

# Modified GA and Data Envelopment Analysis for Multistage Distribution Network Expansion Planning Under Uncertainty

David T.-C. Wang, Luis F. Ochoa, *Member, IEEE*, and Gareth P. Harrison, *Member, IEEE*

**Abstract**—An approach is proposed to solve multistage distribution network expansion planning problems considering future uncertainties, guiding the planner from production of expansion plans, evaluation of the plans under various future uncertain scenarios, to the selection of the best strategy. A new balanced genetic algorithm (BGA) is introduced that improves the intensification of the solution search procedure by trading-off diversification ability. This facilitates searching for the optimal solution, but also the efficient production of suboptimal solutions for the planner to take into consideration. The features of the BGA allow a multistage planning problem to be solved more efficiently; the BGA can consider a set of expansion plans in an early planning stage in a single run and produce planning strategies required to solve network problems in a later stage along the planning horizon. The overall performance of each plan under different uncertain scenarios is evaluated using a modified data envelopment analysis to assist decisions on which solution to adopt. The approach is applied to a multistage ‘greenfield’ distribution network expansion problem considering scenarios for the location of future loads. The results clearly show the advantages of the approach over more conventional methods.

**Index Terms**—Network planning, uncertainties, genetic algorithms, data envelopment analysis.

## I. INTRODUCTION

IN liberalized electricity markets, distribution network companies have an obligation to operate their networks in an efficient, secure and economic way, and to provide unbiased opportunities for the connection of demand and generation customers. Capital intensive infrastructure is required for power networks to satisfy increasing demand and integrate generation while complying with technical regulations.

The creation of optimal network expansion plans, implying that the investment shall be fully justified in the future without stranded assets, is of paramount importance to network companies. The transmission and distribution expansion planning (TDEP) problem is complex and, by its nature,

mixed integer, non-linear and non-convex. There is no consensus on which optimization method is more capable than others in solving the problem. Comprehensive reviews of optimization techniques proposed in recent decades for solving TDEP problems are given in [1-3]. The use of mathematical optimization methods such as linear programming [4], mixed-integer programming [5, 6], non-linear programming [7], Benders decomposition [8] and branch-and-bound techniques [9, 10], have been observed to have potential convergence issues and locally optimal final solutions. While evolutionary algorithms like genetic algorithms (GAs) [11-13], simulated annealing [14], tabu search [15, 16], and particle swarm [17], might, in some cases, produce better solutions, they are subject to stochastic errors due to the algorithms involved.

In addition to the optimization methods, other techniques are required if the TDEP problem takes place under uncertainties related to demand growth, or new load and generation sizes and locations. First, the uncertainties can be modeled as scenarios using the probabilistic approach [18, 19] if the probabilistic data for the input variables are available, otherwise the planner can draw out possible scenarios based on experience and knowledge or, for instance, using fuzzy set theory [20, 21]. Second, a decision support tool, such as those based on the probability choices [22, 23] and risk analysis [22-24], is applied to assign a single index value to a planning solution representing the overall goodness or risk of adopting the solution considering its performance in the different scenarios and their corresponding weighting factors. The weighting factors can be based on the probability of occurrence of the scenario in which the performance of the solution is measured, on the planner’s judgment or using Pareto analysis. Furthermore, a suboptimal solution could be a better choice than the optimum one for the same scenario, if it has a more satisfactory performance across other scenarios. Therefore, adopting optimization methods like the immune-based evolutionary algorithms [19] or improved GAs [25] could be advantageous as they have been credited with producing both optimal and suboptimal solutions that can increase the number of candidate plans for the planner to take into consideration.

Dynamic models [25, 26] or multistage static planning [27, 28] are adopted if the asset connection schedule needs to be defined along the planning horizon. Two approaches are

---

This work is part-funded through the EPSRC Supergen V, UK Energy Infrastructure (AMPerES) grant in collaboration with UK electricity network operators working under Ofgem’s Innovation Funding Incentive scheme – full details on <http://www.supergen-amperes.org/>.

D. T.-C. Wang is with Smarter Grid Solutions Ltd, West Regent Street, Glasgow, G2 2RU, U.K. (email: david.wang@smartergridsolutions.com).

L. F. Ochoa and G. P. Harrison are with the Institute for Energy Systems, School of Engineering, The University of Edinburgh, Edinburgh, EH9 3JL, U.K. (e-mail: luis\_ochoa@ieee.org, gareth.harrison@ed.ac.uk).

generally adopted. The forward approach solves TDEP problems consecutively from the beginning to the end of the planning horizon, while the backward approach first seeks the best expansion solution at the end of the planning horizon which influences expansion planning in earlier years. The backward approach has been said to produce better results as it allows more expensive equipment to be installed in the early years that become truly cost-effective options in solving later TDEP problems [25]. However, this conclusion was made without reference to uncertainty. A reasonable assumption can be made that for multistage planning under uncertainty [29, 30], the uncertainty is greater in the long term than in the short term. The implications are that the backward approach could create problems if the uncertainties are not thoroughly considered, leading to an expansion plan which at the end of the planning horizon is less optimal or even infeasible. The forward approach would at least ensure that effective solutions are adopted for solving the TDEP problems in the early planning years for which the uncertainty could be probabilistically modeled with a degree of confidence. Further complexity can arise in planning distribution networks due to more onerous voltage constraints and the need to enforce radial configurations [31].

This paper introduces a process for solving a multistage, static, distribution expansion planning problem under uncertainty using the forward approach. It harnesses a new balanced genetic algorithm (BGA) to ensure solution diversity as well as modified data envelopment analysis (MDEA) to ensure a flexible and objective comparison of plans by avoiding the need for precise weighting factors for the specific attributes. The multistage aspect of the planning problem requires assessment of multiple later stage planning solutions that evolve from early stages. Examination of all potential future developments of each planning solution is time consuming. However, the BGA is able to speed up this process by exploiting the generational structure of GA search to efficiently generate expansion plans from a set of early stage solutions. The approach is applied to a multistage ‘greenfield’ distribution network expansion problem considering uncertainties surrounding future major load locations. Section II explains the concept and the design of the BGA while Section III introduces MDEA. The case study and results are shown in Section IV, followed by a discussion of results and the conclusions.

