Snapshots of modern mathematics
from Oberwolfach

№ 10/2026

Do the rich get richer? Mathematical models for real-world networks

Bas Lodewijks[1]

Networks can be found all around us. From the friendships we form, both in person and online, to the websites we visit on the Internet, and the telephones we use to send messages and make phone calls. Understanding the behavior and structure of these networks can be a difficult task. Mathematicians use models of such complex real-world networks to provide novel insights from an analytical point of view. In this snapshot, we take a look at why mathematics can be useful in understanding networks and discuss interesting recent developments in the study of a popular class of models known as preferential attachment models, which try to explain why the rich get richer.

1 Introduction: real-world networks

A *network* is a set of interconnected objects. We call the objects in a network *nodes* and whenever two nodes are interconnected, we say that they are connected

[1] Bas Lodewijks is supported by the European Union's Horizon 2022 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101108569, "DynaNet".

by a *bond*. You can think of yourself and your friends as nodes in a network that are interconnected via friendship – or cities that are connected by train tracks, or airports between which planes fly. Think also of connections on social media, or the Internet itself, where two webpages are interconnected when one links to another. When we look around us, we find networks everywhere.

The networks around us influence our everyday lives. They determine how quickly you can reach another place, or how easily goods can be distributed from a factory to stores (via the network of roads, train tracks, ports, and airports). They also affect how easily you can find information on a particular topic on the Internet (via the network of webpages), and how fast a viral disease like the flu can spread through a population (via the network of people and who they interact with). Understanding the structure of these networks gives us insight into how they influence our lives. For example, it can show us how to speed up the distribution of goods, improve search engines on the Internet, and slow down the spread of contagious diseases.

2 Mathematical models for real-world networks: random graphs

But how can we understand the structure of a network? Sometimes, this is a relatively simple task. For example, we have a clear view of the road network in each country, and thus we have a good idea of how quickly we can get from A to B. More often than not, however, it is a difficult task to obtain such a complete picture of a network. For example, there are simply too many neurons in our brain to determine all of the connections between them. Also, social media companies often classify data they have of their social networks.

To counter these difficulties, we use a mathematical representation of a network: a *graph*. A graph is a set of points, also called *vertices*, which are connected by lines, also called *edges*. The vertices represent the nodes in a network, and the edges represent the bonds. We write a graph as $G = (V, E)$, where V and E are the vertex set and edge set, respectively. See Figure 1 for an example of a graph. A graph allows us to capture only the structural information of the network that we are interested in, discarding all other information. For example, when studying a road network, we can represent cities as vertices and the roads connecting them as edges, ignoring details such as distances, travel times, or the surrounding landscape.

We can now create a *random graph*. This is a graph that one constructs using some randomized algorithm. The reason to work with a random graph, rather than with a deterministically constructed graph, is to counter the difficulties we face when trying to understand the structure of real-world networks mentioned above. The randomness in the graph reflects the uncertainty and/or lack of

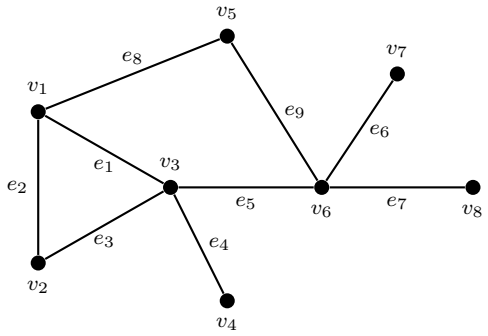

Figure 1: An example of a graph $G = (V, E)$ with vertex set $V = \{v_1, \ldots, v_8\}$ and edge set $E = \{e_1, \ldots, e_9\}$.

information that we face when trying to observe networks in the real world. We often do not know exactly which nodes in a network are connected by bonds, due to missing information or networks being simply too big to map completely. Random graphs mimic this, as we also do not know exactly which vertices in the graph will be connected by edges. Moreover, using a random graph allows us to say something about a large class of "typical" graphs, rather than a single specific one. This can be seen as a way to avoid over-fitting: we do not recreate one network exactly, but instead obtain a graph that looks very similar to a large number of networks. Knowledge of such random graphs can then be applied to a wide range of networks, instead of just one.

