A Genetic Algorithm for Planning WAMS with a Heterogeneous Communication Network

Halil Alper Tokel, Gholamreza Alirezaei, and Rudolf Mathar
Institute for Theoretical Information Technology
RWTH Aachen University, D-52056, Aachen, Germany
{tokel,alirezaei,mathar}@ti.rwth-aachen.de

***Abstract*—The optimal deployment of phasor measurement units (PMUs) and the required communication network topology remains still an open problem. The experiences in PMU deployments so far show that not only the hardware costs but also the cost of the communication network deployment and commissioning is a major driving cost factor. In this work, we, first, present an optimization model which enables the comprehensive integrated planning of a wide-area measurement system, including the optimal locations of measurement units, the optimal locations of phasor data concentrators, and the required optimal communication network topology with possibly multiple communication technologies in a joint optimization framework. Furthermore, a genetic algorithm is presented, which utilizes problem-specific genetic operators and has optimally solved problems with up to 5,000 binary optimization variables and 85,000 constraints. The proposed algorithm provides a basis for further development for finding near-optimal or optimal solutions for very large problem instances. The contribution in this work enables operators of distribution and transmission systems to analyze offers from several telecommunication providers for having a better understanding of possible deployment strategies.**

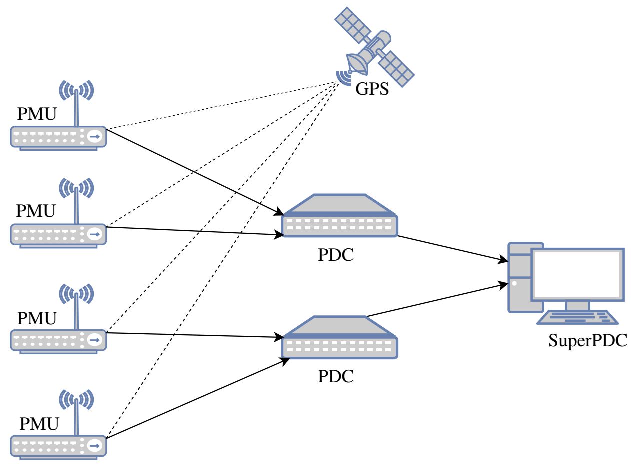

Fig. 1. Hierarchical network architecture of a WAMS. The PMUs send the phasor measurements, that are time-stamped by the GPS signal, to a SPDC over intermediate PDCs [1].

I. Introduction

The placement of measurement units in power systems has traditionally a strategical importance not only due to the possible cost-savings thanks to an optimal design but also because of the stringent requirements for monitoring a power system. It is desirable to acquire as much information about the system as possible under given cost and geographical constraints. The recent trend of decentralization and the increasing availability of various sensing units at ever lower costs are two main driving factors which have accelerated the plans for installation of more and more sensors in power grids to realize the goal of a smart grid for a better controlled and efficient operation.

In this regard, optimal placement of phasor measurement units (PMUs) and the design of the so-called wide area measurement system (WAMS) has become even a more important practical problem. A WAMS consists of *i)* PMUs, which measure the voltage and current phasor values available at the system nodes where they are installed, *ii)* several phasor data concentrator units, called phasor data concentrators (PDCs), and *iii)* a data processing center, called SuperPDC (SPDC). *IEEE Standard for Synchrophasor Data Transfer for Power Systems* [1] lays down the architecture for the communication network in a WAMS as shown in Fig. 1. This architecture postulates a hierarchical transmission of sensor data from PMUs to PDCs, where a preprocessing of the data takes place such as time alignment and consistency check [2]. PDCs send the data to a central unit SPDC, where the measurement data from a larger part of the network are aggregated to execute energy management functions such as state estimation, cf. [3].

Historically, the main concern of the measurement installations had been the minimization of the number of required measurement units in the beginnings due to the limited number of sensors and the low penetration of measurement devices in the lower voltage levels. Today, however, there is an increasing interest in academia and industry to lower the cost of the measurement units and make so-called micro-PMUs a feasible investment in distribution grids, see for example [4], [5]. Therefore, the increasing number of the sensors to be deployed makes the costs for the communication links one of the deciding factors in the planning phase. Recent experiences, for example in the deployment of PMUs in North America, show also that communication network design, commissioning, and deployment is one of the major cost factors [6]. According to [6], "*One utility reported that, absent adequate existing communications, upgrades to communications infrastructure increased the cost of installing PMUs by a factor of seven.*".

The authors gratefully acknowledge the financial support by BMBF (German Federal Ministry of Education and Research) under Promotional Reference 03EK3567B.

Therefore, there has been a growing interest in the joint optimization of the measurement unit locations with the required communication network [2], [7]–[10].

One aspect, which has not closely been considered in most of the planning approaches so far is the optimal use of available multiple communication technologies in the communication network design. In this sense, for example, [11] shows that the use of available power line communication (PLC) links can reduce the investment costs considerably.