## II. BALANCED GENETIC ALGORITHM

### A. Concept, Solution Codification and Quality

Conventional GAs allow newly-generated solutions to have great discrepancies from existing solutions. The high ‘diversification’ increases the chances of avoiding the search being trapped into a local optimum. Consequently, the ability of the GA to search for solutions that resemble existing solutions – ‘intensification’ – is limited. As a result, if an old solution is replaced by a new one, large pieces of information inside the old solution could be eliminated, increasing the risk of losing critical components of the global optimal solution. The balanced GA therefore aims to improve the intensification

of the search by limiting the diversification of conventional GAs such that only a small amount of existing information is replaced at each solution substitution. This means suboptimal solutions will be more thoroughly searched by the BGA than conventional GAs before convergence to an optimum.

As depicted in Fig. 1(a), solutions are represented as chromosomes, comprising a string of genes that contain information that characterizes the solution. In the example of a distribution expansion planning problem shown in Fig. 1(b), the genes indicate the potential network paths. Here, the information denotes the number of new circuits built in each of the paths (e.g., 0, 1, 2).

For the distribution planning problem the quality (or fitness) of the network configuration (solution) is defined by its cost (normally the cheaper the better); and its feasibility in terms of voltage, thermal and connectivity constraints (i.e., that there are no isolated network sections) as determined by power flow analysis.

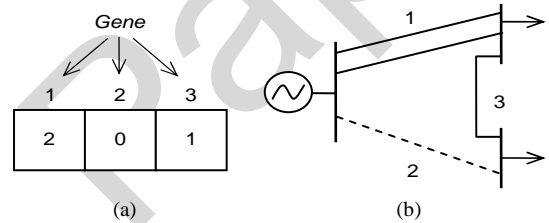


Fig. 1. (a) BGA solution representation in a (b) network expansion plan.

### B. Population, Genetic Operators and Solution Substitution

New solutions are generated by a random process as in a conventional GA and put into a pool if those solutions are feasible and no identical solutions already exist. This solution generation process stops when a specified pool size is reached.

A crossover operator is applied to the parents (two randomly selected solutions from the pool) to produce two new solutions, the offspring. Only the genes that differ between parents are eligible for crossover and, of these, a limited set of genes will be randomly chosen from a predefined maximum number. Fig. 2 shows an example where parents X and Y have five genes that differ between them (indicated by the dashed arrows). In this case where a given maximum of genes (less equal than five) will be able to crossover, only two genes (3, 7) are randomly selected to do so. Since there is strict limitation on the numbers of genes allowed to crossover in the BGA, the structure of the offspring would be very similar to the parent. This is critical in ensuring good information diversity.

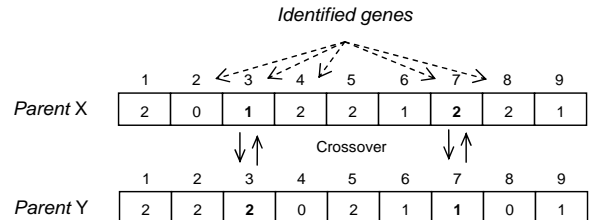


Fig. 2. New chromosomes produced by crossover.

The mutation rate is kept low, typically below 1%. If it occurs, a stochastic process is again applied to select one gene from each offspring and alter the information inside the gene.

In the BGA, offspring can only substitute one of its parents having a similar structure, provided that the offspring is a feasible solution, is better than the parent and that no identical solutions are in the pool. Offspring that fails to meet the requirements are discarded. The novel combination of the process of crossover and solution substitution applied here allows more subtle change to genes in the pool towards optimality. Merely limiting the number of genes in the crossover would not achieve the same aim.

### C. Filtering the Final Population

A filtering method is used to eliminate solutions from the final pool which have similar topological structure but are inferior to other solutions. Filtering avoids the need to consider all suboptimal solutions, by identifying the more worthwhile ones as candidate solutions.

The filtering ranks the planning solutions in the final pool by their performance (from the best to the worst). A term ‘degree of uniqueness’ is defined here as the number of differences (in terms of branches) necessary for the structures of two solutions to be regarded as distinct. Starting from the solution with the best performance as a reference, other solutions that have fewer than the number of differences specified by the degree of uniqueness are said to be similar solutions to the best one and are therefore eliminated. Then the remaining solutions are compared with the solution with the second best performance in the pool and the size of the pool is further reduced by eliminating similar solutions. The process continues until all the configurations left in the pool have been selected as reference and each solution possesses a certain degree of topology uniqueness.

Although no constraints or special methods in the BGA check the radiality of a network solution, there is a high probability that the final distribution network solutions will be radial as the primary objective of minimizing planning costs will promote solutions that have fewer branches than similar parents; these are more likely to be radial. Filtering of the final population eliminates premature meshed networks if they have similar structures to other radial, but cheaper, network solutions.

### D. Efficient Generation of Future Expansion Solutions

Extending the example in Fig. 1, suppose that it is foreseen that the expansion planning at stage 1 will be followed by an additional two loads (stage 2) connected via five additional paths, as shown in Fig. 3(a). With  $n$  planning solutions considered at stage 1, it would be time inefficient to execute separate searches for quality future developments in stage 2 for each expansion plan considered at stage 1. However, with a few modifications the BGA can generate future expansion plans for all candidates together.

