

Two new mathematical models for two level electricity network design with distributed generation

Burçin ÇAKIR ERDENER^{1,*}, Berna DENGİZ¹, Zühal GÜNGÖR², İmdat KARA¹

¹Department of Industrial Engineering, Faculty of Engineering, Baskent University, Ankara, Turkey

²Department of Industrial Engineering, Faculty of Engineering, Gazi University, Ankara, Turkey

Received: 31.05.2022

Accepted/Published Online: 04.11.2022

Final Version: 19.01.2023

Abstract: In the new millennium, traditional electrical power systems have undergone a significant change driven by a set of requirements arising from evolving and changing technology. Thus, fundamental changes have occurred in the way electrical energy is produced, transmitted, and distributed. This situation has revealed the need to expand existing networks or to establish new networks. The available literature revealed that particular attention to the latter one is still limited due to the complexity of the power system. The purpose of this study is to contribute to the body of literature that tries to address the gap at overall design of a power distribution network. Moreover, distributed generation integration is also considered simultaneously with network design. In this paper, a two-level electricity distribution system design with distributed generation (TLEDS_DG) for green-field planning is considered. The TLEDS_DG is defined to find the locations of distribution transformers, to decide the number of distributed generators, and to design underlying two-level network in such a way that demand and capacity constraints are satisfied, and the overall design cost is minimized. Two mathematical models, a node-based and flow-based, are proposed and compared in terms of solution quality and CPU times. Within the flow-based formulation which is a new technique, called node cloning, is used to transform a two-level network into a single-level one. To validate the effectiveness and efficiency of the proposed approach (flow-based formulation with node cloning), we conduct numerical studies based on randomly generated instances up to 300 customers. Additionally, sensitivity analysis is also applied to demonstrate the impacts of initial parameter settings. Computational results on a large suite of test problems show that the proposed approach (flow-based formulation with node cloning) is efficient and highly effective for the generated test problems.

Key words: Two-level networks, facility location

1. Introduction

In today's electricity sector, the design of new network architectures has become a major problem due to the significant increase in electricity demand and growing concern about the climate change impact of fossil fuels. One of the most critical solutions is to build up a decentralized system where more energy is being generated locally, so called distributed generation units (DGs) [1].

DGs are generally connected at the distribution level due to locally accessible resources (renewable and nonrenewable) and their small-scaled sizes. In case of high DG penetration, the generated power of DG units affects not only the power flow in the distribution system, but also the amount of power transmitted from the upper grid. Therefore, deciding the number and the total capacity of DGs in a decentralized distribution system is an important optimization problem [2].

*Correspondence: bcakir@baskent.edu.tr

Decentralized electricity system design can be analyzed as an expansion of the existing system or a green-field deployment. The first approach is related to adding DGs into the existing network as a part of a network expansion problem, and the second one is related to planning the new development zones or building new cities by creating a new distributed network design [3]. Studies in the literature on the integration of DGs to distribution system commonly considers the first group, connection of DGs to an existing network. However, while reconfiguring the network or designing new distribution systems for new demand areas (network expansion), DGs planning and network design decisions should be considered simultaneously. In this study, we considered a new decentralized electricity distribution system design approach which combines two-level network design and DG integration problems. The capacitated location problem of distribution transformers (DTs), integration decision of DGs and the low-voltage (LV) and medium-voltage (MV) network design problem are combined into a single problem and solved in a single optimization framework. Two-level network is designed for greenfield or off-grid planning. We refer to our problem as the two-level electricity distribution system design with distributed generation (TLEDS_DG). To the best of our knowledge, this problem have not been previously addressed in the literature. Given a set of candidate locations and capacities of DTs and also DGs, the problem is to determine the number and locations of DTs and the number of DGs, and to route the underlying two-level network. It should be noted that the problem is not deciding on the locations of the DGs, it just decides which DG can or cannot be installed in already known locations with predefined capacities. This assumption is important as in real life DG locations and capacities are usually determined according to the availability of the primary fuel for the potential region (i.e. wind turbines can be located to a place only if that region has enough wind source capacity). Therefore, the problem deals with the number of DGs that can be installed on potential locations. The contribution of this paper is threefold:

Firstly, a new network design problem is introduced (TLEDS_DG) where two levels of the distribution system and DG integration are jointly considered. The objective is to minimize the total system cost including installation and power loss costs for each level. Secondly, two polynomial-size mixed integer non linear programming (MINLP) formulations such as node based formulation (NBF) and flow based formulation (FBF) are proposed. And finally, in order to reduce the complexity of the model, especially for large sized problems, a new network transformation technique called the node cloning technique (NCT) is adapted from [4] and applied within FBF.

The rest of the paper is organized as follows. Firstly, available literature is provided in Section 2. The TLEDS_DG problem is described in detail in Section 3. Section 4 is devoted to the solution methods starting with the presentation of the NBF followed by the definition of NCT and FBF. Computational experiments are presented in Section 5, and finally, conclusions are drawn in Section 6.

2. Literature review

Various approaches addressing the integration of DGs in distribution networks have existed in the literature [5–7]. Each is a different research study whose objective is to consider an economic or a techno-operational perspective [8]. Total system investment costs, capacity costs, and the cost of power loss are some of their economical optimization criteria. From a techno-operational perspective, fuel savings, carbon emissions, reserve/back-up capacity, minimization of import/export, voltage profile, voltage stability improvement, and elimination of excess power generation can be considered as optimization objectives [2].

Reliability and security of energy is considered for distribution network expansion problem in [9] where the distribution feeders are integrated with the DG units. By incorporating the renewable resources as DG

units in the distribution network expansion problem, the islanding capability in the presence of these units and their uncertainties are considered to get a more accurate solution in [10]. A stochastic multiobjective approach based on the Monte Carlo simulation is applied to solve the distribution network expansion problem in [11] to create a trade-off between the cost and reliability for choosing the size and location of DGs and distribution transformers (DTs). In [12] and [13] two stage optimization methods are proposed for multiyear expansion planning in distribution system to determine the optimal size and site of DGs as well as DTs and feeders. A method for network reconfiguration problem is proposed in [14] by considering the effects of random energy sources and by addressing larger scale problems. In [15], a hybrid simulated annealing approach for expansion planning of a radial distribution network with DG is proposed. The authors considered the installation cost, power loss cost, and cost of energy not supplied due to failure of feeders. An analytical approach for distribution systems, considering a loss sensitivity factor based on equivalent current injection is proposed in [16]. The defined sensitivity factor is used to determine the optimal size and location of DG in order to reduce total power losses. A framework interfacing a mathematical model and MATLAB-based model is developed in [17] to obtain the optimal DG capacity and locations while minimizing total system planning costs. In their study, the total planning cost consists of DG investment, operation and maintenance costs, the cost of power purchased from transmission companies, and system power losses. A multiobjective genetic algorithm is proposed in [18] for the optimum location of DG units in a radial distribution system with multiple voltage dependent load types. An algorithm based on graph theory is presented in [19] to investigate the impact of DGs integration on the feeder routing. A detailed review of the models on expansion planning with DGs can be found in [20]. A more recent approach is to solve the network reconfiguration problem simultaneously with DG integration. In [21], a multiobjective management operation is described based on network reconfiguration and the allocation and sizing of renewable DGs to minimize active power loss, annual operation costs (installation, maintenance, and active power loss costs), and pollutant gas emissions. The changes in wind speed, solar irradiation, and load over time are taken into account. To solve the problem, an effective evolutionary method based on the Pareto optimality is used. A mixed particle swarm optimization (PSO) is introduced in [22], for minimizing active power loss and improving the distribution network's voltage profile. The proposed approach combines binary particle swarm optimization and the conventional PSO algorithms. Initially, the ideal distribution network configuration is identified, and then DG placement and sizing problems are solved. Three load scenarios were assessed during network reconfiguration and DG integration to evaluate the presented approach.

In addition to the various DG integration and network reconfiguration concerns, decentralized energy systems also have management challenges on the demand side particularly due to the smart equipment and automation systems. Demand response allows consumers to play a substantial role in the operation of the electric grid and has been used by recent studies to manage residential hybrid energy systems and to emphasize the efficiency of demand response in frequency regulation. For instance, in [23] a two-step demand response mechanism is used to optimize a hybrid renewable energy system for a single-family residence. By reordering the deferrable loads, demand response technique impacts on lowering the investment cost of the hybrid energy system is showed in [24].