The planning approach, which we have presented in [2], is the only work, to the best of our knowledge, which optimizes the number and location of PMU and PDCs and a heterogeneous communication network jointly in the same framework. This approach provides an optimal strategy to extend the measurement network under consideration of various offers from telecommunication providers or available power lines which can partly or completely owned by the power system operators.

From a modeling point of view, the optimization model used in [2] is based on the concept of multi-commodity flows. Although this approach provides a guarantee for global optimality as an integer linear program, the number of optimization variables escalates due to flow-based formulation. Our contribution in [10] adopts a two-level assignment approach with the restriction that the candidate communication paths can consist of links of only one communication technology.

One major contribution of the present work is an alternative and more efficient formulation, where the possible pre-selected communication paths, which might still consist of multiple communication links of different technologies, are used as main decision variables for the communication network design. The formulation can still be transformed into an integer linear program, which guarantees global optimality when solved by the branch-and-bound algorithms. Furthermore, the path formulation not only leads to a significant reduction in the number of optimization variables and constraints compared to the flow-based formulation, but it also allows a simpler representation of the individual solutions and evolutionary operators in our genetic algorithm approach, which is the major contribution of the current work.

In the following, we first start with the elaboration of our system and optimization model in Section II. In Section III, we discuss the application of genetic algorithms to constrained optimization problems and describe the proposed genetic algorithm for the particular design problem presented in Section II. In Section IV, the performance of the proposed algorithm is evaluated in terms of optimality along with a note on run-time and scalability. Finally, we conclude the paper in Section V with a summary and an outlook.

II. System and Optimization Model

In the following we use the following mathematical notations. Vectors are defined as column vectors and denoted by boldfaced lowercase letters, whereas matrices are denoted by boldfaced uppercase letters. An element of a vector or a matrix is denoted by the same letter but not in boldface and with the relevant index as a subscript. Sets are denoted by calligraphic uppercase letters. We denote an all-one vector of proper length by $\mathbf{1}$, and $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ represents the inner product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ of the same size s, i.e., $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{i=1}^{s} a_i b_i$.

TABLE I
SYMBOL NOTATION, SET DEFINITIONS

Symbol, Domain	Description
$\mathcal{B} = \{0, 1\}$	the binary set
$\mathcal{Z}_+$, $\mathcal{R}_+$	sets of non-negative integer and real numbers
$\mathcal{V}_{\text{pow}},\ n_{\text{bus}} \in \mathcal{Z}_+$	set of power system nodes and its cardinality
$\mathcal{E}_{\text{pow}},\ n_{\text{branch}} \in \mathcal{Z}_+$	set of power system branches and its cardinality
$G_{\text{pow}}(\mathcal{V}_{\text{pow}}, \mathcal{E}_{\text{pow}})$	graph of the power system
$\mathcal{V}_{\text{PDC}},\ n_{\text{pow}} \in \mathcal{Z}_+$	set of power system nodes that can be used as a PDC location and its cardinality
$\mathcal{V}_{\text{com}},\ n_{\text{com}} \in \mathcal{Z}_+$	set of additional communication network nodes and its cardinality
$\mathcal{E}_{\text{com}},\ n_{\text{com}} \in \mathcal{Z}_+$	set of additional communication network links with at least one end in $\mathcal{V}_{\text{com}}$ and its cardinality
$\mathcal{V}_{\text{ext}},\ n_{\text{ext}} \in \mathcal{Z}_+$	set of all communication nodes and its cardinality
$\mathcal{E}_{\text{ext}},\ n_{\text{edge}} \in \mathcal{Z}_+$	directed set of all communication links and its cardinality
$G_{\text{ext}}(\mathcal{V}_{\text{ext}}, \mathcal{E}_{\text{ext}})$	extended directed graph of the communication network
$\hat{\mathcal{E}}_{\text{ext}},\ n_{\text{link}} \in \mathcal{Z}_+$	undirected set of communication links and its cardinality
$\hat{G}_{\text{ext}}(\mathcal{V}_{\text{ext}}, \hat{\mathcal{E}}_{\text{ext}})$	extended undirected graph of the communication network
$\mathcal{P}_{\text{PMU}},\ n_{\text{bus}} \in \mathcal{Z}_+$	set of possible PMU locations and its cardinality
$\mathcal{P}_{\text{PDC}},\ n_{\text{p}} \in \mathcal{Z}_+$	set of all possible PDC locations and its cardinality
$\mathcal{T},\ n_{\text{t}} \in \mathcal{Z}_+$	set of available technologies and its cardinality
$\mathcal{R}_1^{rk},\ p_1^{rk} \in \mathcal{Z}_+$	set of possible paths between PMU-PDC pair (p_r, q_k) and its cardinality
$\mathcal{R}_2^{k},\ p_2^{k} \in \mathcal{Z}_+$	set of possible paths between PDC q_k and SPDC and its cardinality