The first modification is made in the phase of generating initial solutions for stage 2 expansion planning. Every solution generated must carry one of the  $n$  candidates at stage 1 with additional genes indicating which candidate the stage 2 solution is developed from. Fig. 3(b) shows a stage 2 solution developed from the second stage 1 candidate. To create an impartial assessment, there will be an equal number of stage 2

solutions in the pool developed from each of the  $n$  candidates. With a pool size of  $p$ , the pool will initially feature  $p/n$  feasible stage 2 solutions randomly generated from each candidate.

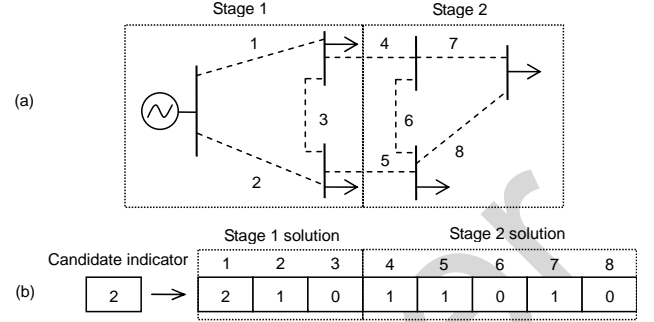


Fig. 3. (a) An example of multi-stage network expansion planning and (b) chromosome representation.

The crossover operator is modified such that the number of single or double lines in a path adopted by the expansion solution at stage 1 within a stage 2 solution cannot be decreased (as this would allow removal of infrastructure) but may be increased. For example, Fig. 4 shows that genes 1 and 6 selected for crossover between parents X and Y developed from two first stage solutions. The information in gene 1 of parent X cannot be substituted by that of parent Y as this would remove a line that has ‘already’ been built. However, parent Y will accept the information from gene 1 of parent X indicating an additional line built as a part of the stage 2 expansion solution.

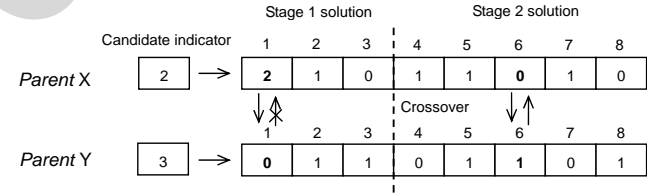


Fig. 4. Further modification of the crossover operator in BGA.

## III. EVALUATING EXPANSION PLANS

Data envelopment analysis (DEA) is a decision support tool developed by Charnes *et al.* [32] to assist decision makers in comparing the performance of a group of candidate solutions and selecting the best one. DEA is a linear programming (LP) procedure that does not require exact weighting values to be assigned to each attribute of a candidate and provides a flexible and objective evaluation. The ‘relative efficiency’ (RE) of each candidate solution (configuration) is defined as the ratio of its aggregate output to its aggregate input; the best candidate has the highest RE. The general form for the RE of candidate  $i$  with multiple inputs and outputs is:

$$\max \text{RE} = \frac{\sum_{t=1}^T B^t Y_i^t}{\sum_{r=1}^R A^r X_i^r} \quad (1)$$

where  $X_i^r$  is the  $r$ -th input,  $Y_i^t$  the  $t$ -th output and  $A^r$  and  $B^t$  are the respective weighting factors assigned by the LP.

In this application, all the planning options at stage 1 need to be evaluated in terms of their ability to be ‘uncertainty-proof’. Hence, for a given configuration at stage 1, the input

will be the corresponding stage 1 planning cost, and the outputs will be the investments needed for each future scenario (stage 2). Thus, the efficiency of the configuration will show its effectiveness in keeping costs low while being able to cater for uncertainties. Hence each of  $n$  candidate configurations has a single input and  $T$  outputs representing each future scenario. To enable solution, the LP is reformulated such that the RE of configuration  $k$  is as follows:

$$\max \text{RE} = \sum_{t=1}^T B^t Y_k^t \quad (2)$$

$$\text{s.t. } AX_k = 1 \quad (3)$$

$$\sum_{i=1}^T B^t Y_i^t - AX_i \leq 0 \quad (4)$$

$$i = 1, \dots, k, \dots, n \wedge A, B^t \geq 0. \quad (5)$$

Constraints (3)-(5) ensure that each candidate's RE will not exceed unity. A candidate with a RE of unity is the best candidate for the specific conditions defined by the weighting factors assigned by the LP. A candidate with RE less than 1 has at least one solution better than itself (akin to dominance). However, where several candidates are assigned the maximum RE, they cannot be distinguished. To allow ranking of these a modified data envelopment analysis (MDEA) was introduced in [33, 34] that allows the RE to exceed unity by relaxing the constraints for each candidate so that aggregate output can exceed aggregate input. In order to maximize the RE, the traditional DEA LP process can assign very large values to one weighting factor and small ones to others. This may conflict with the decision-maker's view of the relative importance or likelihood of occurrence of scenarios. The weighting factors should reflect their judgment and additional constraints inserted to limit the weighting factor values within a credible range that reflects reality. Thus, the MDEA [33] is created by extending the constraints to (6) and using (7) instead of (5):

$$W_{\min}^{pq} \leq B^p / B^q \leq W_{\max}^{pq} \quad ; \quad 1 \leq p, q \leq T \quad (6)$$

$$i = 1, \dots, n, i \neq k \wedge A, B^t \geq 0. \quad (7)$$

Here  $B^p$  and  $B^q$  are the weighting factors for the  $p$ -th and  $q$ -th outputs, whose ratio is constrained by upper and lower limits  $W_{\min}^{pq}$ ,  $W_{\max}^{pq}$ . To ensure consistency these values will be set by the decision maker (say) based on assumptions over likelihood of occurrence of scenarios.

