Metaheuristic-based Design and Optimization of Offshore Wind Farms Collection Systems

Daniel Hermosilla Minguijón, Juan-Andrés Pérez-Rúa, Kaushik Das, Nicolaos A. Cutululis
DTU Wind Energy
Technical University of Denmark
Frederiksborgvej 399, 4000 Roskilde, Denmark
Email: dhmi@dtu.dk, juru@dtu.dk, kdas@dtu.dk, niac@dtu.dk

Abstract—An optimization framework for automated design of offshore wind farms collection systems is proposed in this paper. The core of the framework consists of a metaheuristic algorithm, namely a Genetic Algorithm (GA). The GA is designed for searching high-quality feasible solutions in terms of the capital expenditure (CAPEXcs); a subsequent step runs a power flow in order to calculate electrical power losses for estimating the collection systems share on the Levelized Cost of Energy (LCOEcs). Finally, after several executions of the full framework, the feasible solution bringing the cheapest LCOEcs is selected. The main inputs are the coordinate's location of the wind turbines and the offshore substation (OSS), wind power production time series, and the set of considered cables for the collection system design. The proposed approach offers a full search space exploration for feasible solutions, while taking into account cables capacities and disallowing for cable crossings. The results show that this framework can find feasible solutions improving benchmark methods by 8%.

I. INTRODUCTION

Offshore wind energy represents one of the fastest and most steadily growing renewable technologies. The penetration level has increased almost five times in the last seven years, reaching the impressive globally total installed capacity of nearly 19 GW [1]. This growth is mainly explained by reductions in costs of the technology [2]: the LCOE has dropped recently from 240 USD/MWh to 170 USD/MWh. In terms of the total CAPEX, the electrical infrastructure can represent more than 10%, depending whether turbines are fixed-bottom or floating [3]. One of the components of the electrical infrastructure of Offshore Wind Farms (OWFs) is the collection system, for which finding good designs is becoming increasingly more complex, as OWFs tend to move towards coordinate-based design instead of following grid-based pattern (also called symmetrical designs), thanks to advancements on micrositing optimization techniques. This paper focuses on reducing the costs of the collection system, thus contributing to the overall minimization of economic metrics.

The collection system design and optimization problem has been studied with accentuated focus in the last 10 years. This problem is proved to be mathematically NP-hard [4], which means that there is no certificate on the ability to come up with algorithms to solve it in a polynomial time in function of the problem size. One could cluster all the existent methodologies for tackling this problem as follows: exact solutions, branch exchange solutions, C-MST heuristic solutions, metheuristic solutions, and hybrid solutions. While each of these approaches exhibit advantages and disadvantages, requiring comprenhensive and rigorous studies to determine them, in general, metaheuristic algorithms do not require in-depth mathematical formalities and are a good tool for finding high-quality solutions in combinatiorial problems.

In regard to heuristics, in [5]–[7] the authors focus on the development of deterministic techniques for designing the collection systems, such as Quality Threshold Clustering (QT) algorithm for grouping the wind turbines, and followed by the interconnection of them by applying Dijkstra’s Minimum Spanning Tree (MST) algorithm. In [6], unsupervised learning algorithms for clustering are applied, which depend strongly on initial conditions, after that by adding Steiner splices, Dijkstra is used leading to a reduction of total length. In [7] is proposed a method for avoiding forbidden areas. All the previous works are intended to cover onshore wind farms. OWFs are the focus of [8]–[10]: in [8] a simple Prim algorithm to form the MST is used, getting solutions for different substation locations. Likewise, in [9], and [10], a deterministic algorithm is proposed using a dynamic objective function depending on the calculated LCOE, modifying in each iteration the cable sizing. In OWFs is preferred by the developers to avoid cable crossings (one installed above the other), because of mainly two reasons: hot-spots can be created due to the contact between those cables, requiring extra thermal insulation to avoid failures, and in case of a single failure on the bottom cable, all those cable on the top would have to be removed due to their physical installation. All the previous works mentioned before have the disadvantage of using deterministic techniques that can get trapped in minimals and disregard cable crossings.