TABLE II
SYMBOL NOTATION, INPUT PARAMETERS

Symbol, Domain	Description
$\boldsymbol{A} \in \mathcal{B}^{n_{\text{bus}} \times n_{\text{bus}}}$	connectivity matrix of the power system
$\boldsymbol{L}_{\text{avail}} \in \mathcal{B}^{n_{\text{link}} \times n_{\text{t}}}$	link-technology availability matrix
$\boldsymbol{M}, \boldsymbol{D} \in \mathcal{R}_+^{n_{\text{edge}} \times n_{\text{t}}}$	capacity and delay matrices for technologies
$\boldsymbol{C} \in \mathcal{R}_+^{n_{\text{link}} \times n_{\text{t}}}$	link-technology cost matrix
$\boldsymbol{g} \in \mathcal{R}_+^{n_{\text{bus}} \times 1}$	bandwidth requirements at possible PMU nodes
$c_{\text{WiMAX}} \in \mathcal{R}_+$	fixed cost for WiMaX
$\boldsymbol{E}_1^{rk} \in \mathcal{B}^{n_{\text{edge}} \times p_1^{rk}}$	path-edge matrix for PMU-PDC pair (p_r, q_k)
$\boldsymbol{E}_2^{k} \in \mathcal{B}^{n_{\text{edge}} \times p_2^{k}}$	path-edge matrix for PDC q_k and SPDC

We consider a power system modeled by an undirected graph $G_{\text{pow}}(\mathcal{V}_{\text{pow}}, \hat{\mathcal{E}}_{\text{pow}})$, where $\mathcal{V}_{\text{pow}}$ is the set of power system nodes with $|\mathcal{V}_{\text{pow}}| = n_{\text{bus}}$, and $\hat{\mathcal{E}}_{\text{pow}}$ is the set of power system branches with $|\hat{\mathcal{E}}_{\text{pow}}| = n_{\text{branch}}$. Note that due to the wired communication technologies such as PLC and optical communication, whose topology is assumed to follow the power system topology, each power system branch can be modeled as a candidate link in the communication network design. In addition to the power system nodes, we assume that there are n_{com} communication network nodes modeled by the set $\mathcal{V}_{\text{com}}$, which can be used for the transmission of measurement data, for example as a wireless relay or a data concentrator. Furthermore, $\mathcal{E}_{\text{com}}$, is a symmetric set of ordered pairs of nodes, that denotes possible directed communication links between the nodes in $\mathcal{V}_{\text{pow}}$ and the nodes in $\mathcal{V}_{\text{com}}$, as well as possible links between the nodes in $\mathcal{V}_{\text{com}}$. We define the extended directed communication graph $G_{\text{ext}}(\mathcal{V}_{\text{ext}}, \mathcal{E}_{\text{ext}})$, where $\mathcal{V}_{\text{ext}} = \mathcal{V}_{\text{pow}} \cup \mathcal{V}_{\text{com}}$ with $|\mathcal{V}_{\text{ext}}| = n_{\text{ext}} = n_{\text{bus}} + n_{\text{com}}$, and $\mathcal{E}_{\text{ext}} = \mathcal{E}_{\text{pow}} \cup \mathcal{E}_{\text{com}}$, where $\mathcal{E}_{\text{pow}}$, extension of $\hat{\mathcal{E}}_{\text{pow}}$, is the directed set of power system branches. We denote the set of all possible undirected communication links by $\hat{\mathcal{E}}_{\text{ext}}$, where $|\hat{\mathcal{E}}_{\text{ext}}| = n_{\text{link}} = \frac{n_{\text{edge}}}{2}$. Note that $G_{\text{ext}}(\mathcal{V}_{\text{ext}}, \mathcal{E}_{\text{ext}})$ is the extended directed graph for the communication network,

TABLE III
SYMBOL NOTATION, OPTIMIZATION VARIABLES

Symbol, Domain	Optimization Variable For
$\boldsymbol{x} \in \mathcal{B}^{n_{\text{bus}}}$	PMU locations
$\boldsymbol{y} \in \mathcal{B}^{n_{\text{p}}}$	PDC locations
$\boldsymbol{\rho}_1^{rk} \in \mathcal{B}^{\vert\mathcal{R}_{1,rk}\vert}$	Layer 1 paths from PMU q_r to PDC p_k
$\boldsymbol{\rho}_2^{k} \in \mathcal{B}^{\vert\mathcal{R}_{2,k}\vert}$	Layer 2 paths from PDC p_k to SPDC
$Z_{rk} \in \mathcal{B}$	Assignment of PMU q_r to the PDC p_k
$T_{e\phi} \in \mathcal{B}$	Use of directed link with technology ϕ on $e \in \hat{\mathcal{E}}_{\text{ext}}$
$L_{\hat{e}\phi} \in \mathcal{B}$	Deployment of undirected link with technology ϕ on $\hat{e} \in \hat{\mathcal{E}}_{\text{ext}}$
$t_w \in \mathcal{B}$	Use of WiMaX

whereas $\hat{G}_{\text{ext}}(\mathcal{V}_{\text{ext}}, \hat{\mathcal{E}}_{\text{ext}})$ is the undirected graph for the same network. This differentiation is necessary as the available communication technologies might have different specifications, for example, for uplink and downlink, whereas the selection of a communication technology over a link would contribute to the total costs without consideration of any direction.