#### IV. APPLICATIONS AND RESULTS

##### A. Multistage Distribution Network Expansion Planning

The approach is demonstrated on a two-stage, greenfield 11kV distribution network expansion problem based on [29]. Fig. 5(a) shows the first stage (S1) where two 2.5MW loads known to be locating at nodes 15 and 16 require connection to the existing grid. The dashed lines indicate 34 potential paths. The planner has been informed that three additional load centers will emerge in the future and that this should be planned for as a second stage expansion. The locations of the three 2MW second-stage loads are still uncertain although, based on the information available, the planner has defined

two possible scenarios for their location. These will connect at nodes 17, 18 and 19 although the precise location of these depends on the scenario: in scenario S2-1 (Fig. 5(b)) an additional 10 potential network paths exist while scenario S2-2 (Fig 5(c)) has 8 new paths. Thermal and voltage ( $\pm 6\%$  pu) constraints apply in all cases. Only one type of underground line is used in the analysis:  $0.12 + j0.0765 \Omega/\text{km}$ , with standard capacity of 15MVA and a cost of US\$100k/km. Detailed branch data is given in the Appendix.

The objective is to find an expansion plan in stage 1 which is not only cost-competitive relative to other candidate plans in S1, but will also lead to robust network expansion in both second stage alternatives (S2-1 and S2-2).

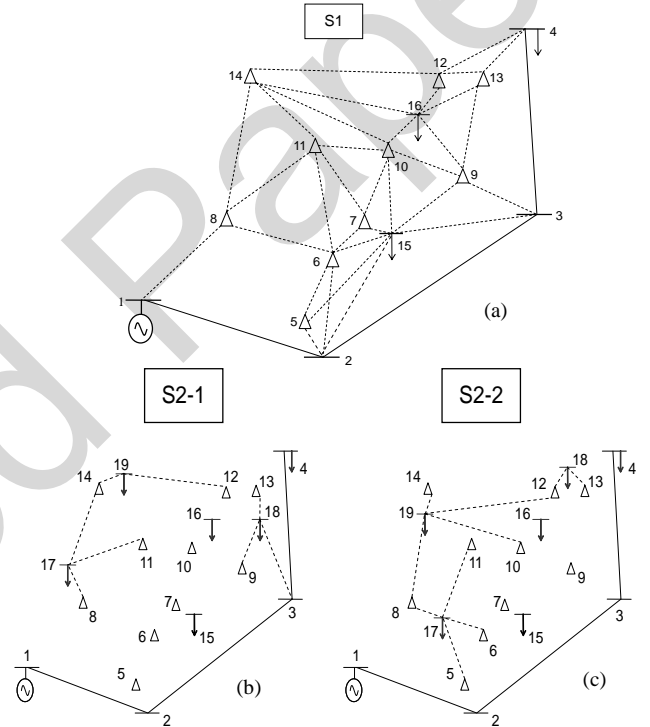


Fig. 5. Greenfield distribution network planning (a) first stage S1; (b) second stage scenario 1 S2-1; (c) second stage scenario 2 S2-2.

##### B. Network Expansion Planning in Stage 1

The BGA is implemented using the computer language Python and interacts with the PSS/E power flow software to model the network and simulate AC power flows. The objective is to search for solutions with minimum costs while subject to voltage, thermal and network connectivity constraints. In order to show the advantages of the BGA, the results are compared with those obtained using a conventional GA with single crossover point and non-generational population substitution technique where if the offspring is feasible and better than the worst solution in the pool, the offspring then replaces it.

The initial population for both algorithms is set to 60 chromosomes. The search stops if no solution substitutions occur after 200 consecutive runs. Ten trials for each algorithm are attempted. For the BGA, the number of genes allowed to crossover is set to 5. It is assumed that crossover will occur every iteration to produce new solutions and the mutation rate

is set to 1%. The filtering procedure is applied to both algorithms, in which the degree of uniqueness is set to 2 such that, after filtering, no network configurations in the pool share more than two routes. The results are summarized in Table I.

The best expansion strategy in stage 1 (S1) costs \$613.6k. BGA found this optimal solution in every trial, while it appears in seven out of the ten trials of the GA. More importantly, after the solution filtering phase, on average 19 solutions remain with the BGA compared with only 7 for the GA. This implies that, for this specific case, when a variety of solutions is needed, the BGA is 2.7 times more efficient than the GA, provided that the latter needs to generate no repeated solutions between each trial – which is very unlikely.

Due to limited information exchange and the substitution method adopted in the BGA, some solutions may converge to local optimal points, producing relatively expensive solutions. However, those solutions may be future-proof and of value to the planner. Despite the wider solution range produced by BGA, the last columns of Table I show that even if only the solutions with costs lower than the most expensive solution found by the GA (\$834.2k) are considered, the BGA search in the vicinity of the optimum is more thorough. In addition, all the solutions have radial configurations with meshed solutions eliminated by the filtering process due to their higher costs. Appendix-B outlines other measures of algorithm efficiency.

Table II shows the top five expansion solutions (P1 to P5) in S1 obtained by the BGA algorithm in one trial. They are assumed here to be the candidates on which future network development will be based, i.e., the candidates for the two stage 2 scenarios (S2-1 and S2-2).

