

Multiple Power Line Outage Detection in Smart Grids: A Probabilistic Bayesian Approach

A. Ahmed, M. Awais, M. Naeem, M. Iqbal, W. Ejaz, A. Anpalagan, and H. Kim

Abstract—Efficient power line outage identification is an important step which ensures reliable and smooth operation of smart grids. The problem of multiple line outage detection (MLOD) is formulated as a combinatorial optimization problem and known to be NP-hard. Such a problem is optimally solvable with the help of an exhaustive evaluation of all possible combinations of lines in outage. However, the size of search space is exponential with the number of power lines in the grid, which makes exhaustive search infeasible for practical sized smart grids. A number of published works on MLOD are limited to identify a small, constant number of lines outages, usually known to the algorithm in advanced. This work applies the Bayesian approach to solve the MLOD problem in linear time. In particular, the work proposes a low complexity estimation of outage detection (EOD) algorithm, based on the classical estimation of distribution algorithm (EDA). Thanks to an efficient thresholding routine, the proposed solution avoids the premature convergence and is able to identify any arbitrary number (combination) of line outages. The proposed solution is validated against IEEE-14 and 57 bus systems with several random line outage combinations. Two performance metrics namely success generation ratio (SGR) and percentage improvement (PI) have been introduced in the paper, which quantify the accuracy as well as convergence speed of proposed solution. The comparison results demonstrate that the proposed solution is computationally efficient and outperforms a number of classical meta-heuristics.

Index Terms—Line outage identification, smart grids, Estimation of distribution algorithm, power networks.

I. INTRODUCTION

Phasor measurement units (PMUs) are devices which provide synchronized, real time, dense and highly accurate measurements of current and voltage phasors [1]. PMUs are widely used for automation and state information collection of wide-area measurement systems [2], [3]. They have been installed by a large number of utility companies across the globe and help to perform key system tasks including protection, observation and control of power networks [4]. The overview of security and privacy related issues in smart grid are discussed in [5].

A. Ahmed, M. Awais, M. Iqbal and M. Naeem are affiliated with the COMSATS Institute of Information Technology Wah, 47040, Pakistan (e-mail: ashfaqahmed, muhammadawais, driqbal @ciitwah.edu.pk)

M. Naeem is also with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: muhammadnaeem@gmail.com).

W. Ejaz and A. Anpalagan are with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, M5B 2K3, Canada (email: waleed.ejaz@ieee.org; alagan@ee.ryerson.ca).

H. Kim is with Department of Electronics Engineering, Sogang University, Seoul, Republic of Korea (Corresponding author) (e-mail: hongseok@sogang.ac.kr).

A. Motivation

PMUs play a vital role in the identification of line outages of geographically large power grids. They are installed on a number of power system buses. One of the PMUs is taken as a reference and the phasor measurements of other PMUs are relative to it. When a single or multiple power lines undergo an outage, the relative phase angles are changed and hence the system's power flow is disturbed. Due to overloading, the cascaded tripping of power lines leads to a complete system blackout [6].

Another smart grid automation technique is supervisory control and data acquisition (SCADA) [7]. However, PMUs have demonstrated a significant advantage over SCADA in a number of ways including measurement resolution, state observation capability and coverage area [8]. Some key differences between SCADA and PMU technologies are highlighted in Table I. In addition to power system protection, PMUs are also used for implementing a load shedding method that consists of simultaneous reactive and active power to address the frequency and voltage stability issues respectively [9]. Moreover, PMUs increase the power service quality by precise analysis, automated correction and disturbance recording of sources of system degradation [10].

B. Related Work

The state-of-the-art protection techniques for power grids can be classified into two broad categories. Firstly, it includes strategies which determine the optimal location and number of PMUs to achieve a grid-wide observability. Secondly, the techniques which exploit rich PMU data for efficient identification of power lines in outage. Some important works are discussed as follows. The work in [11] proposes a two step PMU placement technique. At first, linear programming is used to

TABLE I
SCADA vs. PMU [8].

Parameters	SCADA	PMU
Measurement Resolution	Analog 2-4 samples per cycle	Digital Up to 60 samples per cycle
Observability	Steady state	Dynamic/Transient state
Monitoring Phase Angle Measurement	Local-area No	Wide-area Yes

obtain an optimum upper bound on the number of PMUs. Then, a branch and bound algorithm is used to obtain a near optimal placement solution. An integer linear programming based PMU placement is also considered in [12]. Moreover, the work also studies the effect of a single PMU loss or a single line outage on optimal PMUs placement. The work in [13] proposes a probabilistic approach for PMU placement. Two main objectives are achieved. At first, system observability is ensured with minimum number of PMUs. Secondly, the multi-objective problem is solved with the help of an augmented epsilon constraint method to achieve Pareto optimal points. In [14], an artificial intelligence approach based on support vector machines (SVM) is proposed. Using the PMU data, the work trains all candidate network topologies under several loading conditions in order to detect single line outages. In [15], PMU's phasors and other system topology parameters are used for optimal detection of single line outages. This work is extended in [16], in which the pre- and post-event information of the topology is used to solve the double line outage identification problem.

Mathematically, the approaches used in [15] and [16] are combinatorial in nature. Multiple line outage detection (MLOD) problem is also a kind of combinatorial optimization problem, which requires comprehensive information of all possible network topologies. Therefore, due to drastic increase in the search space, the approaches proposed in [15] and [16] cannot be used. In [17], a solution is proposed which adopts a compressed sensing technique. The solution named as least absolute shrinkage and selection operator (LASSO) is capable to detect outage over multiple lines. However, the accuracy of solution is highly affected by a parameter which depends on a-priori information about perturbed noise and outage probability. The authors in [18] propose a cross entropy optimization (CEO), a stochastic search method for detection of multiple line outages. CEO demonstrates better detection capability as compared to LASSO. In [19], authors suggested an approach named quickest change detection. The authors adopted a linearized model of power system and proposed a statistical method for detection and isolation of line outages. In [20], the authors propose a method which uses internal system measurements in order to track the external network changes. Further, integer linear programming is used for the detection of minor and major network changes.

In [21], the authors proposed a method which uses voltage deviation and overload performance metrics in order to rank line outage contingencies in AC-DC systems. The work uses Jacobian matrix to calculate current compensations for line outages and uses a coupled AC-DC flow model. IEEE 14-bus system is used to test the proposed model and compare it with rigorous AC-DC load flow methods. Authors in [22] proposed a method based on compressive sensing to identify multiple line outages and sudden impedance changes in power system.

The work in [23] simulates the power system state after the event of a sudden outage of an element. The authors adopted the AC power flow model and proposed a simple method to determine the important parameters of the formulated problem. In [24], $N - 1$ criterion is investigated for an evolving power grid along with its impact on cascaded line

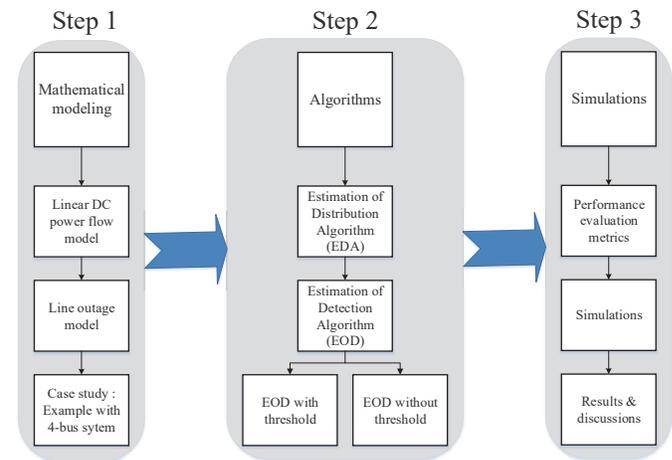


Fig. 1. Schematic overview of the paper.

outages. In [25], the measurement errors in DC power flow are identified by exploiting the singularity of the impedance matrix and sparsity of the error vector. The proposed sparsity-based decomposition-DC power flow approach is able to compute the measurement errors accurately, when the method is verified on the IEEE 118-bus and 300-bus systems. In [26], the work exploits the historical weather data from NCDC, along with the historical interruption information, and predicts the power system interruptions. A minimum variance unbiased estimators (MVUEs) is derived in [27] for active power based on the voltage phase at each node of the power system. It is concluded that power system operations can be more accurately estimated if there is lower correlation between the noise vector elements. In [28], a power loss minimization scheme is presented. The proposed scheme is based on oblivious network design, and is therefore termed as oblivious routing-based power flow method. The approach can be applied for the large scale power loss minimization.

