e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

RANDOM GRAPH AND GROWTH MODELS

SELVARAJAN. M
Research Scholar
M.Phil Mathematics
Bharath Institute Of Higher Education And Research
Mail Id : mselvarajan77@gmail.com
Guide Name: **Dr. N. RAMYA**
Head of the Department, Department Of Mathematics
Bharath Institute Of Higher Education And Research

Address for Correspondence

SELVARAJAN. M
Research Scholar
M.Phil Mathematics
Bharath Institute Of Higher Education And Research
Mail Id : mselvarajan77@gmail.com
Guide Name: **Dr. N. RAMYA**
Head of the Department, Department Of Mathematics
Bharath Institute Of Higher Education And Research

Astract

Our ancestors around the turn of the 20th century put tremendous effort into formulating mathematics as an axiomatic system of thought. The art has since then offered to anyone who wished to cultivate it the luxury of laboratory conditions, free of the imperfections of the surrounding world. Many of the problems mathematicians have worked on since that time have of course been inspired by ”real-world” phenomena, but most results are formulated in a sterile environment, and anyone who wants to apply them rigorously must see to it that the necessary axioms are upheld.

Statistics is far less pure. True, there are beautiful statistical results in the literature, but it is by nature an applied branch of mathematics, and as soon as it is used for investigating a non-trivial problem, several questions arise. How to take an unbiased sample, for instance, can easily turn into a difficult problem, just as the question of what statistics to use for finding relevant patterns in the input data set. One needs only to think about, say, sociological studies to see how these seemingly minor questions influence the

conclusions one might draw through ”scientific” means, and might thus contribute to misunderstandings about the world we live in.

Introduction

That arises most likely from the general perception that statistics is a toolkit for proving

disproving certain statements through numbers. In other words, its main function is pattern matching. That is of course correct if one considers statistical work to be only parameter estimation and hypotheses testing, as it is taught in most schools. But to the more motivated student, it is an art of pattern finding, rather than matching, and one can apply these methods in areas that go far beyond the realm of mathematics. That is very broad, but that is exactly the point: we learn to learn how to find some order in chaos, while relying as little on our (usually unreliable) perceptions as possible.

Some of the hidden patterns are easily guessed if there is some information about the data. For example, financial stock information should likely be affected by company sector and size. Behavioral patterns should relate to an individual’s family background, their ability to speak foreign languages, and so on. But what other, non-trivial hidden patterns are there that might cluster the data? How should we go about finding them? How should we treat them? Several questions, without a universal answer. A particular set of tools may answer a particular mathematical problem, but in practice it is often very difficult to understand the meaning of patterns, when one gets to examine new data.

This study itself concerns investigating random graphs. Our ultimate aim would be to set up a graph lab where all possible random graphs in the world are classified according to some properties deemed generally important. But that is the work of a lifetime, and goes far beyond the scope of this dissertation. We will only list a number of properties that seem relevant for most graphs, and then proceed to a careful investigation of a relatively narrow area.

The amount of previous statistical work on graphs pales in comparison to graph theoretical work, but it does exist. Aldous has been a pioneer in the field (see for instance [1], or more on his web site), with recent results on noisy graphs by Bolla ([5]), on parameter estimation by Bickel and Chen ([4]), on the spectral properties of graphs generated by the threshold model by Ide, Konno and Obata ([35]), and others.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

The structure of the dissertation is as follows. In the next chapter we will introduce some earlier results on plain Erdős-Rényi and power-law graphs, then proceed to investigate methods that are useful for the statistical analysis of graphs. Then from chapter 4 onward we will focus on the following problem (indeed, our main problem): it has been established that if a graph is built up via some so-called ”preferential attachment” mechanism, thenit will have a power-law degree distribution. The reverse implication is generally taken for granted, but it is not so clear. In fact, it is so not clear that we will show in chapters 4 and 5 that it fails completely. In chapter 5 we will introduce some random graph models that produce power-law distributions, but this work will really be crowned by chapter 6, where we build a model that creates uniformly distributed graphs given a degree sequence.

