AUT Journal of Electrical Engineering

AUT J. Elec. Eng., 57(1) (2025) 85-100
DOI: 10.22060/eej.2024.23143.5591

cA Comprehensive Approach to Synthetic Distribution Grid Generation: Erdős–Rényi to Barabási-Albert

Mohammad Shahraeini *

Department of Electrical Engineering, Golestan University, Gorgan, Iran

ABSTRACT: In this extended study, the focus is on advancing the generation of synthetic distribution grids (SDGs) through the introduction of a new algorithm based on the Barabási-Albert random graph model. The initial use of the Erdős model to create SDGs revealed limitations in size and structural adjustability beyond the number of vertices. To address these limitations and push the research forward, the new algorithm utilizes the Barabási-Albert model to provide more control over the structural features of the generated graphs through the introduction of a novel tuning parameter known as the "richness index". The effectiveness of both algorithms in producing SDGs of various sizes is demonstrated by generating SDGs with different sizes, confirming their ability to mimic synthetic radial distribution grids successfully. Additionally, a detailed examination of degree-based parameters and Pearson coefficients for SDGs of sizes from 20 to 1000 uncovers significant patterns. Furthermore, the proposed algorithm is examined in the terms of the variation of richness index in branching rate and μ-PMU placement, confirming the scale-free characteristic of the method. A comparison of the Erdős and Barabási-Albert models shows variations in maximum degree values, branching rates, and mixing patterns. The original Barabási-Albert model tends to have nodes with higher degrees and increased branching rates, which can be adjusted by the richness index. These findings emphasize the ability of the Barabási-Albert model to generate scale-free SDGs with diverse structures by fine-tuning the richness index.

Review History:

Received: Apr. 28, 2024
Revised: Jun. 01, 2024
Accepted: Jul. 09, 2024
Available Online: Jul. 09, 2024

Keywords:

Synthetic Distribution Grids
Random Graph
Erdős–Rényi Random Graph Model
Barabási-Albert Random Graph Model
Power Graph

1- Introduction

Power grids are categorized according to their voltage levels, namely High Voltage (HV), Medium Voltage (MV), and Low Voltage (LV) grids, with the latter two commonly known as Distribution Grids. These grids play a fundamental role in urban and industrial infrastructure by facilitating the transmission of electrical energy to end-users. Distribution grids are essential as they represent the critical final step in the electricity supply chain, guaranteeing the secure and efficient distribution of electricity to residences, commercial enterprises, and industrial facilities [1].

Distribution networks are critical components of smart grids because they facilitate the integration of renewable energy sources, such as wind and solar power, into the electrical grid. This allows for greater flexibility and resilience in meeting electricity demand while reducing greenhouse gas emissions. Moreover, distribution networks play a vital role in demand response programs, which involve real-time communication between utilities and customers regarding electricity consumption patterns and pricing strategies to promote efficient usage and reduce peak loads on the network [2, 3].

The growing importance of electrical distribution networks as a critical component of the smart grid system necessitates the need for better performance and optimization. Therefore, researchers have been exploring innovative methods to analyze these networks recently. One such approach involves modeling and analyzing power grids using complex network theory [4].

1- 1- Power Grids as Complex Networks

Barabási *et al.* initially suggested the idea of examining power grids using complex network analysis, where the authors portrayed the power network as a "scale-free" network [5]. In this framework, the power graph comprises nodes (buses) and edges (lines), representing the connections between them at various voltage levels. At the high voltage (HV) level, buses are considered as vertices and transmission lines as edges. However, in distribution grids (MV and LV levels), buses are represented as vertices and lines as edges. The properties of vertices in distribution grids are defined by their active power supply or demand and voltage, while the properties of lines are determined by their cost, which is influenced by resistance, and capacity, which is determined by their limit [6].

There are several types of distribution grids that vary

*Corresponding author's email: m.shahr@gu.ac.ir

depending on their connection topology, including radial, parallel, ring, and interconnected systems [7]. Radial systems are the simplest and least expensive, consisting of feeders radiating outward from a central substation with power flowing in only one direction. These systems typically use tree graphs.

As previously mentioned, new approaches such as those based on complex network analysis require numerous tests to verify models, but there are currently few available test feeders specifically designed for distribution grids, mostly introduced as IEEE test cases. To address this issue, synthetic test systems have been created to represent diverse real networks. While most existing synthetic systems focus on high-voltage (HV) networks [5], [8-22], some recent research also explores the topological and electrical properties of distribution grids [23-28] and introduces methods for validating these synthetic systems [29-31].

1- 2- Random Graph and Synthetic Power Graph

Generally speaking, a random graph is characterized by starting with a set of *n* unlinked nodes and subsequently adding edges at random among them [32]. Several widely utilized random graph models have been established, including the Erdős-Renyi model, which was put forth by Erdős, Rényi, and Gilbert [33, 34]; the Barabási-Albert (BA) random graph model, suggested by Albert-László Barabási and Réka Albert; and the small-world model, presented by Watts and Strogatz [35]. While these random graph models are suitable for various applications, they have been found to fall short when it comes to accurately portraying power graphs, whether they pertain to transmission or distribution networks [10]. Consequently, researchers devised the "s*ynthetic power graph*" (SPG), an artificial representation of actual power grids based on their statistical features. SPGs may also refer to either "*synthetic transmission grids*" (STGs) or "*synthetic distribution grids*" (SDGs).

In their research on STG generation, [10] presented a revised edition of the small-world model known as RT-nested-Smallworld, a random topology power grid model that considers the features of power graphs. Various other random graph models and algorithms like the Chung-Lu model [20], dual-stage constructed random graph algorithm [15], and Cluster-and-Connect algorithm [36] have also been suggested for creating STGs.

The Erdős–Rényi model has recently been employed to generate STGs. In a previous study [21], STGs of varying sizes were produced by adjusting the input parameters for two versions of the Erdős–Rényi model (Erdős–Rényi itself and the Gilbert model), demonstrating their resemblance to real power grids in the context of optimal PMU placement problem. In a more recent study [22], a modified version of the Erdős–Rényi model was proposed to create STGs of different sizes, while also specifying the locations of zero injection buses (ZIBs). The proposed technique systematically establishes the connectivity of the power graphs by rewiring and renumbering the vertices to achieve consecutive numbering.

1- 3- Review of Related Literature

In the early 21st century, efforts were made to create synthetic power grids at the HV level to study power system blackouts [8] and develop a continuum model of electromechanical dynamics in large power grids [9]. These efforts were not yet official. Previous work on generating synthetic transmission networks was categorized into statistics-based approaches (e.g. [10]) and those utilizing machine learning methods (e.g. [18]), as discussed by the authors of [27].

While the concept of synthetic transmission networks (HV) is relatively new in power system literature and emerged in the 21st century, the unofficial generation of synthetic distribution grids began in the early 2010s with Reference Network Models like RNM-Europe for distribution network planning [37, 38]. This trend continued with models for distribution grids [14, 39]. The official introduction of synthetic distribution grids was documented in [23, 24], where authors used data from a Dutch DSO to create SDGs evaluated by KL divergence. An approach to generate synthetic distribution grids corresponding to existing power distribution systems was proposed in [25], along with a power system planning optimization technique using a multi-objective function. Authors in [26] extended their European reference model (RNM-Europe) to develop algorithms for large U.S. distribution grids, forming the RNM-US model. In [40], a grid synthesis procedure based on Dutch DSO spatial datasets was outlined, incorporating topology recognition, peak load estimation, and line dimensioning heuristics.

Furthermore, a deep learning framework (unbalanced graph generative adversarial network) was introduced in [27] to create synthetic three-phase unbalanced active distribution networks using limited real data. Another framework proposed in [41] leveraged information on interdependent roads and building infrastructures to synthesize distribution grids resembling physical counterparts for specific regions.

In [42], a detailed overview of methods for creating synthetic graphs, encompassing both STGs and SDGs, up to the year of publication (2021) is available. However, the quantity of articles focusing on SDG generation [27, 40, 41, 43-46], and the validation approaches [31, 47] has notably increased since then. The recent increase in publications over the last two years alone underscores the novelty of generating synthetic distribution grids.

