# Distributed Computing Architecture for Optimal Control of Distribution Feeders with Smart Loads

Abolfazl Mosaddegh, *Student Member, IEEE,* Claudio A. Cañizares, *Fellow, IEEE,* Kankar Bhattacharya, *Senior Member, IEEE,* and Hongbing Fan, *Member, IEEE*

***Abstract*—This paper presents a distributed computing architecture for solving a Distribution Optimal Power Flow (DOPF) model based on a Smart Grid Communication Middleware (SGCM) system. The system is modeled as an unbalanced three-phase distribution system, which includes different kind of loads and various components of distribution systems. In this study, fixed loads are modeled as Constant Impedance, Current and Power (ZIP) loads, and Neural Network (NN) models of controllable smart loads are integrated into the DOPF model. A Genetic Algorithm (GA) is used to determine the optimal solutions for controllable devices, in particular Load Tap Changers (LTCs), Switched Capacitors (SCs), and smart loads in the context of an Energy Management System (EMS) for practical feeders, accounting for the fact that smart loads consumption should not be significantly affected by network constraints. Since the number of control variables in a realistic distribution power system is large, solving the DOPF for real-time applications is computationally expensive. Hence, to reduce computational times, a decentralized system with parallel computing nodes based on an SGCM system is proposed. Using a "MapReduce" model, the SGCM system runs the DOPF model, communicates between master and worker computing nodes, and sends/receives data among different parts of parallel computing system. Compared to a centralized approach, the proposed architecture is shown to yield better optimal solutions in terms of reducing energy losses and/or energy drawn from the substation within adequate practical run-times for a realistic test feeder.**

***Index Terms*—Distributed computing, distribution optimal power flow, genetic algorithm, real-time application, smart grid communication middleware system.**

## NOMENCLATURE

*Indices*

| | |
|---|---|
| $h$ | Time interval, $h = 1, ..., 24$. |
| $l$ | Feeder line. |
| $n$ | Nodes. |
| $nd_W$ | Worker-nodes. |
| $p$ | Phases, $p = a, b, c$. |
| $r$ | Set of objective functions, $r = 1, 2$. |

*Parameters*

| | |
|---|---|
| $C^{\max}$ | Total number of capacitor blocks available in capacitor banks. |
| $CR$ | One point cross-over rate. |
| $EN$ | Elite off-spring selection. |
| $G$ | Maximum number of generations in GA. |
| $G_m$ | GA generations. |
| $\overline{I}$ | Maximum feeder current limit [A]. |
| $MR$ | Mutation rate. |
| $Nd$ | Distributed nodes in the Middleware System. |
| $PS$ | Population size. |
| $Ps$ | Simulated active power demand for controllable smart loads [kW]. |
| $R_{max}$ | Maximum allowable value of $R$ [kWh]. |
| $R_{min}$ | Minimum allowable value of $R$ [kWh]. |
| $SV$ | Stall generation value. |
| $\overline{T}$ | Maximum tap changer position. |
| $TL$ | Termination time limit [min]. |
| $\underline{T}$ | Minimum tap changer position. |
| $\overline{V}$ | Maximum voltage limit [kV]. |
| $\underline{V}$ | Minimum voltage limit [kV]. |
| $\alpha$ | Peak demand factor. |
| $\gamma_{max}$ | Maximum allowable value of peak demand cap [kW]. |
| $\gamma_{min}$ | Minimum allowable value of peak demand cap [kW]. |
| $\theta$ | Temperature [°C]. |
| $\sigma$ | Time of Use (TOU) Tariff [\$/MWh]. |

*Variables*

| | |
|---|---|
| $cap$ | Number of capacitor blocks switched on. |
| $F$ | Objective function [kWh]. |
| $I$ | Current phasor [A]. |
| $J_1$ | Energy losses [kWh]. |
| $J_2$ | Energy drawn from the substation [kWh]. |
| $R$ | Sum of peak demand cap signals [kWh]. |
| $tap$ | Tap position. |
| $V$ | Voltage phasor [kV]. |
| $\gamma$ | Peak demand cap [kW]. |

## I. INTRODUCTION

DISTRIBUTION systems have been insufficiently changing in recent years, moving away from traditional systems with low-level control towards smart grids with high-level control, based on improved technologies in communication, monitoring, computation, and real-time control [1], [2]. The need to make the grid more reliable and more efficient, driven by economic growth and climate change concerns, has led to the development of smart grids, which incorporate technologies and advanced applications to optimize, modernize, and transform the electric power infrastructure. In these grids, customers have now the ability to use electricity more efficiently

