

Article

Divisibility Patterns within Pascal Divisibility Networks

Pedro A. Solares-Hernández ^{1,2,3}, Fernando A. Manzano ^{1,2,3} , Francisco J. Pérez-Benito ^{1,4}  and J. Alberto Conejero ^{1,*} 

¹ Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, 46022 València, Spain; pedsoher@doctor.upv.es (P.A.S.-H.); fmanzano@unapec.edu.do (F.A.M.); frapebe@doctor.upv.es (J.-P.B.)

² Departamento de Estadística, Universidad APEC, 10203 Santo Domingo, Dominican Republic

³ Departamento de Ingeniería, Universidad APEC, 10203 Santo Domingo, Dominican Republic

⁴ Instituto de Tecnología Informática (ITACA), Universitat Politècnica de València, 91354 València, Spain

* Correspondence: aconejero@upv.es

Received: 31 December 2019; Accepted: 7 February 2020; Published: 14 February 2020



Abstract: The Pascal triangle is so simple and rich that it has always attracted the interest of professional and amateur mathematicians. Their coefficients satisfy a myriad of properties. Inspired by the work of Shekatkar et al., we study the divisibility patterns within the elements of the Pascal triangle, through its decomposition into Pascal's matrices, from the perspective of network science. Applying Kolmogorov–Smirnov test, we determine that the degree distribution of the resulting network follows a power-law distribution. We also study degrees, global and local clustering coefficients, stretching graph, averaged path length and the mixing assortative.

Keywords: network science; graph theory; divisibility; Pascal matrix

MSC: 05C82; 05C25

1. Introduction

Number theory has been one of the most studied fields of mathematics for centuries. In contrast, network science has emerged as a discipline in the last twenty years. Nevertheless, networks have attracted the interest of many researchers due to their multiple applications to different disciplines, such as biology, telecommunications, social and environmental sciences, as well as systems medicine.

Besides, some networks have emerged from mathematical concepts, such as the divisibility network. This was firstly studied by Zhou et al. in [1]. Here, nodes represent natural numbers and two nodes $n, m \in \mathbb{N}$ are connected by a directed edge if n divides m , denoted by $n|m$. These authors noticed that this network has a large clustering coefficient of approximately 0.34, which is insensitive to the network size. Besides, they showed that: (i) The average distance between a pair of nodes is upper bounded, in contrast to small-world networks, (ii) it posses a hierarchical architecture, and (iii) the degree distribution follows a power-law.

A divisibility network can also be considered as a non-directed one if we connect a pair of nodes $a, b \in \mathbb{N}$ if either $a|b$ or $b|a$. Shekatkar et al. [2] studied it using the framework of a growing complex network. Among other properties, they showed that it is scale-free but has a non-stationary degree distribution, reporting a stretching similarity pattern, and showing how this pattern evolves with the size of the network. Related to divisibility networks, Yan et al. [3] showed that every layer in a multiplex congruence network is a sparse and heterogeneous subnetwork satisfying the scale-free property, providing an insight into the simultaneous congruences problem through the graphical solutions provided there.

All these results have inspired us to look for other ways of consider the divisibility network through a different growing network procedure. Due to the abundance of beautiful and unpredictable properties hidden in the Pascal Triangle (PT), we have studied the non-directed divisibility following a similar approach as Shekatkar et al. did. For this purpose, we have considered squared Pascal symmetric matrices as a covering of growing finite subsets of the PT. These matrices were firstly analyzed by Brawer and Pirovino [4]. We will denote by S_n the Pascal square matrix of order n , that is obtained when taking the square with two orthogonal sides given by the first n ones of both sides of the PT. As an example, we have indicated in bold font the Pascal matrix S_4 in (1). From S_4 we construct a divisibility network whose nodes are $\{1, 2, 3, 4, 6, 10, 20\}$.

Row 0				1												
Row 1				1		1										
Row 2				1		2		1								
Row 3				1		3		3	1							
Row 4				1		4		6		4		1				
Row 5				1		5		10		10		5		1		
Row 6				1		6		15		20		15		6		1

(1)

Pascal matrices present some beautiful properties. As an example, we mention the decomposition of the Pascal matrices proposed by Edelman and Strang [5] who showed how to decompose any Pascal matrix S_n of order n , into the product of two matrices L_n , which is a lower triangular matrix, and U_n , which is an upper triangular matrix, such that, $S_n = L_n U_n$, for every $n \in \mathbb{N}$.

For any given number $k \in \mathbb{N}$, there is some $n_0 \in \mathbb{N}$ such that $k \in S_n$ for all $n \geq n_0$. With this in mind, we will study the divisibility networks provided by these matrices. On the one hand, using these matrices, we can cover all the natural numbers through a growing family of subnetworks. On the other hand, the numbers in these matrices are not consecutive, and when sequentially ordered, we can find big gaps between some of them. This suggests us to study whether the scale-free degree distribution holds and other network properties are still satisfied as to the ones shown in [2].