Metaheuristic techniques encompass not only GA, but also others such as Particle Swarm Optimization (PSO). In [11] and [12] a PSO approach is used to optimize not only the collection system, but also the transmission system; through this technique, the main idea is translated from a purely deterministic fashion, to a probabilistic one, where is more likely to find the global minimum, albeit not formal proof in terms of solution quality and time can be formulated. The main disadvantage of the latest works is that the problem is tackled from a MST perspective, meaning that parallel cables can be placed, being this a not-so-common practice by OWF developers. GAs has been used in works such as [4], [13],
and [14]. In [4] is not very clear, among other things, the penalization strategy and how cable crossings can be avoided, therefore their results can not be reproduced. In [13] the search space is artificially restricted, and it seems that crossings are allowed. Similarly, in [14] only radial solutions can be found and it does not optimize for total costs.

The contribution of this paper is to propose a GA algorithm that solves the optimization problem while considering the full search space of the problem. An implementation of genetic algorithm that accounts for cable capacity and cable crossings constraint is proposed and a detailed description of this method is presented. The full proposed method calculates the LCOE associated to the collection system of cables in each solution. At the end the best solution is output in a ergodic fashion.

The paper is structured as follows. The optimization framework is explained in Section II, then in Section III the methodology is explained in detail, in Section IV two case studies are analyzed, and in Section V the conclusions are presented.

II. Optimization Framework

The optimization framework comprises the cost model, the core optimization model, and the general scheme nesting different methods. In developing the framework, several assumptions have been made, as they are presented in the following.

A. Assumptions

- The OWF is composed of one offshore substation (OSS) and a group of wind turbines (WTs). The location of these elements is known.
- The target size of the wind farm is between fifty and sixty turbines.
- The length of the cables is the horizontal trenching length (i.e. the euclidean distance between two points in the plane).
- The topology of the grid is based on a radial network, star network or a combination of both.
- Each cable segment in the collection system must be sized using the standard [15].
- Cable crossings are not allowed, except at the endpoints of each segment, where both the OSS and WTs are located.

B. Cost Model

The cost of the optimization model is given by the cost of the horizontal trenching length of the cables. Given a list of cables with its corresponding electrical parameters, the unitary cost of each cable (in €/km) is calculated by using the cost function (1) proposed in [16]. This cost function uses the rated voltage of the grid, the rated current of each cable, and a set of constants that have been obtained from an empirical study of cable costs. The cost function is scaled to take into consideration macroeconomics phenomenons, such as inflation and exchange rate.

\[ C_t = A_p + B_p e^{(c_{pt} e^{3pt})} \]  

(1)

\[ S_{nt} = \sqrt{3} V_n I_{nt} \]  

(2)

Where:

- \( C_t \): Cost of cable type \( t \) in €/km.
- \( V_n \): Rated line to line voltage level in \( V \).
- \( I_{nt} \): Rated current of cable type \( t \) in A.
- \( A_p, B_p, C_{pt} \): Cost constants. These values are given as tables depending on the voltage level [16].
- \( S_{nt} \): Rated power of cable type \( t \) in VA.

C. Optimization Model

The aim of the optimization model is to find a high quality solution for the OWF collection system design and optimization problem. The following formulation is based on a Capacitated Minimum Spanning Tree (C-MST) problem class, and includes a constraint to restrict cable crossings.

1) Objective Function: Let \( G = (V, E) \) be a connected undirected graph with vertex-set \( V \) modelling the OSS (root) and WTs locations, and a edge-set \( E \) representing the potential connections between them. Introducing the binary decision variable \( x_e \) for each element \( e \in E \), equal to to 1 if the corresponding edge is selected as part of the solution; the variable \( x_e \) defines unequivocally \( G_S = (V, E_S) \), a graph forming a solution, with \( E_S \subset E \). Following up, for the solution \( G_S \), let \( G_{Sk} \) define the set of maximal subgraphs connected to the root by a single edge, such as \( G_{Sk} = \{ G_{Sk_1}, \ldots, G_{Sk_n} \} \), and \( G_{Sk} = (V_{Sk}, E_{Sk}) \), with \( V_{Sk} \subset V \), \( V_{Sk'} \subset V \), \( V_{Sk} \cap V_{Sk'} = \emptyset \), and \( E_{Sk} \subset E \), \( E_{Sk} \cap E_{Sk'} = \emptyset \).

