ we can also write: $\text{typ}(\acute{z}_k) \neq \text{typ}(\acute{z}_{k+1}), k \in 1, \dots, i+j-1$.

Now we give a lower bound on the shortest path between $\underline{\acute{z}}^w_{x_j}$ and $\underline{\acute{z}}^w_{x_k}$, where $x_j, x_k \in \{0, 1, \ldots, N-1\}, x_j \neq x_k$ with a similar procedure that we have used in Lemma 4.3 and in Lemma 4.7. Notice that $|\underline{\acute{z}}^w_{x_j} \wedge \underline{\acute{z}}^w_{x_k}| = 0$ and $q = r = i + j - 1$, thus we need at least

$$\begin{aligned} r - 1 + q - 1 + 1 = 2i + 2j - 3 \geq 2n_1 + 2 + 2j - 3 = 2n_1 + 2(j-1) + 1 \\ \geq \mathrm{Diam}(G) + 2(j-1) + 1 \geq \mathrm{Diam}(G_j) + 1 \end{aligned}$$

steps on any path between $\underline{\acute{z}}^w_{x_j}$ and $\underline{\acute{z}}^w_{x_k}$. Hence these witnesses must be in distinct l_j boxes, so we need at least N^{i-n_1} l_j-boxes to cover G_{j+i-1}, i.e. substituting $n = j+i-1$ and $k = j$ yields that $B^n_k = B^{k+(n-k)}_k \geq N^{n-k+1-n_1} = N^{n-k} \cdot C$. □

Proof of the lower bound in Theorem 4.2. Using 4.8 we get an upper bound on 4.1 if $n - k \geq n_1$:

$$\begin{aligned} \widetilde{\dim}\left(\{G_n\}_{n\in\mathbb{N}}\right) &= \lim_{k\to\infty}\lim_{n\to\infty} \frac{\log \frac{B^n_k}{|V(G_n)|}}{-l_k} = \lim_{k\to\infty}\lim_{n\to\infty} \frac{\log N^n - \log B^n_k}{\mathrm{Diam}(G) + 2k - 1} \\ &\leq \lim_{k\to\infty}\lim_{n\to\infty} \frac{\log N^n - \log N^{n-k-n_1+1}}{\mathrm{Diam}(G) + 2k - 1} \\ &= \lim_{k\to\infty}\lim_{n\to\infty} \frac{n\log N - (n-k-n_1+1)\log N}{\mathrm{Diam}(G) + 2k - 1} \\ &= \lim_{k\to\infty} \frac{(k+n_1)\log N}{\mathrm{Diam}(G) + 2k - 1} = \lim_{k\to\infty} \frac{k(1+\frac{n_1}{k})\log N}{k\left(2 + \frac{(\mathrm{Diam}(G)-1)}{k}\right)} \\ &= \frac{\log N}{2} = \frac{\log |V(G)|}{2} \end{aligned}$$

In the first row we use the definition of l_k (Definition 4.1) and the bound comes from Lemma 4.8.

Combining this with Lemma 4.1 yields Theorem 4.2, hence:

$$\widetilde{\dim}\left(\{G_n\}_{n\in\mathbb{N}}\right) = \frac{\log |V(G)|}{2}.$$

□

Chapter 5

Summary

The aim of the thesis was to study network science from a mathematical point of view, especially with regard to the connection between networks and fractals. First of all, we overviewed some of the most important results and concepts of network theory. After examining the fascinating features of real-world networks empirically, we described them more precisely. Next, we presented some ways of mathematical modeling of complex networks, we introduced the preferential attachment model and two fractal based models, the Apollonian networks and the hierarchical graph sequence model. Motivated by the observation of the self-similar nature of a few real networks, we described some methods, that were adopted from the study of fractals with the aim of investigating networks. One of these methods is the box counting procedure and the corresponding concept of dimension. As an own new result (joint with J. Komjáthy), we introduced a modified definition of box dimension and gave a rigorous proof for its existence in the case of the hierarchical graph sequence model. The proof consisted of two main parts: we gave an upper and a lower bound on the dimension.

There are several relevant open questions and possible directions for further study. We propose the following problems for future research:

- An interesting direction is to modify the graph sequence model (see Section 2.5) in order to satisfy the property of fractality (see Definition 3.3).
- Another promising research topic is to study the multifractal analysis of some appropriately chosen network models.