An oblivious routing economic dispatch (ORED) algorithm is presented in [29], for congestion in smart power systems. It is claimed that the proposed method works independent of the network topology, and is therefore suitable for large scale economic dispatch problems. The ORED algorithm outperforms the other state-of-the-art economic dispatch algorithms in terms of congestion management and power loss minimization. In [30], fuzzy models are built of overhead power line outages. In [31], Galton-Watson branching process model is proposed in order to estimate the load shed propagation in cascaded line outages. In [32], PMU data is exploited in order to identify the line outage locations within a specified geographical area of the power system. The proposed approach assumes PMUs to be placed within and at the boundary of the area. The proposed approach is unable to identify the line outages outside the boundary of the area.

Table II presents a summary of published solutions for line outage detection (LOD). Broadly, the published solutions are categorized on the basis of line outage detection flexibility. In the first column, there are solutions which are limited to detect only single line outage. In the second column, there

TABLE II
SUMMARY OF SOME PUBLISHED WORKS ON POWER LINE OUTAGE DETECTION.

Ref.	Single line outage	Multiple Fixed line outages	Multiple Variable line outages
[Proposed]			✓
[14]	✓		
[15]	✓		
[16]		✓	
[17]		✓	
[19]	✓		
[20]	✓		
[21]	✓		
[22]		✓	
[23]		✓	
[24]		✓	
[30]		✓	
[31]		✓	
[32]		✓	

are solutions with multiple line outage detection (MLOD) capability. However, they are named as fixed MLOD solutions as they are limited to identify only a fixed (usually small) number of line outages. Almost all published works on outage detection fall in to the first two categories.

C. Contributions

In this work, we have proposed a solution based on a probabilistic Bayesian approach which solves the MLOD problem in linear time. The proposed solution is named as estimation of outage detection (EOD) and based on estimation of distribution algorithm (EDA). The main contribution of this work are:

- The proposed EOD solution has capability to detect and identify any combination and number of lines in outage.
- An efficient thresholding routine is applied to update the probability distribution of candidate solutions. This helps to avoid the local optima. The EOD algorithm with thresholding is thus named as EOD-Threshold.
- The proposed algorithm is simulated for several IEEE bus systems with random line outage combinations. The comparison results show that the proposed EOD solution outperforms a number of other meta-heuristics in terms of success and convergence rates.

D. Paper organization

In Fig. 1, the schematic overview of the paper is shown. The rest of this paper is organized as follows. In Section II, system model is described with MLOD problem formulation. Section III discusses the basics of EDA along with the proposed EOD and EOD-Threshold algorithms. Section IV discusses the simulation results and comparisons. Finally, the conclusions are drawn in Section V.

II. SYSTEM MODEL AND PROBLEM DEFINITION

A. Linear DC Power Flow Model

Consider a power network consisting of L transmission lines and N buses. This network can be modelled in the form of a directed graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, where $\mathcal{N} = \{1, 2, \dots, N\}$ is the set of nodes (buses) and $\mathcal{E} = \{(m, n)\} \subseteq \mathcal{N} \times \mathcal{N}$ denotes the set of possible edges. The incident matrix of \mathcal{G} is a matrix \mathbf{M} of size $N \times L$. An entry at row $n \in \{1, 2, \dots, N\}$ and column $l \in \{1, 2, \dots, L\}$ of \mathbf{M} is given by:

$$M_{nl} = \begin{cases} 1 & , \text{ if line } l \text{ originates from bus } n \\ -1 & , \text{ if line } l \text{ terminates at bus } n \\ 0 & , \text{ otherwise} \end{cases} \quad (1)$$

Fig. 2(a) shows an example of a network with $L = 5$ lines and $N = 4$ buses. For this network, the matrix \mathbf{M} is obtained as:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix}. \quad (2)$$

This work adopts a linear DC power flow model [17]. The conservation of power must be satisfied i.e. for each bus, the total power inflow must be equal to the sum of all power outflows. Mathematically,

$$P_n = \sum_{m \in \mathcal{N}(n)} P_{nm}, \quad (3)$$

where P_n denotes the power flowing in to bus n , P_{nm} is the power flowing out from bus n to bus m and $\mathcal{N}(n)$ denotes the set of all buses which have direct connection with bus n . Moreover,

$$P_{nm} = \frac{1}{x_{mn}} (\theta_n - \theta_m), \quad (4)$$

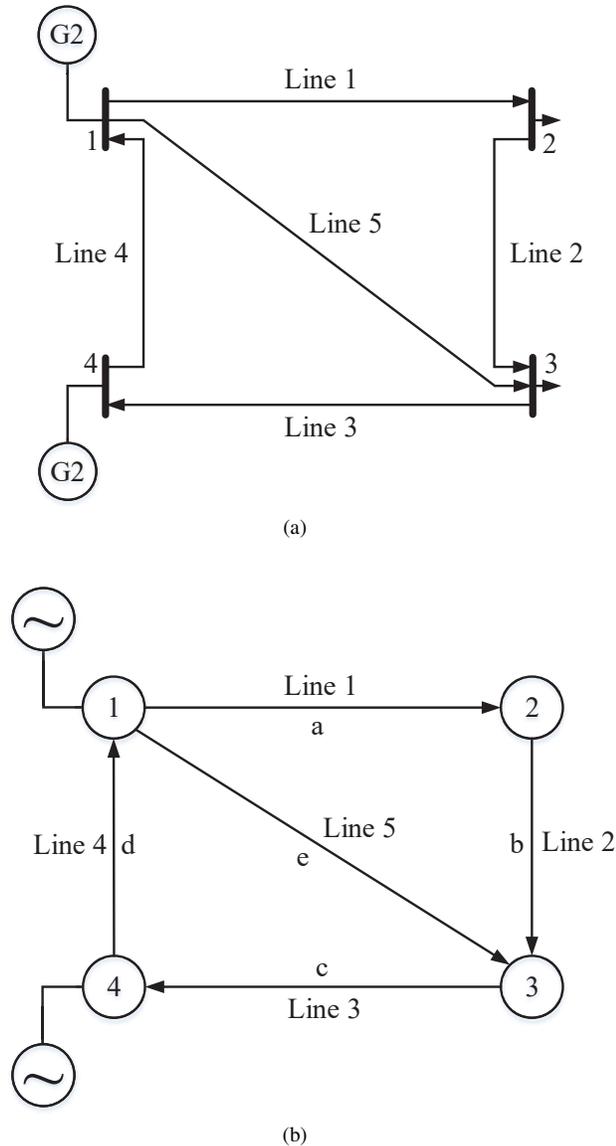


Fig. 2. Transmission line network (a) interconnect grid system and (b) power flow model.