In our very last publication [28], we introduced an algorithm based on the Erdős model to generate synthetic distribution graphs resembling radial distribution networks. The algorithm demonstrated promise in creating small-sized synthetic grids up to 35 vertices, with similarity to actual distribution networks validated through degree centrality metrics.

1- 4- Motivation, Objective, and Contribution

From the literature discussed in the preceding section and its timeliness, it is evident that the development of algorithms for creating synthetic distribution grids remains a current topic. Conversely, our recent algorithm proposal, based on

the Erdős–Rényi model, can indeed generate radial SDGs but is limited in size. Thus, the need for algorithms with the capacity to generate remains an ongoing concern, serving as the primary motivation for our study.

This research has two primary goals. The first objective is to enhance the algorithm in the Erdős–Rényi random graph model as detailed in [28] to produce synthetic distribution grids structured radially. In addressing the limitations of this initial algorithm, our study also introduces a new algorithm based on the Barabási-Albert model to create large and very large radial distribution grids synthetically.

The contributions of this research are twofold: (I) To our knowledge, this work, alongside our earlier study [28], represents the inaugural endeavor to utilize the Erdős–Rényi random graph model for generating SDGs in the form of radial distribution grids. (II) This study marks the first exploration of employing the Barabási-Albert model to generate substantial and extensive SDGs structured as radial distribution grids.

1- 5- Paper Structure

The structure of the remaining sections in this paper is as follows. Section 2 will provide an overview of the graph theory preliminaries. Section 3 will delve into the Erdős–Rényi random graph models, including the Gilbert and Erdős models, as well as explore tree generation. In Section 4, we will introduce our prior algorithm based on the Erdős–Rényi model, emphasizing its constraints in generating synthetic radial distribution grids. Consequently, we will propose a new algorithm in the same section, based on the Barabási-Albert model, aimed at addressing the limitations of the previous approach. Section 5 will showcase simulation results for synthetic distribution grids of various sizes, ranging from 20 to 1000. Finally, Section 6 will conclude the study and outline potential future research directions.

2- Preliminaries

A graph $G = (V, E)$ can be represented by a collection of vertices, $V = \{v_1, v_2, ..., v_n\}$ with a total of n vertices, and a set of edges, $E = \{e_{ij}\} \subseteq V \times V$ with a total of M edges. The undirected edge e_{ij} connecting vertices v_i and v_j is denoted as $e_{ij} = \{v_i, v_j\}$. Two vertices v_i and v_j in V are considered adjacent if $v_i, v_j \in E$, and are referred to as neighbors. An undirected graph is a graph where all its edges are undirected.

A graph $G' = (V', E')$ is defined as a subgraph of another graph $G = (V, E)$ if and only if V' is a subset of V and E' is a subset of E, meaning it includes all the vertices and edges present in the original graph.

A tree is a type of graph that is both connected and acyclic, indicating it does not contain any cycles. It comprises nodes or vertices connected by edges, where each node has at most one edge incident upon it. This implies that there is a unique path between any two nodes in the tree. It is evident that a tree with n nodes has $M = n$-1 edges.

In a graph $G = (V, E)$ with n vertices and M edges, the adjacency matrix **A**, represented as a square matrix, is defined as $\mathbf{A} = [a_{ij}]_{n\times n}$. Here, a_{ij} equals 1 if the i^{th} vertex is connected to the j^{th} vertex, and 0 otherwise. In the case of an undirected graph, **A** is symmetric. **A** connected component in a disconnected undirected graph is a subgraph where any pair of vertices are linked by paths. It is denoted as $C_i = (V_i, E_i)$, where i ranges from 1 to $C(G)$, and the union of all V_i forms the entire vertex set V. The number of components in G is denoted by $C(G)$, with C_1 referred to as the giant component and the others as additional components. Various methods like breadth-first search (BFS), depth-first search (DFS), and the Tarjan algorithm have been suggested to determine graph connectivity and identify connected components. For a detailed explanation of calculating the number of components using DFS, readers are directed to Algorithm 1 in the referenced source [28].

2- 1- Centrality Metrics

The analysis of complex networks involves various metrics categorized as clustering, distance, centrality, and scaling metrics. Within centrality metrics, there are two main types: neighborhood-based and shortest path-based. Degree centrality, a neighborhood-based metric, measures a node's connectivity in a network by counting the number of edges connected to that node. In an undirected graph $G = (V, E)$, degree centrality is defined as:

$$\mathbf{C}_d(v) = \deg(v) = |N(v)| \tag{1}$$

where $N(v)$ represents the direct neighbors of vertex v.

A pendant vertex is a node v linked to the graph by a single edge ($\deg(v) = 1$), while an isolated vertex has no connections ($\deg(v) = 0$). The average degree of graph $G = (V, E)$ is computed as:

$$\langle k \rangle = \frac{2M}{n} \tag{2}$$

where $n = |V|$ and $M = |E|$.

Several metrics based on degree centrality are employed to assess radial-shaped graphs and were introduced in [30] for validating radial-shaped synthetic distribution grids.

The branching rate, denoted as b_r, offers insight into how much a network branches out. It is determined by the ratio of vertices with a degree of 3 or higher to the total number of vertices in the graph:

$$b_r = \frac{\left|\left\{\forall v_i \middle| \deg(v_i) \geq 3\right\}\right|}{n} \tag{3}$$

Assortative mixing in a network examines the correlation between nodes with similar degrees. The degree-based assortativity property is assessed using the Pearson correlation coefficient ρ:

$$\rho = \frac{M^{-1}\sum_i j_i k_i - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}{M^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2} \quad (4)$$

where j_i and k_i represent the degrees of the vertices at the ends of the i^{th} edge, with $i = 1, ..., M$.

Here, negative values indicate a disassortative mixing pattern (nodes with high degrees tend to connect to nodes with low degrees), while positive values indicate an assortative mixing pattern (nodes with similar degrees tend to connect).

3- Erdős–Rényi Random Graph Model

The idea of random graphs emerged as a distinct area in graph theory [32], originating from the pioneering work of Solomonoff and Rapoport in 1951 [48]. They introduced a model for directed graphs with fixed out-degree and randomly chosen connections to other nodes, known as the "Random Net." However, significant progress in the study of random graphs was not made until the late 1950s and early 1960s. Paul Erdős and Alfréd Rényi played a vital role in advancing the theory of random graphs through a series of papers published between 1959 and 1968 [33, 49-54]. Their research delved into various aspects such as evolution, strength, asymmetry, existence, and randomness within random graph models, laying the groundwork for further exploration in this field. Concurrently, Edgar Nelson Gilbert independently introduced his own model for generating random graphs around the same time as Erdős and Rényi's work [34]. Gilbert's model offered an alternative perspective on studying randomness in graph structures. While both models are commonly referred to as Erdős–Rényi, this study will use the term Erdős for the former and Gilbert for the latter.

3- 1- Erdős Random Graph Model

The Erdős model, represented as $G(n, M)$, gives all graphs with exactly M edges the same likelihood of occurring. The total count of different elements in $G(n, M)$ is given by $\binom{L}{M}$, where L is the total number of potential edges that can be formed. Each individual graph within this model has an occurrence probability of $\frac{1}{\binom{L}{M}}$.

To create a random graph utilizing this model, a single loop procedure is employed to generate M edges as outlined in Algorithm 1. The result of this procedure is an adjacency matrix **A**, and the overall count of edges can be determined by tallying the non-zero elements within **A** through the utilization of the SUM function detailed in the algorithm [22].

3- 2- Gilbert Random Graph Model

The Gilbert random graph model is denoted by $G(n, p)$, which is a commonly studied random graph model. In this model, each edge occurs independently with a probability p. The probability of obtaining a random graph with M edges is given by $p^M(1\text{-}p)^{N\text{-}M}$, where N is the number of possible combinations. In this model, the average degree can be approximated as $\langle k \rangle \approx n.p$ [34].