[S] Random graph theory started in 1959 with a series of papers by P. Erd˝os and A. Rényi ([30], [29], [31]). They introduced two models with very interesting properties, but the graphs generated by these models were inadequate for describing some networksobserved in real life. Networks became an area of intense research in the 1990s. Watts and Strogatz introduced a so-called ”small world” model in 1998 ([57]), which aimed to modify the Erdős-Rényi models to better describe existing networks. In 1999, Albert and Barabási proposed a mechanism for building graphs that closely resembled real networks ([2]), notably in the power-law degree distribution. The publication was soon followed by another in similar vein, this time by Albert, Barabási and Jeong ([3]). The term ”scale-free graph” was invented to describe graphs with such a distribution.

In this chapter we will introduce some of the major results in the field of random graph theory, with special attention to Erdős-Rényi (”ER”) and Albert-Barabási (”AB”) model graphs.

Degree distribution

In $G_{n,p}$, the edges leaving a given node v are generated by $n-1$ flips of a p-probabilitycoin, therefore:

$$P(d(v) = k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

For $n \to \infty$ and $np \to \lambda > 0$ constant, the degree distribution is Poisson:

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

$$P(d(v) = k) = \frac{(np)^k e^{-np}}{k!}.$$

Diameter

The diameter of a graph is defined as $diam(G) = max\{d(v, w)|v, w \in V(G)\}$, where $d(v, w)$ is the length of the shortest path between vertices v and w. If G is not connected, $diam(G) = \infty$ by definition.

The diameter is a monotone function on graphs, i.e. if G and H are graphs with $V(G) = V(H)$ and $E(G) \subset E(H)$, then $diam(G) \geq diam(H)$.

Bollobás shows in [7] that for a sufficiently dense graph $G \in G_{n,p}$

$$diam(G) = \frac{\log n + \log\log n + \log 2 + O(1/\log n)}{\log pn}$$

almost surely. Specifically, for $c = \sqrt{d\,(d-1)\exp(p\,n}$ and $\frac{pn}{(\log n)^3} \to \infty$:

2

$$\lim_{n\to\infty} P(diam(G) = d) = e^{-\frac{c}{2}},$$

and

$$c \lim_{n\to\infty} P(diam(G) = d + 1) = 1\ e^{2}.$$

Chromatic number

A proper vertex-coloring, or, for short, ”coloring” of a graph G is a coloring on the graph’s vertices such that any pair of adjacent vertices have a different color. The chromatic number of G is defined as the smallest number of colors required for coloring G, and is usually denoted as $\chi(G)$.

Regarding $G_{n,p}$, we must distinguish between dense and sparse graphs. There is a

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

stronger estimate available for dense graphs, but it fails for low edge densities.

First, the dense case. Suppose $G \in G_{n,p}$, $p > n^{-\lambda}\ \forall \lambda > 0$ and let $b = \frac{1}{1-p}$. McDiarmid's 1989 result ([47]) shows that in this case

$$\chi(G) = \frac{n}{2\log_b n - 2\log_b \log_b n + O_C(1)}.$$

This result was an improvement upon earlier results of Shamir and Spencer ([53] and Bollobás ([8]).

In case p tends to zero too fast, the considerations that led the above result are not applicable. Using Frieze's argument ([32]), Łuczak was able to provide bounds in the sparse case in his 1991 paper ([44]).

$\forall G \in G_{n,p}, \exists C_0$ such that $\forall p = p(n)$ with p enclosed in the range $\frac{C0}{n}$

$\leq p \leq \log^{-7} n,$

then asymptotically almost surely

$$\frac{np}{2\log np - 2\log\log np + 1} \leq \chi(G) \leq \frac{np}{2\log np - 40\log\log np}.$$

Note that for higher values of p, the previous estimate gives stronger bounds.