All of the papers mentioned above, considers the integration of DGs into the existing grid considering different aspects. However, greenfield planning is also an important area as the traditional approaches that are not expected to meet future economic and environmental sustainability standards. Moreover, these studies considers DG integration by determining the location and size of the units and/or routing of only medium-level network of the distribution system, whereas the distribution system consists of two-level network with two

different voltage levels, namely medium voltage (MV) and low voltage (LV). One of the major drawbacks of studies in the literature is that the planning of each network separately will lead to a low-level of accuracy design problem in practice [25] as one-level design reduces the probability of reaching an optimal solution for the main problem consisting two levels. Therefore, in an electricity network design problem, integration of DGs, routing of feeders and locating DTs should be determined considering two-level network elements in a combined manner.

Although the number is limited, there are a few studies which consider two levels of the distribution system simultaneously in the literature [25–33]. A problem called “two level network design with intermediate facilities” is presented in [25] where the two levels correspond to medium and low voltage, respectively, and intermediate facilities correspond to the distribution transformers. Authors have introduced an exact formulation for the problem ignoring the installation fixed costs for low voltage level and power loss costs for medium level. They propose a hybrid decomposition-based approach which limits the candidate DT nodes to reduce the computational complexity of the exact method. In [26], the electricity distribution system is defined as two-level network design problem and three mathematical models are presented to find the location and size of DTs. In order to solve the subproblems, a heuristic solution approach based on k-median, minimum spanning tree, and shortest path is used. In [27], an evolutionary algorithm method over a specific tree is proposed to design an electricity distribution system. In this algorithm, each tree in a forest represents an LV level and is fed by one DT. The DTs among trees are connected by a minimum spanning tree that represents MV level. The cost function for each subtree is optimized by dynamic programming taking into account feeder losses, maintenance, voltage drops, and the cost functions of feeders and DTs. In [28], the authors have proposed an integral methodology for the optimization of two level electrical distribution network optimization simultaneously. The proposed methodology takes into account an important factor that can only be considered by means of an integral approach, such as the incidence in the costs due to the use of common or shared routes and structures by the MV and LV feeders. The horizon distribution network planning problem and optimal distribution system model formulation are described in two papers [29, 30]. The proposed model and optimization formulation in these referenced studies provide a generalized horizon planning approach and introduces a fully functioning comprehensive horizon planning model. In [31], a reference network model (RNM) is proposed for a large-scale distribution planning tool that can help regulators to estimate efficient costs in the context of incentive regulation applied to distribution companies. Authors present the main features of an RNM for planning distribution networks from scratch, greenfield planning, or incrementally from an existing grid. The main properties of the model are highlighted as the simultaneous planning of both MV and LV networks by using simultaneity factors; and the layout of feeders, taking into consideration the street map, which is automatically generated by the model. A new discrete particle swarm optimization algorithm is presented for the optimal planning of two level distribution system in [32]. The location and capacities of DTs, as well as the routes and types of MV and LV feeders, are obtained by this algorithm. None of the cited studies above dealing with the two-level distribution network has considered DG integration, mainly because of the increasing complexity of this problem. The most recent study on greenfield planning can be found in [33]. A two-level distribution network design is considered to meet a number of demand points with a tree-star network approach. Primary facilities at first level that have source capabilities feed secondary facilities through tree networks. Both the primary and secondary facilities at the lower level provide star network service to customers within a coverage area. In order to reduce the distribution cost, an optimization problem is introduced that identifies the quantity, varieties, and locations of the facilities as well as the lower and higher level networks. A heuristic approach

which solves a discrete counterpart and then enhances its result by rearranging facilities on the continuous space is proposed to solve the problems.

As we will consider the two-level network design problem, it is worth mentioning that two level network design problems are also very popular in the design of telecommunication networks, in which two types of customers requiring two different levels of service are considered. The purpose of telecommunication network architecture is to supply broadband services including high-speed Internet, phone, and cable TV. Due to the availability of various technologies such as optical fibers and coaxial lines, intermediate facilities are required to connect both technologies. Typically, demand vertices are served by coaxial lines from intermediate facilities. Optical fibers connect the intermediate facilities to a central root vertex. The types of intermediary facilities employed depend on the capacities and functions necessary. The goal of the design problem is to reduce the total cost of edge and vertex facilities. Recent studies related with two level network design in telecommunication networks can be found in [34–37].

3. Problem description

Traditional power distribution system is an important application of two level network design problem, where networks with two different voltage levels (MV and LV) are built. MV network transmits the power from main transmission grid to DTs, where the MV is stepped down to the LV level. LV network is responsible to distribute the power to final consumers. In a typical two-level power distribution network, the upper level is called the primary network (PN), and connects the transmission substations (S-TS) to DTs. The lower level network is called secondary network (SN) and they connect the DTs to final consumers. The DTs are vertices containing a costly and capacitated facility where primary flow is converted in secondary flow. In this paper, decentralized network structure is considered by integrating DGs to the PN. The number of DGs directly effects the amount of power transmitted from the upper grid.

This problem is named TLEDS_DG. The objective of the problem is (a) to select the capacity of S-TS, (b) to route PN and SN, (c) to select the locations and capacity levels of DTs, and (d) to determine the number of DGs while minimizing the total system design cost.

The schematic representation of the TLEDS_DG facilities is shown in Figure 1.

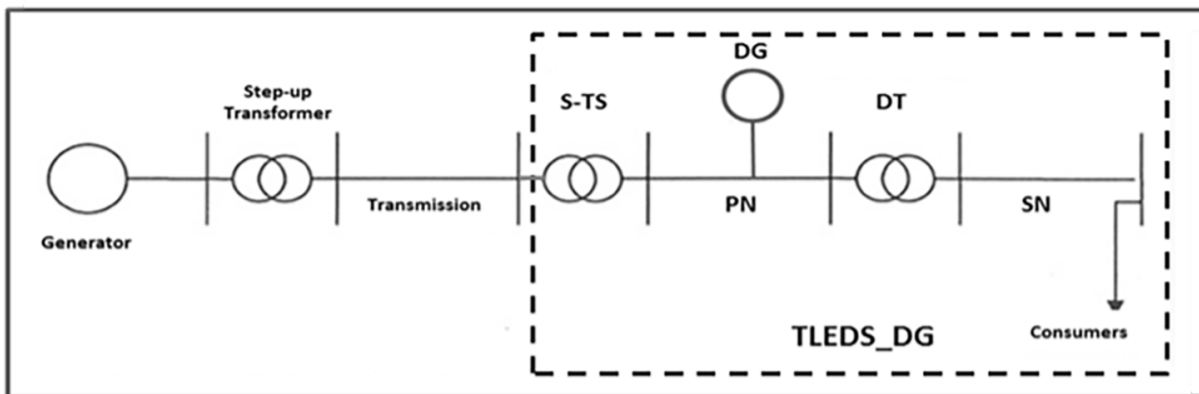


Figure 1. The facilities of the TLEDS_DG.

The assumptions of the TLEDS_DG problem are given below:

- The S-TS can be connected only to DTs or TNs with PN links,
- DNs can be connected to DNs, or DTs with SN links,
- DTs can be located on DNs or TNs,
- DGs can only be located in PN on TNs,
- Only one S-TS may be located to root node,
- Only one type of DT may be located per node,
- Only one type of DG may be located per node,
- Both PN and SN are radial (tree shaped),
- Demand and DG production are assumed to be deterministic.