In particular, we have considered the evolution of several network measures along with the size of Pascal matrices, such as the average degree $\langle k \rangle$, the histogram of the connectivity degree distribution; the local and global clustering coefficients C_i and C_Δ , the assortativity index r ; and the average path length $\langle d \rangle$. We will recall the definition of these notions in the next section.

The degree distribution is studied in Section 2.1, where we show that it is scale-free. We also study an example of how to compute the fitting parameters. In order to study the structure of this network, we have studied how the clustering evolves with the size of the network. The local and global clustering coefficients are presented in Sections 2.2 and 2.3. The tendency of nodes to connect to nodes of similar degree is analyzed through the assortative coefficient in Section 2.4, and how nodes are separated respect to the others through the average path length, see Section 2.5.

We refer the reader to the books of Barabási [6], Estrada [7], and Newman [8], as basic references of Network Science.

2. Network Analysis

For any arbitrary $n \in \mathbb{N}$, we consider the divisibility network $M_n = (V_n, E_n)$, associated to the Pascal matrix S_n of order n . This network has V_n as its set of nodes and E_n as its set of edges. We point out that we exclude the number 1 of V_n since it will be linked with any other number, which does not provide useful information for studying the evolution of the network properties of these matrices.

We denote $\text{card}(V_n)$ by N_n , with $N_n \in \mathcal{O}(n^2)$, and $\text{card}(E_n)$ by L_n . Once fixed the set of nodes V_n with the non-repeated elements of S_n , we recall that an arbitrary pair of elements $a, b \in V_n$ are linked by an edge if, and only if, $a|b$ or $b|a$. Figures 1a,b are generated from matrices S_6 and S_7 .

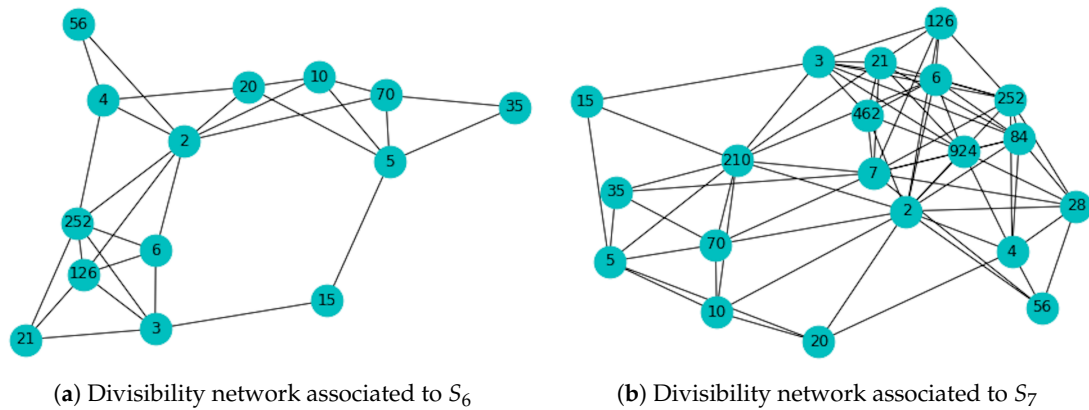


Figure 1. Examples of divisibility networks obtained from Pascal matrices S_6 and S_7 . From S_6 , we have that the network M_6 has 14 nodes and 29 edges, see Figure 1a, and from S_7 we have that M_7 has 20 nodes and 72 edges, see Figure 1b.

2.1. Degree Distribution

Given the sequence of nodes from a network M_n , with $n \in \mathbb{N}$, we first analyze the evolution of the degree distribution $p(k) = n_k/N_n$, where n_k is the number of coefficients with degree k in M_n , along with n . First, we check if the scale-free property holds, which will result into the existence of many nodes with only a few edges and a few nodes with a large number of edges, that are called hubs, see [6,9].

In Figure 2, we illustrate the asymptotic growth of the cumulative and the averaged cumulative degree distributions for the network M_{2^6} , that has a total of 2001 nodes and 51147 edges. The nodes are indexed in increasing order respect to the term of S_{2^6} that they represent. We appreciate how both measures tend to stabilize when adding the last nodes of each network.