where, $x_{nm}=x_{mn}$ is the reactance of line between buses n and m and θ_n and θ_m are their respective voltage phasor angles. With real power injections $\mathbf{p} = [p_1, \dots, p_N]^T \in \mathbb{R}^N$ and voltage phasor angles $\boldsymbol{\theta} \in [\theta_1, \dots, \theta_N]^T \in \mathbb{R}^N$, the above equations can be stacked in vector notation as

$$\mathbf{p} = \mathbf{B}\boldsymbol{\theta}, \quad (5)$$

where $\mathbf{B} = [B_{nm}] \in \mathbb{R}^N$ is an $N \times N$ matrix with entry at m -th row and n -th column given by,

$$B_{mn} = \begin{cases} -x_{mn}^{-1} & , \text{ if } (m, n) \in \mathcal{E} \\ \sum_{n \in \mathcal{N}(m)} x_{mn}^{-1} & , \text{ if } m = n \\ 0 & , \text{ otherwise} \end{cases} \quad (6)$$

The above equation shows that \mathbf{B} is dependent upon the network topology as well as line reactance values x_{mn} . The

TABLE III
DESCRIPTION OF MAIN NOTATIONS USED IN THE SYSTEM MODEL

Symbol	Details
L	Number of power lines in the network
N	Number of power buses in the network
\mathcal{G}	Weighted graph of network
\mathcal{E}	Set of edges in \mathcal{G}
\mathcal{L}	Set of lines in the network
\mathcal{N}	Set of buses in the network
\mathbf{M}	Bus line incidence matrix of \mathcal{G}
\mathbf{p}	Real vector of length N of power injections
\mathbf{D}_x	$L \times L$ diagonal matrix with line reactances as diagonal entries
$\boldsymbol{\eta}$	Noise perturbation vector
\mathbf{b}	A $1 \times L$ binary vector. A bit 0 denotes a normal line whereas, 1 denotes an outaged line
x_l	Reactance of line l that belongs to set \mathcal{L} or \mathcal{E}
P_n	Total power flowing into bus n
P_{nm}	Power flowing out from bus n to m
θ_n	Pre-outage voltage phasor angle measured at bus n
\mathbf{B}	Laplacian matrix of pre-outage event
θ'_n	Post event phasor angle measured at bus n
$\tilde{\mathbf{B}}$	Laplacian matrix of post-outage
\mathbf{m}_l	Incidence vector of single line l
$\tilde{\mathbf{B}}$	$\mathbf{B}' - \mathbf{B}$
B_{mn}	Single reactance between bus m and n

topological changes in network as a result of line outages affect the voltage phasor angle vector $\boldsymbol{\theta}$ in (5). This topological effect can be modelled by expressing \mathbf{B} as a Laplacian matrix of weighted graph \mathcal{G} and representing it as [17]

$$\mathbf{B} = \mathbf{M}\mathbf{D}_x\mathbf{M}^T = \sum_{l=1}^L \left(\frac{1}{x_l} \mathbf{m}_l \mathbf{m}_l^T \right), \quad (7)$$

where the diagonal matrix \mathbf{D}_x has line reactance x_l^{-1} as its l -th diagonal entry. For the power network for Fig. 2 (a), the power flow graph is represented in Fig.2 (b) The matrix \mathbf{B} takes the form

$$\mathbf{B} = \begin{bmatrix} \frac{1}{a} + \frac{1}{d} + \frac{1}{e} & \frac{-1}{a} & \frac{-1}{e} & \frac{-1}{d} \\ \frac{-1}{a} & \frac{1}{a} + \frac{1}{b} & \frac{-1}{b} & 0 \\ \frac{-1}{e} & \frac{-1}{b} & \frac{1}{b} + \frac{1}{c} + \frac{1}{e} & \frac{-1}{c} \\ \frac{-1}{d} & 0 & \frac{-1}{c} & \frac{1}{c} + \frac{1}{d} \end{bmatrix} \quad (8)$$

where real numbers a, b, \dots, e denote respectively, the reactance of the lines 1 to 5.

Table III lists all notations used in this section.

B. Line Outage Model

Given a linear model of DC power flow (5), vector \mathbf{p} of real power injections and vector $\boldsymbol{\theta}$ of pre-event phase angles are related to each other by the topology dependent susceptance matrix \mathbf{B} . As a result of network topological changes caused by an outage over one or multiple lines, the pre-outage directed

graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ is transformed into a post-outage directed graph $\mathcal{G}' = \{\mathcal{N}, \mathcal{E}'\}$, where $\mathcal{E}' = \mathcal{E} \setminus \tilde{\mathcal{E}}$ is the set of lines in outage. The post outage directed graph remains connected and grid is assumed to be in quasi-stable state. The DC power flow following the outage event is expressed as

$$\mathbf{p}' = \mathbf{B}'\boldsymbol{\theta}' = \mathbf{p} + \boldsymbol{\eta}, \quad (9)$$

where \mathbf{B}' is the post outage susceptance matrix, $\boldsymbol{\theta}'$ is the post outage phasor vector and $\boldsymbol{\eta}$ is the noise vector which contains perturbations between \mathbf{p} and \mathbf{p}' . The value $\boldsymbol{\eta}$ contains the DC approximation errors and is usually modeled a vector of Gaussian distribution with covariance matrix $\sigma_{\boldsymbol{\eta}}^2 \mathbf{I}$ and zero mean [33]. Substituting (5) in (9) yields

$$\mathbf{B}\boldsymbol{\theta} + \boldsymbol{\eta} = \mathbf{B}'\boldsymbol{\theta}' = \mathbf{B}\boldsymbol{\theta}' - \tilde{\mathbf{B}}\boldsymbol{\theta}', \quad (10)$$

where $\tilde{\mathbf{B}} = \mathbf{B} - \mathbf{B}'$, is the weighted Laplacian of outaged lines in $\tilde{\mathcal{E}}$ and expressed as,

$$\tilde{\mathbf{B}} = \sum_{l \in \tilde{\mathcal{E}}} \frac{1}{x_l} \mathbf{m}_l \mathbf{m}_l^T = \mathbf{M} \mathbf{D}_{\mathbf{x}} \{\mathbf{b}\} \mathbf{M}^T, \quad (11)$$

where \mathbf{b} is an L -dimensional binary vector whose element b_l can be written as:

$$b_l = \begin{cases} 1 & , l \in \tilde{\mathcal{E}} \\ 0 & , \text{otherwise} \end{cases}. \quad (12)$$

Considering the phase angle change vector $\tilde{\boldsymbol{\theta}} \triangleq \boldsymbol{\theta}' - \boldsymbol{\theta}$ and substituting in difference model (9) yields

$$\mathbf{y} \triangleq \mathbf{B}\tilde{\boldsymbol{\theta}} = \tilde{\mathbf{B}}\boldsymbol{\theta}' + \boldsymbol{\eta}. \quad (13)$$

Substituting (11) into (13) and applying the equality $\{\mathbf{b}\} \mathbf{M}^T \boldsymbol{\theta}' = \{\mathbf{M}^T \boldsymbol{\theta}'\} \mathbf{b}$ yields

$$\mathbf{y} = \mathbf{M} \mathbf{D}_{\mathbf{x}} \text{diag} \{\mathbf{M}^T \boldsymbol{\theta}'\} \mathbf{b} + \boldsymbol{\eta}, \quad (14)$$

Defining the notation

$$\mathbf{A}_{\boldsymbol{\theta}'} \triangleq \mathbf{M} \mathbf{D}_{\mathbf{x}} \text{diag} \{\mathbf{M}^T \boldsymbol{\theta}'\}. \quad (15)$$

yields the following equation.

$$\mathbf{y} = \mathbf{A}_{\boldsymbol{\theta}'} \mathbf{b} + \boldsymbol{\eta}, \quad (16)$$

where the matrix $\mathbf{A}_{\boldsymbol{\theta}'} \in \mathbb{R}^{N \times L}$ depends on $\boldsymbol{\theta}'$. It turns out that, if the pre-event network topology (i.e. \mathbf{M} and $\mathbf{D}_{\mathbf{x}}$) and the post-event phasor angle vector $\boldsymbol{\theta}'$ are given, then $\mathbf{A}_{\boldsymbol{\theta}'}$ can be computed via (15).