Algorithm 1. Erdős Random Graph Model

Input: n : number of vertices
M : number of edges
Output: **A** : adjacency matrix of generated random graph

function ER(n, M)
  Initialize: $\mathbf{A} = \left[\ 0\ \right]_{n\times n}$
  while SUM(SUM(**A**)) $< 2M$ **do**
    $i, j \sim \lceil \mathrm{U}(0, n) \rceil$ ▷ random integers $\in [1, n]$
    if $i = j \ \vee \ a_{ij} > 0$ **then**
      continue
    end if
    $a_{ij}, a_{ji} \leftarrow 1$
  end while
  return A
end function

ADDITIONAL FUNCTIONS

S=SUM(**A**): S is the sum of array elements. If **A** is a matrix, S will be a row vector of the sums of each column. If **A** is a vector, S will be the sum of the elements.

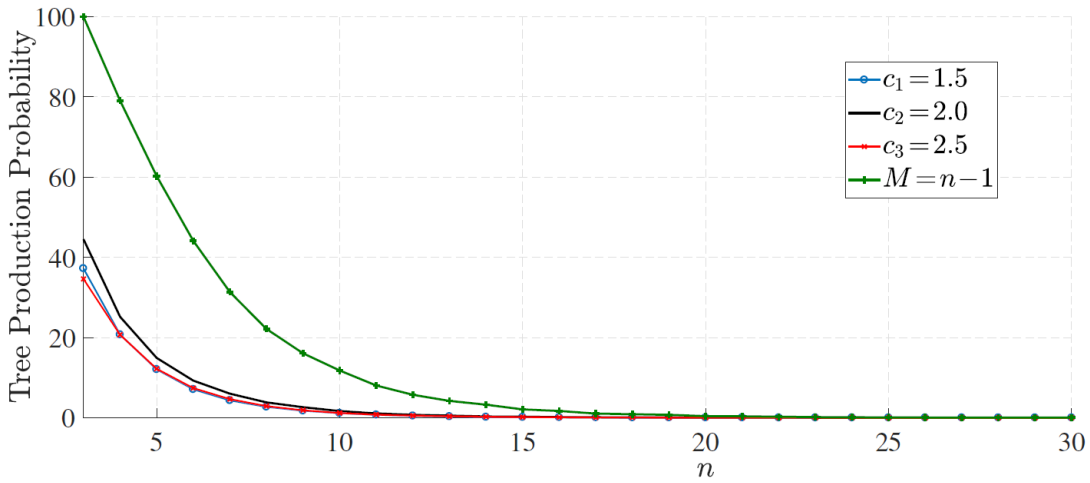


Fig. 1. Probability of producing connected tree by G(n, p=c/n) model with c=1.5, 2.0, and 2.5, and by G(n, M=n − 1) [28].

The behavior of $G(n, p)$ in its evolution depends on the value of $c \triangleq n.p$. In the subcritical regime, if $n.p$ <1, the components of $G(n, p)$ are likely to be small trees. In the critical regime, if $n.p$=1, the largest component is likely to have size $n^{2/3}$. In the supercritical regime, if $n.p$ >1, there is likely to be a unique giant component. Furthermore, in the sparse regime, which occurs as n increases, the properties of the components of $G(n, p)$ differ from those in the subcritical, critical, and supercritical regimes. It is important to highlight that the various regimes outlined in the Gilbert model can also be extended to this model, where $c = 2M/n$ [22].

The Gilbert model generates graphs by selecting edges between vertices using two loops and a random number generator. If the generated random number θ is less than or equal to p, an edge is selected. Please refer to Algorithm 2 in the [28] for further details.

3- 3- Analyzing Two Erdős–Rényi Models in Tree Production

The comprehensive analysis in [28] focuses on generating random graphs in tree structures. It emphasizes that the Erdős model, using G(n, M=n-1) as input parameters, has a higher likelihood of generating trees compared to the Gilbert model using G(n, 2/n). This contrast is illustrated in Figure 1.

We generated 10,000 random graphs for three values of c (1.5, 2.0, and 2.5) and n ranging from 3 to 30 to analyze the tree generation probability. Similarly, for the Erdős model with adjustment $G(n, M=n-1)$, we examined 10,000 random graphs for n ranging from 3 to 30. The results, depicted in Figure 1, show that the highest probability of generating a tree in the Gilbert model $G(n, p)$ occurs at $c = 2.0$. For $n > 15$, the tree generation probability in this model drops to zero, while in the Erdős model $G(n, M)$, this happens for $n > 20$. The probability of forming a tree in the Erdős model is notably higher than the best case of the Gilbert model ($c = 2$) due to the equal probability distribution of graphs with M edges in the Erdős model, particularly when $M = n - 1$, compared to the approximate edge determination in the Gilbert model.

3- 4- SDG Generation with Erdős Model

The objective of [28] was to create random graphs resembling distribution networks using the Erdős–Rényi random graph model. To accomplish this, we initially analyzed the structure of distribution networks and extracted their topology. Subsequently, leveraging the insights from previous sections and the network structures, we examined the potential of the Erdős–Rényi random graph models (Erdős itself and Gilbert) in replicating radial distribution network structures. Finally, an algorithm based on the Erdős model was introduced in [28] to generate synthetic distribution graphs.

Algorithm 2 illustrates the proposed algorithm, with the sole input being n, representing the number of nodes in the synthetic distribution grid. The algorithm iterates the process of creating an Erdős random graph with the parameters $G(n, M=n-1)$ until a connected graph is established. The connectivity verification is conducted using Algorithm 1 of [28], which assesses the number of components. Specifically, if $C(G)$ equals one, indicating the formation of a tree, the generation process concludes.

The data presented in Fig. 1 suggests that the Erdős model can potentially construct a tree up to n=20. However, upon implementing Algorithm 2 in [28], the algorithm demonstrated success in constructing trees up to n=35. In the study, various SDGs of sizes 15, 20, and 35 were created using Algorithm

Algorithm 2. Proposed Algorithm to Generate SDGs based on Erdős Random Graph Model

Input: n : number of vertices
Output: $\mathbf{A}$: adjacency matrix of generated SDG

Initialize: $\mathbf{A} = [\,0\,]_{n\times n}$
$\mathbf{catch_{tree}} \leftarrow$ `false`
while $\neg\mathbf{catch_{tree}}$ **do**
  $\mathbf{A} \leftarrow \mathrm{ER}(n, M=n-1)$ ▷ generate random graph $G(n, M=n-1)$
  $C(G) \leftarrow \mathrm{COMPONENT}(\mathbf{A})$
  if $C(G) = 1$ **then**
    $\mathbf{catch_{tree}} \leftarrow$ `true`
  end if
end while
return $\mathbf{A}$

2. The average degree, maximum degree, branching rate, and clustering coefficient of these synthetic grids were then compared with real-world scenarios, revealing significant similarities.

The research in [28] primarily focused on generating SDGs using the established Erdős model, which, despite its widespread use, has size limitations. Moreover, within this model, there are no additional parameters beyond *n* to adjust the structure of the resultant tree, such as maximum degree and branching rate. To address these limitations and build upon the findings of [28], a new algorithm based on the renowned Barabási-Albert random graph model is proposed in this study. This novel algorithm, detailed in the subsequent section, aims to enhance the generation of SDGs by introducing additional control over the structural characteristics of the generated graphs.

4- Proposed Method: SDG Generation with the Barabási-Albert Model

High-level description: The primary aim of this study is to introduce algorithms capable of generating Synthetic Distribution Grids (SDGs) in the form of trees, i.e., synthetic radial distribution grids. The initial algorithm, as presented in [28] and fully described in the previous section, relies on the Erdős random model. While effective, this algorithm is constrained, particularly in terms of SDG size. To address this limitation, our novel algorithm is based on the Barabási-Albert model. It provides a solution without the size restrictions of the previous method and includes a tuning parameter to modify the structure of the generated SDGs.