There is not a single way of creating a random graph; in fact, there are many different approaches to it. The aim is to do it in a way that reflects the properties of the networks we find in the real world. This may, at first, seem like an impossible task. As networks can be found in vastly different contexts, one might expect them to be structurally quite different. This turns out not to be the case, however. Despite coming from different contexts, they often have very similar properties. Two properties that can often be found in such networks are the following:

Scale-free property: The *degree* of a node is defined as the number of nodes it is connected to. In a scale-free network, the proportion of nodes with degree k decreases like $\frac{1}{k^\tau}$ as k becomes large, where $\tau > 1$. This means that scale-free networks contain many nodes that have few connections, but also a small amount of nodes that have a very large number of connections. One can think of very popular people in a social (media) network, or a popular webpage such as Google on the Internet.

In Figure 2, we see the proportion of nodes with degree k in a collaboration network of condensed matter physicists. In this network, the scientists form the nodes, and two nodes are connected by a bond if the corresponding scientists have co-authored a scientific paper. The figure shows a *log-log plot*, in which both axes have logarithmic scales. When depicted this way, the points almost form a straight line for large degree k, meaning that the proportion decays like a polynomial. One can also see that there is a very small amount of scientists that have a large amount of collaborators.

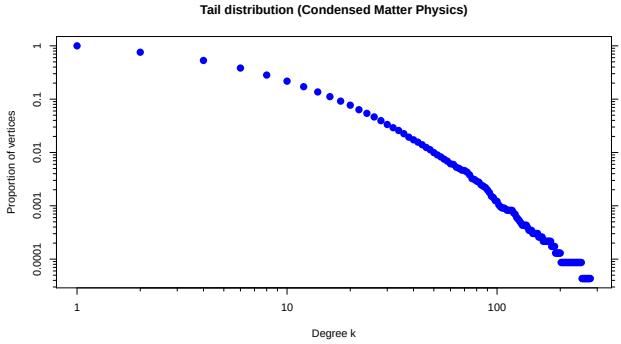

Figure 2: Log-log plot of the proportion of vertices that have a degree larger than a certain value k, based on a collaboration network of condensed matter physicists.

Small-world property: The *graph distance* between two nodes u and v in a network is the shortest length of any path of bonds from u to v. The small-world property states that, for a network with n nodes, the average graph distance is of order $\log n$, or even $\log \log n$. This means that even if the network is very large, the average distance between two nodes is still quite small. An example can be found in Figure 3, where we see the distribution of graph distances between pairs of nodes in a (very large) collaboration network of condensed matter physicists.

The small-world property was first empirically observed in an experiment conducted by the psychologist Stanley Milgram in 1967 [6]. He gave about 100 people in the USA a package and told them it needed to reach a certain person that lived far away. They were only allowed to send the package to someone they personally knew. Milgram was interested in how many intermediate people it would take before the packages would reach their desired destination. Though only 18 packages actually made it to the destination, it turned out that no path of recipients was longer than 11, and that the average number of intermediaries

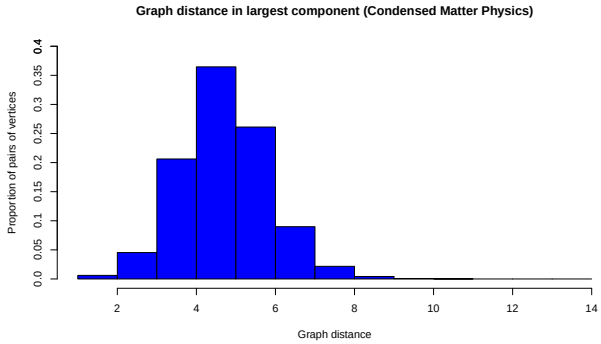

Figure 3: Histogram of the graph distance of pairs of nodes in a collaboration network of condensed matter physicists. From [4].

required was about 6. This later came to be known as the "six degrees of separation". Milgram later repeated his experiment on a slightly larger scale and found that the average was even lower this time, namely, somewhere between 4 and 5. An analysis carried out by Facebook on its social media network revealed an even smaller distance between 3 and 4 [7].

It turns out that many of such networks are very well connected, and this is exactly the reason why it can be so easy to find someone on social media, to find the right information online, or why viruses such as COVID-19 can spread so quickly throughout the population.

3 Preferential attachment graphs

Now that we have a rough idea of what real-world networks look like from a structural point of view, we can start to think of a mathematical model for such networks in terms of random graphs. For this, we shall use "preferential attachment random graph models" (PA models). These models were popularized in the late 1990s by the physicists Barabási and Albert, who used it as a way to explain the formation of the Internet [1]. They are based on the intuition that as a network grows and new nodes are added to it, the well-connected nodes in the network are most likely to connect to these new nodes.