Subgraphs

The famous 1960 paper by Erdős and Rényi ([30]) discusses the subgraph containment problem at length. They provide threshold conditions for a random graph for containing certain small subgraphs. Namely, they show that:

If $k \geq 2$ and $k - 1 \leq l \leq \binom{k}{2}$ are positive integers, and $\beta_{k,l}$ is an arbitrary non-empty class of connected graphs with k vertices and l edges, then $\forall G \in G_{n,m}$ (or $G \in G_{n,p}$ with $|E(G)| = m$):

$$P(\exists H \in \beta_{k,l} : H \subset G) = \begin{cases} 0, & \text{for } m = o(n^{2-\frac{k}{k}}) \\ 1, & \text{otherwise} \end{cases}$$

up to isomorphism over $\beta_{k,l}$.

Simple substitution yields that

- the threshold that G contains a tree of order k is $n^{\frac{k-2}{k-1}}$,
- the threshold that G contains a connected subgraph H with $|V(H)| = |E(H)| = k$ for arbitrary $k \geq 3$ is n,
- the threshold that G contains a cycle of order k for any $k \geq 3$ is n,
- the threshold that G contains a clique of order k for any $k \geq 3$ is $n^{\frac{2k-4}{k-1}}$.

The Albert-Barabási model

A quote from the 1960 Erdős-Rényi paper ([30]): 'It seems plausible that by considering the random growth of more complicated structures (e.g. structures consisting of different sortsof "points" and connections of different types) one could obtain fairly reasonable models of more complex real growth processes (e.g. the growth of a complex communication net

consisting of different types of connections, and even of organic structures of living matter, etc.).'

It seems the authors understood that the random graph models they worked with had their limitations as far as practical applications were concerned. Over the next decades, "real-life" networks were occasionally studied in detail, but it took until the 1990s that research sped up dramatically.

It became clear in many of these cases that the Erdős-Rényi model cannot provide an adequate description. The reason is instantly obvious: the degrees in the ER models have identical expected values, with a very narrow degree distribution. Many real-life networks, however, seem to carry a few degree-rich nodes and many degree-poor ones. So it seemed natural to investigate broader distributions, such as

$$P(deg(v) = k) \sim k^{-\lambda} \text{ for some } \lambda.$$

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

Albert, Barabási and others found that this so-called ”power-law” distribution held well across a host of different real-life networks, for example:

- Collaboration graph for movie actors. Actors are represented by the nodes, and two nodes are connected by an edge if the corresponding actors have collaborated in a film. In this graph, $\lambda_{actor} = 2.3 \pm 0.1$.
- The World Wide Web. Each vertex represents a document, with a (directed) edge runs between them if one document has a link to the other. They found that $\lambda_{WWW} = 2.9 \pm 0.1$.
- The electrical power grid of the western United States. The vertices are generators, transformers, substations and the edges are the high-voltage lines between them.$\lambda_{power} \approx 4$ fit well.
- Scientific citations. The vertices are publications, the (directed) edges are references between them. It was found that $\lambda_{citation} = 3$.
- The number of sexual partners on a Swedish sample by Liljeros et al. ([43]). The study settled for $\lambda_{men} \approx 1.6$ and $\lambda_{women} \approx 2.1$. Their data suggests (again!) that men have more partners than women. And here we are, writing about graphs whoseedges are supposed to have two end vertices...

All the above results are cited from Albert and Barabási ([2]) or a reference therein, unless otherwise noted.

Albert and Barabási pointed at two observations about such networks. First, that their vertex set generally grows over time, so any model that aims to describe them should consider the mechanism of how a new vertex is introduced to the system. And second, that a newly added vertex is more likely to connect to vertices that are already degree-rich. These two considerations motivated them to develop a model that leads naturally toa power-law distribution.