C. Line outage identification problem

Given the noisy observation \mathbf{y} in (14), the objective of optimization problem is to obtain the vector \mathbf{b} i.e. identify the lines that belong to $\tilde{\mathcal{E}}$. Formally, the optimization problem for power line outage identification can be formulated as,

$$\text{P1 : } \hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in \{0,1\}^L} \mathcal{F}_1(\mathbf{b}; \mathbf{y}, \mathbf{A}_{\boldsymbol{\theta}'}), \quad (17)$$

where \mathcal{F}_1 is the objective function of \mathbf{b} in the model (16) which is defined as

$$\text{P1 : } \hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in \{0,1\}^L} \|\mathbf{y} - \mathbf{A}_{\boldsymbol{\theta}'} \mathbf{b}\|_2^2. \quad (18)$$

The above objective function is evaluated on the basis of l_2 norm of least square criterion. For a network consisting of L lines, the number of possible combinations of line outages is 2^L . Hence, exhaustive enumeration of all possible line outages is not feasible for practical sized networks. In order to solve this problem in linear time, this work considers the probabilistic Bayesian evolutionary approach which is meta-heuristic in nature and computationally efficient as compared to exhaustive search.

III. PROPOSED ESTIMATION OF OUTAGE DETECTION (EOD) AND EOD-THRESHOLD ALGORITHMS.

In order to provide a better insight into the proposed solution, this section begins by discussing the main features of classical EDA. The discussion is extended later to the proposed EOD and EOD-Threshold algorithms for MLOD problem.

A. EDA Bayesian Algorithm

Estimation of distribution algorithms (EDAs) are the iterative, population based stochastic optimization methods. They are based on probabilistic modelling of candidate solutions to the problem. Unlike the classical evolutionary algorithms e.g. genetic algorithm, mutation or crossover operations are not used to generate the new population. Instead, the individuals of new population are generated based on sampling of probability distribution of best selected candidates of previous population. Moreover, while in other evolutionary algorithms, the new candidates are generated by an *implicit distribution* defined by variation variables, in EDAs the interrelations between the candidates are expressed through an *explicit probability distribution* encoded by a multivariate normal distribution, Bayesian network or another model class. The binary EDA is formally outlined by the following steps.

1. Generation of Initial Population. The EDA begins with a generation of an initial population, a matrix of dimensions $P_s \times L$, where P_s denotes the number of candidate solutions, i.e., the population size and L denotes the size of each individual candidate. Each candidate individual is represented by a row vector $X = (x_1, x_2, x_3, \dots, x_L)$, where $x_i \in \{0, 1\}$. The initial population is represented as,

$$\Lambda_0 = \begin{pmatrix} X^1 \\ X^2 \\ \vdots \\ X^{P_s} \end{pmatrix} = \begin{pmatrix} x_1^1 & x_2^1 & \cdots & x_L^1 \\ x_1^2 & x_2^2 & \cdots & x_L^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{P_s} & x_2^{P_s} & \cdots & x_L^{P_s} \end{pmatrix} \quad (19)$$

The initial population is typically obtained by sampling over an initialization model that represents uniform (equally likely) distribution over admissible solutions.

2. Fitness Evaluation. After generating the initial population, the iterative process begins until some termination criterion is not met. During each iteration, all individuals of current population are first evaluated for a given objective/fitness

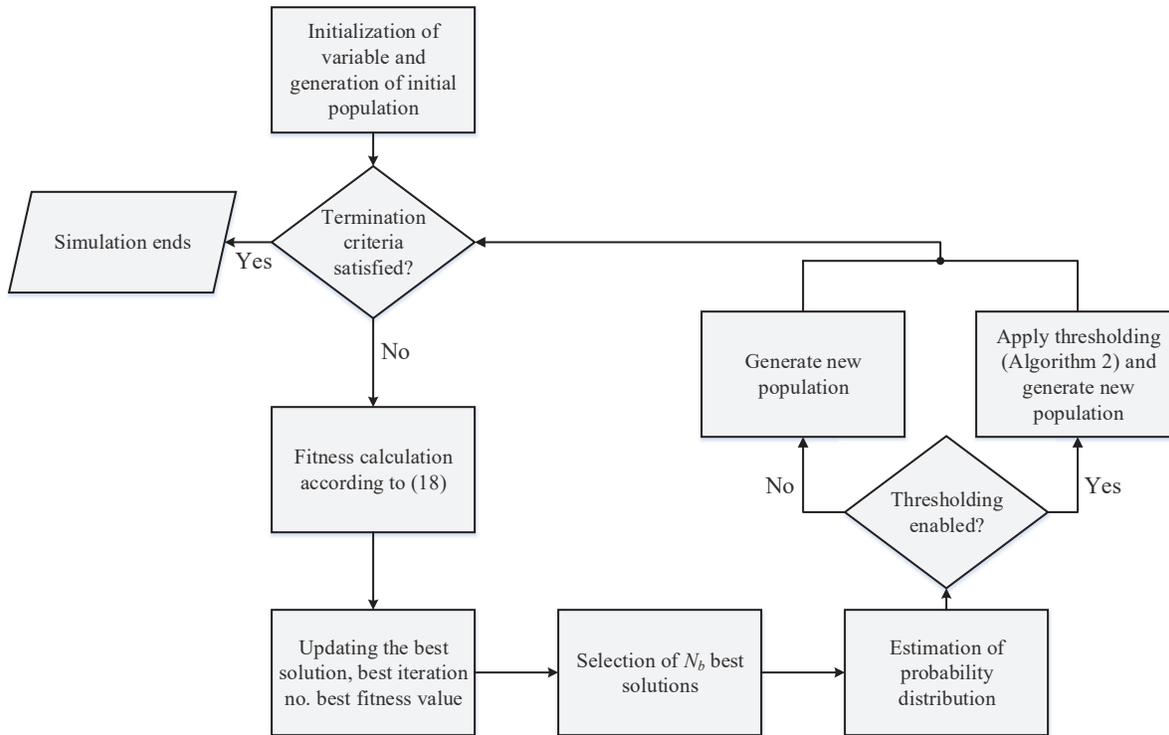


Fig. 3. Overview of Estimation of Distribution Algorithm.

function \mathcal{F} . In our MLOD problem, the fitness function is given in (18). In the next step, the candidates of current population are sorted in terms of their fitness values.

3. Selection of healthy candidates. At this step, the best $N_{bp} = \rho_s \times P_s$ individuals are selected from the sorted current population, where ρ_s is the probability of selection given as an input to the algorithm.

4. Probability model estimation for next generation. Estimate the probability distribution $P(\theta_1, \theta_2, \dots, \theta_L)$ on the basis of N_{bp} best candidate solutions. The probability distribution denotes the probability of ‘1’ in each column. We denote this estimation by:

$$Pr_1 = P(\theta_1, \theta_2, \dots, \theta_L | N_{bp}). \quad (20)$$

5. Generation of new population. New $(P_s - N_{bp})$ individuals are generated using the probability distribution Pr_1 . These individuals are combined with the previous best individuals to generate a new population for next generation (iteration).