We can heuristically verify this intuition in a variety of contexts. In a social context, someone who joins a social group is more likely to befriend the popular individuals than those who do not have many friends. In a professional environment, a new PhD student is more likely to work with a successful academic than with someone who has not supervised many PhD students. And

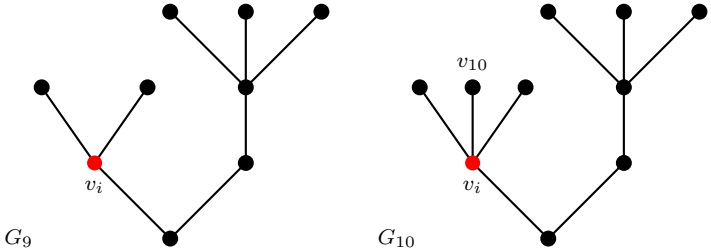

Figure 4: The graph G_{10} is created from the graph G_9 by selecting a vertex v_i according to the probability distribution in Equation 1 and connecting a new vertex v_{10} to the selected vertex.

a well-cited scientific paper is more likely to be cited by new papers than scientific papers that do not have many citations.

To obtain a *preferential attachment random graph*, we construct a sequence of graphs $(G_1, G_2, G_3, \dots)$ with $G_n = (V_n, E_n)$ in the following way: We start with the graph $G_1 = (\{v_1\}, \varnothing)$ that has a single vertex v_1 and no edges. Then, given G_n for any $n \geq 1$, we construct G_{n+1}. First, select a vertex v_i from the vertex set $V_n = \{v_1, \dots, v_n\}$ of G_n with probability

$$\frac{\deg_n(v_i)}{\deg_n(v_1) + \cdots + \deg_n(v_n)}. \tag{1}$$

Here, $\deg_n$ denotes the degree of a vertex in the graph G_n (recall that the degree of a vertex equals the number of vertices it is connected to in the graph). This means that the higher the degree of v_i in G_n, the more likely it is to be chosen. Then, define the vertex set of G_{n+1} as $V_{n+1} = V_n \cup \{v_{n+1}\}$ and the edge set of G_{n+1} as $E_{n+1} = E_n \cup \{(v_i, v_{n+1})\}$. In words, we add a vertex v_{n+1} and connect it to v_i by a new edge. A step of this construction is visualized in Figure 4.

In practice, the probability that a vertex connects to a new one might not be exactly proportional to its degree, as in Equation 1. To model a given network more realistically, we therefore adjust the probability by defining an *attachment function* f. This function assigns to each degree $\deg_n(v)$ a positive value $f(\deg_n(v))$. The probability that the vertex v is chosen from V_n then is

$$\frac{f(\deg_n(v))}{f(\deg_n(v_1)) + \cdots + f(\deg_n(v_n))}. \tag{2}$$

When the attachment function f is increasing, we see from Equation 2 that vertices with a larger degree have an even higher probability of being connected

to a new vertex than in Equation 1. This strong preference to connect to high-degree vertices explains the name of these models.

Albert and Barabási were the first to analyze PA models. Later, this approach was expanded to a variety of such models that incorporate many different features. These models are celebrated due to the fact that the randomized algorithm which is used for their construction produces graphs with both the scale-free and small-world properties. This contrasts with certain other types of random graph models, where such properties must be imposed by parameter choices. In the PA models, the intrinsic randomized dynamics of the model are sufficient. As such, the preferential attachment construction could be an explanation of how real-world networks form, and has hence attracted much attention in mathematics, physics, and network science.

The behavior observed here is that vertices with a high degree are more likely to make new connections and thus further increase their degree. This in turn makes it even more likely for them to make new connections. This feedback mechanism is often called the *rich-get-richer effect*. The "rich" vertices with a high degree are most likely to increase their degree even further and get even "richer". Moreover, as the graph grows over time, the "old" vertices (that were added to the graph at an early stage) have an advantage in obtaining a larger degree than "young" vertices (that were added to the graph much later). Such old vertices simply have had many more opportunities to increase their degree compared to young vertices. Combined with the rich-get-richer effect, old vertices are thus much more likely to have a large degree compared to young vertices. In fact, *persistence* of the largest degree can occur. Persistence means that there exists some number N and a vertex $v \in V_N = \{v_1, \ldots, v_N\}$ such that v has the largest degree in the graphs G_n for all steps $n \geq N$. In other words, from step N on, v keeps being the vertex with the largest degree of all vertices. We say that the maximum degree *persists at* v.