4. Proposed models

Mathematically the TLEDS_DG can be defined on a graph: $G = (N, A)$, where $N = N_R \cup N_D \cup N_T$ set of nodes in which N_R represents the S-TS locations which has the smallest index from a given number of nodes), N_D and N_T represent the demand nodes (DNs) and transportation nodes (TNs) with zero demand, respectively, and $A = \{(i, j) : i, j \in N\}$ is the set of arcs. Each customer $i \in N_D$ has a demand d_i (MW), with $0 < d_i$. Each arc $(i, j) \in A$ with a distance l_{ij} (km), has a fixed capacity which is M^1 (MW) for PN links and M^2 (MW) for SN links. Each arc $(i, j) \in N$ has also a nonnegative cost function associated with fixed installation cost and power loss cost per year. The fixed installation cost and power loss cost coefficient are f^1 (\$/km) and c^1 (\$/(MW)² km), for PN, and f^2 (\$/km) and c^2 (\$/(MW)² km) for SN, respectively. The cost for purchasing the energy from upper grid (which corresponds to total demand minus total installed DG capacity) per year is f^{up} (\$/MW). There are also potential DT nodes, $P_{DT} \in (N_D \cup N_T)$, and potential DG nodes, $P_{DG} \in N_T$. There are K types of DT which has a capacity s^k (MW) and a fixed cost f^k (\$) where $k \in K$. The D different types DG has a fixed capacity s^d (MW) and a fixed cost f^d (\$) where $d \in D$. Finally, there are E types of STS which has a capacity s^e (MW) and a fixed cost f^e (\$) where $e \in E$. Under all definitions and assumption given above, the problem considered in this study is to determine the S-TS capacity, the location of the DT nodes and their capacities, the number of the DGs, as well as routing PN and SN, thereby constructing a distribution network to meet all of the demand while minimizing total system design cost. Fixed costs are incurred for constructing nodes and links and variable costs are incurred for power loss cost due to the transportation of electricity. Figure 2 illustrates an initial graph and also a feasible solution of the TLEDS_DG problem.

Two model formulations such as node-based and flow-based proposed to solve the problem addressed is described below.

4.1. Node-based model formulation

Decision variables:

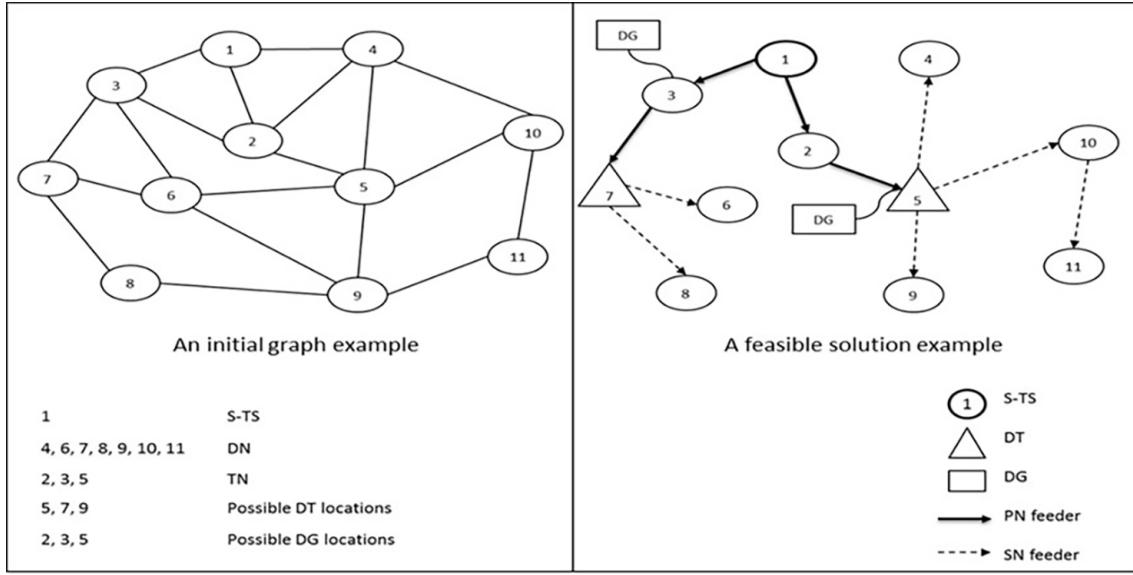


Figure 2. (a) An initial graph, (b) a feasible solution for TLEDS_DG problem.

$$y_{ij}^1 = \begin{cases} 1, & \text{if } arc(i, j) \text{ exist on PN} \\ 0, & \text{ow} \end{cases}$$

$$y_{ij}^2 = \begin{cases} 1, & \text{if } arc(i, j) \text{ exist on SN} \\ 0, & \text{ow} \end{cases}$$

$$z_i^k = \begin{cases} 1, & \text{if } k \text{ type DT is located on node } i \\ 0, & \text{ow} \end{cases}$$

$$z_i^e = \begin{cases} 1, & \text{if } e \text{ type S-TS is located on node } 1 \\ 0, & \text{ow} \end{cases}$$

$$z_i^d = \begin{cases} 1, & \text{if } d \text{ type DG is located on node } i \\ 0, & \text{ow} \end{cases}$$

x_{ij}^1 : the amount of power transported from node i to node j with PN links (MW),

x_{ij}^2 : the amount of power transported from node i to node j with SN links (MW),

x^{up} : the amount of power purchased from upper grid (MW).

Proposed NBF is as follows:

$$\begin{aligned} \min \quad & \sum_{(i,j) \in A} f^1 l_{ij} y_{ij}^1 + \sum_{(i,j) \in A} c^1 l_{ij} (x_{ij}^1)^2 + \sum_{(i,j) \in A} f^2 l_{ij} y_{ij}^2 + \sum_{(i,j) \in A} c^2 l_{ij} (x_{ij}^2)^2 \\ & + \sum_{e \in E} f^e z_1^e + \sum_{k \in K} \sum_{i \in P_{DT}} f^k z_i^k + \sum_{d \in D} \sum_{i \in P_{DG}} f^d z_i^d + f^{up} x^{up} \end{aligned} \quad (1)$$

s.t.

$$\sum_{j \in N} (x_{ij}^1 + x_{ij}^2) - \sum_{j \in N} (x_{ij}^1 + x_{ij}^2) = \begin{cases} -x^{up} & i = 1 \\ d_i - \sum_{d \in D} s^d z_i^d & i \neq 1 \end{cases} \quad \forall i \in N, i \neq j \quad (2)$$

$$- \sum_{d \in D} s^d z_i^d \leq \sum_{j \in N} x_{ji}^1 - \sum_{j \in N} x_{ij}^1 \leq \sum_{k \in K} s^k z_i^k - \sum_{d \in D} s^d z_i^d \quad \forall i \in N, i \neq 1, i \neq j \quad (3)$$

$$- d_i \leq \sum_{j \in N} x_{ji}^2 - \sum_{j \in N} x_{ij}^2 \leq \sum_{k \in K} s^k z_i^k - d_i \quad \forall i \in N, i \neq 1, i \neq j \quad (4)$$

$$0 \leq \sum_{j \in N} x_{ij}^1 \leq \sum_{e \in E} s^e Z_1^e \quad i = 1, i \neq j \quad (5)$$

$$\sum_{k \in K} z_i^k \leq 1 \quad \forall i \in P_{DT} \quad (6)$$

$$\sum_{d \in D} z_i^d \leq 1 \quad \forall i \in P_{DG} \quad (7)$$

$$\sum_{e \in E} z_i^e \leq 1 \quad \forall i = 1 \quad (8)$$

$$x^{up} = \sum_{i \in N_D} d_i - \sum_{i \in P_{DG}} \sum_{d \in D} s^d z_i^d \quad (9)$$

$$x_{ij}^1 \leq M^1 y_{ij}^1 \quad \forall (i, j) \in A \quad (10)$$

$$x_{ij}^2 \leq M^2 y_{ij}^2 \quad \forall (i, j) \in A \quad (11)$$

$$y_{ij}^1 + y_{ji}^1 \leq 1 \quad \forall (i, j) \in A \quad (12)$$

$$y_{ij}^2 + y_{ji}^2 \leq 1 \quad \forall (i, j) \in A \quad (13)$$

$$\sum_{j \in N} y_{ij}^1 \leq 1 \quad \forall i \in N, i \neq j \quad (14)$$

$$\sum_{j \in N} y_{ij}^2 + \sum_{k \in K} z_i^k \leq 1 \quad \forall i \in N, i \neq j \quad (15)$$

$$x_{ij}^1 \geq 0, x_{ij}^2 \geq 0 \quad \forall (i, j) \in A \quad (16)$$

$$y_{ij}^1 \in \{0, 1\}, y_{ij}^2 \in \{0, 1\}, \quad \forall (i, j) \in A \quad (17)$$

$$z_1^e \in \{0, 1\}, z_i^k \in \{0, 1\}, z_i^d \in \{0, 1\}, \quad \forall i \in N, e \in E, k \in K, d \in D \quad (18)$$