Additionally, let \( T \) be the set of cables considered for the optimization problem. Each cable type \( t \in T \) has an associated unitary cost \( C_t \), and an a maximum capacity \( U_t \), expressed in terms of the nodes (turbines) that supports downstream, where \( U_{max} = \max U_t \).

\[ C_t^t = C_t \cdot \text{length}(e) \]  

(3)

Each edge \( e \in E \) has a cost \( C_{et}^t \) that is defined by (3), in function of the cable type and the edge trenching length. Therefore, let \( x_e^t \) model the binary decision variable, equal to 1 if the edge \( e \) is active through a cable type \( t \), and 0 conversely.

\[ \min \left( \text{CAPEX}_{et} = \sum_{t \in T} \sum_{e \in E} C_{et}^t \cdot x_e^t \right) \]  

(4)

The objective function minimizes the total cable layout cost as defined in (4).
2) Constraints:

\[ \text{deg}(v) \geq 1 \quad \forall v \in V \quad (5) \]

\[ x_e = \sum_{t \in T} x_e^t \quad \forall e \in E \quad (6) \]

\[ \sum_{e \in E} x_e = \|V\| - 1 \quad (7) \]

\[ \|V_{S_k}\| \leq U_{max} \quad \forall G_{S_{k_i}} \in G_{S_k} \quad (8) \]

\[ x_q + x_w \leq 1 \quad \forall \{q, w\} \in C \quad (9) \]

\[ x_e \in \{0, 1\} \quad (10) \]

\[ x_e^i \in \{0, 1\} \quad (11) \]

The enforcement of connectivity between vertices and that the graph is a tree is achieved by satisfying Constraints (5) and (7) simultaneously. The first indicates that all nodes having a degree of at least one. The second one forces the graph to have the same amount of active edges as the total number of nodes minus one. Constraint (6) enforces that each active edge has one and only one type of cable. To abide the cable capacity restrictions, Constraint (8) is implemented. Each set of nodes of the subgraphs \( G_{S_{k_i}} \) has to have a smaller number of nodes than the maximum capacity cable from the set of cables. Finally, the cable crossings restriction is shown in Constraint (9): It is defined by considering a generic input set, \( C \subseteq E \times E \), of pairs of crossing edges with the property that two crossing edges \( q \) and \( w \) cannot be active simultaneously.

After the convergence of the GA, a LCOE\textsubscript{cs} metric is used in order to assess the final quality of the obtained solution, as expressed by (12), where \( r \) is the rate of return, an economic factor regarding bank interest and inflation rate (assumed as 5% in this paper), \( N_y \) is the project lifetime (assumed as 30 years), and \( E_{\text{net}} \) is the produced annual energy accounting for power losses (in MWh) [10].

\[
\text{LCOE}_{\text{cs}} = \frac{\text{CAPEX}_{\text{cs}} r (1 + r)^{N_y}}{(1 + r)^{N_y} - 1} E_{\text{net}} \quad (12)
\]

D. General framework

The optimization framework is composed by two main processes: the solver and a power flow analysis. The solver is based on GAs and it is described in detail in this paper. The power flow analysis consists on a power flow calculation in order to find the losses associated with the solution given by the solver, allowing the calculation of the LCOE\textsubscript{cs}.