B. Estimation of Detection (EOD) Algorithm

This section discusses in detail, the proposed solution to MLOD problem of Section II. The proposed EOD algorithm is derived from the classical EDA discussed above. The computational steps of proposed EOD solution are shown in Algorithm 1 and its important variables and functions are listed in Table IV. A graphical overview of EOD (Algorithm 1) is shown in Fig. 3.

The algorithm takes as input, the binary vector $b = \{b_1, \dots, b_L\}$, which represents a possible line outage combination and noise corrupted observation vector y . A zero at

bit position b_i of b indicates a normal line whereas, a one indicates an outaged line. The optimization goal is to minimize the fitness function (18). During the initialization phase, the main parameters of the algorithm are initialized which include number of lines L in the system, population size P_s , number of algorithm generations G and the best selection probability ρ_s etc. The random binary population matrix Pop of size $P_s \times L$ is generated similar to the one represented in (19). The population is generated by sampling the uniform distribution i.e. the probability $Pr(b_{ij} = 0) = Pr(b_{ij} = 1) = 0.5$, where b_{ij} represents a bit at row i and column j of Pop .

The second phase is the execution phase which runs for G iterations. During each iteration, all individuals of Pop are evaluated by applying the objective function \mathcal{F} of (18) and corresponding fitness values are saved in vector Λ_g . In the next step, the function \mathcal{S} sorts the fitness values of Λ_g in ascending order and the resulting sorted values are stored in vector Γ_f along with their indexes in vector Γ_g^\uparrow . Since, the optimization goal is minimization of objective function \mathcal{F} therefore, the first value of Γ_f is the minimum fitness value of current population. This value is named as the population best and denoted by Ω_p . During each iteration, the population best fitness Ω_p is compared with the global best fitness Ω_g , if Ω_g is greater than Ω_p , it is replaced by Ω_p . Next step is the selection of $N_b = \lceil \rho_s \times P_s \rceil$ best individuals of Pop . To do so, the individuals of Pop whose indexes are stored in the first N_b entries of Γ_g^\uparrow are selected and stored back at first N_b rows of Pop whereas, the remaining individuals of Pop are discarded (i.e. set to zero). In the next step, the joint probability of best selected candidates is obtained. For this purpose, the vector Pr_{bs} of length L is computed such that the value at index $l = \{1, \dots, L\}$ of

Pr_{bs} is obtained by taking the sum of binary values in the l -th column of Pop and dividing the sum by the number N_b .

In order to avoid the premature convergence of proposed solution, an efficient thresholding is applied to Pr_{bs} vector and the EOD algorithm is thus named as EOD-Threshold. The threshold function routine is given in Algorithm 2 and is enabled by the f_{th} flag. The global best fitness value Ω_g is monitored and if it remains same for a pre-defined number of iterations N_{th} , then the probability values of Pr_{bs} are investigated. If the probabilities exceed P_{ht} , they are clipped to P_h value. Similarly, if they go below P_{lt} , they are clipped to P_l . The thresholds values are obtained through extensive simulations and are presented in the initialization phase of Algorithm 1. After applying the threshold, the probability matrix for the whole population is generated by applying the replication operator Rep which generates P_s copies of Pr_{bs} vector and cascades them in vertical order. Finally, the next population Pop is obtained for the next generation. The output of the algorithm are the global best individual \hat{b} and its fitness value Ω_g .

IV. PERFORMANCE EVALUATION

The proposed solution is implemented in Matlab. The simulations are performed for IEEE-14 and 57 bus systems. The MatPower [34] simulation tool is used to extract the topology information and other necessary parameters for these systems. To compare with the proposed EOD solution, a number of other meta-heuristics are also implemented namely; the genetic algorithm (GA), binary particle swarm optimization (BPSO) and cross entropy optimization (CEO). Simulation results are shown for up to five outages; however, the EOD solution is capable to blindly identify any combination and number of line outages. The simulation results are the average of a number of Monte Carlo runs of the above mentioned algorithms. In each Monte Carlo iteration, the algorithm executes its G generations and outputs its global best solution \hat{b} , an estimated line outage vector.

In order to quantify the results and comparison, two performance metrics have been adopted in this work. The first one is named as percentage improvement (PI). The PI of proposed algorithm over another algorithm X is defined as

$$PI = \frac{S_{EOD} - S_X}{S_X} \times 100, \quad (21)$$

where S_{EOD} and S_X denote respectively the number of successes of proposed EOD solution and another solution X . An algorithm success occurs when the input line outage vector b is exactly equal to the estimated line outage vector B_G . A failure occurs when these vectors differ by at least one bit. The positive value of PI indicates that the proposed solution achieves a greater number of successes as compared to the other solution. PI provides a good picture of algorithm's overall performance. However, it may be possible that an algorithm achieves a high success rate but at the cost of large number of generations. To account for this factor, another metric is adopted and named as success generation ratio (SGR). The SGR of a solution X is defined as:

Algorithm 1 : Estimation of Detection (EOD) Algorithm for MLOD Problem.

1: **Initialize Parameters:**
2: $L \leftarrow$ Number of Lines in Power System,
3: $\rho_s \leftarrow 0.4$, $P_{ht} \leftarrow 0.85$, $P_{lt} \leftarrow 0.15$,
4: $P_h \leftarrow 0.75$, $P_l \leftarrow 0.25$, $N_{th} \leftarrow 5$,
5: $P_{val} \leftarrow 0$, $\Omega_g \leftarrow \infty$, $\Omega_p \leftarrow \infty$, $B_p(1, 1 : L) \leftarrow 0$,
6: $B_G(1, 1 : L) \leftarrow 0$, $N_c \leftarrow 0$
7: $N_b \leftarrow \lceil \rho_s \cdot P_s \rceil$
8: $Pr_G(1 : G, 1 : L) \leftarrow 0.5$
9: $Pr_{bs}(1, 1 : L) \leftarrow 0$
10: $Pop \leftarrow \text{random}(1 : P_s, 1 : L) < 0.5$
11: $g \leftarrow 1$
12: $f_{th} \in \{0, 1\}$
13: **Execution:**
14: **while** $g \leq G$ **do**
15: $\Lambda_g \leftarrow \mathcal{F}(Pop)$
16: $[\Gamma_f \Gamma_g^*] \leftarrow \mathcal{S}(\Lambda_g)$
17: $\Omega_p \leftarrow \Gamma_f(1)$
18: $B_p \leftarrow Pop(\Gamma_g^*(1), 1 : L)$
19: **if** $\Omega_p < \Omega_g$ **then**
20: $\Omega_g \leftarrow \Omega_p$
21: $B_G \leftarrow B_p$
22: **end if**
23: $\Gamma_g^s \leftarrow \Gamma_g^*(1 : N_b)$
24: $Pop(1 : N_b, 1 : L) \leftarrow Pop(\Gamma_g^s, 1 : L)$
25: $Pop(N_b + 1 : P_s, 1 : L) \leftarrow 0$
26: $Pr_{bs}(1, l) \leftarrow \frac{\sum Pop(1:N_b, l)}{N_b} \quad \forall l = 1 : L$
27: **if flag then**
28: $Pr_{bs} \leftarrow \text{Threshold}(P_{val}, \Omega_g, N_c, N_{th}, Pr_{bs}, P_{ht}, P_{lt})$
29: **end if**
30: $Pr_{pop} \leftarrow Rep(P_s, Pr_{bs}, 1)$
31: $P_{temp} \leftarrow \text{random}(1 : P_s, 1 : L) \in [0, 1]$
32: $Pop(1 : P_s, 1 : L) \leftarrow (P_{temp} < Pr_{pop}) \in [0, 1]$
33: $Pr_G(g, :) \leftarrow Pr_{bs}$
34: $g \leftarrow g + 1$
35: **end while**
36: $\hat{b} \leftarrow B_G$
37: **Outputs:** \hat{b} , Ω_g