In this formulation, objective function (1) minimizes the total system design cost including S-TS, DTs, DGs, PN, and SN links fixed installation costs and variable power loss cost (VPLC) on PN and SN links which are dependent on the amount of power transported on the link, as well as the cost of energy purchasing from the upper grid. Constraint (2) ensures the flow conservation between the entire level of the network. Constraints (3) assure that if a facility (DT and/or DG) installed on a node in PN, the capacity of the facility is respected. Constraint (4) is the same as explained in Constraint (3) for SN where only DTs can be located as facilities. Additionally, those two constraints (3 and 4) ensure the flow transformation is always from PN to

SN. Constraint (5) guarantees the capacity constraint for S-TS. While constraint (6) ensures that only one type of DT must be installed to one node, constraint (7) and constraint (8) guarantees that only one type of the DGs and S-TS can be installed to one node, respectively. Constraint (9) shows the amount of electricity that must be purchased from the upper grid. Constraints (10) and (11) ensure the capacity constraints of PN and SN feeders, respectively. Constraints (12) and (13) forbid negative flows. The radial (tree) network structure for PN and SN networks are ensured by constraints (14) and (15), respectively. While constraint (16) defines positive variables, constraints (17) and (18) state integer variables.

The power loss coefficients (c^1 and c^2) formulation is given in Equation (19) [38].

$$c^1 \text{ or } c^2 = \frac{8,76R}{(KV)^2} c_e P_l \quad (19)$$

where KV is the voltage level of the link, c_e indicates the energy transportation cost per MWh , R states the resistance of the link per km , and P_l is the power loss factor. An important issue in electricity distribution system design is the consideration of voltage drop levels. Since the objective function minimizes the power losses which have positive correlation with voltage drops, the sum of the voltage drops in all nodes will be under control, preventing excessive voltage drops. Thus, an acceptable voltage profile for overall network will be provided.

The model of TLEDS_DG is nonlinear due to the convex function of power loss costs in the objective function. In order to enhance the computational efficiency, the nonlinear objective function is linearized by piecewise linearization technique.

4.2. Flow based model formulation

The complexity of the TLEDS_DG problem arises from integration of two hard problems such as; capacitated facility location and two-level network routing. The dimension of the problem will increase rapidly with the number and type of nodes and arcs. Moreover, specific to the TLEDS_DG problem formulated with NBF, some nodes have more than one role, called multirole nodes, such as DG and TN; DT and TN; DG and DT; DT and DN; and DG, TN, and DT. Two types of arcs (PN and SN feeders) can be installed between two nodes if at least one of them is a multirole node. The number of multirole nodes and total number of arcs directly affect the number of constraints and variables since arcs connected to multirole nodes have the potential for both primary and secondary feeder installation.

The nature of the problem requires the use of an effective solution approach to solve the large sized problems. In this section, we introduce a new solution method, called NCT, based on the transformation of the multirole nodes into one-role nodes. Once the network is transformed into the associated one by NCT, new proposed flow-based mathematical formulation is used to solve the problem.

4.2.1. Node cloning technique

The approach in the node cloning technique is pioneered in [4] where the network transformation was implemented for the connected facility location problem for telecommunication networks. The authors consider two-level network structures including general Steiner tree-star problems that combine the facility location decisions with connectivity requirements. The technique mainly provides the transformation of the network by duplicating dual-role nodes and creating an additional copy. The dual-role node is defined as the node, which can be both a facility and a demand node. In our study this approach is adapted to the TLEDS_DG problem.

Different from the technique used in [4], the approach in this paper is adapted for tree-tree network structure that requires additional network transformation rules. TLEDS_DG problem contains multirole nodes (dual and triple roles) due to the different type of nodes (DNs and TNs) and facilities (DTs and DGs) that can be located on the same node. Moreover, the adapted NCT approach used in this study allows considering different capacities for each facility on the transformed network. Possible three states of multirole nodes can be listed as below:

State 1. While $P_{DG} \in N_T$, and $P_{DT} \in (N_D \cup N_T)$, and $N_T \cap N_D = \emptyset$, if $P_{DT} \cap P_{DG} = \emptyset$, the network contains dual role nodes with DG-TN roles.

State 2. While $P_{DG} \in N_T$, and $P_{DT} \in (N_D \cup N_T)$, and $N_T \cap N_D = \emptyset$, if $P_{DT} \cap P_{DG} \neq \emptyset$, the network contains triple role nodes with DG-DT-TN roles.

State 3. While $P_{DT} \in (N_D \cup N_T)$, and $N_T \cap N_D \neq \emptyset$, the network contains dual role nodes with DT-DN; and DT-TN.

For TLEDS_DG problem, NCT is used as follows: the original network is transformed to an associated new one to address the node sets that are not disjoint. The technique is based on creating augmented virtual copies of the multirole nodes where each copy has one single role. For instance, a dual role node will have one virtual copy node, where the original node has one role and the virtual copy has the other role. The original node and its virtual copies are connected with virtual arcs having zero cost. The original node and its virtual copies are connected to the rest of the network with arcs having different costs depending on the role of the nodes. For instance, a demand node can only be connected to a demand node and/or a distribution transformer by a secondary feeder. In the transformed network each node has a unique role; therefore, each arc has a unique cost associated with the roles of the end nodes of the arc. The procedure used in NCT is as follows:

1. Create the virtual copies of the multirole nodes such that each node takes only one role in the final representation and add a virtual arc with zero cost between original node and its virtual copy.
2. Add arcs between virtual node and original nodes depending on the arcs adjacent to the original node.
3. Replace every arc $\{i, j\}$ by two directed arcs (i, j) and (j, i) , where $i \neq 1$. If $i = 1$, add one directed arc (i, j) .
4. Delete unfeasible directed arcs; i.e (i, j) is unfeasible if $i \in N_D$ and $j \in P_{DT}$, or (i, j) is unfeasible if $i \in N_T$ and $j \in P_{DG}$.
5. If $j \in N_D$, assign feeder cost $a_{ij} = f^2 l_{ij} y_{ij}^2 + c^2 l_{ij} (x_{ij})^2$ (SN feeder installation and loss cost function between node i and j), and where x_{ij} is the amount of power flow from node i to node j .
6. If $j \in N_T$, assign feeder cost $b_{ij} = f^1 l_{ij} y_{ij}^1 + c^1 l_{ij} (x_{ij})^2$, (PN feeder installation and loss cost function between node i and j), and where x_{ij} is the amount of power flow from node i to node j .
7. If $j \in P_{DT}$, assign cost $b_{ij} + f^k$ (PN feeder installation cost and loss cost function for PN plus installation cost of DT of type k).
8. Create an artificial root node (super node), N_s , and add a directed arc from N_s to root node (N_R) with costs f^e , where f^e is the installation cost for S-TS, type e .

9. Add a directed arc from N_s to every node $i \in P_{DG}$ with costs f^d , where f^d is the installation cost for DG, type d .

The implementation of NCT for a simple network that has 4 nodes transformation can be shown in Figure 3, following the procedure described above.