TABLE I  
OPTIMAL SOLUTIONS FOR S1 WITH BGA AND GA (A = AVERAGE)

Trial	Best solution US\$K		Number of solutions		Solution range US\$K		Solutions under US\$834.2k	
	BGA	GA	BGA	GA	BGA	GA	BGA	GA
1	613.6	613.6	22	5	613.6 - 1209.7	613.6 - 717.3	17	5
2	613.6	613.6	20	11	613.6 - 1363.7	613.6 - 801.5	18	11
3	613.6	662.4	16	6	613.6 - 1050.0	662.4 - 804.5	15	6
4	613.6	630.3	17	10	613.6 - 1526.0	630.3 - 807.4	16	10
5	613.6	613.6	16	8	613.6 - 792.1	613.6 - 804.8	16	8
6	613.6	613.6	24	6	613.6 - 1494.6	613.6 - 705.3	21	6
7	613.6	613.6	18	6	613.6 - 849.2	613.6 - 804.8	17	6
8	613.6	630.3	25	6	613.6 - 1200.0	630.3 - 834.2	21	6
9	613.6	613.6	18	8	613.6 - 1226.5	613.6 - 804.8	16	8
10	613.6	613.6	18	5	613.6 - 1383.6	613.6 - 705.3	17	5
	10/10	7/10	A=19	A=7			A=17	A=7

TABLE II  
TOP FIVE PLANNING CANDIDATES IN STAGE 1

Planning Candidate	Network Configuration	Cost (US\$K)
P1	2-15, 10-15, 10-16	613.6
P2	3-9, 9-15, 9-16	630.3
P3	2-15, 4-12, 12-16	662.4
P4	2-15, 7-10, 10-16	663.0
P5	2-6, 6-15, 10-15, 10-16	688.6

### C. Network Expansion Planning in Stage 2

Using the BGA the performance of the five stage 1 planning strategies (P1–P5) are examined for each stage 2 scenario. Fifty feasible solutions for scenarios S2-1 and S2-2 are generated from each of the five candidates in stage 1,

resulting in 250 initial solutions in each pool. The search stops if no substitutions occur in the pool after 500 iterations. The crossover, mutation and degree of uniqueness are the same.

Ten trials are attempted for each scenario and the top solutions found are shown in Table III. Overall, the final planning cost is considerable in meeting the requirements of scenario S2-2 over S2-1. If S2-1 happens, the best planning strategy is developed from P5, which is the most expensive among all the five candidates in S1. In this case, four of the top S2-1 solutions are based on P3. As for S2-2, the best expansion option along with two other options is found to contain the structure of P2.

TABLE III  
TOP SOLUTIONS FOUND IN S2-1 AND S2-2 (US\$K)

Rank	S2-1	Developed from	S2-2	Developed from
1	1298.3	P5	1516.0	P2
2	1372.0	P1	1524.4	P3
3	1386.7	P4	1549.1	P1
4	1387.3	P5	1559.4	P2
5	1424.3	P3	1560.4	P1
6	1431.9	P3	1592.1	P4
7	1458.1	P3	1611.9	P2
8	1459.3	P2	1624.1	P5
9	1466.9	P2	1624.8	P4
10	1470.3	P3	1633.8	P1

### D. Selecting the Best Planning Candidate in Stage 1

The relative efficiency (RE) of the five S1 strategies (P1–P5) in terms of their performance under both scenarios is calculated by both conventional DEA and MDEA. One input and two outputs are taken to measure the efficiency of each candidate. The input X is the cost of a candidate in S1. The two outputs are related to the costs of future planning solutions in S2-1 and S2-2. The idea is to maximize the efficiency which, in this case, the costs being cheaper. Therefore, the outputs, Y1 and Y2 are, respectively, the reciprocals of the cost (US\$K<sup>-1</sup>) of the best solutions found in S2-1 and S2-2 (Table IV) that developed from the candidate, multiplied by 10<sup>6</sup>. Table IV shows the values of input and outputs of each strategy.

TABLE IV  
FOR EACH SOLUTION P1 – P5: INPUT AND OUTPUT VARIABLES; DEA AND MDEA EVALUATION

Candidate in S1	Input X	Outputs		DEA		MDEA	
		Y1	Y2	RE	Rank	RE	Rank
P1	613.6	728.9	645.5	1.00	1	1.04	1
P2	630.3	685.3	659.6	1.00	1	0.98	2
P3	662.4	702.1	656.0	0.94	2	0.93	3
P4	663.0	721.1	628.1	0.92	3	0.91	4
P5	688.6	770.2	615.7	0.94	2	0.89	5

Here, it is assumed that the planner can judge that the probability of scenario S2-1 occurring is between 0.4 and 0.8. The scenarios are mutually exclusive and the weighting factors for the outputs depend on the relative probability of the scenarios  $W_{\min}^{12}$  and  $W_{\max}^{12}$ . The ratio of weighting factors assigned to Y1 and Y2 are within the range 0.67 (=0.4/0.6) to 4 (=0.8/0.2). The RE of each strategy is calculated by the DEA and MDEA with the results presented in Table IV.

Both stage 1 plans P1 and P2 yield the maximum RE calculated by DEA. Since the planning cost under S2-1 is

cheaper than S2-2 for all candidates, DEA tends to assign a much larger weighting factor to Y1 than Y2; as a consequence P5 appears to be an attractive strategy to adopt due to its excellent performance in S2-1. However, MDEA is able to further distinguish between P1 and P2, indicating P1 as the better strategy. By limiting the ratio of the output weighting factors based on the relative probability of occurrence of S2-1 and S2-2 using the MDEA, P5 becomes the least efficient strategy to adopt. As a result, P1 is the best network planning strategy in this case and could be adopted to solve the problem in S1. Fig. 6 shows the best expansion planning choices for S2-1 and S2-2 based on first stage plan P1 (thicker lines).

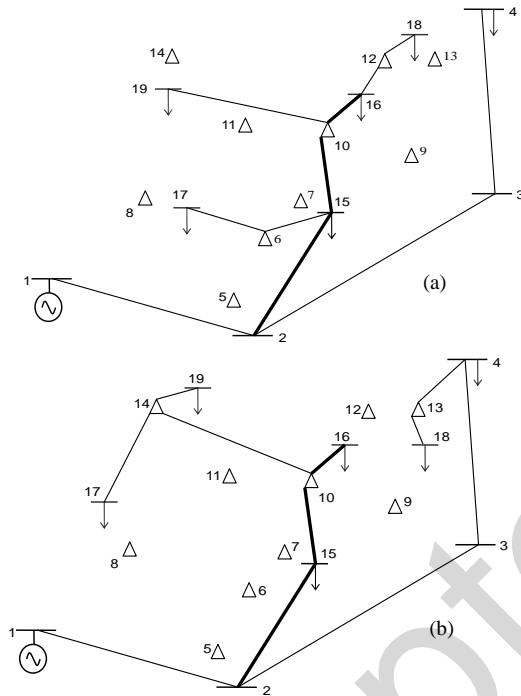


Fig. 6. The best expansion plans in (a) S2-1 (b) S2-2 based on P1.