We will now describe the original Albert-Barabási model, and reproduce the authors' heuristic computations. In the next chapters we will provide a paraphrased version and analysis.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Merits and dangers - politics

The popularity of the model inspired a lot of research, most of it outside mathematics.It is perhaps not impolite to say that an entire industry has been built up on scale-free networks. Numerous scale-free models have been introduced since Albert and Barabási's first paper on the subject, not the least by these same authors. These models are slight modifications to the original, to explain phenomena not contained in earlier versions, like how a late-comer might be able to attract neighbors at an exceptionally high rate, and so on.

The existence of such broadly applicable mathematical theory provides strong guidelines to researchers across many fields, and thus inspires work. This is welcome. True, much of what is understood by those who apply it is just a formal version of common sense.

One does not need highly developed theories to understand that the spread of signals in relatively centralized real-life networks depends largely on the central nodes, which means the Internet's backbone hubs need to be strongly protected, information / diseases spreads faster on them, etc. But having such models can be helpful in quantifying properties of such networks.

The process is not without dangers, however. The models tend to work with the assumption that scale-free distributions express something about a hidden law of nature. In the quest of understanding and describing the world, one is often tempted to turn to simplifications, and the existence of such universal assumptions might lead one to explain phenomena within the scale-free framework even where it is inappropriate. E.F. Keller ([39]) and others find this worrying. It is hard to argue against some of the criticism: forcing a scale-free uniform on too wide a range of networks should not become a world religion. But even if some think these models are weak or irrelevant, they do describe the growth of certain networks quite well. We will show that the model has well-defined connection properties, and since says something meaningful about certain real networks,its existence is justified.

Thus, it inspires the student to hammer out the model's weaknesses, rather than to discard it altogether.

Let us now look at some of the basic properties of scale-free graphs.

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

Research Paper

Connectedness

Denote the number of components of a graph G by $C(G)$. In the AB model, the addition of a new vertex does not increase the number of components, $C(G_0) \geq C(G_t)$, $\forall t \in \{1, 2, \dots\}$. Thus, if G_0 is connected, so will be G_t.

Suppose $C(G_0) = k$. For $m = 1$, $\forall t \in \{1, 2, \dots\}$, $C(G_t) = k$, since each new vertex connects to exactly one component. For $m > 1$: let $\{H_i\}$ denote the components with $\sum_{v \in H_1} d_v \geq \cdots \geq \sum_{v \in H_k} d_v \geq 1$. Let $A_j(t)$ be the event that H_1 is connected to H_j at time t, and A_j the event that they are ever connected ($j = 1, \dots, k$). These two components would be connected only if a vertex connected to them both, thus

$$P(A_j(t)) \geq \frac{1}{j}\left(\frac{mt}{j} \cdot \frac{1}{ct}\right)^{m-1} = \frac{k}{j}(2k)^{-m}.$$

$$P(A_j) \leq \prod_{t=t_0}^{\infty} \frac{ct+a}{ct} \cdot \frac{ct-1}{ct+a} \leq \prod_{t=t_0}^{\infty} \left(\frac{ct-1+a}{ct+a}\right)^{1/c} \to 0.$$

Denote $c = \frac{m}{2k}$. Then

That holds for all $j = 2, \dots, k$, thus if the graph is large enough, it will be connected.Indeed, an AB graph is disconnected only if $m = 1$ and $C(G_0) > 1$.

$$\frac{2mt}{k} \qquad \frac{2mt}{t}$$

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Degree distribution

Albert and Barabási's estimate ($P\,(deg(v) = k) \sim \lambda^{-3}$. Dorogovtsev, Mendes and Samukhin show in [24] that for an AB graph with m edges added in each step:

$$P\,(d_V = k) = \frac{2m(m+1)}{k(k+1)(k+2)}.$$

Bollobás, Riordan, Spencer and Tusnády have a similar, but stronger result for $0 \geq k \geq n^{1/15}$ in [12].

The popular notation for the above is

$$P\,(deg(v) = k) \sim k^{-3},$$

for a randomly selected vertex.

Methods, considerations

[T] In this chapter we will give an overview of some of the considerations for evaluating graphs. We cannot afford to go into great detail about all of them – some of what is written here can simply provide the basis or motivation for future work.