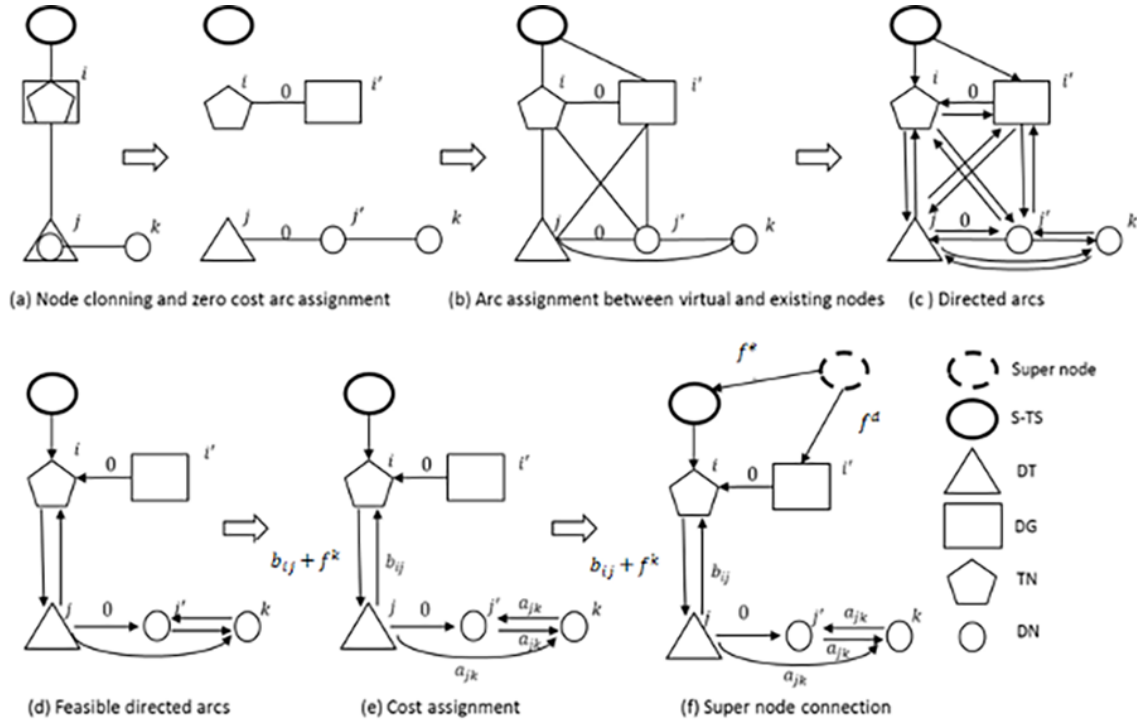


Figure 3. Network transformation with the node cloning and cost assignment procedure.

The network has two nodes i and j , and each node has multiroles. Multirole node i is connected to the root node (S-TS), and multirole node j is connected to node k , which has one role. Figure 3a represents the node cloning for multirole nodes. For multirole node i , $i \in \{N_T \cap P_{DG}\}$, a virtual copy is created as (i^{prime}), where original node is treated as a TN and the copy as a DG. Node j also has two roles, $j \in \{P_{DT} \cap N_D\}$, and a copy of node j is created as (j'), where j is treated as a DT and j' as a DN node. The cost of each new arc between any original node and its virtual copy nodes is 0. Figure 3b shows the extended network that consists of arc assignments between the original nodes and the virtual nodes. Arcs assigned according to the adjacent nodes to the original node. Every virtual copy is also connected to the original node if the original node has already connected to that node. Thus, the existing network in Figure 3b is converted to a network with two directed arcs given in Figure 3c. The arcs between root node (S-TS) and (i)/(i') are not two directed as reverse flow to S-TS is not allowed. After deleting the unfeasible arcs in the network in Figure 3c, a new network is obtained as shown in Figure 3d containing only feasible arcs. For instance, (i, j) is unfeasible if $i \in N_D$ and $j \in P_{DT}$, or if $i \in N_T$ and $j \in P_{DG}$. Following feasible arc determination, all edge and facility costs are assigned to the arcs depending on the end points role of the arc. For instance, if there is a DT at the end point of arc(i, j), then DT must be connected to a PN feeder. In this case, the cost of the arc(i, j) is the computed as the sum of installation and loss cost of PN feeder, b_{ij} , and installation cost of DT of type k , f^k .

The end point k , on the arc (j', k) , is a demand node and it can only be connected with an SN feeder, with cost a_{jk} . Thus, Figure 3e is obtained after all related cost values are assigned on arcs in Figure 3d. Finally, Figure 3f shows the super node connection to define the installation costs of ST-S and DGs considering their types.

A new flow based formulation (FBF) is needed to solve the TLEDS_DG problem once the original network is transformed by NCT as explained above. In this formulation, all costs (installation and loss costs) are assigned to the arcs as mentioned above. In the transformed network, the defined sets are as follows.

$G = (N, A)$ is directed network where $N = N_S \cup N_R \cup N_D \cup N_T$ is the set of nodes, so N_S , N_R , N_D , and N_T represent super node, the S-TS location, demand nodes (DNs), and transportation nodes (TNs) with zero demand, respectively. $A = \{(i, j) : i, j \in N\}$ is the set of arcs. $A = SSF \cup DGS \cup PF \cup TF \cup SF$, where SSF represents the S-TS feeders, DGS represents the DGs-related feeders, and PF , TF and SF correspond to DTs, TNs and DN related feeders, respectively. The parameters are the same as those in the NBF. Decision variables specific to FBF are as follows:

$$y_{ij}^e = \begin{cases} 1, & \text{if } e \text{ type SSF feeder is installed on arc}(i, j) \\ 0, & \text{ow} \end{cases}$$

$$y_{ij}^k = \begin{cases} 1, & \text{if } k \text{ type PF feeder is installed on arc}(i, j) \\ 0, & \text{ow} \end{cases}$$

$$y_{ij}^d = \begin{cases} 1, & \text{if } d \text{ type DGS feeder is installed on arc}(i, j) \\ 0, & \text{ow} \end{cases}$$

$$y_{ij}^t = \begin{cases} 1, & \text{if TF feeder is installed on arc}(i, j) \\ 0, & \text{ow} \end{cases}$$

$$y_{ij}^s = \begin{cases} 1, & \text{if SF feeder is installed on arc}(i, j) \\ 0, & \text{ow} \end{cases}$$

e^e : spare capacity of the S-TS with type e (MW)

x_{ij} : the amount of power transported from node i to node j (MW)

FBF for TLEDS_DG problem where its network is transformed with NCT is given below:

$$\begin{aligned} \min \quad & \sum_{(i,j) \in SSF} \sum_{e \in E} (f^e y_{ij}^e) + \sum_{(i,j) \in DGS} \sum_{d \in D} (f^d y_{ij}^d) + \sum_{(i,j) \in PF} \sum_{k \in K} (f^k y_{ij}^k + f^1 l_{ij} y_{ij}^k + c^1 l_{ij} (x_{ij})^2) \\ & + \sum_{(i,j) \in TF} f^1 l_{ij} y_{ij}^1 + c^1 l_{ij} (x_{ij})^2 + \sum_{(i,j) \in SF} f^2 l_{ij} y_{ij}^2 + c^2 l_{ij} (x_{ij})^2 + \sum_{(i,j) \in SSF} \sum_{s \in SS} f^{up} (f^e y_{ij}^e - e^e) \end{aligned} \quad (20)$$

$$\sum_{(i,j) \in A} x_{ji} - \sum_{(i,j) \in A} x_{ij} = \begin{cases} -\sum_{i \in N} D_i & i \in N_S \\ d_i & i \in N_D \\ 0 & ow \end{cases} \quad \forall i \in N \quad (21)$$

$$\sum_{e \in E} (s^e y_{ij}^e - e^e) - x_{ij} = 0 \quad \forall (i, j) \in SSF \quad (22)$$

$$\sum_{d \in D} (s^d y_{ij}^d) - x_{ij} = 0 \quad \forall (i, j) \in DGS \quad (23)$$

$$\sum_{k \in K} s^k y_{ij}^k - x_{ij} \geq 0 \quad \forall (i, j) \in PF \quad (24)$$

$$x_{ij} \leq M^1 y_{ij}^t \quad \forall (i, j) \in \{PF \cup TF\} \quad (25)$$

$$x_{ij} \leq M^2 y_{ij}^s \quad \forall (i, j) \in SF \quad (26)$$

$$\sum_{e \in E} (s^e y_{ij}^e - e^e) = \sum_{i \in N_D} D_i - \sum_{(i,j) \in DGS} \sum_{d \in D} (s^d y_{ij}^d) \quad \forall (i, j) \in SSF \quad (27)$$

$$\sum_{e \in E} y_{ij}^e = 1 \quad \forall (i, j) \in SSF \quad (28)$$

$$\sum_{k \in K} y_{ij}^k \leq 1 \quad \forall (i, j) \in PF \quad (29)$$

$$y_{ij}^e \in \{0, 1\}, y_{ij}^k \in \{0, 1\}, y_{ij}^t \in \{0, 1\}, y_{ij}^s \in \{0, 1\}, y_{ij}^d \in \{0, 1\} \quad \forall (i, j) \in A \quad (30)$$

$$x_{ij} \geq 0 \quad \forall (i, j) \in A \quad (31)$$

Here, equation (20) is the objective function that minimizes the overall arc costs, including the installation costs for DTs, DGs as well as PN- and SN-related feeders, loss costs that occur during the transportation of power through PN and SN feeders, and the cost of energy purchased from upper grid. Constraint (21) is the flow constraint for all nodes in the network. For a DN, the constraint ensures that the difference between the input and output flow is equal to the demand of that node. Additionally, the total output flow from the super node is equal to the total demand. Constraint (22) represents the spare capacity of the S-TS located on the root node (N_R). Constraints (23) and (24) ensure that the capacity constraints are respected for the DGs and DTs. Constraints (25) and (26) guarantee the capacity constraints for PN and SN feeders. Constraint (27) represents the amount of the power purchased from the upper grid. While constraint (28) ensures that only one type of S-TS can be assigned to one node, constraint (29) guarantees the same condition explained in constraint (28) for DTs. Constraints (30) states integer variables and finally constraint (31) forbid negative flows. Same as in the NBF, piecewise linearization technique is used to linearize the nonlinear cost functions in the objective function.