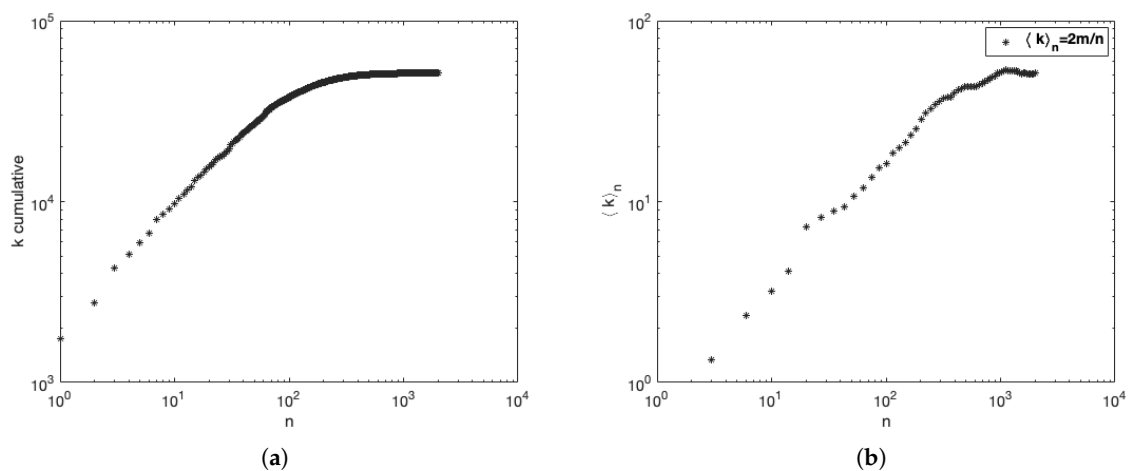


Figure 2. Cumulative degree and average degree distributions for the network M_{2^6} . Figure 2a shows the cumulative degree as the coefficients of S_{2^6} grow in value. Figure 2b plots the evolution of the average degree when the coefficients of S_{2^6} are progressively added. For $n = 2^6$ the average degree is $\langle k \rangle_n = 51.114$.

Applying the Maximum Likelihood method [2,10], we confirm that the divisibility network of Pascal matrices satisfies a scale-free law. This means that the network degree distribution follows, at least asymptotically, a power-law of the form $p(k) = C \cdot k^{-\gamma}$ for all $k \geq k_{\min}$. We recall that when the power-law governs a process, this usually occurs from what we determinate the “minimum value”

k_{min} , which is exactly the point where one can start to observe the fall of the heavy tail. In Figure 3a, we represent again the degree distribution for M_{211} with logarithmic binning in the degree values.

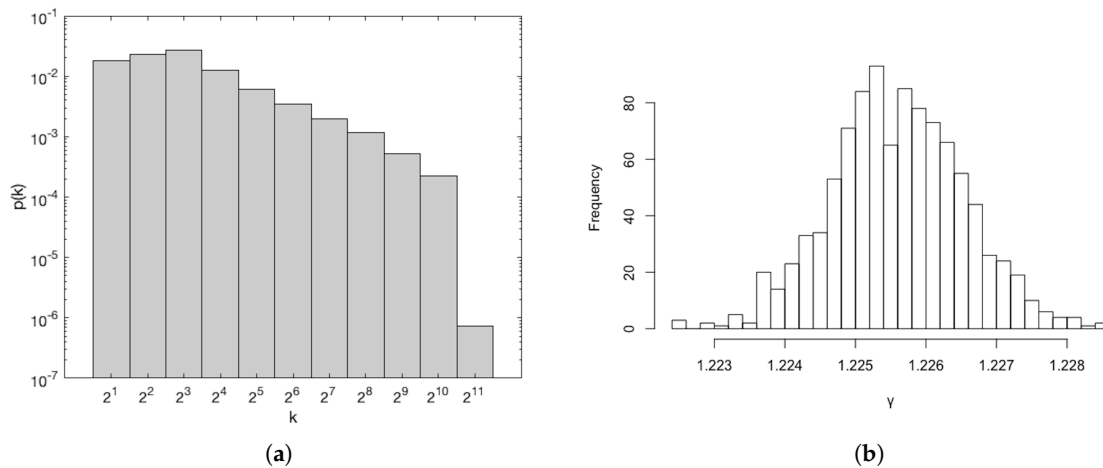


Figure 3. In Figure 3a, the sizes of the bins are equal to successive positive powers of 2, and the count in each bin is normalized dividing by the bin width, like in ([2], Figure 2). In Figure 3b, we characterize the uncertainty in the parameter fitting to a power law using 1000 bootstraps for M_{26} .

We characterize the power-law of the degree distribution of M_{26} through bootstrapping, using 1000 iterations of bootstrapping. The obtained fitting parameters were $k_{min} = 3$ and $\gamma = 1.2256$. In Figure 2a, we plot the histogram of the values obtained for the γ parameter. For $\gamma = 1.2256$, the Kolmogorov–Smirnov coefficient was 0.976, see Figure 3b, which shows the goodness of the estimation. It is worth to mention that network with similar degree distributions could have a different inner structure, see [11].

We also show the degree distributions of M_{25} , M_{26} , M_{27} and M_{28} in Figure 4. We observe a high concentration of nodes (plateau) in the region of high values of the degree. This suggests a bias in the linear model fitting due to the comparatively large number of nodes with low degrees respect to the fewer number of nodes with large degrees [6,12].