We will use some of these methods in the later, more numerical part of the dissertation, here we wish to focus more on the philosophy. We believe the learning process is like chasing butterflies: butterflies are fast and elusive. Even if we manage to catch one, we might kill it, and what is the value of a dead butterfly? Still, we learn from the process, and that might be more important than finding answers. Over time, the answers will start coming...

The list we provide below is of course incomplete, but we hope the reader might find some of it not only helpful, but also inspiring for further work.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Models, likelihood, Neyman-Pearson

Statistics education tends to focus on two areas: parameter estimation and hypothesis test- ing. What is often not mentioned, because it is implicit in the former, is model estimation. Statistical data is always assumed to be generated by some model, and the parameters are basically parameters for that model. True, for simple models this distinction may not be necessary, but it would help students learn to be more critical with data.

So once we have settled for a model, we can look at the graph realization and try to derive its likelihood as per the model. This can be a very difficult task, if the model is complex (and graphs are reasonably complex), but there are great rewards if the attempt bears fruit. First, different realizations are comparable, and second, by the Neyman-Pearson lemma, different models are comparable. In practice, this means that we can make perturbations on the original data, and if the likelihood function changes significantly asa result of the perturbation, we have found something through a test of maximum power.

While this is not always achievable, it is a nice ideal to follow.

Degree distribution

The degree distribution is a property typical of every graph. There are numerous ways of comparing the empirical distribution and theoretical distributions. In practice these methods often aim at transforming the data in a way that a line should fit it well. For instance, the following approach is apparent in the original Albert-Barabási paper ([2]):

Suppose $P(d = k) \sim k^{-\lambda}$. Then

$$\log P(d = k) = -\lambda \log k + b,$$

so $\log P(d = k)$ should form a line with slope $-\lambda$ as a function of $\log k$.

In practice, many of the real-life graphs cited as scale-free show this linear relationship only in mid-range, and fail to fit near the head and tail of the distribution. More seriously, while the expected degree distribution associated with a random graph model might be known, the probability distribution of individual vertex degrees might not be identical, so as much as a nice picture might lead us to think we found a fitting, we should remain skeptical. So we are tempted to find some alternative methods for fitting distributions.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Labeling

Some graphs, like those from an AB model, evolve in time, and therefore have a natural labeling. These models may assign edge probabilities conditionally independently in every step, and thus reconstructing the original labeling would mean we can attain an accurate likelihood value for the graph. This motivates us to examine the problem of labeling.

The most natural idea, to consider all possible labelings and pick one according to some constraints, say, maximum likelihood, are computationally unrealistic. So it is clear that for a general degree distribution, we are helpless about this problem.

So let us consider scale-free graphs, and say our graph $G(n, m)$ is from the AB model. The general tendency in these graphs is that the earlier a vertex is connected, the higher its degree. Móri shows in [49] that even as $n \to \infty$, the maximum degree node's identity will change only a finite number of times with probability 1. Naturally, this must hold true for the subgraph generated by vertices $\{i, \ldots, n\}$, regardless of the choice of i. This implies that for any fixed i and an arbitrary $j > i$, only $O(1)$ of these pairings will be such that $d_i < d_j$, out of $O(n)$ choices for j as $n \to \infty$. Thus, labeling the vertices in decreasing order by their degree is a negligible deviation from the original labeling.

We will use this as a general method: anytime we are to examine a scale-free graph whose nascent vertex order is not known, we will start with sorting the vertices by their degree. Note that this method is not applicable in general, but it does work for scale-free graphs.

3.1 Preferential attachment

[S] The term preferential attachment refers to a graph construction mechanism where every time a new vertex is joined to the graph, its neighbors are selected by some kind of preference function. Namely, if we wish to join vertex v to graph G to form G^{J}, we will have

$$\forall w \in V(G) : P((w, v) \in E(G^{J})) \sim f(deg(w)).$$

This is a paraphrased form of the following urn process: let urn k (A_k) represent the degree of v_k. In step t of the process, some balls are picked from urns $1, \ldots, t-1$ with

probabilities proportional to $f(A_1), \ldots, f(A_{t-1})$, and exactly as many balls are added to urn t.