5. Computational results

In this section, quite a large computational analysis is performed for the proposed two models with NBF and FBF, respectively, on a wide range test problems.

5.1. Generation of test problems

Since there are no benchmark test problems in the related literature for the TLEDS_DG, we randomly generate small-, medium-, and large-sized test problems, similar to the one in [25], but considering DG locations. Firstly, small-sized problems are generated by selecting the number of nodes between 5 and 30. For medium-sized problems, the number of nodes is between 35 and 90. Finally, 10 large-sized problems are generated with node numbers between 100 and 410. The nodes are generated on a 100 by 100 grid. The percentage of TN nodes is between 20% and 30% randomly, and the remaining nodes defined as DN nodes. Once the number of nodes and their locations are determined, Prim algorithm [39] is used to form a spanning tree. After generating a tree network, the number of arcs is increased until meeting the desired arc number. The nodal demands are assigned uniformly between $U \sim (5, 15)$. It is assumed that the S-TS location is always present on node with lowest index. The number of candidate DT nodes and their locations are generated randomly unless the number respects the minimum number that necessary for a feasible solution. The minimum number is computed by dividing the total demand by the maximum capacity of DT types ($DTN \geq \sum_{i \in N_D} d_i / \max_{k \in K} (f^k)$), where DTN is the total number of the DTs for a test problem. It is assumed that DGs are connected only to the PN; therefore, candidate nodes for DGs are selected among the TN nodes randomly. The capacities of the possible DGs are generated randomly with the assumption that the total DG capacity of a problem cannot exceed 40% of the total demand. In addition to the generated test problems, two test problems given in [17] are adapted. The two real-life problems are extended considering candidate DG location nodes and their capacities and named $R1_{ext}$ and $R2_{ext}$.

5.2. Comparisons and discussions of NBF and FBF

Two polynomial sized mathematical formulations based on NBF and FBF, are presented for the TLEDS_DG problem that is introduced in this study. In order to evaluate the performances of the models, a total of 52 problem instances (50 randomly generated and 2 adapted from the literature as explained in 4.1) are solved with the CPLEX 12.6. optimizer, with a processor Intel(R) Core(TM) i7-3630QM CPU 2.40 GHz and 8 GB of RAM. In each run the default settings of the solver is used and the CPU time is limited to 7200 s.

Tables 1-3 show the computational results. The name of the problem, the number of nodes, the number of the arcs, the number of multirole nodes, and the best solution value are given from the first to the fifth column in each table, respectively. The CPU times and the linear relaxation deviation percentage value (LRDP) are placed in the sixth and seventh columns for both NBF and FBF, respectively. LRDP is calculated with $LRDP = 100 * (Z^* - Z^{LR}) / (Z^*)$, where Z^* is the best value obtained by each model and Z^{LR} is the value obtained by the relaxation of the binary variables.

In Table 1, the results show that small-sized problems can be solved in very short times using both models. For medium-sized problems, seen in Table 2, while the proposed model based on FBF with NCT has found the optimal solutions for all problems, the model based on NBF finds optimal solution for only 5 of 20 problems. For larger test problems, in Table 3, excessive longer computational times are needed compared to the medium-sized test problems. Here, it is seen that while the model based on NBF cannot find any solution for 12 test problems, the model based on FBF with NCT finds 7 optimal solutions among 12. When LRDP column is analyzed in each table respectively, we see that the average value of LRDP obtained by the model based on NBF is 16.25%, by the other model based on FBF is 7.54%, for small-sized instances. Similarly, in medium-sized problems, the model based on FBF outperforms the model based on NBF for medium-sized test

Table 1. Computational results for small-sized problems.

Problem	N	A	MRN	Best solution (\$/year)		CPU (s)		LRDP (%)	
				NBF	FBF	NBF	FBF	NBF	FBF
Ss1	5	5	0	168	168	2.17	1.45	4.63	3.02
Ss2	5	8	3	179	179	2.95	1.51	4.75	3.35
Ss3	7	9	0	256	256	3.01	1.92	5.86	3.52
Ss4	7	12	5	229	229	2.86	4.15	6.15	4.38
Ss5	10	15	0	689	689	3.68	3.55	12.57	4.83
Ss6	10	20	4	612	612	6.57	9.31	10.49	5.63
Ss7	10	22	6	660	660	8.59	8.57	14.88	6.18
Ss8	15	20	0	735	735	20.79	7.74	13.5	6.84
Ss9	15	22	9	790	790	37.48	10.12	15.56	6.35
Ss10	22	30	0	825	825	175.59	15.44	19.81	8.3
Ss11	22	32	15	1163	1163	567.28	49.74	17.38	9.61
Ss12	25	30	0	2386	2386	469.01	38.17	18.01	8.89
Ss13	25	35	15	2692	2692	1628.76	57.52	16.45	9.57
Ss14	25	45	20	3185	3185	2733.45	53.65	21.58	9.79
Ss15	27	35	0	3270	3270	2594.12	61.3	19.84	7.03
Ss16	27	50	17	2794	2794	3145.8	74.75	22.43	10.97
Ss17	27	55	20	4196	4196	2984.96	75.2	20.49	11.06
Ss18	30	45	0	4005	4005	5495.65	66.08	23.48	9.56
Ss19	30	60	20	4594	4594	6751.19	88.43	25.96	10.34
Ss20	30	65	22	7605	7605	6643.78	97.49	31.15	11.49
Average						1663.88	36.30	16.25	7.54

problems. The average values of LRDP for this case are 24.64% and 14.53%, respectively. The average values of LRDP for large-sized problems for two models are 32.29% and 22.29%, respectively. Moreover, in terms of the CPU times, the model based on FBF outperforms the model based on NBF for all test problems. Average CPU time in solution with the model based on FBF is less than 1 min for small-sized and less than 25 min for medium-sized problems. For large-sized problems, the maximum CPU time limit was enough to solve 7 of 12 problems.

5.3. Sensitivity analysis for FBF

In order to see the effects of the initial parameter values of the model based on FBF, the performance of the model is investigated by sensitivity analysis. Different parameter sets are considered and 50 test problems are solved again (excluding the 2 real problems within the large sized instances). The parameter tuning scenarios considered for sensitivity analysis are explained as follows.

- High/low feeder capacity levels: For capacity levels of the feeders, $M^1 = \lambda * \max_{k \in K} (s^k)$ and $M^2 = \lambda * \max_{i \in N_D} (d_i)$ values are obtained, where λ is 4 for high-capacity setting ($M \uparrow$) and 1.5 for low-capacity setting ($M \downarrow$). (λ was 2 for the original test problems).
- High/low DT installation costs: DT installation costs are taken as 1.25 times the value in the original problem for high-cost setting ($f^k \uparrow$) and 0.50 times the value in the original problem for low-cost setting ($f^k \downarrow$).

Table 2. Computational results for medium-sized problems.