Whether persistence of the maximum degree occurs depends on the growth rate of the attachment function f. Though old vertices are more likely to have a large degree compared to young vertices, there are also many more young vertices than old vertices. If f does not grow "fast enough", one of these young vertices may get lucky and acquire an exceptionally large degree, even larger than the degrees of the old vertices. On the other hand, if f grows sufficiently fast, then the degrees of the old vertices are so large that no young vertex obtains an even larger degree, no matter how lucky they are.

4 Preferential attachment graphs with vertex death

As we have seen, preferential attachment graphs model growing networks and can reproduce properties observed in real-world networks. One shortcoming

is that the graph is only allowed to grow: at every step, we add vertices and edges to the graph. This, however, is not always realistic. For example, people can create social media accounts, causing a social media network to grow, but they can also delete their account and thus remove it from the network, causing it to decrease in size. In social networks, people may move to a different city and join a new social network, leaving their old one behind. And a network of supermarkets and suppliers can change when a store goes out of business or a new branch opens up. Hence, real-world networks are exposed to both growth and shrinking; something that is not reflected in the PA models.

Models that incorporates this shrinking are the *preferential attachment with vertex death models* (PAVD models), introduced by Deijfen [2]. In these models, we differentiate between "alive" and "dead" vertices. To do this, we specify for each graph G_n a set of *alive vertices* A_n, which is a subset of the set of all vertices V_n of G_n. We begin with the graph $G_1 = (\{v_1\}, \varnothing)$ consisting of one alive vertex $A_1 = \{v_1\}$. Given G_n, we choose an alive vertex a_i from $A_n = \{a_1, \ldots, a_k\}$. This time, the probability of being chosen should not only depend on an attachment function f (as in the PA models), but also on a *removal function* d. Precisely, the probability that a_i gets chosen is given by

$$\frac{f(\deg_n(a_i)) + d(\deg_n(a_i))}{f(\deg_n(a_1)) + d(\deg_n(a_1)) + \cdots + f(\deg_n(a_k)) + d(\deg_n(a_k))}. \tag{3}$$

Then, we kill a_i with probability

$$\frac{d(\deg_n(a_i))}{f(\deg_n(a_i)) + d(\deg_n(a_i))}, \tag{4}$$

that is, we remove it from the set of alive vertices A_n and consider it to be "deceased". If a_i is not killed, we add a vertex v_{n+1} to the graph and connect it to a_i. We continue this process either indefinitely, or until there are no alive vertices left. One step of this construction is visualized in Figure 5.

For example, consider an alive vertex a_i with a high degree, meaning that it is connected to a lot of other vertices. Such a vertex is more likely to attract new connections, so that the value of the attachment function should be high in this case. At the same time, because of its importance within the network, the vertex is less likely to disappear, meaning that the value of the removal function should be small. Accordingly, the probability that a_i is selected in Equation 3 is high, while the probability of a_i being killed in Equation 4 is quite low. On the other hand, if a_i has a high value on the removal function, it will still be chosen frequently, but has a high probability of being killed.

By introducing a "death function", we create a much larger family of models that is more likely to mimic the behavior of real-world networks. This is because, firstly, it incorporates the concept of removing vertices from the graph, and,

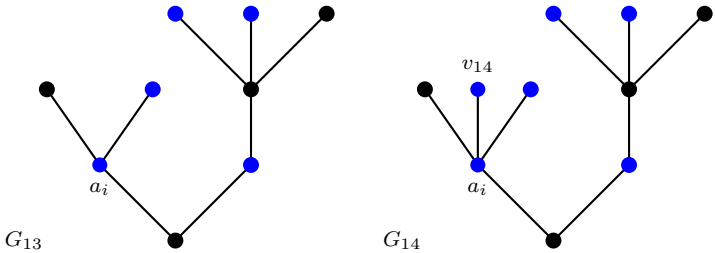

Figure 5: The graph G_{14} with set of alive vertices A_{14} is obtained from G_{13} and A_{13}. The black vertices are the "dead" ones, so A_{13} consists of the blue vertices on the left. First, we choose a blue vertex a_i in G_{13} according to the probability distribution in Equation 3. Then, with the probability in Equation 4, we remove the selected vertex from A_{13} to form A_{14} (that is, we color it black). Otherwise, we add a new alive vertex v_{14} to the graph and connect it to a_i. Here, we have done the latter.

secondly, we have a much larger family of models to choose from, so that it is more likely that we can use a model that fits well with data from real-world networks.