Our planning approach is based on the following assumptions: First, we assume that the location v_{SPDC} of SPDC is predetermined and known beforehand. We denote the set of possible PMU locations by $\mathcal{P}_{\text{PMU}} \subseteq \mathcal{V}_{\text{pow}}$. However, without loss of generality, we will assume in this work that a PMU can be installed at any $v_i \in \mathcal{V}_{\text{pow}}$, i.e. $\mathcal{P}_{\text{PMU}} = \mathcal{V}_{\text{pow}}$. Furthermore, a subset of $n_{\text{pow}} \leq n_{\text{bus}}$ power system nodes are available as candidate PDC locations in addition to the n_{com} communication nodes which can be selected as a PDC location. The set of all possible PDC locations are denoted by $\mathcal{P}_{\text{PDC}}$ with $|\mathcal{P}_{\text{PDC}}| = n_{\text{p}} = n_{\text{pow}} + n_{\text{com}}$. Furthermore, we assume that only one communication technology among all possible ones can be deployed between any two communication nodes.

For each possible PMU-PDC pair (p_r, q_k), the set $\mathcal{R}_1^{rk}$ with cardinality $p_1^{rk} \in \mathcal{Z}_+$ denotes the set of possible paths. Similarly, $\mathcal{R}_2^{k}$ with cardinality $p_2^{k} \in \mathcal{Z}_+$ denotes the set of possible paths between PDC q_k and SPDC location v_{SPDC}. Note that a path might consist of multiple directed links. The matrix $\boldsymbol{E}_1^{rk} \in \mathcal{B}^{n_{\text{edge}} \times p_1^{rk}}$ holds in its non-zero entries the indices of the directed links which belong to the paths between PMU-PDC pair (p_r, q_k), whereas $\boldsymbol{E}_2^{k} \in \mathcal{B}^{n_{\text{edge}} \times p_2^{k}}$ holds the indices for the PDC q_k in the same fashion.

The objective of the design problem is to find the minimum cost network infrastructure by determining the optimal numbers and locations of PMU and PDC units, and the optimum heterogeneous communication network topology under the constraints of power system observability and data communication requirements. Note that variations of this base problem with other objectives can similarly be formulated, e.g., with the goal of minimum-delay, multi-objective optimization considering several objectives, or considering given cost constraints.

For the decision variables, first we define the optimization variables for PMU locations and PDC locations as $\boldsymbol{x} \in \mathcal{B}^{n_{\text{bus}}}$ and $\boldsymbol{y} \in \mathcal{B}^{n_{\text{p}}}$, respectively. The entries of vector optimization variables $\boldsymbol{\rho}_1^{rk}$ and $\boldsymbol{\rho}_2^{k}$ represent the selection of the corresponding paths between the PMU-PDC pair (q_r, p_k) and the paths between PDC p_k and the SPDC, respectively. Furthermore, we define $T_{e\phi} \in \mathcal{B}$ for the selection of the technology τ_ϕ over the directed link $e \in \mathcal{E}_{\text{ext}}$. The optimization variable $L_{\hat{e}\phi} \in \mathcal{B}$ is defined for the decision of the technology τ_ϕ over the undirected link $\hat{e} \in \hat{\mathcal{E}}_{\text{ext}}$. Note that the variable $T_{e\phi}$ is for the directed link $(v_i, v_j) \in G_{\text{ext}}$, whereas $L_{\hat{e}\phi}$ is for the undirected links in $\hat{\mathcal{E}}_{\text{ext}}$. Table I, Table II, and Table III show the used notation for the set definitions, input parameters, and the defined optimization variables, respectively.

In the following, we introduce the constraints, and then the objective function of the proposed optimization model.

A power system is observable, if the voltage values of all system nodes can be calculated or accurately estimated by using the available measurement set [12]. In the case where the measurement set consists of PMU measurements only, the vector $\boldsymbol{x}$, whose nonzero entries denote the locations of PMUs, should satisfy

$$\boldsymbol{A}\boldsymbol{x} \succeq \boldsymbol{1}, \tag{1}$$

where $\boldsymbol{A}$ is the adjacency matrix of the power system. Note that the observability constraint in (1) is adopted for sake of simplicity in this work and can be modified trivially, for example, to achieve N-1 redundancy or to exclude the system buses with zero injection. For a more detailed discussion on observability constraints, please refer to [12].