In order to correct the non-uniform sampling seen with the linear binning, we show these degrees distributions with logarithmic binning again, see Figure 5.

As n grows, the γ coefficient of the power-law fitting decreases. Increasing the number of nodes of the network does not guarantee a better fit of the linear model. It is worth to mention that these indicators do not fully reflect the network robustness due to the inherent bias of the particular sampling process, see for instance [13].

We also have wondered how important is the role of the lowest numbers, which have the highest degrees, in determining the scale-free nature of the network. In this line, we have removed the hubs corresponding to nodes associated with numbers 2 to 6. These new degree distributions also follow a power-law, as we can see in Figure 6.

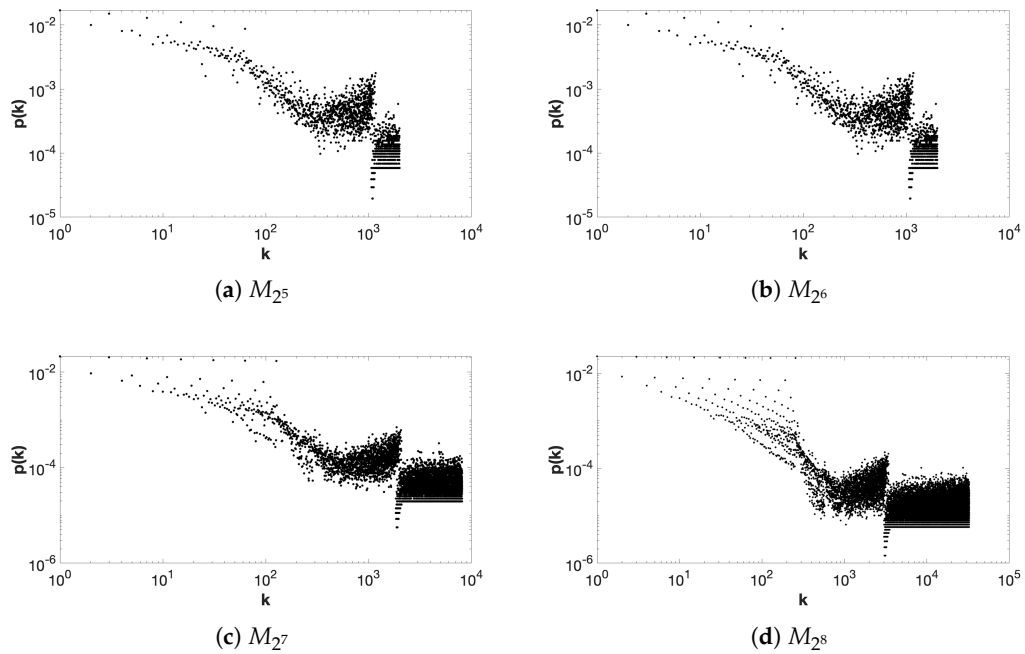


Figure 4. Representation of the degree distribution for M_{25} , M_{26} , M_{27} , and M_{28} .