Krapivsky, Redner and Leyvraz showed in [41] that for

- $f(k) = k^a$, with $a < 1$: $P(deg(v) = k) \approx ck^{-a} \exp(-ck^{1-a})$,
- $f(k) = k^a$, with $a > 1$: the resulting graph will have one vertex with degree of order t, others $O(1)$,
- $f(k) = a + k$, with $a > -1$: $P(deg(v) = k) \sim ck^{(-3+a)}$.

It is a task in itself to tell if the model that built a graph included preference considerations, so estimating f might be very difficult. But without a proper estimate for f , we might not be able to have an accurate likelihood estimate. For likelihood computations, we will assume the simplest of preferential attachment models, namely $f(k) = k$, and show that it is not too great a sacrifice. But for this reason we will be using the term ”pref- erential attachment” loosely, pointing to a sequential model’s property that higher degree vertices have greater strength in attracting neighbors, and not require f to be linear.

Clustering

[T] Clustering on a graph is, in the broadest sense, a partition on the vertex set. The term ”clustering” is typically used for a partitioning where vertex groups are formed froma set of vertices that are in some sense similar. It is important to note that this is onlyone of the possible ways to approach the subject of clustering.

One can also define partitioning in terms of cross-cluster behavior. The criterion for a vertex to belong to a cluster could be its distance to vertices in other clusters (according to some distance metric), or by some other regularity property.

To illustrate the difference, consider a graph with some distance metric defined over it, say, a real-world map. Clustering by geographic location is of the first type, and by thevertex

degree is of the second type.

Let us briefly discuss these two approaches.

Clusters of ”similar” vertices

This approach has very rich literature, so there is no need to go into technical details. The only consideration we can add is how to define a distance metric that is relevant for our graph. Hamming distance is one metric that comes to mind automatically, another one isthe one implied by the process of drawing up the graph, where a metric space is set up automatically. The list could go on.

If we do not want to make a conscious decision on the metric space, we might want to set up some constrained process, and let the vertices find their own place, hoping there is a stationary state. For instance, we could randomly distribute the vertices along a spherical surface, and view edges as gravitational forces that pull the end vertices together. The surface is automatically metric, so once the stationary state is reached, clustering becomesa trivial problem.

There is no immediate summary of what this might do to a particular graph, it depends on the surface and the dynamic process, but once we have applied the method to a wide range of graphs, some of their properties might just stand out. And over time, one will develop an intuition for the features that seem relevant.

Microstructures

”The devil is in the details” - the saying goes. One approach to understanding a graphis through its small subgraphs, which are not always diabolic. These are easy to grasp both analytically and computationally. Borgs, Chayes, Lovász, T. Sós and Vesztergombi have two elegant papers about the subject ([14], [13]), where they introduce a notion of convergent graph sequences based on small subgraphs.

If we are able to compute some asymptotic subgraph properties of a hypothesized graph model, we might use this as part of a test on the graph we are examining. This so-called ”containment problem” is often an easy task, since it involves small subgraphs. Section 2.1.6 has a demonstration of some basic results for Erdős-Rényi graphs.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Distances, metrics

Given two general graphs, we might want to know how ”similar” they are. This becomesa very difficult problem if these graphs are not labeled. There are some fast algorithms finding isomorphism between trees, and that means one can come up with a tree-to-tree distance metric with some effort, but the same does not hold for general graphs.

The graph isomorphism problem is known to belong to NP, but it is currently unknown whether it is NP-complete or not. The subgraph isomorphism problem is in fact NP-complete ([56]). That means coming up with a practical graph distance algorithm is by no means a trivial matter.