The PMU-PDC assignments and the selection of PDC locations are guaranteed by the constraints

$$\sum_{k=1}^{n_{\text{p}}} Z_{rk} = x_r, \quad \forall q_r \in \mathcal{P}_{\text{PMU}}, \tag{2}$$

$$y_k = \max\{Z_{rk} \mid q_r \in \mathcal{P}_{\text{PMU}}\}, \quad \forall p_k \in \mathcal{P}_{\text{PDC}}, \tag{3}$$

respectively. Furthermore, the communication paths between PMU-PDC pairs and between PDC locations and the SPDC are selected by the constraints

$$\langle \boldsymbol{\rho}_1^{rk}, \boldsymbol{1} \rangle = Z_{rk}, \quad \forall (q_r, p_k) \in \mathcal{P}_{\text{PMU}} \times \mathcal{P}_{\text{PDC}}, \tag{4}$$

$$\langle \boldsymbol{\rho}_2^{k}, \boldsymbol{1} \rangle = y_k, \quad \forall p_k \in \mathcal{P}_{\text{PDC}}, \tag{5}$$

whereas the selection of communication technology on directed links are governed by

$$\boldsymbol{T}\boldsymbol{1} \geq \boldsymbol{E}_1^{rk} \boldsymbol{\rho}_1^{rk}, \quad \forall (q_r, p_k) \in \mathcal{P}_{\text{PMU}} \times \mathcal{P}_{\text{PDC}}, \tag{6}$$

$$\boldsymbol{T}\boldsymbol{1} \geq \boldsymbol{E}_2^{k} \boldsymbol{\rho}_2^{k}, \quad \forall p_k \in \mathcal{P}_{\text{PDC}}. \tag{7}$$

The decision variable $L_{\hat{e}\phi}$ for the installation of the technology on the link is governed by

$$L_{\hat{e}\phi} = \max\{T_{e\phi} \mid e \in \hat{e}\}, \quad \forall \hat{e} \in \hat{\mathcal{E}}_{\text{ext}}, \tag{8}$$

under the availability of the communication technologies on each link by

$$\boldsymbol{L} \leq \boldsymbol{L}_{\text{avail}}, \tag{9}$$

where the element $L_{\text{avail},\hat{e}\phi}$ of $\boldsymbol{L}_{\text{avail}}$ is 1 if the technology τ_ϕ is available on undirected link $\hat{e} \in \mathcal{E}_{\text{ext}}$ and 0 otherwise. This is due to the fact that some communication technologies might not be possible on certain links because of the link distance or other factors. Note further that constraints (6) and (7) require indeed an equality to the maximum of the related sets. The formulations in (6) and (7) lead, however, still to the optimum solution since the objective function is minimized with the least number of selected technologies. Finally, the communication links are restricted by their available data rates which are considered by

$$\sum_{p_k \in \mathcal{P}_{\text{PDC}}} \sum_{q_r \in \mathcal{P}_{\text{PMU}}} g_r \boldsymbol{E}_1^{rk} \boldsymbol{\rho}_1^{rk} + \sum_{p_k \in \mathcal{P}_{\text{PDC}}} \sum_{q_r \in \mathcal{P}_{\text{PMU}}} g_r Z_{rk} \boldsymbol{E}_2^{k} \boldsymbol{\rho}_2^{k} \leq \boldsymbol{s} \tag{10}$$

under consideration of required data rates g_r by the individual PMUs, where the i^{th} row of the righthand side $\boldsymbol{s} \in \mathcal{R}^{n_{\text{edge}} \times 1}$ is written as the inner product of i^{th} rows of $\boldsymbol{T}$ and $\boldsymbol{M}$ as $s_i = < \boldsymbol{T}_i, \boldsymbol{M}_i >$.

Note that in this model the delay constraints are not included explicitly since this aspect is already handled by the preselection of paths in the optimization model. As also discussed in [11], one major concern for the measurement delays is the number of hops on paths due to the processing delays. However, further delay constraints can still be trivially added to this model as shown in [2] and [10].

The optimization problem for the integrated planning of the WAMS can finally be written as

$$\begin{aligned} \underset{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{T}, t_w, \mathbf{L}}{\text{minimize}} \quad & F(\boldsymbol{x}, \boldsymbol{y}, \mathbf{L}, t_w), \\ \text{subject to} \quad & (1) - (10), \end{aligned} \tag{11}$$

where the objective function $F(\boldsymbol{x}, \boldsymbol{y}, \mathbf{L}, t_w)$ is formulated as

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{y}, \mathbf{L}, t_w) = & \underbrace{\boldsymbol{c}_x^T \boldsymbol{x}}_{\text{PMU costs}} + \underbrace{\boldsymbol{c}_y^T \boldsymbol{y}}_{\text{PDC costs}} \\ & + \underbrace{\sum_{\hat{e} \in \hat{\mathcal{E}}_{\text{ext}}} \sum_{\phi=1}^{n_{\text{t}}} C_{\hat{e}\phi} L_{\hat{e}\phi} + t_w c_{\text{W}}}_{\text{Communication Costs}}, \end{aligned} \tag{12}$$

where $\boldsymbol{c}_x \in \mathcal{R}_+^{n_{\text{bus}}}$ and $\boldsymbol{c}_y \in \mathcal{R}_+^{n_{\text{p}}}$ are given cost vectors for the PMU and PDCs, respectively, and the third term is the total cost for the communication network, which includes the costs for each installed link as well as the fixed cost for WiMAX. Note that c_{W} is the assumed fixed license fee for WiMAX and t_w is an additional binary optimization variable with the constraints

$$t_w = \max\{t_{e\phi} \mid \forall e \in \mathcal{E}_{\text{ext}}, \tau_\phi = \text{WiMAX}\}. \tag{13}$$