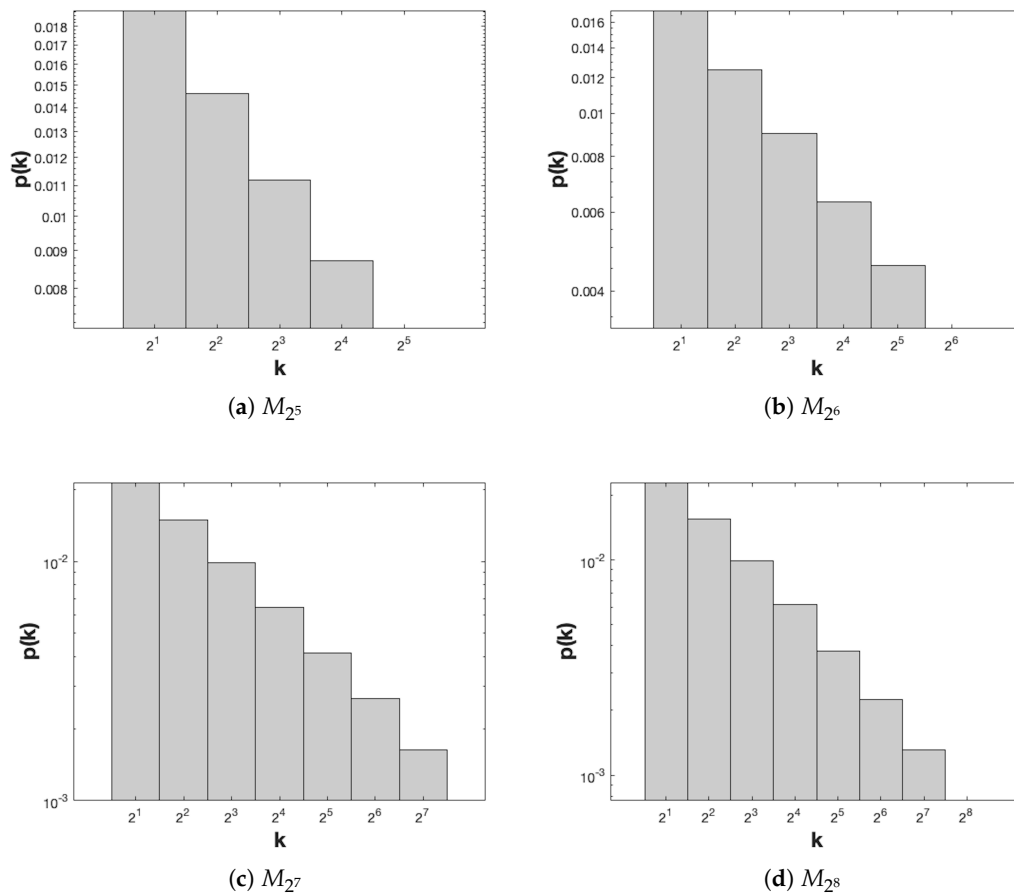


Figure 5. Representation of the degree distribution for M_{25} , M_{26} , M_{27} , and M_{28} , with logarithmic binning.

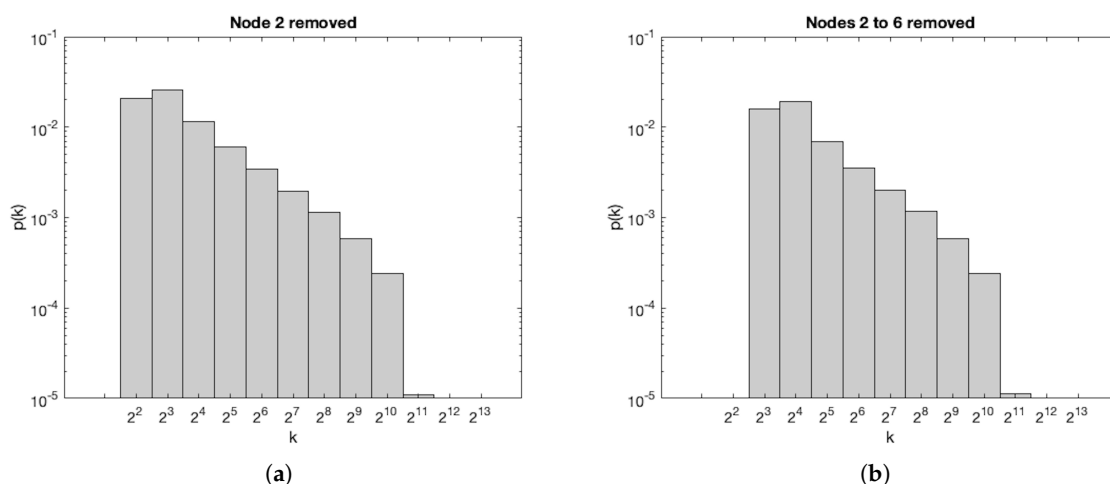


Figure 6. Degree distributions of the Pascal network $M_{2^{11}}$ with logarithmic binning. In Figure 6a we have removed node 2 and in Figure 6b we have removed nodes 2 to 6.

2.2. Local Clustering Coefficient

Many scale-free networks also display a high degree of clustering. This is the result of a hierarchical organization in which small groups of nodes organize into increasingly larger ones while preserving the scale-free property [14].

Since all the natural numbers appear along the first row/column of Pascal matrices, and the rest of their elements are composite numbers, there will be many connections from the first elements to the second ones. Besides, as numbers in the inner columns/rows of the matrices grow pretty fast, there will be fewer edges among them. Therefore, we have analyzed how clustered the network is. The main two measures for analyzing the clustering are the local and the global clustering coefficients. The global version provides an indicator of the clustering in the network, whereas the local one gives an indicator of the connection between the adjacent nodes to a given one.

We recall that given $M_n = (V_n, E_n)$ and a node $v_i \in V_n$ with degree k_i , we denote by \tilde{v}_i the set of nodes adjacent to v_i . Then, the local clustering coefficient C_i is the number of pairs of adjacent nodes to v_i that are connected between them by an edge, divided by the number of admissible neighbor pairs of v_i [8,15,16] that is:

$$C_i = \frac{2 \left| \left\{ l_{jk} : v_j, v_k \in \tilde{v}_i \right\} \right|}{k_i(k_i - 1)}, \text{ where } l_{jk} \text{ denotes an edge linking } v_j \text{ and } v_k. \tag{2}$$

From Equation (2), we generate some stretching graphs to show how the clustering coefficient evolves when adding the nodes one by one to the networks M_{2^6}, M_{2^7} and M_{2^8} , see Figure 7.

To see this evolution, we can also represent the difference between the local clustering of two consecutive nodes, namely $\Delta C = C_i - C_{i+1}$. This can be compared respect to ([2], Figure 5). Figure 8 shows a symmetrical pattern with respect to the value $\Delta C = 0$, as the network grows.