$$SGR_X = \frac{S_X/I_X}{\bar{G}_X/G}, \quad (22)$$

where I_X denotes the total number of Monte Carlo runs of algorithm X , G denotes the generations per Monte Carlo iteration and \bar{G}_X denotes the mean of best generations. \bar{G}_X is defined as

$$\bar{G}_X = \frac{\sum_{i \in \{1, \dots, I_X\}} g_b(i)}{I_X}. \quad (23)$$

For each Monte Carlo iteration $i \in \{1, \dots, I_X\}$, the best generation value $g_b \in \{1, \dots, G\}$ is noted which corresponds to the generation number beyond which the algorithm no more converges. The mean best generation \bar{G}_X is obtained by adding the g_b values for all Monte Carlo iterations dividing the sum by I_X . The numerator of (22) corresponds to the success rate of algorithm whereas, the denominator corresponds to the

Algorithm 2 : Threshold Function Routine

```

1: Inputs :  $P_{val}, \Omega_g, N_c, N_{th}, Pr_{bs}, P_{ht}, P_{lt}$ 
2: if  $P_{val} \neq \Omega_g$  then
3:    $P_{val} \leftarrow \Omega_g$ 
4:    $N_c \leftarrow 0$ 
5: else
6:    $N_c \leftarrow N_c + 1$ 
7: end if
8: if  $N_c > N_{th}$  then
9:    $N_c \leftarrow 0$ 
10: for  $l = 1 : L$  do
11:   if  $Pr_{bs}(l) \geq P_{HT}$  then
12:      $Pr_{bs}(l) \leftarrow P_H$ 
13:   else if  $Pr_{bs}(l) \leq P_{LT}$  then
14:      $Pr_{bs}(l) \leftarrow P_L$ 
15:   end if
16: end for
17: end if
18: Outputs:  $Pr_{bs}$ 

```

convergence rate of algorithm. Smaller value of denominator means a fast convergence rate. A high value of SGR means a high success rate with fast convergence rate.

In Figs. 4 and 5, the results are plotted for IEEE-14 bus system for P_s ranging from 20 to 100 and G ranging from 50 to 200. First, Fig. 4 plots the PI of EOD-Threshold over other algorithms including the EOD without thresholding. For a scenario consisting of small values of $P_s = 20$ and $G = 50$, Fig. 4 (a) demonstrates that the EOD-Threshold achieves a very large PI over other algorithms. This is due to the fact that EOD-Threshold achieves a very high number of successes thanks to the efficient thresholding technique, while other algorithms get stuck in local optima. Furthermore, the plots of Figs. 4 (b) and (c) reveal that although the PI of EOD-Threshold decreases with increase in P_s and G values, still the PI is fairly large and positive. Moreover, the EOD-Threshold achieves a maximum PI over CEO. Overall, the results of Fig. 4 demonstrate the advantage in terms of accuracy of the proposed EOD-Threshold solution. The SGR results for same scenarios are reported in Fig. 5. For $P_s = 20$ and $G = 50$, Fig. 5 (a) reveals that maximum SGR is obtained by EOD-Threshold. This trend changes in Figs. 5 (b) and (c) where the SGR performance is dominated by simple EOD algorithm, as the aim of thresholding is to increase the accuracy (PI) of solution at the cost of a greater number of generations. It is worth to mention here that, by adjusting the values of thresholds P_{HT} and P_{LT} , one can trade between accuracy and convergence rate of solution. If accuracy (PI) is the prime objective, one can apply the EOD with thresholding. Otherwise, simple EOD algorithm can be applied to achieve a high SGR. As a matter of fact, setting the values of $P_{HT} = 1$ and $P_{LT} = 0$, the EOD-Threshold algorithm becomes the simple EOD algorithm.

For a considerably larger network such as IEEE-57 bus system, the simulation results are demonstrated in Figs. 6 and 7. Here, considerably large P_s and G values have been selected and PI plot of Fig. 6 demonstrates the accuracy of

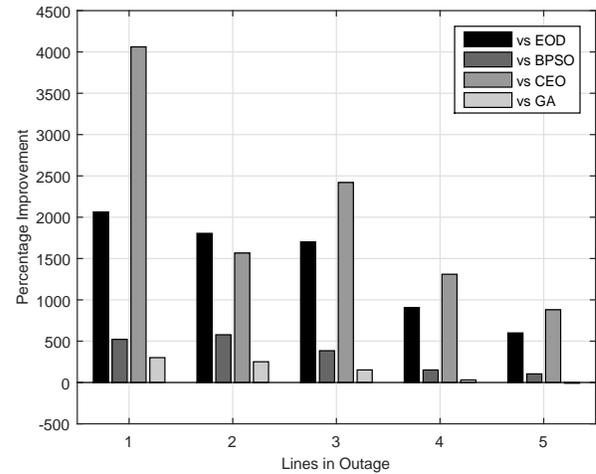
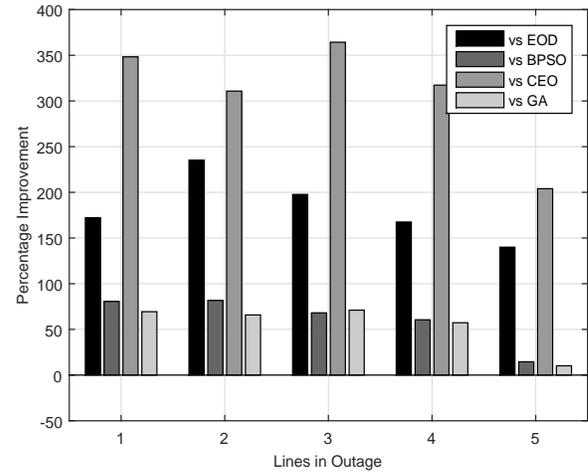
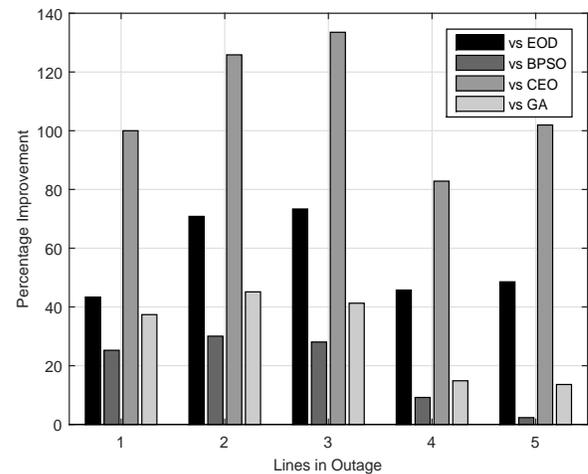
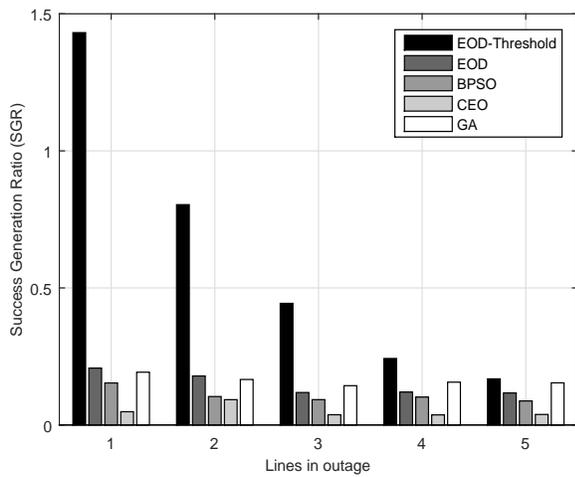
(a) $P_s = 20, G = 50$ (b) $P_s = 50, G = 100$ (c) $P_s = 100, G = 200$