The Barabási-Albert (BA) random graph model, proposed by Albert-László Barabási and Réka Albert in 1999, revolutionized the study of complex networks [55]. This model aimed to explain the emergence of scale-free networks, where a few nodes have a disproportionately large number of connections compared to the majority of nodes with only a few connections. In the BA model, the network grows over time by adding nodes that preferentially attach to existing nodes with high degrees. This "rich get richer" mechanism leads to the formation of hubs, highly connected nodes that play a crucial role in the network's structure. As new nodes join the network, they are more likely to connect to well-connected nodes, reinforcing the network's scale-free nature. The BA model captures the idea that real-world networks, such as social networks, the World Wide Web [56], power grids [55], and biological networks [57], exhibit a power-law distribution of node degrees. This means that while most nodes have only a few connections, a small number of nodes act as hubs with a large number of connections. The algorithm for generating a Barabasi-Albert random graph can be described as follows:

1. Start with a small number of nodes, typically m_0, fully connected to each other to form a seed graph.

2. Add a new node at each time step t and connect it to $m \leq m_0$ existing nodes. The probability p_i that the i-th existing node will be chosen for connection by the new node is proportional to its degree $\mathbf{C}_d(i)$:

$$p_i = \frac{\mathbf{C}_d(i)}{\sum_{k=1}^{n} \mathbf{C}_d(k)} \tag{5}$$

3. Repeat step 2 until the desired number of nodes is reached.

The resulting graph is a scale-free network, meaning that the degree distribution follows a power-law distribution, as follows:

$$P(k) \propto k^{-\gamma} \tag{6}$$

where γ is a constant that represents the degree exponent.

The Barabasi-Albert model offers flexibility in generating various graph structures, including trees. To produce a tree shape graph in the first step, only two nodes, v_1 and v_2, are inserted in m_0. Subsequently, new nodes are added and connected to existing nodes using the preferential attachment described in Eq. 5. In each iteration, a new node is connected to existing nodes based on their connectivity, determined by the degree vector calculation. This iterative process fosters the growth of a tree structure where nodes are hierarchically connected like branches of a tree. Indeed, the initialization step and the connection of new nodes from the third node onwards contribute to shaping the resulting graph into a tree-like structure.

In the original version of BA model described above, the probability of connecting to a node is influenced by its current degree, favoring nodes with higher degrees, described in Eq. 5. Following this approach, there is no any control on the structure of generated tree. Hence, we have introduced a new parameter, named "richness index" (denoted by *ri*) to modify the preferential attachment mechanism, as follow:

$$p_i = \frac{\left(\mathbf{C}_d(i)\right)^{ri}}{\sum_{k=1}^{n} \mathbf{C}_d(k)} \tag{7}$$

where, *ri* is the richness index and can be varied from 0 to 1.

By introducing this tuning parameter, i.e. richness index, the proposed B-A model can be adapted to generate tree structures by modifying the preferential attachment mechanism. The richness index influences the probability of connecting to existing nodes during each step of the process, allowing for control over the growth patterns of the resulting SDG. Algorithm 3 outlines this process, which is described in 5 steps as follows:

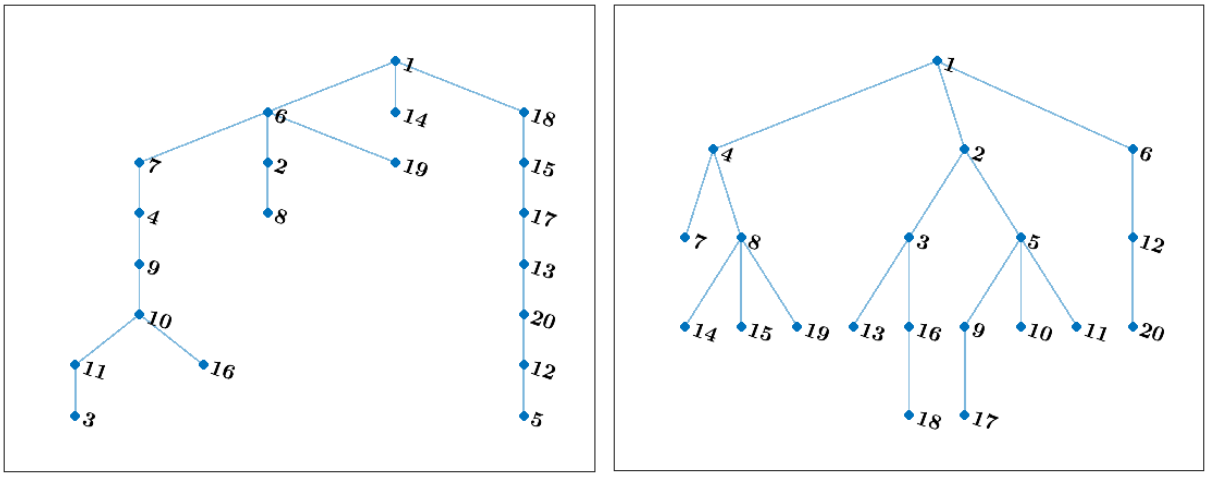
1. Initialization: Begin by initializing an adjacency matrix **A** of size $n \times n$ filled with zeros. Connect the first two nodes (v_1 and v_2) in the graph to start the network.

2. Iterative Addition of Nodes: For each new node from 3 to n, calculate the degree vector $\mathbf{C}_d$ by summing the rows of

Algorithm 3. Proposed Algorithm to Generate SDGs based on Barabási -Albert Random Graph Model

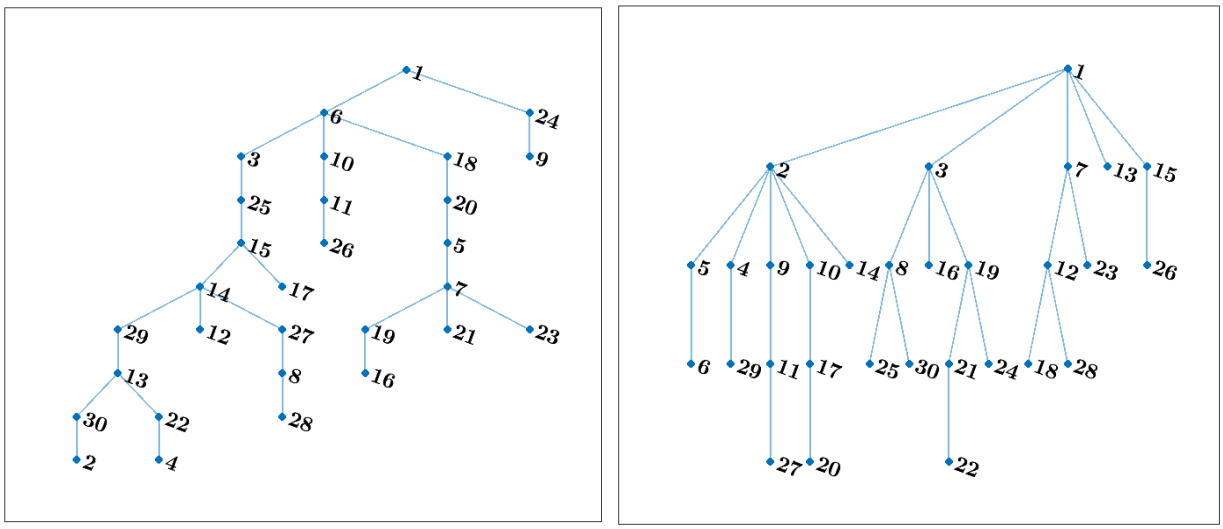
Input:
n : number of vertices;
ri : richness index
Output:
A : adjacency matrix of generated random graph

function GENERATE-SDN-BA(n, ri)
  Initialize: $\mathbf{A} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{n \times n}$ ▷ initialize adjacency matrix
  $\mathbf{A}(1,2) \leftarrow 1$; $\mathbf{A}(2,1) \leftarrow 1$ ▷ connect initial nodes
  for $n_{new} \leftarrow 3$ to n **do**
    $\mathbf{C}_d \leftarrow \text{SUM}(\mathbf{A}, 1)$ ▷ calculate degree vector
    $p_v \leftarrow \frac{\left(\mathbf{C}_d(v)\right)^{ri}}{\sum_{i=1}^{n} \mathbf{C}_d(i)}$ ▷ probability of connecting to node v
    $N = \{1, 2, \ldots, n_{new} - 1\}$ ▷ set of candidates
    $p = \{p_1, p_2, \ldots, p_{n_{new}-1}\}$ ▷ selection probability
    $n_{conn} \sim \text{Mult}(N; p)$
    ▷ select n_{conn} based on preferential attachment
    $\mathbf{A}(n_{new}, n_{conn}) \leftarrow 1$
    ▷ connect n_{new} to selected existing node n_{conn}
    $\mathbf{A}(n_{conn}, n_{new}) \leftarrow 1$
    ▷ ensure symmetry of the adjacency matrix
  end for
  return A
end function



(a) (b)

Fig. 2. Two versions of SDG with the size of n=20 generated by two proposed algorithms (a) Erdos and (b) Barabási-Albert.