This work was supported by Hydro One Networks, Energent Inc., Milton Hydro Distribution, the Ontario Power Authority (OPA), and the Ontario Centres of Excellence (OCE) under the Energy Hub Management System (EHMS) (http://www.energyhub.uwaterloo.ca/) and SGCM (http://bohr.wlu.ca/hfan/rinlab/sgcm/) projects, and also by the Natural Sciences and Engineering Research Council (NSERC) Smart Microgrid Research Network (NSMG-Net).

A. Mosaddegh, C. A. Cañizares, and K. Bhattacharya ({amosaddegh, ccanizares, kankar}@uwaterloo.ca) are with University of Waterloo, Waterloo, ON, Canada, and H. Fan (hfan@wlu.ca) is with Wilfrid Laurier University, Waterloo, ON, Canada.

through smart metering and intelligent controls, based on new communication infrastructure, information management, and automated control technologies that empower power grids to operate in more efficient, more intelligent, and faster ways [3].

One of the basic components of smart grids is Distributed Automation (DA), which includes "smart meters" to provide two-way communication between utilities and customers. These meters have the ability to measure energy consumption in real-time and also provide usage information. With this capability, both utilities and customers can communicate, observe, and control their specific needs [4]. Voltage and reactive power control are two important aspects of DA, which have been traditionally carried out by Load Tap Changers (LTCs), Switched Capacitors (SCs), fixed capacitors, and step-voltage regulators. To this effect, Mixed-Integer Non-Linear Programming (MINLP) models have been proposed to determine the discrete settings of LTCs and SCs with the goals of reducing power losses and improving voltage profiles [5]–[8]. Rule-based methods, which use real-time measurements, have been applied to control discrete devices such as SCs and LTCs, wherein the key factors for determining the discrete settings are rules and operator experience [9]. In [10], the authors propose a rule-based solution with different levels of measurement and communication infrastructure in the distribution system. Because rule-based methods do not guarantee an optimal solution, for systems with local measurements, the authors in [11] propose regular changes in the settings of controllers to account for seasonal load variations.

Network-model-based methods are based on different parameters such as the topology of a distribution system, system configuration, measurements, and statistical data [9], plus Distribution Load Flow (DLF) and Distribution Optimal Power Flow (DOPF) models. For example in [12], a quasi-Newton method is implemented to obtain an optimal DLF of the distribution system in real-time, using a Non-Linear Programming (NLP) model with continuous controls. To solve network-model-based Voltage/Var Control (VVC) problems for larger systems and to reduce computational costs, heuristic methods such as Neural Network (NN) techniques [6], computational cost reduction [11], and decoupling of the VVC problem into sub-problems [8] have been proposed. In [13] and [14], the authors consider the DLF model as a constraint in the DOPF model, and relax the MINLP problem into an NLP problem by rounding the discrete variables to their closest integer values. Reducing the complexity of the DOPF model reduces the computational burden and makes it suitable for real-time applications. However, the drawback of rounding continuous variables into discrete variables is that the solution may be sub-optimal. Therefore, in order to solve these MINLP problems without approximations, Evolutionary Algorithms (EAs) such as Genetic Algorithms (GAs) have been proposed. For example, in [15], heuristic and GA-based solution methods to solve the DOPF problem are compared, noting that although the GA-based method yields a better solution, the computational burden is quite significant, making the model unsuitable for real-time applications.

The Smart Grid Communication Middleware (SGCM) system is a software system designed for smart grid applications including data acquisition, monitoring, and distributed computing. This system is built on top of a reconfigurable interconnection network computing platform [16], and provides reliable, secure, and fast performance two-way communication among nodes of utility servers, energy hubs [17]–[19], and smart sensors/controllers. Similar to Hadoop [20]–[22], a distributed computing platform for "MapReduce" models [23], the SGCM uses proprietary protocols on top of the Transmission Control Protocol (TCP)/Internet Protocol (IP) for inter-communication of computing components. However, unlike Hadoop, the SGCM uses programmable data paths to distribute/collect data to/from computing nodes running distributed computers, while Hadoop uses a chained pipe to distribute files to different data nodes in a cluster. Hence, the SGCM has the same functionality for solving MapReduce problems as Hadoop does, but in a customized approach.