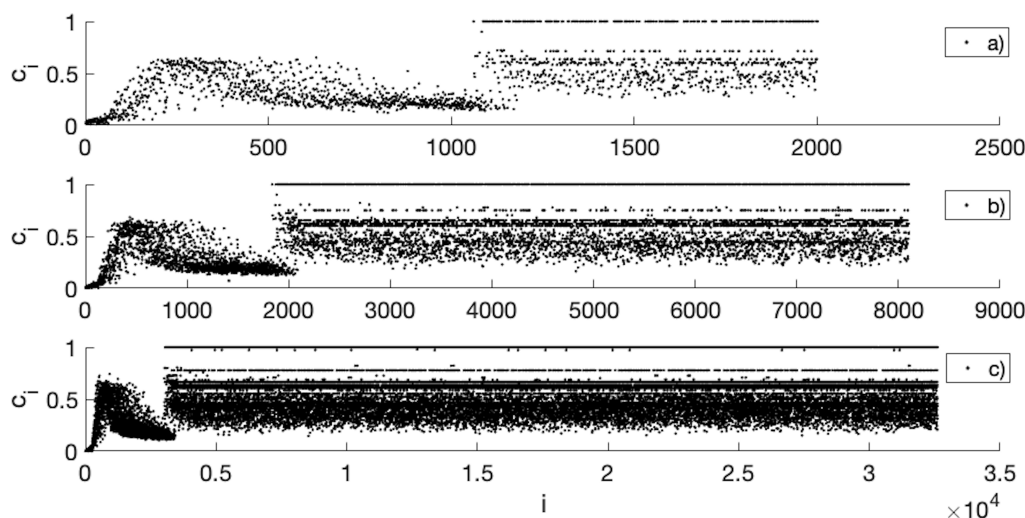


Figure 7. Local clustering coefficient in terms of the node index for different network sizes: (a) M_{26} , (b) M_{27} , (c) M_{28} . Nodes are indexed in increasing order of the number to which they are associated. This graph shows a stretch of similarity due to the divisibility between the coefficients of the Pascal matrix. The stretch is the same, regardless of the size of the network, as was shown in [2].

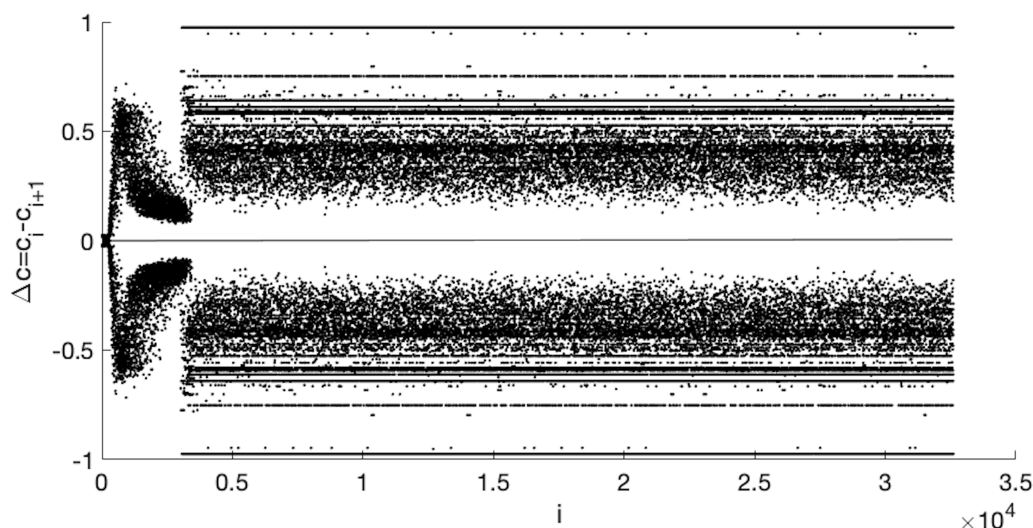


Figure 8. Behavior of the evolution of the difference between the clustering coefficient C_i and the clustering coefficient C_{i+1} .

2.3. Global Clustering Coefficient

The global clustering coefficient is based on ordered triplets of connected nodes, that can be linked by 2 (open triplet) or 3 edges (closed triplet). With this, a triangle graph has 3 closed triplets: That is, a triangle ABC has 3 triplets associated to it: ABC, BCA, and CAB. With this, the global clustering coefficient can be computed as

$$C_{\Delta} = \frac{3 \times \text{number of triangles}}{\text{number of connected triplets}} \tag{3}$$

As we can see in Figure 9 the global clustering coefficients decreases when the size of M_n grows, approaching 0 for high values of n , which represents that the networks M_n become more and more sparse when increasing its size n .