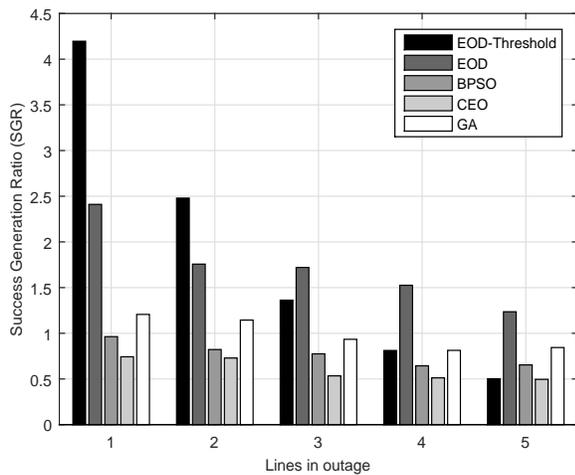
Fig. 4. PI of the proposed EOD-Threshold v.s. other algorithms for IEEE-14 bus system

TABLE IV
VARIABLES DESCRIPTION OF ALGORITHMS 1 AND 2.

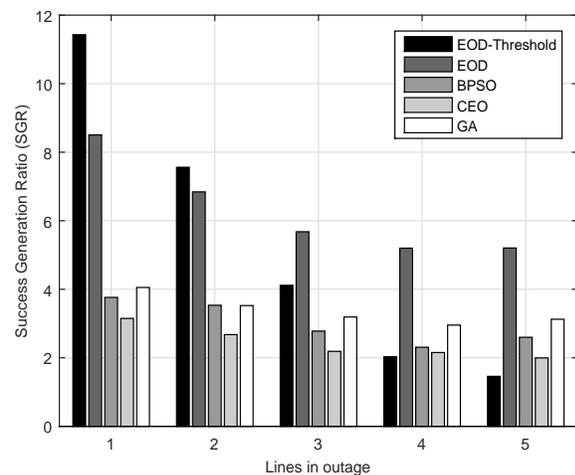
Variable	Detail
G	Number of generations of algorithm
L	Number of transmission lines in the power grid
P_s	Size of population
ρ_s	Probability of selection of best candidates
P_{ht}, P_{lt}	Upper and lower threshold limits respectively of probability
P_h	Upper probability value
P_l	Lower probability value
N_{th}	No. of generations after which thresholding is applied
P_{val}	temporary variable for holding population best probability
Ω_g	Global best fitness value
Ω_p	Local (population) best fitness value
B_p	Population best individual
B_G	Global best individual
\hat{b}	Output of algorithm, the estimated line outage vector
N_b	Number of best selected candidates
Pr_{bs}	The vector of size $P_s \times L$ containing joint probability distribution for best selected candidates
Pr_G	The matrix containing generation wise Pr_{bs} vectors
Pop	The population matrix
f_{th}	binary flag to control the threshold function on Pr_{bs}
$S(x)$	Sorting function: sorts the vector x in ascending order
$Rep(a, b, 1)$	Matrix replication function. Produces a matrix which consists of a copies of vector b concatenated in vertical order
$\lceil \cdot \rceil$	Ceil function
Λ_g	Column vector of size P_s stores the fitness values of all individuals of Pop
Γ_g^\uparrow	Indexes of sorted values of Λ_g
Γ_f	Actual values of Λ_g after sorting
Γ_g^s	Indexes of best selected solutions
\mathcal{F}	objective function (18)
$X(1:r, 1:c)$	rows from 1 to r and columns from 1 to c of matrix X



(a) $P_s = 20, G = 50$



(b) $P_s = 50, G = 100$



(c) $P_s = 100, G = 200$

Fig. 5. Success Generation Ratio (SGR) results for IEEE-14 bus system

EOD-Threshold with respect to other algorithm. For all values of lines in outage, the EOD-Threshold achieves a positive, fairly large PI value and hence shows a better success rate as compared to other solutions. The plot of Fig. 7 demonstrates that the SGR performance is dominated by the simple EOD solution. However, the EOD-Threshold also achieves an SGR fairly larger than other solutions and comparable to simple EOD algorithm. Therefore, for such case, the EOD-Threshold solution is a good option in terms of accuracy and convergence

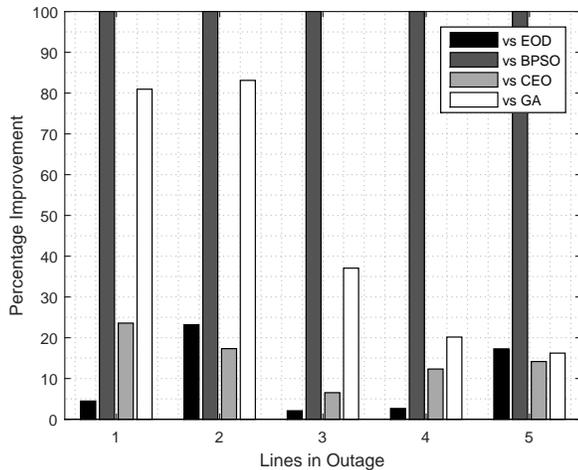


Fig. 6. Percentage improvement of the proposed EOD-Threshold for IEEE-57 with $P_s = 500$, $G = 200$.

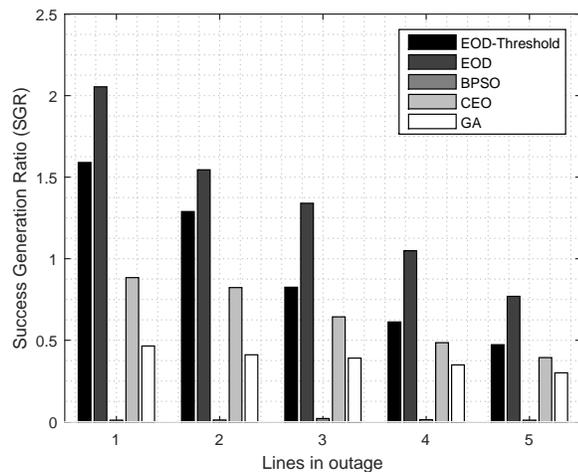


Fig. 7. Success Generation IEEE-57 with $P_s = 500$, $G = 200$

rate.

V. CONCLUSIONS

To avoid the grid-wide blackouts, an accurate and prompt line outage identification method is required. For a system consisting of L lines, the complexity of exhaustive search is 2^L , which is clearly infeasible for practical sized smart grids. The published works on line outage identification are limited to identify only a fixed (usually small) number of lines in outage. As a key contribution, this work applies the Bayesian probability theory to solve the MLOD problem in linear time. The proposed solution named as estimation of outage detection (EOD) is accurate to identify any arbitrary number and combination of line outages. Moreover, the solution avoids the local optima, thanks to an efficient thresholding routine applied to update the probabilities of candidate solutions. The simulations of proposed solution are carried out for IEEE-14 and -57 bus systems. The results are compared with a number of meta-heuristics. To have an unbiased comparison,

two performance metrics have been introduced in the paper namely the success generation ratio (SGR) and performance improvement (PI). These metrics quantify respectively the convergence rate and accuracy of solution. For a number of random line outage scenarios, the simulation results demonstrate that the proposed solution achieves a better PI and SGR as compared to other meta-heuristics. This confirms the validity of proposed approach as well as its applicability to practical smart grids.