(a) (b)

Fig. 3. Two versions of SDG with the size of n=30 generated by two proposed algorithms (a) Erdos and (b) Barabási-Albert.

the adjacency matrix.

3. Preferential Attachment: Calculate the probability p_v of connecting to a node v based on its degree raised to the power of the richness index ri. Normalize these probabilities to create a selection probability p for existing nodes.

4. Connect New Node: Randomly select an existing node to connect to based on preferential attachment. Update the adjacency matrix **A** by connecting the new node to the selected existing node and ensuring symmetry in the matrix.

5. Return Result: Return the final adjacency matrix **A**, representing the generated SDG following the Barabasi-Albert model with the specified parameters.

In Algorithm 3, through the adjustment of the richness index, the algorithm ensures that new nodes are added to the tree in a strategic manner that promotes a SDG, while still preserving the scale-free properties characteristic of B-A model. This fine-tuning allows for the creation of scale-free SDGs with desired characteristics. We will show this capability in the next sections.

Table 1. Degree Metrics and Pearson Coefficient for Produced SGDs

SDG Size	Model	Average Degree	Max Degree	Branching Rate	Pearson Coefficient
20	Erdős	1.9000	4	0.1500	-0.2160
	B-A	1.9000	4	0.3000	-0.4074
30	Erdős	1.9333	4	0.1667	-0.2959
	B-A	1.9333	6	0.2333	-0.0755
50	B-A	1.9600	6	0.2800	-0.1765
100	B-A	1.9800	7	0.2600	-0.0195
500	B-A	1.9960	10	0.2380	-0.0279
1000	B-A	1.9980	11	0.2390	+0.0091

5- Examples of Applications

In this section, to demonstrate the effectiveness of the proposed B-A algorithm in the generation of SDGs, we have used this model and our previous Erdős model in three different applications. Firstly, we assess the performance of both algorithms (i.e. Erdős and B-A) in creating SDGs of varying sizes, highlighting their capabilities in scaling effectively. Furthermore, we explore the versatility of the B-A algorithm in constructing SDGs with diverse structures by leveraging the richness index (*ri*) parameter. Through this exploration, we demonstrate how adjusting the *ri* value allows for the generation of SDGs with different topological characteristics, showcasing the flexibility and adaptability of the B-A algorithm in capturing a wide range of network structures.

5- 1- SDG Generation with Different Sizes

In this section, 5 SDGs are created by the proposed algorithms, displayed in Fig. 2 to Fig. 4, with the B-A model set at *ri*=1, that is, the proposed B-A model is considered to be the original one. Results confirm the successful generation of synthetic radial distribution grids by both algorithms.

The degree-based parameters and Pearson coefficients of these SDGs, alongside three others of sizes 100, 500, and 1000, are detailed in Table 1. The average degree values align closely with the theoretical formula for tree structures, indicating the effective generation of tree-like graphs by both models. Variations in maximum degree values show the original version of the B-A model tends to produce higher-degree nodes, likely due to its preferential attachment mechanism set at *ri*=1.

The branching rate increases with tree size for both models, with the original B-A model consistently exhibiting higher rates, indicating a larger proportion of nodes with degrees ≥ 3. Pearson coefficients vary across sizes: smaller grids show negative coefficients, shifting to more neutral or slightly positive values with increased size, particularly in very large trees like n=1000 where most neighbor nodes have degree values of 2.

5- 2- Proportion of Branching Rate (b_r) and Richness Index (*ri*)

In this section, we analyze the impact of varying the richness index (*ri*) on the branching rates of SDGs. By generating 10,000 SDGs for sizes 25 and 100, with *ri* ranging from 0 to 1 in steps of 0.01, we observe changes in the branching rates displayed in Figure 5 by calculating the mean value of 10,000 obtained values of b_r for each B-A-SDG(*n*,*ri*). The results indicate that an increase in *ri* corresponds to a decrease in branching rate (b_r). This inverse relationship can be attributed to the equal probability of selecting nodes at *ri*=0, resulting in more nodes having higher degrees. In other words, the same probability for every node in the tree leads to the form of more nodes with higher degrees. For instance in Figure 7 (a), the uniform distribution of 2*M*=48 connections between *n*=25 nodes (*ri* set to 0) leads to the form of only 5 nodes with higher degrees (v_1 to v_4 and v_6), while the preferential attachment mechanism (*ri*=1) lead to create 6 high degree nodes (v_1, v_3, v_6, v_8, v_{14}, v_{19}) in Figure 7 (b). In summary, the introduction of the new parameter (*ri*) in the proposed B-A algorithm enables the transformation of generated SDGs from "*rural*" to "*urban*" configurations.

Additionally in SDGs of smaller sizes (*n*=25), we observe lower branching rates at *ri*=0, due to limited candidates for nodes with high degrees. Conversely, in SDGs following the original preferential attachment mechanism at *ri*=1, larger sizes (*n*=100) lead to increased branching rates as there are more candidates available for attachment mechanisms.

5- 3- Proportion of μ-PMU Coverage (%PMU) and Richness Index (*ri*)

The μ-PMU device, equipped with the ability to measure voltage and current phasors across a bus in the distribution grid, synchronized with GPS clock, enables observation of the bus and its neighboring buses. The μ-PMU placement problem

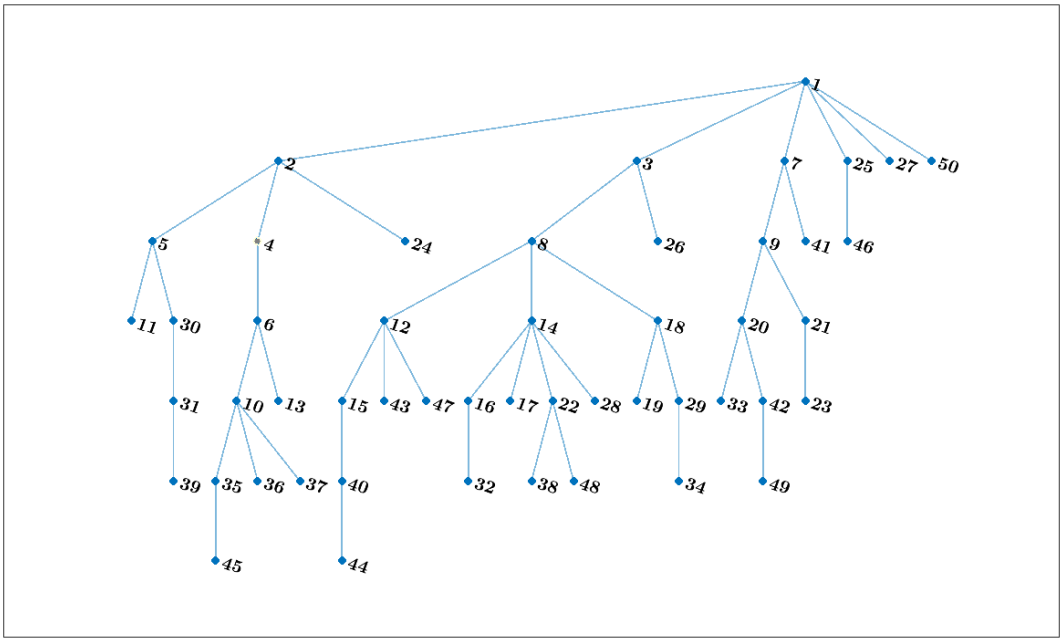


Fig. 4. The SDG with the size of n=50 generated by Barabási-Albert model.