The SGCM provides a framework for MapReduce model based on distributed computing applications. The framework contains templates to build client, master, and worker components by adding application specific data structures, data file information, "Map" or "Reduce" processes, and work-flow logic. The client component runs on a console to distribute, install, and configure master and worker components on available nodes; it distributes data files to the master-node (i.e., the node with master components installed), and calls master functions to run. The master-node creates jobs in forms of files, sends jobs to worker-nodes (i.e., nodes with worker components installed) for Map or Reduce processes, and tracks jobs. Worker-nodes call Map or Reduce programs to process job files, and send result files to the master-node or other worker-nodes according to the work-flow logic. Since the SGCM framework allows to accommodate more complicated work-flow logic, it has been used in this work as a template to implement a recursive MapReduce model for the GA-based DOPF solution proposed in this paper.

In this paper, in order to reduce the computational burden of the DOPF model with a GA-based solution and make it suitable for real-time applications, the SGCM system is chosen as the preferred platform for a recursive MapReduce model due to its fast performance when large number of nodes are involved; it is light, reliable, and flexible, and is capable of running multiple smart grid applications simultaneously. Hence, based on these discussions, the main contributions of the present work are as follows:

- Modify the DLF model of an unbalanced three-phase distribution system in the OpenDSS simulation tool [24] to include an NN-based model of new controllable smart loads, Constant Impedance, Current and Power (ZIP) loads, and other components of distribution systems. The DLF model is a stand-alone executable program that can be used on different platforms (i.e., Windows, Linux, and Mac).
- Develop an enhanced DOPF model with a GA-based solution to find optimal decisions regarding switching of control devices such as LTCs and SCs, and Local Distribution Company (LDC) imposed peak demand caps on controllable smart loads.
- Develop a recursive MapReduce model, as discussed

in [25], using a distributed computational approach on an SGCM system, to reduce the computational time of the aforementioned GA-based model to make it suitable for real-time applications. The performance of the proposed distributed computing approach is demonstrated and compared with a centralized DOPF solution using a practical test feeder.

The rest of the paper is organized as follows: Section II presents the unbalanced three-phase DOPF model, which includes the DLF model of the unbalanced three-phase distribution system. Section III describes the architecture of the SGCM system, its applications to distributed computing, and the implementation of the DOPF model on the proposed distributed computing platform, using a MapReduce model. In Section IV, the results obtained from different realistic case studies and scenarios are discussed. The main conclusions and contributions of the presented work are highlighted in Section V.

## II. Unbalanced Three-Phase DOPF Model

In the DOPF model, the decision variables from the LDC's perspective are the optimal set of LTC tap positions, the number of capacitor blocks switched on, and the demand caps on controllable loads, with the objective of minimizing energy losses and/or energy drawn from the substation. On the other hand, since customers tend to maximize their comfort and convenience, which may lead to high peaks, the proposed DOPF model should capture this behavior of controllable customers in its objective function. Hence, the proposed objective function is as follows:

$$F_r = J_r - \alpha_r R \quad (1)$$

where all variables and parameters in this and other equations are defined in the Nomenclature Section, with the first term representing the feeder energy losses and/or energy drawn from the substation, and the second term:

$$R = \frac{1}{24}\sum_h \sum_n \sum_p \gamma_{p,n,h} \quad (2)$$

corresponding to caps in the smart loads' peak demand, which are maximized to represent customers' behavior. The parameter $\alpha_r$ allows to control the relative weight given to the two conflicting terms in the objective function (1).

Common distribution system components are modeled in the OpenDSS simulation tool [24]. Thus, series components model conductors/cables, transformers, switches, and LTCs; and shunt components represent loads, fixed capacitors, SCs, and Distributed Generations (DGs). Polynomial ZIP load models are used to represent unbalanced three-phase loads. In this tool, a current injection model, Kirchhoff's Voltage Law (KVL), and Kirchhoff's Current Law (KCL) equations are used to compute per-phase feeder power flows, as in [14] and [15].