Problem	N	A	MRN	Best solution (\$/year)		CPU (s)		LRDP (%)	
				NBF	FBF	NBF	FBF	NBF	FBF
Ms1	35	45	0	9236*	8967	7200	88.69	10.34	9.34
Ms2	35	48	25	8562	8562	4625	90.56	11.08	8.08
Ms3	37	45	0	9019*	8691	7200	112.17	13.79	5.93
Ms4	37	58	30	7159	7159	5183	159.5	20.53	10.26
Ms5	40	55	0	8671*	7128	7200	152.84	21.07	11.67
Ms6	40	60	32	6314	6314	4369	258.93	28.06	12.95
Ms7	40	62	36	5501	5501	6782	364.8	33.9	17.53
Ms8	45	55	0	6807	6807	5972	272.01	31.54	12.66
Ms9	45	61	38	8795*	4921	7200	475.84	35.38	19.36
Ms10	57	75	0	12521*	7306	7200	628.95	28.49	14.06
Ms11	57	79	40	13669*	10748	7200	837.21	25.76	12.64
Ms12	57	90	43	11027*	7911	7200	1139.53	28.27	14.84
Ms13	65	90	0	12792*	9572	7200	1035.9	19.05	15.79
Ms14	65	110	45	16448*	11842	7200	1459.33	22.69	24.47
Ms15	70	120	0	15875*	10748	7200	2649.53	31.64	19.42
Ms16	70	135	48	21055*	13955	7200	2967.27	25.79	15.73
Ms17	75	150	0	45698*	25896	7200	3165.94	17.95	16.9
Ms18	80	138	0	35478*	22071	7200	3070.31	29.53	18.43
Ms19	80	160	56	40279*	25038	7200	4000.76	21.38	12.57
Ms20	90	170	60	39522*	23852	7200	4196.56	36.54	17.93
Average						6746.55	1356.33	24.64	14.53

*Best solution obtained in given CPU time.

Table 3. Computational results for large-sized problems.

Problem	N	A	MRN	Best solution (\$/year)		CPU (s)		LRDP (%)	
				NBF	FBF	NBF	FBF	NBF	FBF
Bs1	100	154	61	25689*	17965	7200	4159.34	32.54	21.22
Bs2	100	159	70	22452*	15975	7200	4800.57	31.56	18.67
Bs3	150	300	85	45063*	28647	7200	5005.89	44.6	20.51
Bs4	150	279	140	51289*	34922*	7200	7200	32.07	25.9
Bs5	220	415	0	64208*	41298	7200	6891.07	37.17	21.17
Bs6	220	500	70	86820*	68163*	7200	7200	29.19	25.24
Bs7	300	496	48	76962*	51489	7200	7157	43.89	27.53
Bs8	320	579	170	80994*	64211*	7200	7200	26.78	22.05
Bs9	320	853	200	110598*	87103*	7200	7200	26.75	19.55
Bs10	410	9186	250	125033*	99835*	7200	7200	29.98	25.22
R1	100	100	40	36824*	25323	7200	3472.02	25.76	18.76
R2	173	192	45	43946*	31294	7200	4105.55	27.21	21.70
Average						7200	3596.417	32.29	22.29

*Best solution obtained in 2 h CPU time.

- High/low SN feeders installation costs: SN feeder installation cost is assumed as \$ 3 for high-cost setting ($f^2 \uparrow$), and \$ 1.75 for low-cost setting ($f^2 \downarrow$). It was \$ 2.25 for the original test problems).

- High/low DT installation costs together with high/low SN feeders installation costs (as explained above).
- High/low MRN number: For multirole nodes, the high values are considered to be $MRN = N$, (MRN \uparrow), and low values considered to be $MRN = 0$, (MRN \downarrow). The values considered in the original problem are given in Tables 1–3.
- High number of possible DT locations: The possible locations of the DTs are increased by multiplying the original values by 1.3, $DTN \uparrow = \lfloor DTN * 1.3 \rfloor$.

Table 4 shows the results of sensitivity analysis. In the table, the second column represents the average CPU times for the first (CPU₁), second (median) (CPU₂), and third quartiles (CPU₃). And finally, the last column depicts the number of optimal solutions (OS) obtained among 50 problems. The reference row (FBF^+) represents the reel values obtained by the model based on FBF, and the rest of the values given for CPU times show the deviation from the reference values when the parameter scenario given in the first column is used.

Table 4. Sensitivity analysis on the performance of the model based on FBF.

	CPU quartiles (s)			# OS
	CPU ₁	CPU ₂	CPU ₃	
FBF^+	230.4	510.7	786.0	45
M \uparrow	-4.7	-25.0	-24.4	47
M \downarrow	14.8	20.2	28.5	37
$f^k \uparrow$	8.2	27.9	29.5	41
$f^k \downarrow$	-5.0	-19.3	-22.3	46
$f^2 \uparrow$	20.1	-10.5	5.1	40
$f^2 \downarrow$	-1.8	-5.3	-3.7	45
$f^k \uparrow f^2 \downarrow$	5.4	20.2	30.8	41
$f^k \downarrow f^2 \uparrow$	15.4	-13.5	-9.0	45
MRN \uparrow	22.1	41.5	63.9	28
MRN \downarrow	-6.2	-38.5	-49.8	48
DTN \uparrow	40.4	35.2	48.6	35

When Table 4 is analyzed, it is clear that the capacity of feeders have effects on the performance of the algorithm. As feeder capacities are increased, both the CPU times and total numbers of optimal solutions obtained are increased and when capacities are decreased, it has an opposite effect on the performance of the model.

The high value setting for DT installation costs ($f^k \uparrow$) has a negative impact on the performance of the FBF, the low value setting ($f^k \downarrow$) has a positive impact in terms of CPU time and the number of optimal solutions. When the effects of SN feeder installation cost is analyzed ($f^2 \uparrow$), the increase in the cost parameter results in slightly longer CPU time for the third quartile and less number of optimal values. However, the low-cost parameter setting for SN feeder installation ($f^2 \downarrow$) does not have a considerable effect on the performance values. This result supports that the installation cost for SN feeders does not have big impacts on the overall network design since the installation cost of PN feeders are relatively higher. The SN feeders are more important when it comes to loss costs since the voltage levels are lower, resulting in higher power loss, in that part of the network. The combination of the DT installation cost scenarios and SN feeder installation cost scenarios

($f^k \uparrow f^2 \downarrow$ and $f^k \downarrow f^2 \uparrow$) show similar results with scenarios ($f^k \uparrow$ and $f^k \downarrow$), supporting the results that SN feeder installation cost does not have a significant effect on the overall performance of the formulation. When the number of multirole nodes are analyzed, the increased number of MRN (MRN \uparrow) has a negative impact on the performance of the FBF since it increases the complexity of the model. And for the same reason observed in MRN, the increased number of DTs (DTN) results in higher complexity in the model and worse quality results in terms of CPU times and number of optimal solutions obtained. It can be concluded from Table 4 that the changes in the cost settings affect the performance of the model only about 30%. This value is similar when the capacity constraints of the feeders are changed (max. 25%). On the other hand, the number of MRNs and DTs can be quite effective on the performance of the FBF in terms of CPU time and the number of optimal solutions.

6. Conclusion

This paper proposes two new polynomial-sized model formulations to design power distribution networks for green-field deployments, as well as existing network expansion in order to meet the growing demand.

A new realistic problem is introduced, called TLEDS_DG, in which location, assignment, and two-level network design tasks are jointly tackled. In order to solve the problem, a new node-based formulation (NBF) and a new flow-based formulation (FBF) are proposed. We have also proposed a new technique called node cloning technique (NCT) integrated with the FBF formulation in order to reduce the complexity of the original network by copying multirole nodes and defining augmented virtual nodes and arcs in an associated transformed network. Finally, we have presented computational results conducted on 50 new and 2 real-life-based test instances extended for this problem. A comparison shows that the proposed model based on NBF is a viable approach to solve small- and medium-sized TLEDS_DG problems, and the other proposed model based on FBF with NCT is a more effective method to solve small-, medium-, and large-sized test problems. Furthermore, the effects of the initial parameter setting on the performance of the FBF is analyzed by using sensitivity analysis and changing the capacities of the feeders, the installation costs for DTs and the number of multirole nodes and DTs are proved to be effective in the performance of the algorithm. In terms of future research directions, the output of the DGs can be considered stochastic due to variability of the renewable generators power output.