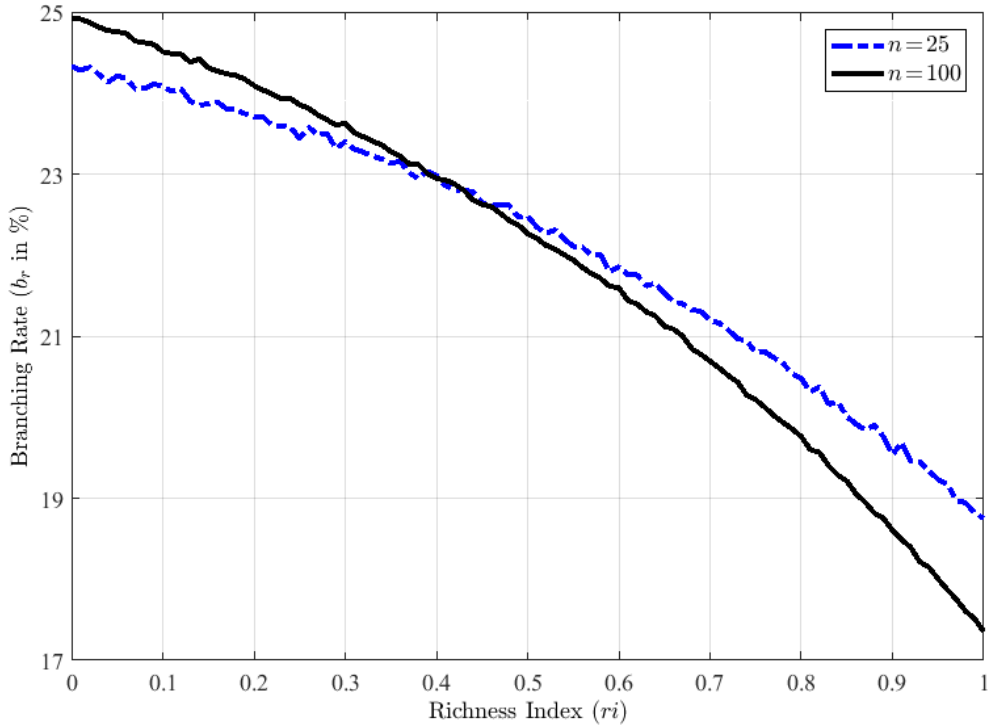


Fig. 5. The changes in branching rate (br) based on the richness index (ri) in the B-A model for SDGs with two different sizes, n=25 (blue-dashed line) and n=100 (black line).

focuses on determining the minimum number of μ-PMUs required for complete observability of the distribution grid, known as μ-PMU coverage (%**PMU**) [58]. This placement problem can be formulated as a linear programming problem as follows [21, 22]:

$$\min \sum_{i=1}^{n} x_i \quad \text{s.t.} \begin{cases} \mathbf{A}^{+}.\mathbf{X} \geq \hat{1} \\ x_i \in \{0,1\} \end{cases} \tag{8}$$

where, **X** is a Boolean decision vector used to specify the location of a μ-PMU, with x_i=1 indicating that the μ-PMU will be installed in the i-th bus. The generalized adjacency matrix $\mathbf{A}^{+}$ is defined as $\mathbf{A}^{+}=\mathbf{A}+\mathbf{I}_{n\times n}$.

While our recent research extensively explores the impact of transmission grid structures on %**PMU** values [22], this section delves into investigating how the structure of radial distribution grids influences μ-PMU coverage, represented as %**PMU**.

To investigate how structural changes affect the %PMU value in radial grids of sizes 25 and 100, we generated 100 SDGs with varying values of the parameter *ri* (ranging from 0 to 1 in steps of 0.01). We then calculated the %**PMU** for each B-A-SDG(n, ri) by using linear programming as described in Eq. (8) and reported the mean value of 100 obtained values of %**PMU** in Figure 6. The figure illustrates that as *ri* increases from 0 to 1, there is a decrease in the %**PMU**, indicating a reduction in the number of μ-PMUs as the richness index increases. In light of the findings from Figure 5, it can be inferred that an increase in the branching rate leads to a higher %**PMU**, which may initially seem counterintuitive.

This observation can be explained by the fact that in urban SDGs with higher branching rates, the number of branches with depths of 2 or 3 expands in the main branches (feeders), necessitating a greater number of μ-PMUs for full observability. This phenomenon is illustrated in Figure 7; Figure (a) represents a rural SDG with *ri*=0, requiring 8 μ-PMUs for full observability, while Figure (b) depicts an urban SDG with *ri*=1, necessitating 9 μ-PMUs for full observability.

As a final conclusion, disregarding the minor differences between the graphs of n=25 and n=100 shown in Figures 5 and 6, it can be deduced that our proposed B-A model demonstrates the characteristics of a scale-free model with respect to the variation in the richness index.

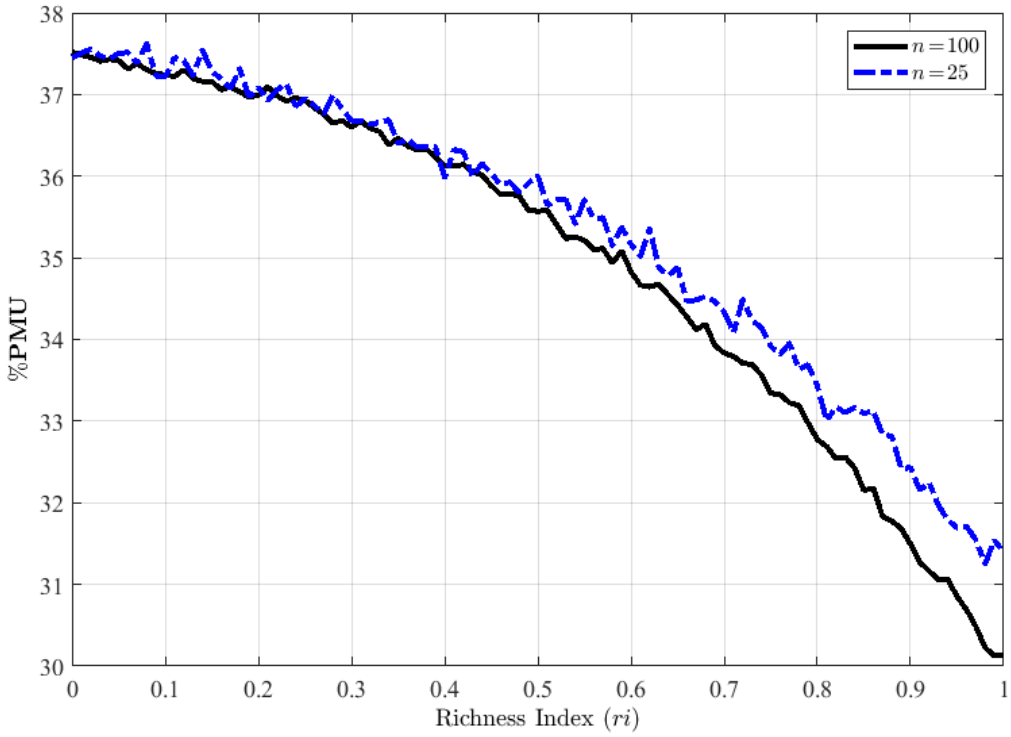
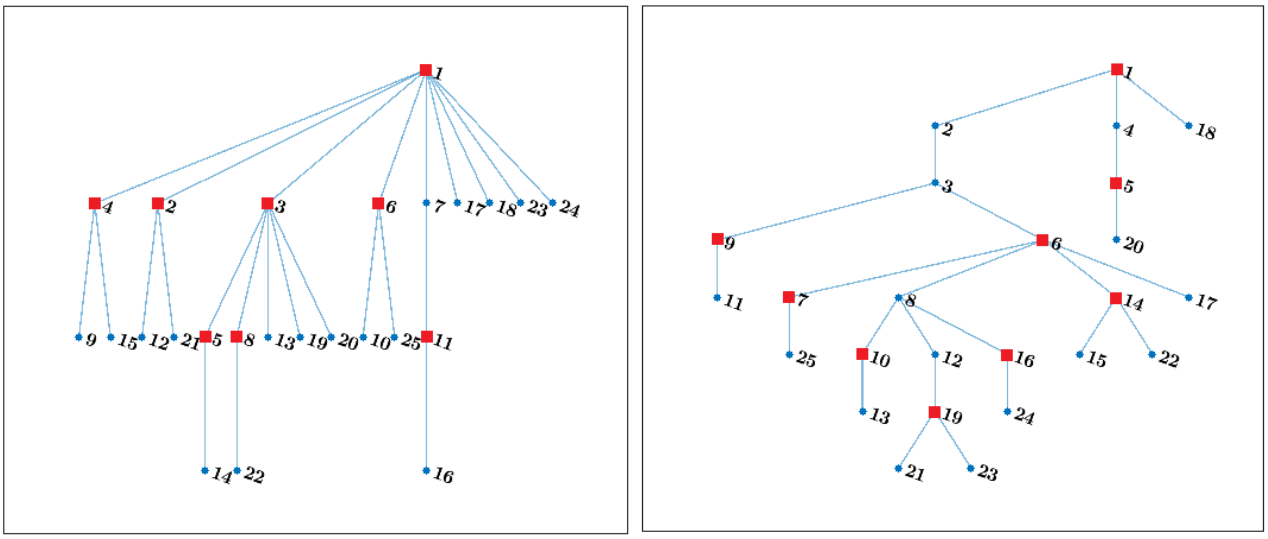