The optimization framework workflow is described in Figure 1. Since GAs are non-deterministic and do not guarantee finding the optimal solution, the process is ran \( n \) times, comparing the LCOE\textsubscript{cs} of the different solutions in the end.

\[ \text{LCOE}_{\text{cs}} = \frac{\text{CAPEX}_{\text{cs}} r (1 + r)^{N_y}}{(1 + r)^{N_y} - 1} E_{\text{net}} \quad (12) \]

III. METHODOLOGY

A. Genetic algorithm

Genetic algorithms (GAs) are iterative processes in which a population of candidate solutions, also known as individuals, are evolved towards a better solution. Each individual has a chromosome, which is a set of properties that can be altered through a predefined set of genetic operators. Each individual is a solution to the problem with the variables representing the existence of a cable over an edge of the complete graph using a binary encoding. A more comprehensive explanation of GAs can be found in [17].

\[ \text{LCOE}_{\text{cs}} = \frac{\text{CAPEX}_{\text{cs}} r (1 + r)^{N_y}}{(1 + r)^{N_y} - 1} E_{\text{net}} \quad (12) \]

B. Fitness assessment

The general workflow of the GA can be seen in Figure 2. The genetic operators implemented are reproduction, crossover and mutation. Reproduction, simulating a roulette
the total population size. The individuals are then ranked according to their fitness value. Finally, the population is truncated to the original population size. This approach allows the algorithm to converge faster as it does not need to re-discover solutions discarded in previous generations.

Each individual goes through a fitness assessment that determines the cost of the individual. The fitness assessment implements the mathematical formulation proposed in Section II as a fitness function, (13). The fitness function represents the objective function by assessing the cost of the layout of each individual and penalizing its cost if the constraints are violated.

$$Z = \sum_{i=1}^{n} c_i w_i x_i + P$$  \hspace{1cm} (13)$$

Where:

- $Z$ is the fitness of the individual. If no penalizations apply it is also the final cost of the collection system.
- $c$ is the cost of the cable selected, if no cable has been selected the value is 1.
- $w$ is the trenching length of each edge of the complete graph.
- $x$ is the binary string of variables that represent whether an edge of the complete graph is active of not.
- $P$ is the sum of the penalizations that apply to the individual being assessed.

The implementation of constraints from the mathematical formulation is achieved in two ways. Equation (6) is implemented implicitly through the use of binary variables on the $x$ vector, this only allows for only one cable to be built over an edge. The rest of the constraints are implemented via the use of penalizations.

**Total connectivity** deals with (5), the implementation is achieved doing a depth-first search and checking the connectivity of all nodes. **Tree graph**, checks that the graph is indeed a tree, the condition is described in (7). **Cable capacity** is the implementation of the constrain described in (8). First the tree is transformed into a directional tree rooted at the substation. Then a depth-first search is performed and the number of nodes accessible from each node has to be smaller than the maximum capacity cable available. **Cable crossings restriction**, (9), is implemented by checking intersections in a set composed of all possible pairs of active edges in the solution.

A hierarchy is established among the penalizations for the constraints through the use of conditional functions and through the cost given to each penalization. The need for a hierarchy is two-fold, the first reason being that in order to calculate the number of nodes per branch the graph needs to be a tree. The second reason is that certain functions used to analyze the cable capacity and cable crossing constraints use large amounts of computational power and it is unproductive to run them if the solution is already known to be unfeasible.

In terms of the order of the constraint assessment, first both the total connectivity and tree graph constraints are determined. If both constraints are met then capacity constraint is calculated. Again if this constraint is satisfied then the cables crossing restriction constrained is analyzed. In this way time is not wasted in computational intensive tasks.

In terms of the penalization cost they are ranked by importance, with the most important constraints having the higher costs. Each penalization is assigned a base cost differentiated by several orders of magnitude from the others, ensuring that, in the case of proportional penalizations the constraints do not interfere with each other.

![Fitness assessment flowchart](image)

The total connectivity penalization is assigned the highest cost to ensure that all elements are connected. The tree constraint penalization is proportional to the number of extra edges that impede the formation of a tree graph. The cable constraint penalization is proportional to the number of cables that do not meet the constraint. Finally, the cable crossing restriction penalization has the lowest cost and it is proportional to the number of crossings detected.

After each of the constraint analysis the fitness cost is calculated which includes the addition of the corresponding penalizations.
penalizations.