VI. ACKNOWLEDGEMENT

This research was supported in part by the Korea Electric Power Corporation (KEPCO) (*CX72166553 – R16DA17*) of the Republic of Korea.

REFERENCES

- [1] A. Phadke, "Synchronized phasor measurements in power systems," *IEEE Computer Applications in Power*, vol. 6, no. 2, pp. 10–15, Apr. 1993.
- [2] D. Ghosh, T. Ghose, and D. Mohanta, "Communication feasibility analysis for smart grid with phasor measurement units," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1486–1496, Aug. 2013.
- [3] A. Moreno-Munoz, V. Pallares-Lopez, J. J. G. de la Rosa, R. Real-Calvo, M. Gonzalez-Redondo, and I. M. Moreno-García, "Embedding synchronized measurement technology for smart grid development," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 52–61, Feb. 2013.
- [4] H. Morais, P. Vancraeyveld, A. H. B. Pedersen, M. Lind, H. Jóhannsson, and J. Østergaard, "SOSPO-SP: Secure operation of sustainable power systems simulation platform for real-time system state evaluation and control," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2318–2329, Nov. 2014.
- [5] K. G. Boroojeni, M. H. Amini, and S. S. Iyengar, *Overview of the Security and Privacy Issues in Smart Grids*. Cham: Springer International Publishing, 2017, pp. 1–16.
- [6] K. Yamashita, S.-K. Joo, J. Li, P. Zhang, and C.-C. Liu, "Analysis, control, and economic impact assessment of major blackout events," *European Transactions on Electrical Power*, vol. 18, pp. 854–871, 2008.
- [7] D. J. Gausshell and H. T. Darlington, "Supervisory control and data acquisition," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1645–1658, Dec. 1987.
- [8] B. Singh, N. Sharma, A. Tiwari, K. Verma, and S. Singh, "Applications of phasor measurement units (PMUs) in electric power system networks incorporated with FACTS controllers," *International Journal of Engineering, Science and Technology*, vol. 3, no. 3, pp. 64–82, 2011.
- [9] J. Tang, J. Liu, F. Ponci, and A. Monti, "Adaptive load shedding based on combined frequency and voltage stability assessment using synchrophasor measurements," *IEEE Transactions on power systems*, vol. 28, no. 2, pp. 2035–2047, May 2013.
- [10] R. O. Burnett, M. M. Butts, and P. S. Sterlina, "Power system applications for phasor measurement units," *IEEE Computer Applications in Power*, vol. 7, no. 1, pp. 8–13, Jan. 1994.
- [11] H. Zhu and Y. Shi, "Phasor measurement unit placement for identifying power line outages in wide-area transmission system monitoring," in *HICSS*, Jan 2014, pp. 2483–2492.
- [12] A. Enshae, R. A. Hooshmand, and F. H. Fesharaki, "A new method for optimal placement of phasor measurement units to maintain full network observability under various contingencies," *Electric Power Systems Research*, vol. 89, no. 0, pp. 1 – 10, 2012.
- [13] J. Aghaei, A. Baharvandi, A. Rabiee, and M.-A. Akbari, "Probabilistic PMU placement in electric power networks: an MILP-based multiobjective model," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 332–341, Apr. 2015.
- [14] A. Abdelaziz, S. Mekhamer, M. Ezzat, and E. El-Saadany, "Line outage detection using support vector machine (SVM) based on the phasor measurement units (PMUs) technology," in *IEEE Power and Energy Society General Meeting*. San Diego, CA: IEEE, Jul. 2012, pp. 1–8.
- [15] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, Nov. 2008.

- [16] —, “Double line outage detection using phasor angle measurements,” in *IEEE Power & Energy Society General Meeting*. Calgary, AB: IEEE, Jul. 2009, pp. 1–5.
- [17] H. Zhu and G. B. Giannakis, “Sparse overcomplete representations for efficient identification of power line outages,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.
- [18] R. Y. Rubinstein and D. P. Kroese, *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media, 2013.
- [19] T. Banerjee, Y. C. Chen, A. D. Dominguez-Garcia, and V. V. Veeravalli, “Power system line outage detection and identification-A quickest change detection approach,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 3450–3454.
- [20] R. Emami and A. Abur, “Tracking changes in the external network model,” in *North American Power Symposium (NAPS)*. Arlington, TX: IEEE, Sep. 2010, pp. 1–6.
- [21] P. Bijwe, J. Nanda, and K. Puttabuddhi, “Ranking of line outages in an AC-DC system causing overload and voltage problems,” in *IEE Proceedings C-Generation, Transmission and Distribution*, vol. 138, no. 3. IET, May 1991, pp. 207–211.
- [22] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Ching, and Z. Fan, “Smart grid health monitoring via dynamic compressive sensing,” in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*. Lyngby, Denmark: IEEE, Oct. 2013, pp. 1–5.
- [23] M. M. El-Arini, S. Soliman, and A. Al-Kandari, “A fast technique for outage studies in power system planning and operation,” *Electric power systems research*, vol. 32, no. 3, pp. 203–211, Mar. 1995.
- [24] H. Ren, I. Dobson, and B. A. Carreras, “Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid,” *IEEE Transactions on power systems*, vol. 23, no. 3, pp. 1217–1225, Aug. 2008.
- [25] M. H. Amini, M. Rahmani, K. G. Boroojeni, G. Atia, S. S. Iyengar, and O. Karabasoglu, “Sparsity-based error detection in dc power flow state estimation,” in *2016 IEEE International Conference on Electro Information Technology (EIT)*, May 2016, pp. 0263–0268.
- [26] A. I. SARWAT, M. AMINI, A. DOMIJAN, A. DAMNJANOVIC, and F. KALEEM, “Weather-based interruption prediction in the smart grid utilizing chronological data,” *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 2, pp. 308–315, 2016.
- [27] M. Amini, A. I. Sarwat, S. S. Iyengar, and I. Guvenc, “Determination of the minimum-variance unbiased estimator for dc power-flow estimation,” in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, Oct 2014, pp. 114–118.
- [28] K. G. Boroojeni, M. H. Amini, and S. S. Iyengar, “An oblivious routing-based power flow calculation method for loss minimization of smart power networks: A theoretical perspective,” in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec 2016, pp. 641–645.
- [29] K. G. Boroojeni, M. H. Amini, S. S. Iyengar, M. Rahmani, and P. M. Pardalos, “An economic dispatch algorithm for congestion management of smart power networks,” *Energy Systems*, pp. 1–25, 2016.
- [30] W. Li, J. Zhou, and X. Xiong, “Fuzzy models of overhead power line weather-related outages,” *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1529–1531, Aug. 2008.
- [31] J. Kim and I. Dobson, “Propagation of load shed in cascading line outages simulated by OPA,” in *Complexity in Engineering (COMPENG)*. Rome, Italy: IEEE, Feb. 2010, pp. 1–6.
- [32] H. Schwail and I. Dobson, “Locating line outages in a specific area of a power system with synchrophasors,” in *North American Power Symposium (NAPS)*. Champaign, IL: IEEE, Sep. 2012, pp. 1–6.
- [33] A. Schellenberg, W. Rosehart, and J. Aguado, “Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and gamma distributions,” *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 773–781, May 2005.
- [34] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.