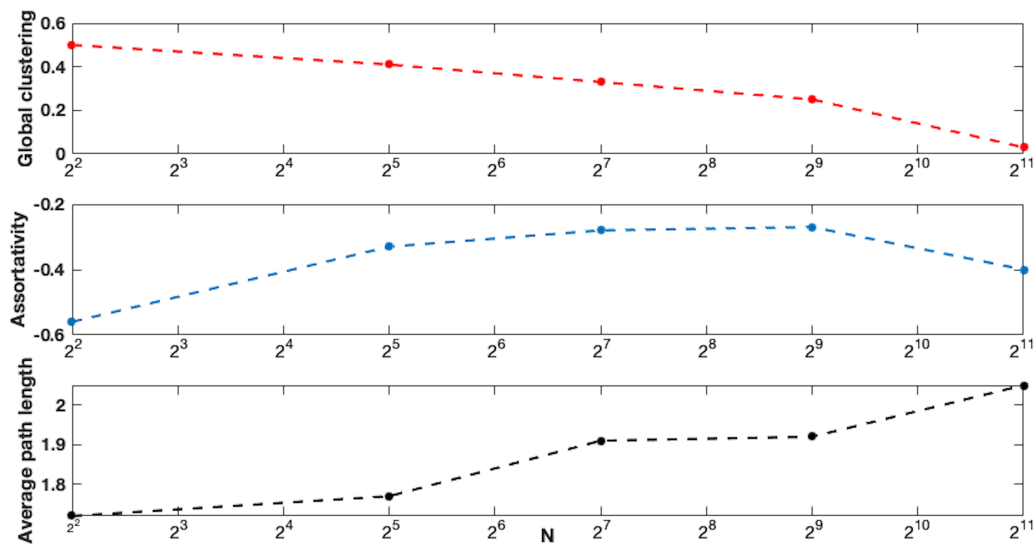


Figure 9. Evolution of the global clustering, assortativity, and average path length for $M_{2^2}, M_{2^5}, M_{2^7}, M_{2^9},$ and $M_{2^{11}}$.

2.4. Assortative Coefficient

The assortativity coefficient r the correlation coefficient of the degrees of adjacent nodes. It shows the preference of nodes to connect with other nodes of comparable degree [8,17].

Given a network $G_n = (V_n, E_n)$, it is defined as follows:

$$r = \frac{\sum_{ij}(A_{ij} - k_i k_j / 2L_n)k_i k_j}{\sum_{ij}(k_i \delta_{ij} - k_i k_j / 2L_n)k_i k_j} \tag{4}$$

where k_i is the degree of v_i , A_{ij} is the $(i, j)^{th}$ element of the adjacency matrix associated to the eventual connection between v_i and v_j , L_n is the total of links in the network, and δ_{ij} is the Kronecker delta. However, determining the assortativity from (4) supposes a high computational cost, and therefore it is suggested to approximate the assortativity by means of the next expression ([8], Section 10.7):

$$r = \frac{S_1 S_e - S^2}{S_1 S_3 - S_2^2} \quad \text{with} \quad S_e = \sum_{ij} A_{ij} k_i k_j = 2 \sum_{l_{(i,j)}} k_i k_j, \quad \text{and} \quad S_m = \sum_{i=1}^{N_n} k_i^m \quad \text{for } m = 1, 2, 3, \tag{5}$$

where we have introduced $l_{(i,j)}$ for referring to all unordered pairs of nodes connected by an edge and N_n is the total number of nodes of a network M_n .

Assortativity coefficient presents values ranging from -1 to 1 . If $r = 1$, we called the network to be fully assortative. In case of $r = 0$ the network is said to be not assortative, while if $r = -1$, the network is called disassortative [15,17,18]. The divisibility networks M_n approaches to be disassortative as long as n increases, see Figure 9, which reflects that nodes that have a high degree tend to connect with low-grade nodes.

2.5. Average Path Length

The network average path length, denoted by $\langle d \rangle$ is the averaged distance between all pairs of nodes. For each pair, the distance between them is given by the shortest path connecting them. Clearly, if two nodes belong to different connected components, the distance will be ∞ . For computing the average path length we have excluded pairs of nodes belonging to different connected components.

$$\langle d \rangle = \frac{1}{N(N-1)} \sum_{\substack{i,j=1,N \\ i \neq j}} d_{i,j} \quad (6)$$

We can see that, if connected, the averaged path length is short. In Figure 9, we show it for several networks. It increases with the size of the network, but even for M_{211} it is 2.05. This means that the biggest connected component is closed to be a bipartite network.

3. Conclusions

PT and Pascal matrices present multiple properties that have fascinated mathematicians for ages. Here, we have studied them from the perspective of Network Science. We have considered Pascal matrices as a ground for constructing a growing divisibility network. Its structure has been studied, following the approach of [2]. The interest of this choice lies in the fact that PT contains all the natural numbers on each of their sides. The way in which we increasingly construct the divisibility network by taking the Pascal matrices provides a different arrangement of the natural numbers, with gaps between some of them. Nevertheless, the scale-free property for the degree distribution also holds.