Fig. 6. The changes in PMU coverage (%PMU) based on the richness index (ri) in the B-A model for SDGs with two different sizes, n=25 (blue-dashed line) and n=100 (black line).



(a) ri=0 , b_r=0.2 , No. of PMUs=8 (32%)

(b) ri=1, b_r=0.24 , No. of PMUs=9 (36%)

Fig. 7. Two versions of B-A SDGs with the size of n=25 generated by B-A algorithm (a) ri=0, br=0.2, %PMU=32% - and (b) ri=1, br=0.24, %PMU=36%. The µ-PMU locations are specified by red squares.

6- Conclusion

In this extended study, we delve into the generation of synthetic distribution grids (SDGs) using innovative algorithms based on the Erdős and Barabási-Albert (B-A) random graph models. Building upon prior work, our research aims to overcome the size limitations and lack of adjustable structural parameters in the Erdős model by introducing the B-A model, offering enhanced control over the structural characteristics of the generated graphs through the introduction of a novel tuning parameter known as the "richness index". Through the production of five SDGs of varying sizes and the analysis of degree-based parameters and Pearson coefficients, we showcase the efficacy of both algorithms in generating synthetic radial distribution grids. The comparison between the Erdős and the original B-A models reveals variations in maximum degree values, branching rates, and mixing patterns, with the original B-A model demonstrating a preference for nodes with higher degrees. Notably, as tree size increases, both models exhibit differing Pearson coefficients, reflecting changes in mixing patterns from disassortative to more neutral or slightly positive. Moreover, the proposed B-A model is analyzed concerning the variation of richness index in branching rate and µ-PMU placement. The results indicate that as the richness index increases, both the branching rate and PMU coverage decrease. These two analyses also serve to validate the scale-free characteristic of the method, demonstrating its effectiveness in creating a well-connected SDG. Future research directions may delve into the unexplored area of vertex renumbering to enhance the refinement of algorithms for generating SDGs.

Terms and Symbols

In the context of this paper, italic letters denote integers and scalars, such as n and p. Matrices and vectors are symbolized by capital, bold, and non-italic letters, for example, $\mathbf{A}$ and $\mathbf{C}_d$. Lowercase, italic, and indexed letters indicate arrays of matrices, like a_{ij}. Additionally, $\mathbf{C}_d(v)$ specifically designates the v-th element of the vector $\mathbf{C}_d$.

References

[1] G. Kalt, P. Thunshirn, H. Haberl, A global inventory of electricity infrastructures from 1980 to 2017: Country-level data on power plants, grids, and transformers, Data in Brief, 38 (2021) 107351.

[2] M. Shahraeini, S. Alishahi, A survey on information and communication technology (ICT) applications in distribution systems, in: 21st International Conference and Exhibition on Electricity Distribution, 2011.

[3] M. Shahraeini, Z. Farmani, Designing of Communication Systems in Advanced Metering Infrastructure (AMI) using Wi-Fi Offloading Technology, in: 2019 9th

International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, 2019.

[4] G.A. Pagani, M. Aiello, The Power Grid as a complex network: A survey, Physica A: Statistical Mechanics and its Applications, 392(11) (2013) 2688-2700.

[5] R. Espejo, S. Lumbreras, A. Ramos, A Complex-Network Approach to the Generation of Synthetic Power Transmission Networks, IEEE Systems Journal, 13(3) (2019) 3050-3058.

[6] H. Labrini, Graph-based model for distribution systems: Application to planning problem, University of Waterloo, 2015.

[7] H. Willis, Power distribution planning reference book, New York: M, in, Dekker, 2004.

[8] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Critical points and transitions in an electric power transmission model for cascading failure blackouts, Chaos: An interdisciplinary journal of nonlinear science, 12(4) (2002) 985-994.

[9] M. Parashar, J.S. Thorp, C.E. Seyler, Continuum modeling of electromechanical dynamics in large-scale power systems, IEEE Transactions on Circuits and Systems I: Regular Papers, 51(9) (2004) 1848-1858.

[10] Z. Wang, A. Scaglione, R.J. Thomas, Generating statistically correct random topologies for testing smart grid communication and control networks, IEEE transactions on Smart Grid, 1(1) (2010) 28-39.

[11] S. Pahwa, C. Scoglio, A. Scala, Abruptness of Cascade Failures in Power Grids, Scientific Reports, 4(1) (2014) 3694.

[12] J. Hu, L. Sankar, D.J. Mir, Cluster-and-connect: An algorithmic approach to generating synthetic electric power network graphs, in: 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2015, pp. 223-230.

[13] S.H. Elyas, Z. Wang, Improved synthetic power grid modeling with correlated bus type assignments, IEEE Transactions on Power Systems, 32(5) (2016) 3391-3402.

[14] K.M. Gegner, A.B. Birchfield, T. Xu, K.S. Shetye, T.J. Overbye, A methodology for the creation of geographically realistic synthetic power flow models, in: 2016 IEEE Power and Energy Conference at Illinois (PECI), IEEE, 2016, pp. 1-6.

[15] S. Ma, Y. Yu, L. Zhao, Dual-stage constructed random graph algorithm to generate random graphs featuring the same topological characteristics with power grids, Journal of Modern Power Systems and Clean Energy, 5(5) (2017) 683-695.

[16] A.B. Birchfield, T. Xu, T.J. Overbye, Power flow convergence and reactive power planning in the creation of large synthetic grids, IEEE Transactions on Power Systems, 33(6) (2018) 6667-6674.

[17] E. Schweitzer, A. Scaglione, A mathematical programming solution for automatic generation of synthetic power flow cases, IEEE Transactions on Power Systems, 34(1) (2018) 729-741.

[18] S. Soltan, A. Loh, G. Zussman, A learning-based method for generating synthetic power grids, IEEE Systems Journal, 13(1) (2018) 625-634.