## V. DISCUSSION

Approaches, such as this, that provide better understanding of the robust cost-effective expansion plans are highly desirable to distribution planners. In addition, the use of industry-standard power flow software driven by advanced use of its programming capability provides a robust method that lends itself to direct application in network planning. While the analysis here is limited to checking voltage and thermal limit feasibility using the power flow engine, additional aspects such as fault level checks, can be considered. The method lends itself to capturing the uncertainties brought about by incorporating distributed generation (DG) within the expansion planning process. This is particularly the case as availability of resources, planning and financial incentives drive the location and size of DG plant. Work is ongoing to consider DG within this framework.

This work demonstrates the effectiveness of using a balanced genetic algorithm combined with modified data envelopment analysis to provide the distribution planner with a minimum-cost set of network configurations that are robust

to uncertainty. While the method is applied here to a greenfield cases, it is expected that problems involving highly developed distribution networks with smaller search spaces due to limited new paths can also benefit, particularly using the MDEA approach. While only a pathway search was demonstrated here, the methodology can handle alternative expansion options, such as reconductoring, needed to cope with local demand growth. With few modifications the approach is also applicable to transmission planning problems.

There is no fundamental restriction to the use of the approach on larger scale problems. Additionally, the solution filtering and MDEA process are applicable as an add-on to other meta-heuristic or classical expansion planning methods. It is important to note that in common with all meta-heuristic or classical optimization methods, the BGA cannot universally guarantee a globally optimal solution to a non-linear problem.

The novelty in the work arises from (1) the development of the balanced GA that uses strict limits on crossover and solution substitution to promote higher intensity searching; (2) the filtering of solutions and excluding those that are insufficiently unique; and (3) the combination of a multistage approach with MDEA to allow differentiation between alternatives under uncertainty.

It would be interesting to compare BGA with other algorithms designed to improve the quantity of optimal and suboptimal solutions. However, the results showed that BGA only requires simple modification from a conventional GA to achieve the purpose. Nevertheless, BGA will suffer from a certain degree of stochastic errors and could be still trapped into a local optimum, despite more subtle information addition and elimination in the crossover and solution substitution phases. To counter this potential disadvantage, a large pool size would be normally required to ensure good information diversity in the initial solutions. However, the technique used in [25] can be used in the phase of initial solution generation in the BGA for effectively generating limited numbers of initial solutions with good information diversity. Another improvement perceived from the current BGA is the limit on the number of genes allowed to crossover ( $k$ ) can be flexibly changed to adjust the search intensification through the process to yield good results in a satisfactory time frame.

The solution filtering process not only assists in identifying distinct solutions but also promotes radial network solutions. The degree of uniqueness setting is crucial: if the value is too large few very different solutions pass through, increasing the chance of missing competitive solutions; if the value is too small, large numbers of solutions (including meshed) would increase the computational time. For a multi-objective problem in which a better solution does not necessary mean the cheapest with fewest branches, a constraint of radiality may still be required to be inserted into the algorithm.

## VI. CONCLUSIONS

The paper presents an approach to tackle network expansion planning under uncertainty, guiding the planner from generation of expansion plans, through evaluation of the plans under various future uncertain scenarios, to the selection of the

best strategy. The balanced genetic algorithm (BGA) and modified data envelopment analysis (MDEA) were introduced to generate optimal planning strategies and evaluate the performances against identified uncertainties. Here, uncertainties are treated as different future scenarios. The approach is applied to a green-field distribution expansion planning problem considering possible future demand sites. By enhancing the intensification of the searching procedure, BGA produces better range of quality planning solutions than a conventional GA. MDEA allows a distinction to be made between two high performing solutions based on their probability of occurrences and the judgment of the planner, allowing more flexible evaluation of each planning candidate compared with conventional methods.

## VII. APPENDIX

### A. Line Characteristics

TABLE A.I  
CHARACTERISTICS OF LINE OPTIONS AT STAGE 1 (S1)

From	To	Distance (km.)	From	To	Distance (km.)
1	8	2.802	7	11	1.931
6	8	2.530	10	11	1.503
5	6	1.709	11	14	2.247
2	5	1.063	3	9	2.247
2	6	2.309	9	10	1.616
5	15	2.900	10	14	3.360
2	15	3.046	14	16	3.448
6	15	1.487	12	14	3.622
6	7	1.140	3	13	3.384
6	11	2.508	9	13	2.600
8	11	3.106	9	16	1.836
8	14	3.863	13	16	2.062
3	15	8.932	12	16	0.671
9	15	2.220	12	13	1.612
7	15	5.39	4	12	2.907
10	15	1.924	4	13	1.803
7	10	1.879	10	16	1.166

TABLE A.II  
CHARACTERISTICS OF LINE OPTIONS AT STAGE S2-1 (LEFT) AND S2-2 (RIGHT)

From	To	Distance (km.)	From	To	Distance (km.)
5	17	2.508	8	17	1.603
6	17	1.746	11	17	2.702
8	17	0.806	14	17	2.500
11	17	2.546	14	19	0.985
13	18	0.707	12	19	2.816
12	18	0.949	3	18	2.773
8	19	3.138	9	18	1.924
10	19	2.731	13	18	0.707
12	19	3.437			
14	19	0.922			

### B. Efficiency of BGA vs GA

Other measures of algorithm efficiency include the number of feasible solutions generated in each trial and how many of these are different. Table B.1 shows that on average the BGA visited 4218 different solutions per trial compared with 1494 solutions in GA; this shows that the BGA explored the solution space more thoroughly. The GA experiences rapid degradation in information diversity in the solution pool, limiting the potential for new solutions and increasing the chance of finding existing solutions. The BGA's greater intensification and restrained randomness creates a more consistent search towards final solutions. Fig. B.1 shows the frequency of occurrence of the top 10 solutions across all trials. Apart from the best solution, the chance of getting other

solutions is up to 50% with the GA, compared with at least 70% for the BGA.