Since in PAVD models, vertices are removed, the concept of persistence of the maximum degree can be interpreted in various ways. Firstly, what is meant by "degree"? Does it mean the number of connections, not distinguishing between alive or dead vertices, or do we want the degree to only consider connections with alive vertices? In the definition provided, we interpret this to be the former, though one could find arguments for either case from a real-world perspective. Secondly, when we determine which vertex attains the largest degree, do we want to consider all vertices, or alive vertices only? Again, a case can be made for either, but for our purposes, we shall only consider alive vertices.

When we discussed persistence of the maximum degree for PA models, we considered whether there exists some vertex $v \in V_N$ that has the largest degree of all vertices in the graphs G_n with $n \geq N$. Now that we only take into account alive vertices when determining the maximum degree, such a vertex would now not only have to attain the largest degree, but also to remain alive the whole time. One can imagine that this may not occur and that, instead, every vertex that we add to the graph will also be removed and no longer considered alive after a finite, though random, number of steps. In such a case, we can no longer speak of persistence as we previously did and have to define a more general notion.

To this end, let O_n denote the oldest alive vertex in G_n and let I_n denote the alive vertex in G_n with the largest degree. We then say that persistence of the maximum degree occurs when I_n was introduced in the graph at around the same time as O_n. On the other hand, lack of persistence means that I_n is introduced to G_n at a much later step, compared to O_n. Despite this disadvantage, I_n has managed to obtain a larger degree than older vertices such as O_n.

One can imagine that the analysis of the PAVD models is more difficult compared to that of PA models. Indeed, there is not much known about persistence of the maximum degree in full generality. Recent developments can be found in [3, 5], which consider some particular classes of PAVD models.

5 Future outlook

PAVD models present a way of modeling real-world networks that evolve over time in a more realistic way, compared to the more classical PA models. The presence of the removal of nodes also complicates their analysis, however. Their behavior is not as evident, and understanding the full breadth of these models will require time and effort. Furthermore, there are many other properties of these graphs that are interesting to study, not just persistence of the maximum degree. With time, as we gain a deeper theoretical understanding of these models, it will also be possible to delve into using statistical analysis to start applying these models to real-world problems. This will enable us to really start using these kinds of models, applying them to datasets, fitting models to real-world networks, and using them to simulate the formation of networks.

Image credits

Figure 2 Bas Lodewijks, *Evolving random graphs in random environment*, PhD thesis, University of Bath, 2021.

Figure 3 Bas Lodewijks, *Evolving random graphs in random environment*, PhD thesis, University of Bath, 2021.

References

[1] A. L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science 286, no. 5439, 509–512, 1999.

[2] M. Deijfen, *Random networks with preferential growth and vertex death*, Journal of Applied Probability 47, no. 4, 1150–1163, 2010.

[3] M. Heydenreich and B. Lodewijks, *Preferential attachment trees with vertex death: Lack of persistence of the maximum degree*, arXiv preprint arXiv:2503.08675, 2025.

[4] B. Lodewijks, *Evolving random graphs in random environment*, PhD thesis, University of Bath, 2021.

[5] B. Lodewijks, *Preferential attachment trees with vertex death: Persistence of the maximum degree*, arXiv preprint arXiv:2505.06187, 2025.

[6] S. Milgram et al., *The small world problem*, Psychology Today 2, no. 1, 60–67, 1967.

[7] L. Backstrom et al., *Four degrees of separation*, Proceedings of the 4th Annual ACM Web Science Conference, 2012.

Bas Lodewijks *is a postdoctoral researcher in probability theory at the University of Augsburg.*

Mathematical subjects
Probability Theory and Statistics

Connections to other fields
Computer Science, Engineering and Technology, Humanities and Social Sciences, Physics

DOI
10.14760/SNAP-2026-010-EN

Snapshots of modern mathematics from Oberwolfach provide exciting insights into current mathematical research. They are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the interested public worldwide. All snapshots are published in cooperation with the IMAGINARY platform and can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

ISSN 2626-1995

Junior Editors
Elisabeth aus dem Siepen
and Victor Bloch
junior-editors@mfo.de

Senior Editor
Anja Randecker
senior-editor@mfo.de

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9–11
77709 Oberwolfach
Germany

Director
Gerhard Huisken

Mathematisches
Forschungsinstitut
Oberwolfach

IMAGINARY
open mathematics