Note that the optimization problem in (11) is a binary non-linear problem due to the multiplication of optimization variables in (10) and the maximum over the binary set of optimization variables as in (3), (8). Fortunately, these constraints can be linearized by additional binary variables and additional constraints to obtain a binary linear program [13]. However, the exact optimal algorithms for solution of binary linear programs can hardly provide solutions to very large problem instances due to memory issues. One way to alleviate this problem can be to cluster large networks into sub-graphs, to solve these smaller instances and combine the results, which would prohibit finding the global optimum. As an alternative, meta-heuristic approaches can be used to obtain accurate solutions, which we aim to do in this work. In the next section, we present the proposed genetic algorithm.

III. Genetic Algorithm for WAMS Design

Genetic algorithms are powerful population-based metaheuristics which are used in many different fields for solving hard optimization problems [14]. They owe their capability in finding very near-optimal and even optimal solutions of difficult optimization problems to the mathematical modeling of key evolutionary concepts, such as survival of the fittest, recombination, reproduction, and gene mutations. The generic genetic algorithm, as described in Algorithm 1, can be applied to any unconstrained optimization problem in a simple manner. However, constrained optimization problems, such as the one we have under consideration, require problem-specific approaches in order to design a powerful and robust algorithm which delivers solutions with an acceptable accuracy and precision [14].

Algorithm 1 Simple Generic Algorithm

```
p ← InitializePopulation(popSize, problemSize)
EvaluateFitness(p)
S_best ← GetBestSolution(p)
repeat
    ChildrenCO ← ApplyCrossover(p)
    ChildrenM ← ApplyMutation(p)
    EvaluateFitness(ChildrenCO, ChildrenM)
    p ← AddToPopulation(ChildrenCO, ChildrenM)
    p ← ApplySelection(p)
    S_best ← GetBestSolution(p)
until Stopping criteria
return S_best
```

Obviously, any constrained optimization problem can be transformed into an unconstrained one by introducing penalty coefficients for the violated constraints. Furthermore, the problem structure can also be exploited to generate feasible solutions, or solutions satisfying a part of the constraints, and to create problem-specific genetic operators which preserve feasibility and repair the solutions in case of infeasiblity [14].

A look at the literature shows several works which took a similar approach to solve constrained ILPs, see for example [15], [16]. The optimization problem in (11) is a highly constrained binary optimization problem, which is similar to a multi-level assignment or time scheduling problem. One aspect which differs from the previously studied problems, for example compared to [15], is the significantly higher number of constraints than the number of optimization variables, which make search much more challenging. Therefore the results presented in the present paper can provide some insight about the performance of genetic algorithms in such highly constrained problems. In the following, we present our approach with a genetic algorithm for solving the problem in (11).

A. Generation of the Initial Population

In approaching this problem with a genetic algorithm, we use a random initialization strategy with partial feasible solutions which satisfy the constraints (1) to (9). The violation of capacity constraints (10) is penalized as an additional term in the objective function.

The generation of the initial population follows a hierarchical flow, where the optimization variables are randomly generated starting by $\boldsymbol{x}$ satisfying constraint (1) followed by $\boldsymbol{y}$, $\boldsymbol{z}$ and the others until all the constraints up to (9) are satisfied.

B. Problem-Specific Genetic Operators

In the design of a heuristic optimization technique, two main concepts, namely diversification and intensification, play a crucial role [17]. The former one relates to the capability of the optimization algorithm in searching wide regions in the search space of the problem, whereas the latter one is related to its capability for further intensifying the search around possible regions of interest for even higher-quality solutions. The convergence behavior and also the performance of the designed algorithm are mainly determined by the balance between these two factors [17]. In the following, we describe the problem-specific operators that we adopted in this work along with their function and parametrization in relation to the aspects of diversification and intensification. An overview of the complete algorithm is provided in Algorithm 2.

1) PMU-PDC Location Crossover: This crossover is introduced to bring more diversity in the population and to extend the search space by a crossover on the PMU locations and PDC locations. Basically, two new children are generated from two randomly selected parents by a one-point cross-over of PMU locations included in variable $\boldsymbol{x}$. For each PMU location coming from a parent, the generated children take over the corresponding PDC locations, communication paths, and links from the related parent. Any infeasible child is repaired by adding further PMU locations and communication paths.