In terms of the computations performance, using an Intel Core2 2.13 GHz PC, the BGA solved the stage 1 problem in around 9 minutes (with the GA slightly faster). The larger stage 2 problem with its stricter stopping criteria saw the duration increase to 45 minutes.

TABLE B.I  
FEASIBLE SOLUTIONS AND DIFFERENT FEASIBLE SOLUTIONS GENERATED IN EACH OF THE 10 TRIALS IN BGA AND GA

Algorithm	Feasible solutions			Different feasible solutions		
	High	Low	Average	High	Low	Average
BGA	8564	5627	7237	4430	4053	4218
GA	11990	7991	9724	1805	1148	1494

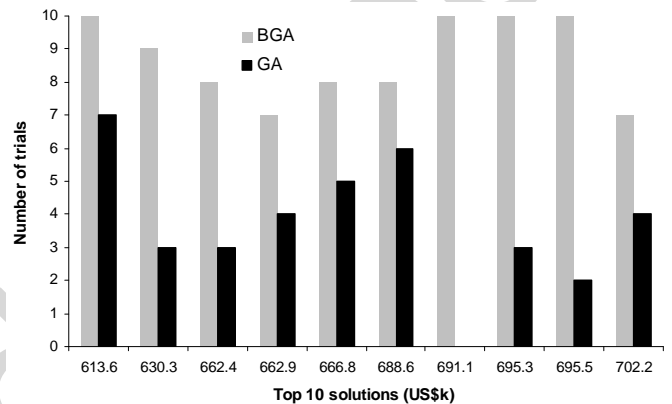


Fig. B.1. Number of trials in which GA and BGA find the top 10 solutions.

## REFERENCES

- [1] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas, "Classification of publications and models on transmission expansion planning," *IEEE Trans. on Power Systems*, vol. 18, no. 2, pp. 938-946, May 2003.
- [2] R. C. Bansal, "Optimization Methods for Electric Power Systems: An Overview," *International Journal of Emerging Electric Power Systems*, vol. 2, no. 1, 2005.
- [3] C. W. Lee, S. K. K. Ng, J. Zhong, and F. F. Wu, "Transmission expansion planning from past to future," in *Proc. 2006 Power Systems Conference and Exposition (PSCE)*, pp. 257-265.
- [4] S. H. M. Hashimoto, R. Romero, and J. R. S. Mantovani, "Efficient linear programming algorithm for the transmission network expansion planning problem," *IEE Proceedings Generation, Transmission & Distribution*, vol. 150, no. 5, pp. 536-542, Sep. 2003.
- [5] L. Bahiense, G. C. Oliveira, M. Pereira, and S. Granville, "A mixed integer disjunctive model for transmission network expansion," *IEEE Trans. on Power Systems*, vol. 16, no. 3, pp. 560-565, Aug. 2001.
- [6] N. Alguacil, A. L. Motto, and A. J. Conejo, "Transmission expansion planning: a mixed-integer LP approach," *IEEE Trans. on Power Systems*, vol. 18, no. 3, pp. 1070-1077, Aug. 2003.
- [7] Z. M. Al-Hamouz and A. S. Al-Faraj, "Transmission expansion planning using nonlinear programming," in *Proc. 2002 IEEE/PES Transmission and Distribution Conference and Exhibition Asia Pacific*, vol. 1, pp. 50-55.
- [8] S. Binato, M. V. F. Pereira, and S. Granville, "A new benders decomposition approach to solve power transmission network design problems," *IEEE Trans. on Power Systems*, vol. 16, no. 2, pp. 235-240, May 2001.
- [9] S. Haffner, A. Monticelli, A. Garcia, J. R. S. Mantovani, and R. Romero, "Branch and bound algorithm for transmission system expansion planning using a transportation model," *IEE Proceedings Generation, Transmission & Distribution*, vol. 147, no. 3, pp. 149-156, May 2000.
- [10] M. J. Rider, A. V. Garcia, and R. Romero, "Branch and bound algorithm for transmission network expansion planning using dc model," in *Proc. 2007 IEEE/PES PowerTech*, pp. 1350-1355.
- [11] E. L. Da Silva, H. A. Gil, and J. M. Areiza, "Transmission network