Either in [2] or here, both growing networks present similar structures and characteristics to real based networks. This can be noticed when looking at the degree distribution, the global and local clustering coefficients, the assortativity, and the average path length. This work fits within our interest in studying divisibility networks constructed from subsets of the natural numbers, and to see how network measures can help us to describe them and how to find hidden structures and hierarchies [19]. In future works, we will study the divisibility networks provided by other arrangements of the natural numbers, and the divisibility networks of other countable sets of numbers such as the rational numbers in the unit interval [20].

The results concerning the local clustering coefficient are similar to the ones given by [2]. Besides, we have seen that when the size of the network grows, the global clustering coefficient tends to 0, and the divisibility networks approach more and more to being disassortative. Both results agree with the average path length that, despite low, it increases as the network size grows and indicates that the network is close to being bipartite.

Author Contributions: Conceptualization, P.A.S.-H. and J.A.C.; software P.A.S.-H, F.J.P.-B. and F.A.M.; validation and formal analysis, P.A.S.-H. and J.A.C.; writing—original draft preparation, writing—review and editing, P.A.S.-H., F.J.P.-B., F.A.M. and J.A.C. All authors have read and agreed to the published version of the manuscript.

Funding: J.A.C. was funded by MEC grant number MTM2016-75963-P. P.A.S.-H. acknowledges the support of MESCyT-RD and Casa Brugal for his PhD grants.

Conflicts of Interest: The authors declare no conflict of interest. The funding institutions had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Zhou, T.; Wang, B.H.; Hui, P.; Chan, K. Topological properties of integer networks. *Phys. A Stat. Mech. Appl.* **2006**, *367*, 613–618. [[CrossRef](#)]
2. Shekatkar, S.M.; Bhagwat, C.; Ambika, G. Divisibility patterns of natural numbers on a complex network. *Sci. Rep.* **2015**, *5*, 14280. [[CrossRef](#)] [[PubMed](#)]
3. Yan, X.Y.; Wang, W.X.; Chen, G.R.; Shi, D.H. Multiplex congruence network of natural numbers. *Sci. Rep.* **2016**, *6*, 23714. [[CrossRef](#)] [[PubMed](#)]
4. Brawer, R.; Pirovino, M. The linear algebra of the Pascal matrix. *Linear Algebr. Its Appl.* **1992**, *174*, 13–23. [[CrossRef](#)]
5. Edelman, A.; Strang, G. Pascal matrices. *Am. Math. Mon.* **2004**, *111*, 189–197. [[CrossRef](#)]
6. Barabási, A.-L. *Network Science*; Cambridge University Press: Cambridge, UK, 2016.
7. Estrada, E. *The Structure of Complex Networks: Theory and Applications*; Oxford University Press: Oxford, UK, 2012.

8. Newman, M. *Networks*; Oxford University Press: Oxford, UK, 2018.
9. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)] [[PubMed](#)]
10. Clauset, A.; Shalizi, C.; Newman, M. Power-law distributions in empirical data. *SIAM Rev.* **2009**, *51*, 661–703. [[CrossRef](#)]
11. Shang, Y. Distinct clusterings and characteristic path lengths in dynamic small-world networks with identical limit degree distribution. *J. Stat. Phys.* **2012**, *149*, 505–518. [[CrossRef](#)]
12. Gillespie, C. Fitting heavy tailed distributions: The poweRlaw Package. *J. Stat. Softw.* **2015**, *64*, 1–16. [[CrossRef](#)]
13. Shang, Y. Subgraph robustness of complex networks under attacks. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *49*, 821–832. [[CrossRef](#)]
14. Ravasz, E.; Barabási, A. Hierarchical organization in complex networks. *Phys. Rev.* **2003**, *67*, 026112. [[CrossRef](#)] [[PubMed](#)]
15. Dorogovtsev, S.; Mendes, J. Evolution of networks. *Adv. Phys.* **2002**, *51*, 1079–1187. [[CrossRef](#)]
16. Newman, M. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [[CrossRef](#)]
17. Newman, M. Mixing patterns in networks. *Phys. Rev.* **2003**, *67*, 026126. [[CrossRef](#)] [[PubMed](#)]
18. Newman, M. Assortative mixing in networks. *Phys. Rev.* **2002**, *89*, 208701. [[CrossRef](#)] [[PubMed](#)]
19. Bunimovich, L.; Smith, D.; Webb, B. Finding hidden structures, hierarchies, and cores in networks via isospectral reduction. *Appl. Math. Nonlinear Sci.* **2019**, *4*, 225–248. [[CrossRef](#)]
20. Hernández-Solares, P.; Conejero, J.A. Divisibility patterns of the rational numbers on the interval on a complex network. *Preprint* **2020**.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).