Non-trivial does not mean unsolvable: Frieze and Kannan introduced a cut or rectangle distance in [33]. This distance satisfies all conditions for being a metric. The graph convergence idea developed by the Borgs, Chayes, Lovász, T. Sós and Vesztergombi articles mentioned in the previous section lay the groundwork for metricization. These ideas are very complex computationally, but the idea of a metric appears there. That implies a small rank graph that can be partitioned into a low number of clusters can be easily metricized in practice.

We have now introduced some general methods for examining graphs. Let us proceedto our main problem.

Conclusion

In the process of coming up with these results, we stumbled upon some open questions, which might be interesting subjects for further research. These are:

- Edge swap mixing times.

  P.L. Erdos, I. Miklós and L. Soukup proved in [26] that graph $G(n)$ approaches uniform distribution on its degree sequence class under the swap operation in $P\ oly(n)$ time, if it is bipartite semi-regular. We think the statement is expandable, quite possible to all non-directed graphs, but currently this is not proven.

- Uniform generation.

  There are many publications on how to generate graphs with a specific degree sequence with uniform distribution. To the best of our knowledge, this all use some

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

branching process, or edge mixing (e.g. swaps). Now, given a graphic sequence $\{d_1, \ldots, d_n\}$, model V will generate all graphs with this expected degree sequence with identical probability. This implies that it might be possible to use this model for generatic graphs with a prescribed degree sequence from the uniform distribution

from that degree sequence class. This might involve some iterative process that can remove as well as add edges. Currently we do not see the distribution of the individual degrees generated by the model, and do not have a solution for this problem, but we find it inspiring.

- The boundaries of the convex hull.

  We have stated that the maximum likelihood equation is solvable on the interior of the convex hull identified in section 6.1. What happens on the boundary?

- Perturbation vs. edge swaps.

  For models I through IV/IVa, the graphs' likelihood changed more under urn matrix perturbations than edge swaps. This implies that given a graph, there is a limit to how much its preference structure can be altered by rewiring the edges, and suggests that there is a lower bound on how much information the degree sequence containsabout the graph's connection structure. The question requires further study.

- Urn likelihood and swaps.

  Given the fact that swaps on graph G had a notably stronger effects on the value $S(G)$ when G was directed, one would naturally expect to see the same with the urn likelihood. But the results begged to differ. This is counterintuitive, and should be revisited.

- Applications.

  So-called "heavy-tail" or "fat-tail" distributions have gained much popularity in recent years, in areas as far apart as finance and biology. In some of these cases, thereis an underlying network model, but far from always. History shows that graph the- ory

was able to produce some beautiful solutions for problems that had nothing to do with graphs. It is a natural idea to investigate whether one can devise a random graph mechanism for a heavy-tail distribution even in the absence of a network. Examining Zipf's law about the frequency of words in a language might be an interesting starting problem.

Bibliography

[1] ALDOUS, D. A stochastic complex network model. *Electron. Res. Announc. Amer. Math. Soc. 9* (2003), 152–161.

[2] BARABÁSI, A.-L., ALBERT, R. Emergence of scaling in random networks. *Science 286* (1999), 509–512.

[3] BARABÁSI, A.-L., ALBERT, R., AND JEONG, H. Scale-free characteristics of random networks: The topology of the world-wide web. *Physica A 281* (2000), 69–77.

[4] BICKEL, P., AND CHEN, A. A nonparametric view of network models and Newman-Girvan and other modularities. *Proc. Natl. Acad. Sci. USA 106* (2009), 21068–73.

[5] BOLDI, M. Noisy random graphs and their Laplacians. *Discrete Mathematics 308*, 18 (2008), 4221–4230.

[6] BOLLA, M., AND TUSNÁDY, G. Spectra and optimal partitions of weighted graphs. *Discrete Mathematics 128* (1994), 1–20.

[7] BOLLOBÁS, B. The diameter of random graphs. *Trans. Amer. Math. Soc. 267* (1981), 41–52.

[8] BOLLOBÁS, B. The chromatic number of random graphs. *Combinatorica 8* (1988), 49–56.