To study the impact of demand caps on controllable smart loads, the load profile $Ps_{p,n,h}$ for a smart load, proposed, implemented, and deployed in [26], are estimated using an NN with Time of Use (TOU) tariff ($\sigma_h$), ambient temperature

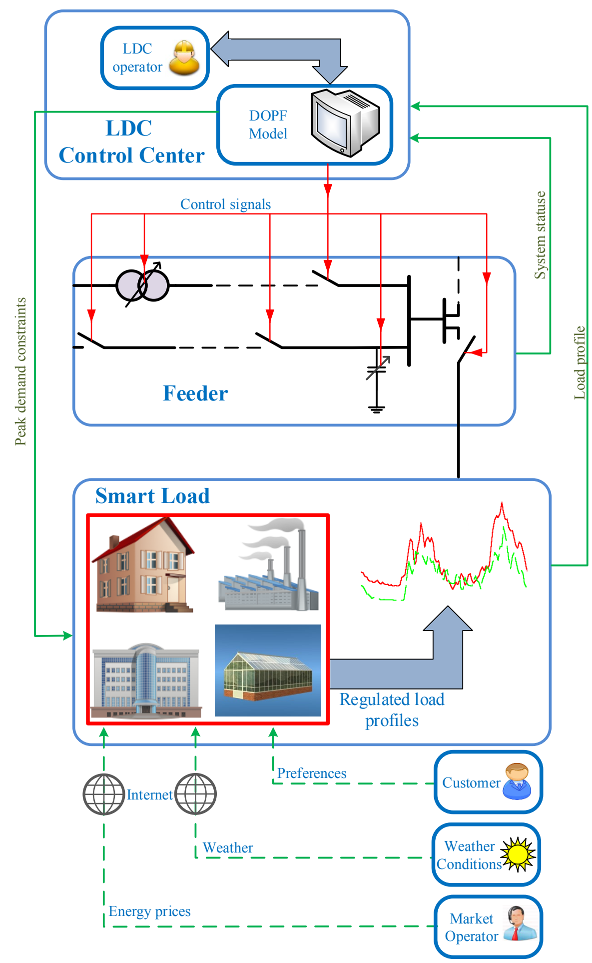



Fig. 1. Schematic of the proposed optimal feeder control [14], [28].

($\theta_h$), time ($h$), and peak demand cap ($\gamma_{p,n,h}$) as inputs as follows [27]:

$$Ps_{p,n,h} = f(\sigma_h, \theta_h, h, \gamma_{p,n,h}) \quad (3)$$

Hence, without loss of generality, $\gamma$ is assumed to remain constant in a 24-hour timeframe. A fixed power factor is considered for all controllable loads, modeling them as polynomial ZIP loads in order to better represent the various appliances accounted in smart loads. The estimated smart load profile is integrated within the DOPF as a function of the LDCs peak demand cap signal, while the other parameters in (3) are assumed known for the day ahead. Thus, the load profile of each smart load is controlled by peak demand cap.

Figure 1 presents the architecture of the proposed feeder and load control approach [14], [28]. In this framework, customers have the ability to receive real-time information from weather stations and market operators, so that their smart load controllers internally optimize their energy consumption and a load profile is obtained. These load profiles are collected by the LDC from the smart loads together with the distribution system status, and used by the LDC control center in real-time to solve the DOPF model. The optimal controls including the peak demand caps obtained from the DOPF solution are fed back to smart loads to regulate their load accordingly, considering the following operating feeder constraints:

$$\underline{V} \leq V_{p,n,h} \leq \overline{V} \qquad \forall\, p, n, h \tag{4a}$$

$$I_{p,l,h} \leq \overline{I} \qquad \forall\, p, h \tag{4b}$$

$$\underline{T}_{p,n} \leq tap_{p,n,h} \leq \overline{T}_{p,n} \qquad \forall\, p, n, h \tag{4c}$$

$$0 \leq cap_{p,n,h} \leq C^{\max}_{p,n} \qquad \forall\, p, n, h \tag{4d}$$

$$\gamma_{\min p,n,h} \leq \gamma_{p,n,h} \leq \gamma_{\max p,n,h} \qquad \forall\, p, n, h \tag{4e}$$

where (4a) represents the nodal voltage operating limits; (4b) the feeder current operating limits; (4c) the LTCs operating limits; (4d) the SCs operating limits; and (4e) the peak demand cap limits. Observe that, since the demand varies over the day, the LTCs and SCs would switch frequently in order to keep voltages within limits; however, maintenance costs of LTCs and SCs are considerable. Therefore, limits on the number of switching operations per day (5 times) and limits on maximum step changing over consecutive hours ($\pm$ 3 steps) for LTCs and SCs are also considered here.