[19] T. Xu, A.B. Birchfield, T.J. Overbye, Modeling, tuning, and validating system dynamics in synthetic electric grids, IEEE Transactions on Power Systems, 33(6) (2018) 6501-6509.

[20] S.G. Aksoy, E. Purvine, E. Cotilla-Sanchez, M. Halappanavar, A generative graph model for electrical infrastructure networks, Journal of complex networks, 7(1) (2019) 128-162.

[21] M. Shahraeini, A. Alvandi, S. Khormali, Behavior analysis of random power graphs for optimal PMU placement in smart grids, in: 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, 2020, pp. 107-112.

[22] M. Shahraeini, Modified Erdős–Rényi Random Graph Model for Generating Synthetic Power Grids, IEEE Systems Journal, (2023).

[23] V.K. Krishnan, B.S. Palmintier, B.S. Hodge, E.T. Hale, T. Elgindy, B. Bugbee, M.N. Rossol, A.J. Lopez, D. Krishnamurthy, C. Vergara, Smart-ds: Synthetic models for advanced, realistic testing: Distribution systems and scenarios, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017.

[24] E. Schweitzer, A. Scaglione, A. Monti, G.A. Pagani, Automated generation algorithm for synthetic medium voltage radial distribution systems, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 7(2) (2017) 271-284.

[25] A. Trpovski, D. Recalde, T. Hamacher, Synthetic distribution grid generation using power system planning: Case study of singapore, in: 2018 53rd International Universities Power Engineering Conference (UPEC), IEEE, 2018, pp. 1-6.

[26] C. Mateo, F. Postigo, F. de Cuadra, T.G. San Roman, T. Elgindy, P. Duenas, B.-M. Hodge, V. Krishnan, B. Palmintier, Building large-scale US synthetic electric distribution system models, IEEE Transactions on Smart Grid, 11(6) (2020) 5301-5313.

[27] R. Yan, Y. Yuan, Z. Wang, G. Geng, Q. Jiang, Active distribution system synthesis via unbalanced graph generative adversarial network, IEEE Transactions on Power Systems, 38(5) (2022) 4293-4307.

[28] M. Shahraeini, An Algorithm for Generating Synthetic Distribution Grids based on Edrös-Rényi Random Graph Model, in: 2023 13th Smart Grid Conference (SGC), IEEE, 2023, pp. 1-8.

[29] S. Abeysinghe, J. Wu, M. Sooriyabandara, M. Abeysekera, T. Xu, C. Wang, Topological properties of medium voltage electricity distribution networks,

Applied Energy, 210 (2018) 1101-1112.

[30] V. Krishnan, B. Bugbee, T. Elgindy, C. Mateo, P. Duenas, F. Postigo, J.-S. Lacroix, T.G. San Roman, B. Palmintier, Validation of synthetic US electric power distribution system data sets, IEEE Transactions on Smart Grid, 11(5) (2020) 4477-4489.

[31] M. Ali, K. Prakash, C. Macana, M. Raza, A.K. Bashir, H. Pota, Modeling synthetic power distribution network and datasets with industrial validation, Journal of Industrial Information Integration, 31 (2023) 100407.

[32] B. Bollobás, Random Graphs, Cambridge University Press, Cambridge, UK, (2001).

[33] P. Erdos, A. Renyi, On Random Graphs. Publicationes Mathematicae (Debrecen), Vol. 6, (1959).

[34] E.N. Gilbert, Random graphs, The Annals of Mathematical Statistics, 30(4) (1959) 1141-1144.

[35] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world'networks, nature, 393(6684) (1998) 440-442.

[36] X. Hu, Z.-L. Zhang, Towards a graph signal processing framework for modeling power system dynamics, in: 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, 2019, pp. 1-5.

[37] C.M. Domingo, T.G. San Roman, A. Sanchez-Miralles, J.P.P. Gonzalez, A.C. Martinez, A reference network model for large-scale distribution planning with automatic street map generation, IEEE Transactions on Power Systems, 26(1) (2010) 190-197.

[38] T. Gómez, C. Mateo, Á. Sánchez, P. Frías, R. Cossent, Reference network models: A computational tool for planning and designing large-scale smart electricity distribution grids, High performance computing in power and energy systems, (2013) 247-279.

[39] E. Schweitzer, K. Togawa, T. Schloesser, A. Monti, A Matlab GUI for the generation of distribution grid models, in: International ETG Congress 2015; Die Energiewende-Blueprints for the new energy age, VDE, 2015, pp. 1-6.

[40] A. Bidel, T. Schelo, T. Hamacher, Synthetic distribution grid generation based on high resolution spatial data, in: 2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), IEEE, 2021, pp. 1-6.

[41] R. Meyur, A. Vullikanti, S. Swarup, H.S. Mortveit, V. Centeno, A. Phadke, H.V. Poor, M.V. Marathe, Ensembles of realistic power distribution networks, Proceedings of the National Academy of Sciences, 119(42) (2022) e2205772119.

[42] M.H. Mohammadi, K. Saleh, Synthetic benchmarks for power systems, IEEE Access, 9 (2021) 162706-162730.

[43] F.B. Dos Reis, R. Tonkoski, B.P. Bhattarai, T.M. Hansen, A real-world test distribution system with appliance-level load data for demand response and transactive energy studies, IEEE Access, 9 (2021) 149506-149519.

[44] L. Wang, J. Halvorsen, S. Pannala, A. Srivastava, A.H. Gebremedhin, N.N. Schulz, CP-SyNet: A tool for generating customised cyber-power synthetic network for distribution systems with distributed energy resources, IET Smart Grid, 5(6) (2022) 463-477.

[45] R.M. Hoff, M.V. Chester, Preparing infrastructure for surprise: fusing synthetic network, interdependency, and cascading failure models, Environmental Research: Infrastructure and Sustainability, 3(2) (2023) 025009.

[46] A. Trpovski, Synthetic Power Grid Development: Optimal Power System Planning-based Approach and Its Application for Singapore Case Study, Technische Universität München, 2023.

[47] R. Meyur, K. Lyman, B. Krishnamoorthy, M. Halappanavar, Structural Validation of Synthetic Power Distribution Networks Using the Multiscale Flat Norm, in: International Conference on Computational Science, Springer, 2023, pp. 55-69.

[48] R. Solomonoff, A. Rapoport, Connectivity of random nets, The bulletin of mathematical biophysics, 13 (1951) 107-117.

[49] P. Erdos, A. Rényi, On the evolution of random graphs, Publ. math. inst. hung. acad. sci, 5(1) (1960) 17-60.

[50] P. Erdos, On the evolution of random graphs, Bulletin of the Institute of International Statistics, 38 (1961) 343-347.

[51] P. Erdős, A. Rényi, On the strength of connectedness of a random graph, Acta Mathematica Hungarica, 12(1) (1961) 261-267.

[52] P. Erdős, A. Rényi, Asymmetric graphs, Acta Mathematica Academiae Scientiarum Hungarica, 14(3) (1963) 295-315.

[53] P. Erdős, A. Rényi, On the existence of a factor of degree one of a connected random graph, Acta Math. Acad. Sci. Hungar, 17(359-368) (1966) 192.

[54] P. Erdos, &A. Renyi-On random matrices, II, Studia Sci. Math. Hungar, (1968) 3,459-464.

[55] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, science, 286(5439) (1999) 509-512.

[56] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Graph structure in the web, Computer networks, 33(1-6) (2000) 309-320.

[57] A.-L. Barabási, Z.N. Oltvai, S. Wuchty, Characteristics of biological networks, Complex networks, (2004) 443-457.

[58] K. Hassini, A. Fakhfakh, F. Derbel, Optimal Placement of μ PMUs in Distribution Networks with Adaptive Topology Changes, Energies, 16(20) (2023) 7047.

HOW TO CITE THIS ARTICLE

M. Shahraeini. cA Comprehensive Approach to Synthetic Distribution Grid Generation: Erdős–Rényi to Barabási-Albert. AUT J. Elec. Eng., 57(1) (2025) 85-100.

DOI: 10.22060/eej.2024.23143.5591