Bibliography

[1] José S Andrade Jr, Hans J Herrmann, Roberto FS Andrade, and Luciano R da Silva. Apollonian networks: simultaneously scale-free, small world, euclidean, space filling, and with matching graphs. *Physical review letters*, 94(1):18702, 2005.

[2] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[3] Albert-László Barabási, Erzsébet Ravasz, and Tamás Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3):559–564, 2001.

[4] Vladimir Batagelj and Andrej Mrvar. Some analyses of erdos collaboration graph. *Social Networks*, 22(2):173–186, 2000.

[5] Béla Bollobás* and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[6] Béla Bollobás, Oliver Riordan, Joel Spencer, Gábor Tusnády, et al. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms*, 18(3):279–290, 2001.

[7] Stefan Bornholdt, Heinz Georg Schuster, and John Wiley. *Handbook of graphs and networks*, volume 2. Wiley Online Library, 2003.

[8] Reuven Cohen and Shlomo Havlin. Scale-free networks are ultrasmall. *Physical Review Letters*, 90(5):58701, 2003.

[9] ROBIN Dunbar. Youve got to have (150) friends. *The New York Times, The Opinion Pages*, 2010.

[10] Robin IM Dunbar. Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences*, 16(4):681–693, 1993.

[11] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 5:17–61, 1960.

[12] Casper Goffman. And what is your erdos number? *American Mathematical Monthly*, pages 791–791, 1969.

[13] K-I Goh, Giovanni Salvi, Byungnam Kahng, and Doochul Kim. Skeleton and fractal scaling in complex networks. *Physical review letters*, 96(1):18701, 2006.

[14] Jean-Loup Guillaume and Matthieu Latapy. A realistic model for complex networks. *arXiv preprint cond-mat/0307095*, 2003.

[15] Hawoong Jeong, Bálint Tombor, Réka Albert, Zoltan N Oltvai, and A-L Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.

[16] Frigyes Karinthy. Chain-links. *Everything is the Other Way*, 1929.

[17] Júlia Komjáthy and Károly Simon. Generating hierarchial scale-free graphs from fractals. *Chaos, Solitons & Fractals*, 44(8):651–666, 2011.

[18] Jeffrey C Lagarias, Colin L Mallows, and Allan R Wilks. Beyond the descartes circle theorem. *The American mathematical monthly*, 109(4):338–361, 2002.

[19] Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

[20] Mark EJ Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.

[21] Mark EJ Newman. Power laws, pareto distributions and zipf's law. *Contemporary physics*, 46(5):323–351, 2005.

[22] O Shanker. Defining dimension of a complex network. *Modern Physics Letters B*, 21(06):321–326, 2007.

[23] Chaoming Song, Shlomo Havlin, and Hernan A Makse. Self-similarity of complex networks. *Nature*, 433(7024):392–395, 2005.

[24] Chaoming Song, Shlomo Havlin, and Hernán A Makse. Origins of fractality in the growth of complex networks. *Nature Physics*, 2(4):275–281, 2006.

[25] GJ Toomer. Apollonius of perga. *Dictionary of scientific biography*, 1:66–80, 1970.

[26] Remco van der Hofstad. Random graphs and complex networks. *Lecture notes, available online at: http://www. win. tue. nl/ rhofstad/Notes-RGCN2011. pdf*, 2011.

[27] Marc Vidal, Michael E Cusick, and Albert-Laszlo Barabasi. Interactome networks and human disease. *Cell*, 144(6):986–998, 2011.

[28] Duncan Watts and S Strogatz. The small world problem. *Collective Dynamics of Small-World Networks*, 393:440–442, 1998.

[29] Duncan J Watts, Peter Sheridan Dodds, and Mark EJ Newman. Identity and search in social networks. *science*, 296(5571):1302–1305, 2002.

[30] Zhongzhi Zhang, Lili Rong, and Francesc Comellas. High-dimensional random apollonian networks. *Physica A: Statistical Mechanics and its Applications*, 364:610–618, 2006.

[31] Zhongzhi Zhang, Lili Rong, and Shuigeng Zhou. Evolving apollonian networks with small-world scale-free topologies. *Physical Review E*, 74(4):046105, 2006.