Algorithm 2 Generic Algorithm for WAMS Planning

$pd \leftarrow$ GetProblemData(...)
$p \leftarrow$ InitializePopulation($popSize$, pd)
EvaluateFitness(p)
$S_{\text{best}} \leftarrow$ GetBestSolution(p)
repeat
    ChildrenCO $\leftarrow$ ApplyPMUCrossover(p, pd)
    ChildrenM1 $\leftarrow$ ApplyPMUMutation(p, pd)
    ChildrenM2 $\leftarrow$ ApplyPDCMutation(p, pd)
    ChildrenM3 $\leftarrow$ ApplyPathMutations(p, pd)
    ChildrenM4 $\leftarrow$ ApplyTechnologyMutations(p, pd)
    $p \leftarrow$ AddToPopulation(ChildrenCO, ChildrenM1, ... ChildrenM4)
    EvaluateFitness(p)
    $p \leftarrow$ ApplySelection(p)
    $S_{\text{best}} \leftarrow$ GetBestSolution(p)
until Stopping criteria
return S_{best}

2) PMU / PDC Path Mutations: In these mutations, as well as in the other mutations below, the main goal is to boost intensification in the neighborhood of the individual population members. Furthermore, depending on the number of mutated variables, the mutation operators can also diversify the population. We aim to achieve this effect by starting with a high number of mutated genes and decreasing it depending on the iteration number.

In path mutations, the selected path for a selected PMU/PDC is mutated and the communication links are updated if an update is necessary to satisfy the constraints.

3) PMU / PDC Location Mutations: In this case, one or more PMU / PDC locations are mutated. If necessary for the observability or for the assignment between PMUs and PDCs, further PMUs or PDCs are added and their communication paths are selected to preserve the feasibility regarding the constraints (1) to (9).

4) Communication Technology Mutations: In these mutations, the selected technology variables over one or more selected links are mutated.

IV. Performance Evaluation

In order to test the designed algorithm, several benchmark instances of the formulated model presented in Section IV are used, where medium-voltage distribution network generator in [18] is used to generate realistic power grid topologies. An overview of benchmark models is provided in Table IV with the related parameters and problem sizes. Note that the problems A and B, and C and D have different number of optimization variables and constraints due to their different power system topologies although they have the same number of nodes and the same number of possible PDC locations.

In the scope of this work, we focus on the performance of the proposed algorithm in terms of optimality observed in the test cases. An analysis of the possible cost-savings due to the modeling approach adopted in Section II is provided in [2] based on realistic cost assumptions and technology specifications. Interested readers may refer to [2] for details and cost assumptions.

The problem-specific genetic algorithm is implemented in MATLAB and applied to the benchmark model instances with 30 initializations and different seeds of the random number generator. In addition to the reported population sizes in Table V and Table VI, we use the following parameters for the genetic operators. The number of generated offsprings is set to 50% of the population size for the crossover and 40% for each mutation type. In each iteration 80% of the new population is selected with the best members and the rest 20% is filled randomly from the remaining off-springs to keep the diversity in the population. The optimization is continued until the global optimum has been reached or there has not been any improvement in the last 30 iterations.

Table V provides a summary of the deviation (best and worst achieved values) from the optimal objective function value in each of the cases.

We can see that the presented and implemented algorithm can attain the optimal solution of all the benchmark models, where for the largest benchmark model the algorithm delivers near-optimal solutions in some of the trials. As usual, the

TABLE IV
BENCHMARK MODELS

Id	Region Size	n_{bus}	n_{pow}	n_{com}	Number of Variables	Number of Constraints
A	5×5 km	10	4	2	398	3487
B	5×5 km	10	4	2	572	4502
C	15×15 km	25	6	4	1139	22423
D	15×15 km	25	6	4	1183	24555
E	30×30 km	50	8	4	4777	84203

TABLE V
SUMMARY OF RESULTS - OPTIMALITY

Id	Population Size	Deviation From Optimum (%)			
		Min	Max	Mean	Var
A	200	0	0	0	0
B	200	0	0	0	0
C	1000	0	0	0	0
D	1000	0	0	0	0
E	3000	0	0.89	0.39	0.09

TABLE VI
SUMMARY OF RESULTS - RUN-TIME

Id	Population Size	Run-time (s)			
		Min	Max	Mean	Var
A	200	1.17	5.19	2.2	0.77
B	200	1.85	7.9	3.31	1.94
C	1000	13.50	78.56	43.26	221.7
D	1000	34.13	115.6	77.13	459.5
E	3000	1616	4419	3403	3.04×10^5

selection of parameters plays a crucial role in the performance of genetic algorithms [14]. However, the investigation of the parameter settings for this particular problem is devoted to future work.

Table VI presents the best and worst run times over all trials. The reported CPU times are obtained using an Intel(R) Core(TM) i7-4790S processor with a clock rate of 3.20GHz and 16 GB of random-access memory (RAM). Note that we have observed in further experiments, in fact, a better run-time when the linearized models of the problem (11) are solved by the powerful ILP solver Gurobi version 7 [19] on the same machine by using 8 CPUs in parallel. However, our experience with solving the model presented in [2] shows that it is not possible to solve problems with more than around 500,000 optimization variables on the same machine due to memory issues. This has been our main motivation for this work.