[9] BOLLOBÁS, B. *Random graphs, second edition.* Cambridge University Press, 2001.

[10] BOLLOBÁS, B., AND ERDŐS, P. Cliques in random graphs. *Math. Proc. Cambridge Phil. Soc. 80* (1976).

[11] BOLLOBÁS, B., AND RIORDAN, O. The diameter of a scale-free graph. *Combinatorica 24* (2004), 5–34.

[12] BOLLOBÁS, B., RIORDAN, O., SPENCER, J., AND TUSNÁDY, G. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms 18* (2001), 279–290.

[13] BORGS, C., CHAYES, J., LOVÁSZ, L., T. SÓS, V., AND VESZTERGOMBI, K. Con- vergent sequences of dense graphs II: Multiway cuts and statistical physics. *Preprint, available at http: // research. microsoft. com/ ~ borgs/ Papers/ ConRight. pdf* (2007).

[14] BORGS, C., CHAYES, J., LOVÁSZ, L., T. SÓS, V., VESZTERGOMBI, K. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties andtesting. *Advances in Math. 219* (2008), 1801–1851.

[15] CHATTERJEE, S., DIACONIS, P., AND SLY, A. Random graphs with a given degree sequence. *arXiv, available at http: // arxiv. org/ PS_ cache/ arxiv/ pdf/ 1005/ 1005. 1136v3. pdf* (2010).

[16] CHUNG, F., AND LU, L. The average distance in a random graph with given expecteddegrees. *Internet Mathematics 1* (2002), 15879–15882.

[17] CHUNG, F., AND LU, L. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics 6,* 2 (2002), 125–145.

[18]CHUNG, F., LU, L., AND VU, V. Eigenvalues of random power law graphs. *Annalsof Combinatorics 7* (2003), 21–33.

[19]CHUNG, F., LU, L., AND VU, V. The spectra of random graphs with given expecteddegrees. *Internet Mathematics 1*, 3 (2003), 257–275.

[20]CSISZÁR, V., HUSSAMI, P., KOMLÓS, J., MÓRI, T., REJTŐ, L., AND TUSNÁDY,

G. When the degree sequence is a sufficient statistic. *To appear* (2010).

[21]CSISZÁR, V., REJTŐ, L., AND TUSNÁDY, G. Statistical inference on random struc- tures. In *Horizons of combinatorics, Balatonalmádi, Bólyai Society Mathematical Studies, 17. Ed., E. Gyori, G.O.H. Katona, L. Lovász* (2008), Springer Verlag, Berlin, pp. 37–66.

[22]DOROGOVTSEV, S., AND MENDES, J. *Evolution of networks: from biological nets tothe Internet and WWW*. Oxford University Press, 2003.

[23]DOROGOVTSEV, S., MENDES, J., AND SAMUKHIN, A. Structure of growing networkswith preferential linking. *Phys. Rev. Lett. 85* (2000), 4633–4636.

*[24]*DURRETT,R. *Random graph dynamics (Cambridge series in statistical and proba-*

bilistic mathematics). Cambridge University Press, 2007.

[25]ERDŐS, L. P., MIKLÓS, I., AND SOUKUP, L. Towards random uniform sampling of bipartite graphs with a given degree sequence. To appear.

[26]ERDŐS, P. On cliques in graphs. *Israel J. Math. 4* (1966), 233–234.

*[27]*ERDŐS, P., AND GALLAI, T. Gráfok előírt fokú pontokkal. *Mat. Lapok 11* (1960), 264–274.

*[28]*ERDŐS, P., AND RÉNYI, A. On random graphs, I. *Publ. Math. Debrecen 6* (1959), 290–297.

[29]ERDŐS, P., AND RÉNYI, A. On the evolution of random graphs. In *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* (1960), pp. 17–61.

[30]ERDŐS, P., AND RÉNYI, A. On the evolution of random graphs. *Bull. Inst. Internat. Statist. 38*, 4 (1961), 343–347.