References

- [1] Frías P, Gómez T, Cossent R, Rivier J. Electrical Power and Energy Systems Improvements in current European network regulation to facilitate the integration of distributed generation. *International Journal of Electrical Power and Energy Systems* 2020; 31: 445–51. doi: 10.1016/j.ijepes.2009.03.001
- [2] Abdmouleh Z. et al. Review of optimization techniques applied for the integration of distributed generation from renewable energy sources. *Renewable Energy*, 2017, 113: 266-280.
- [3] Audring D, Hable M, Lerch E, Schilling K, Ag S et al. *Greenfield Planning of modern Metropolises* 2009: 1–6.
- [4] Bardossy MG, Raghavan S. Dual-based local search for the connected facility location and related problems. *INFORMS J Comput* 2010; 22: 584–602. doi: 10.1287/ijoc.1090.0375
- [5] Parihar SS, Malik N. Analysing the impact of optimally allocated solar PV-based DG in harmonics polluted distribution network. *Sustainable Energy Technologies and Assessments*, 2022, 49: 101784.
- [6] Hassan AS et al. Optimal integration of distributed generation resources in active distribution networks for techno-economic benefits. *Energy Reports*, 2020, 6: 3462-3471.

- [7] Diaaeldin IM et al. Large-scale integration of distributed generation in reconfigured distribution networks considering load uncertainty. In: *Uncertainties in Modern Power Systems*. Academic Press, 2021. p. 441-484.
- [8] Østergaard PA. Reviewing optimisation criteria for energy systems analyses of renewable energy integration 2009; 34: 1236–45. doi: 10.1016/j.energy.2009.05.004.
- [9] Aghaei J, Muttaqi KM, Azizivahed A, Gitizadeh M. Distribution expansion planning considering reliability and security of energy using modified PSO (Particle Swarm Optimization) algorithm. *Energy* 2014; 65: 398–411. doi: 10.1016/j.energy.2013.10.082
- [10] Bagheri A, Monsef H, Lesani H. Renewable power generation employed in an integrated dynamic distribution network expansion planning. *Electr Power Syst Res* 2015; 127: 280–96. doi: 10.1016/j.epsr.2015.06.004
- [11] Tarôco CG, Takahashi RHC, Carrano EG. Multiobjective planning of power distribution networks with facility location for distributed generation. *Electr Power Syst Res* 2016; 141: 562–71. doi: 10.1016/j.epsr.2016.08.020
- [12] Ahmadigorji M, Amjady N. A multiyear DG-incorporated framework for expansion planning of distribution networks using binary chaotic shark smell optimization algorithm. *Energy* 2016; 102: 199–215. doi: 10.1016/j.energy.2016.02.088
- [13] Ahmadigorji M, Amjady N, Dehghan S. A novel two-stage evolutionary optimization method for multiyear expansion planning of distribution systems in presence of distributed generation. *Appl Soft Comput J* 2017; 52: 1098–115. doi: 10.1016/j.asoc.2016.09.020
- [14] Cavalheiro EMB, Vergílio AHB, Lyra C. Computers and Operations Research Optimal configuration of power distribution networks with variable renewable energy resources. *Comput Oper Res* 2018; 96: 272–80. doi: 10.1016/j.cor.2017.09.021
- [15] Kerleta VD, Popovi DS. Hybrid simulated annealing and mixed integer linear programming algorithm for optimal planning of radial distribution networks with distributed generation 2014; 108: 211–22. doi: 10.1016/j.epsr.2013.11.015
- [16] Gözel T, Hocaoglu MH. An analytical method for the sizing and siting of distributed generators in radial systems 2009; 79: 912–8. doi: 10.1016/j.epsr.2008.12.007
- [17] Porkar s, Pourec p, Abbaspour-Tehrani-farda SS. A novel optimal distribution system planning framework implementing distributed generation in a deregulated electricity market. *Electr Power Systems Research* 2010; 80: 828–837.
- [18] Singh D, Singh D, Verma KS. Multiobjective Optimization for DG Planning With Load Models 2009; 24: 427–36.
- [19] Kumar D, Samantaray SR, Joos G. Electrical Power and Energy Systems A reliability assessment based graph theoretical approach for feeder routing in power distribution networks including distributed generations. *Int J Electr POWER ENERGY Syst* 2014; 57: 11–30. doi: 10.1016/j.ijepes.2013.11.039
- [20] Sedghi M, Ahmadian A, Aliakbar-golkar M. Assessment of optimization algorithms capability in distribution network planning : Review , comparison and modification techniques. *Renew Sustain Energy Rev* 2016; 66: 415–34. doi: 10.1016/j.rser.2016.08.027
- [21] Hamida, Imen Ben and Salah, Brini and Msahli, Faouzi and Mimouni, Mohamed Faouzi. Optimal network reconfiguration and renewable DG integration considering time sequence variation in load and DGs. *Renewable Energy* 2018; 121: 66-80.
- [22] Essallah S, Khedher A. Optimization of distribution system operation by network reconfiguration and DG integration using MPSO algorithm. *Renewable Energy Focus*, 2020, 34: 37-46.
- [23] Wang X, Palazoglu A, El-Farra NH. Operational optimization and demand response of hybrid renewable energy systems. *Applied Energy*, 2015, 143: 324-335.
- [24] He Y. et al. The multi-stage framework for optimal sizing and operation of hybrid electrical-thermal energy storage system. *Energy*, 2022, 245: 123248.

- [25] Costa AM, Franc PM, Lyra C. Two-level network design with intermediate facilities : An application to electrical distribution systems \$ 2011; 39: 3–13. doi: 10.1016/j.omega.2010.01.005
- [26] Carneiro MS, Franca PM. Long-range planning of power distribution systems secondary networks. *Computer Electrical Engineering* 1996; 22: 179–91.
- [27] Diaz-Dorado E, Pidre JC. GEM. Planning of large rural low-voltage networks using evolution strategies. *IEEE Trans Power Systems* 2003; 18: 1594–1600.
- [28] Paiva PC, Member S, Khodr HM, Yusta JM, Urdaneta AJ et al. Integral Planning of Primary – Secondary Distribution Systems Using Mixed Integer Linear Programming 2005; 20: 1134–43.
- [29] Fletcher RH, Member S, Strunz K. Planning – Part I : Formulation 2007; 22: 791–9.
- [30] Fletcher RH, Member S, Strunz K. Planning – Part II : Application 2007; 22: 862–70.
- [31] Domingo CM, Gómez T, Román S, Member S, Sánchez-miralles Á et al. A Reference Network Model for Large-Scale Street Map Generation 2011; 26: 190–7.
- [32] Ziari I, Ledwich G, Ghosh A. Optimal integrated planning of MV – LV distribution systems using DPSO 2011; 81: 1905–7. doi: 10.1016/j.epsr.2011.05.015
- [33] Gokbayrak K. A two-level off-grid electric distribution problem on the continuous space. *Computers & Operations Research* 2022; 144: 105853.
- [34] Yaman H, Carello G. Solving the hub location problem with modular link capacities. *Comput Oper Res* 2005; 32: 3227–45. doi: 10.1016/j.cor.2004.05.009.
- [35] Carello G, Della Croce F, Ghirardi M, Tadei R. Solving the hub location problem in telecommunication network design: A local search approach. *Networks* 2004; 44: 94–105. doi: 10.1002/net.20020
- [36] Fischetti M, Polo C, Scantamburlo M. A local branching heuristic for mixed-integer programs with 2-level variables, with an application to a telecommunication network design problem. *Networks* 2004; 44: 61–72. doi: 10.1002/net.20017
- [37] Chamberland S, Sanso B, Marcotte O. Topological design of two-level telecommunication networks with modular switches. *Oper Res* 2000; 48: 745–60. doi: 10.1287/opre.48.5.745.12412
- [38] Ponnavaikko M, Prakasa Rao KS, Venkata SS. Distribution System Planning Through a Quadratic Mixed Inger Programming Approach. *IEEE Power Eng Rev* 1987; PER-7: 55. doi: 10.1109/MPER.1987.5526753
- [39] Prim RC. Shortest Connection Networks And Some Generalizations. *Bell System Technical Journal* 1957; 36: 1389–401. doi: 10.1002/j.1538-7305.1957.tb01515.x