The cable selection process is a method for choosing the appropriate cable from a list for each segment. This method is only done once the cable capacity constraint is met, as it requires that at least the highest capacity cable of the list is supported in the solution. The process assigns the smallest cross section possible to each segment according to the number of wind turbines being supported upstream. In doing so the cost is also minimized, as the size of the cross section is correlated to the cost of the cable.

C. Termination criteria

The genetic algorithm finishes the optimization process and outputs a solution when either of the following conditions are met:

- Iteration number: The process stops after a set number of iterations \( it_{max} \).
- Stall of the fitness value: The process stops if the fitness value does not change for a fixed amount of iterations \( it_{stall} \), shown in (14) where \( i \) is a set that varies each generation according to the iteration number \( it \).

\[
Z_i = Z_{i+1} \quad \forall i \in \{it - it_{stall} + 1, it - 1\} \tag{14}
\]

IV. CASE STUDIES

Two case studies are analyzed to exemplify the usefulness of the optimization framework here presented, the two OWFs are Ronne Bank North (RBN) and Ronne Bank South (RBS). These are proposals part of the Baltic InteGrid Project and show a grid pattern with bigger spacing between wind turbines on the horizontal axis than in the vertical axis. Each case is presented with three collection system designs; one empirical design (ED) and two solutions given by the GA (GA1 and GA2). Due to the stochasticity of the GA, it converges into different solutions for GA1 and GA2.

Table I contains the list of cables considered for the optimization process, with the Capacity given in terms of the number of WTs rated at 8 MW supported by each cable in a 66 kV grid. Table II summarizes the characteristics of each OWF. Table III shows the key results.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>LIST OF AVAILABLE CABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross section [mm²]</td>
<td>95 120 150 185 240 300 400 500 630 800 1000</td>
</tr>
<tr>
<td>Capacity [Number of WT]</td>
<td>4 4 5 6 6 7 8 9 10 11 11</td>
</tr>
<tr>
<td>Price [€/km]</td>
<td>270 287 303 327 362 395 440 497 558 628 694</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>WIND FARM CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBN</td>
<td>RBS</td>
</tr>
<tr>
<td>WTs [number]</td>
<td>27 53</td>
</tr>
<tr>
<td>Rated power per WT [MW]</td>
<td>8 8</td>
</tr>
<tr>
<td>Grid voltage [kV]</td>
<td>66 66</td>
</tr>
<tr>
<td>AEP [MWh]</td>
<td>905,675 1,777,806</td>
</tr>
</tbody>
</table>

The first case study, dealing with RBN, shows that the GA solutions improve both the active power losses and the investment cost, leading up to a 6% decrease compared to the empirical design in terms of the LCOEcs. Figure 5 shows that the GA solution prefers vertical connections between the wind turbines as these are shorter than the horizontal ones. The GA uses higher capacity cables by clustering the WT in bigger groups, effectively reducing the number of cables coming out of the OSS. The empirical design, Figure 4, has five cables entering the OSS while the GA solution, Figure 5, has four.
GA shows a tendency of using the vertical connections instead of the horizontal ones as they are shorter. It also shows the same behaviour as in the previous case by clustering the wind turbines in groups of 5-8 WT. In areas that are difficult to reach areas it opts for a higher capacity cable, such as the subtree that starts in the OSS and branches at turbine number 36.

The collection system design of offshore wind farms is becoming increasingly more complex, thus finding economical based solutions has become a more arduous task. This article presents an optimization framework based on GA that is successful in solving the collection system design for a typical OWFs. The framework includes the exploration of the full search space of problem and several realistic constraints. The effectiveness is proven in two case studies showing reductions on the LCOEcs of up to 8% when compared to empirical solutions. Future work includes the analysis of the computational performance of the optimization framework and the inclusion of a reliability analysis to solve looped and meshed topologies.

**ACKNOWLEDGMENT**

This research has received funding from the Baltic InteGrid Project (http://www.baltic-integrid.eu/).

**REFERENCES**