V. CONCLUSION

In this work, we have presented an optimization model which enables minimum-cost planning of a WAMS with a heterogeneous communication network. The advantage of the presented formulation in comparison to our previous work [2] is the lower number of optimization variables in large problem instances due to the path-based formulation. Furthermore, we proposed a customized genetic algorithm for the solution of this problem which have provided optimal solutions to ILP models with up to around 5000 binary variables and 85000 constraints with the current selected algorithm parameters. We believe that further improvements can be achieved by optimizing the algorithm parameters and by introducing control mechanisms in crossover and mutations in order to bring diversity based on some observations of algorithm flow rather than simply relying on random generation.

REFERENCES

[1] "IEEE Standard for Synchrophasor Data Transfer for Power Systems," *IEEE Std C37.118.2-2011*, pp. 1–53, Dec 2011.

[2] H. A. Tokel, G. Alirezaei, and R. Mathar, "Integrated network design for measurement and communication infrastructures in smart grids," in *26th International Telecommunication Networks and Applications Conference (ITNAC 2016)*, Dunedin, New Zealand, Dec. 2016, pp. 267–273.

[3] ——, "Modern heuristical optimization techniques for power system state estimation," in *The 2016 International Conference on Swarm Intelligence Based Optimization*, June 2016, pp. 78–85.

[4] A. Angioni, G. Lipari, M. Pau, F. Ponci, and A. Monti, "A low cost PMU to monitor distribution grids," in *2017 IEEE International Workshop on Applied Measurements for Power Systems (AMPS)*, Sept 2017, pp. 1–6.

[5] Y. Sun, X. Chen, S. Yang, Rusli, K. J. Tseng, and G. Amaratunga, "Micro PMU based monitoring system for active distribution networks," in *2017 IEEE 12th International Conference on Power Electronics and Drive Systems (PEDS)*, Dec 2017, pp. 518–522.

[6] "Factors affecting PMU installation costs," US Department of Energy, October 2014, Accessed : 30.03.2018. [Online]. Available: "https://www.smartgrid.gov/files/PMU-cost-study-final-10162014_1.pdf"

[7] A. Ghasemkhani, H. Monsef, A. Rahimi-Kian, and A. Anvari-Moghaddam, "Optimal design of a wide area measurement system for improvement of power network monitoring using a dynamic multiobjective shortest path algorithm," *IEEE Systems Journal*, vol. PP, no. 99, pp. 1–12, 2015.

[8] M. B. Mohammadi, R. A. Hooshmand, and F. H. Fesharaki, "A new approach for optimal placement of PMUs and their required communication infrastructure in order to minimize the cost of the WAMS," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 84–93, Jan 2016.

[9] M. H. F. Wen and V. O. K. Li, "Optimal phasor data compression unit installation for wide-area measurement systems–an integer linear programming approach," *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1–10, 2015.

[10] H. A. Tokel, G. Alirezaei, S. Baig, and R. Mathar, "An optimization framework for planning of WAMS with a heterogeneous communication network," in *2016 IEEE International Conference on Smart Grid Communications*, Sydney, Australia, Nov. 2016.

[11] K. V. Katsaros, B. Yang, W. K. Chai, and G. Pavlou, "Low latency communication infrastructure for synchrophasor applications in distribution networks," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, Nov 2014, pp. 392–397.

[12] A. Phadke and J. Thorp, *Synchronized Phasor Measurements and Their Applications*, ser. Power Electronics and Power Systems. Springer US, 2008.

[13] H. P. Williams, *Model Building in Mathematical Programming, 5th Edition*. John Wiley & Sons, Inc, 2013.

[14] X. Yu and M. Gen, *Introduction to evolutionary algorithms*. Springer Science & Business Media, 2010.

[15] K. Deb and K. Pal, "Efficiently solving: A large-scale integer linear program using a customized genetic algorithm," in *Genetic and Evolutionary Computation – GECCO 2004*, K. Deb, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 1054–1065.

[16] A. Colorni, M. Dorigo, and V. Maniezzo, "Metaheuristics for high school timetabling," *Computational Optimization and Applications*, vol. 9, no. 3, pp. 275–298, Mar 1998. [Online]. Available: https://doi.org/10.1023/A:1018354324992

[17] F. Rothlauf, *Design of Modern Heuristics: Principles and Application*. Springer Publishing Company, Incorporated, 2011.

[18] E. Schweitzer, K. Togawa, T. Schloesser, and A. Monti, "A Matlab GUI for the generation of distribution grid models," in *International ETG Congress 2015; Die Energiewende - Blueprints for the new energy age; Proceedings of*, Nov 2015, pp. 1–6.

[19] Gurobi Optimization Inc., "Gurobi optimizer reference manual," 2015. [Online]. Available: http://www.gurobi.com