- expansion planning under an improved genetic algorithm," *IEEE Trans. on Power Systems*, vol. 15, no. 3, pp. 1168-1174, Aug. 2000.
- [12] X. Feng, Y. Liao, J. Pan, and R. E. Brown, "An application of genetic algorithms to integrated system expansion optimization," in *Proc. 2003 IEEE/PES General Meeting*, vol. 2, pp. 741-746.
- [13] R. Romero, M. J. Rider, and I. d. J. Silva, "A metaheuristic to solve the transmission expansion planning," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 2289-2291, Nov. 2007.
- [14] R. Romero, R. A. Gallego, and A. Monticelli, "Transmission system expansion planning by simulated annealing," *IEEE Trans. on Power Systems*, vol. 11, no. 1, pp. 364-369, Feb. 1996.
- [15] E. L. Da Silva, J. M. A. Ortiz, G. C. De Oliveira, and S. Binato, "Transmission network expansion planning under a tabu search approach," *IEEE Trans. on Power Systems*, vol. 16, no. 1, pp. 62-68, Feb. 2001.
- [16] A. Sadegheh and P. R. Drake, "System network planning expansion using mathematical programming, genetic algorithms and tabu search," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1557-1566, Jun. 2008.
- [17] Y.-X. Jin, H.-Z. Cheng, J.-y. Yan, and L. Zhang, "New discrete method for particle swarm optimization and its application in transmission network expansion planning," *Electric Power Systems Research*, vol. 77, no. 3-4, pp. 227-233, Mar. 2007.
- [18] G. Celli, S. Mocci, F. Pilo, and R. Cicoria, "Probabilistic optimization of mv distribution network in presence of distributed generation," in *Proc. 2002 14th Power Systems Computation Conference (PSCC'02)*, pp. 1-7.
- [19] E. G. Carrano, F. G. Guimaraes, R. H. C. Takahashi, O. M. Neto, and F. Campelo, "Electric distribution network expansion under load-evolution uncertainty using an immune system inspired algorithm," *IEEE Trans. on Power Systems*, vol. 22, no. 2, pp. 851-861, May 2007.
- [20] J. Choi, A. A. El-Keib, and T. Tran, "A fuzzy branch and bound-based transmission system expansion planning for the highest satisfaction level of the decision maker," *IEEE Trans. on Power Systems*, vol. 20, no. 1, pp. 476-484, Feb. 2005.
- [21] I. d. J. Silva, M. J. Rider, R. Romero, and C. A. F. Murari, "Transmission network expansion planning considering uncertainty in demand," *IEEE Trans. on Power Systems*, vol. 21, no. 4, pp. 1565-1573, Nov. 2006.
- [22] V. Miranda and L. M. Proenca, "Probabilistic choice vs. risk analysis-conflicts and synthesis in power system planning," *IEEE Trans. on Power Systems*, vol. 13, no. 3, pp. 1038-1043, Aug. 1998.
- [23] G. Carpinelli, G. Celli, F. Pilo, and A. Russo, "Distributed generation siting and sizing under uncertainty," in *Proc. 2001 IEEE/PES PowerTech*, pp. 7.
- [24] R. Fang and D. J. Hill, "A new strategy for transmission expansion in competitive electricity markets," *IEEE Trans. on Power Systems*, vol. 18, no. 1, pp. 374-380, Feb. 2003.
- [25] A. H. Escobar, R. A. Gallego, and R. Romero, "Multistage and coordinated planning of the expansion of transmission systems," *IEEE Trans. on Power Systems*, vol. 19, no. 2, pp. 735-744, May 2004.
- [26] H. K. M. Youssef, "Dynamic transmission planning using a constrained genetic algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 23, no. 8, pp. 857-862, Nov. 2001.
- [27] V. A. Levi and M. S. Calovic, "A new decomposition based method for optimal expansion planning of large transmission networks," *IEEE Trans. on Power Systems*, vol. 6, no. 3, pp. 937-943, Aug. 1991.
- [28] S. Binato and G. C. Oliveira, "A heuristic procedure to cope with multi-year transmission expansion planning," in *Proc. 1995 IEEE/PES PowerTech*.
- [29] P. M. S. Carvalho, L. A. F. M. Ferreira, F. G. Lobo, and L. M. F. Baruncho, "Distribution network expansion planning under uncertainty: a hedging algorithm in an evolutionary approach," *IEEE Trans. on Power Delivery*, vol. 15, no. 1, pp. 412-416, Jan. 2000.
- [30] S. Vassena, P. Mack, P. Rouseaux, C. Druet, and L. Wehenkel, "A probabilistic approach to power system network planning under uncertainties," in *Proc. 2003 IEEE/PES PowerTech*, vol. 2, pp. 6.
- [31] E. M. Carreno, R. A. Romero, and A. Padilha-Feltrin, "An efficient codification to solve distribution network reconfiguration for loss reduction problem," *IEEE Trans. on Power Systems*, vol. 23, no. 4, pp. 1542-1551, Nov. 2008.
- [32] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429-444, Nov. 1978.
- [33] P. Andersen and N. C. Petersen, "A procedure for ranking efficient units in data envelopment analysis," *Management Science*, vol. 39, no. 10, pp.

1261-1264, Oct. 1993.

- [34] F. Yang, L. Lujie, H. Rong, and C. Jialin, "Modified interval DEA method for economic evaluation of distribution network," in *Proc. 2007 2nd International Conference on Innovative Computing, Information and Control (ICICIC)*, pp. 598.

## BIOGRAPHIES



**David T.-C. Wang** received a BEng degree from the University of Auckland, New Zealand, in 2005, and MSc and PhD degrees from the University of Edinburgh, U.K. in 2006 and 2010, respectively. He recently joined Smarter Grid Solutions Ltd, Glasgow, U.K.

His interests include distribution network planning and analysis considering the impact of distributed generation.

Dr Wang is a member of the Institution of Engineering and Technology (IET).



**Luis F. Ochoa** (S'01-M'07) is a Research Fellow in the School of Engineering, University of Edinburgh, U.K. He obtained a BEng degree from UNI, Lima, Peru, in 2000, and the MSc and PhD degrees from UNESP, Ilha Solteira, Brazil, in 2003 and 2006, respectively.

His current research interests include network integration of distributed energy resources and distribution system analysis.

Dr. Ochoa is a member of the IET.



**Gareth P. Harrison** (M'02) is a Senior Lecturer in Energy Systems in the School of Engineering, the University of Edinburgh, Edinburgh, U.K.

His research includes network integration of distributed generation and analysis of the impact of climate change on the electricity industry.

Dr. Harrison is a Chartered Engineer and member of the IET.