The proposed feeder and load models are implemented on the MATLAB® environment, which allows to generate an executable file that can be run on a multitude of commonly available platforms, without the need to have MATLAB® at each computing node.

## III. Proposed Computing Architecture and Implementation

### A. SGCM System Architecture

The core components of the SGCM system are the nodes and the console. The node is a client and a server program, which can be installed and run on commodity computers and smart devices. As a client, the node is able to make TCP connections to other nodes, and as a server, it waits for a TCP connection and provides services. The node is a virtual router which can forward packets to other nodes. Also, the node is a service provider, on which application modules can be installed. The node is a processor that executes groups of node instructions, which consist of: net instructions for creating structured data path of nodes; file instructions for file Input/Output (I/O) and file transfer; execution instructions for remote function calls; bundle instructions for install/uninstall application modules; and cluster instructions for node cluster discovery and management. The console is a client program used to connect to a node and to launch network commands. The console is a command line interface equipped with script language extended with the command functions. Thus, application deployment, installation, execution, and management can be done by running script console applications.

A typical smart grid application involves multiple computing components on different devices that need to exchange data during the execution of the application. In the SGCM system, computing components are installed as modules on the nodes. Inter-communications of the computing components are realized by structured data paths, which are computed and established by the reconfiguration of nodes. The communication pattern can be one-to-one, one-to-many, many-to-one, or many-to-many. For a typical one-to-many communication

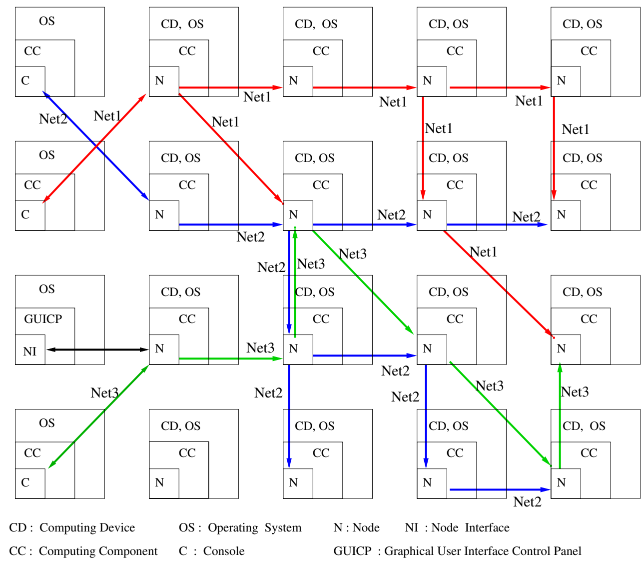



Fig. 2. Multiple applications running simultaneously on an SGCM system [16].

request, routing algorithms are used to find an optimal network topology (i.e., a Steiner tree), and a structured data path of the topology is established through reconfiguration of the nodes. The data transfer for the one-to-many communication request is through the data path with an identical net Identification (ID), which avoids data paths interfering with each other when they overlap; thus, the SGCM system supports simultaneous multiple applications. Also, the SGCM system has flexibility to set different fault tolerance levels with available computing resources. For high level of fault tolerance, the system can be configured to run the same job simultaneously on different groups of nodes, and for low level of fault tolerance, the task tracker on the master-node monitors the execution of Map or Reduce process on worker-nodes, and in case of failure, it assigns the task to alternative nodes. Figure 2 illustrates the architecture of SGCM with nodes, a console, and programmable data paths of applications. This architecture and its communication protocols are based on reconfigurable interconnection network computing platform [16].

### B. SGCM for Distributed Computing

Complex smart grid applications involve a large number of smart computing devices in data acquisition and processing, monitoring, control, and intelligent computing, and have high requirements on performance, reliability, and security. Due to the constraints of cost and distributed computing resources, a distributed approach is used in the SGCM system, which provides a platform for solving distributed optimization problems, particularly, problems formulated by a MapReduce model or more complicated work-flow model. The following are the two main steps of the MapReduce model [23]:

- *Map*: The master-node divides the input into smaller sub-problems and distributes them between the nodes. The worker-nodes process the sub-problems and return the solutions to the master-node. If the system has a tree structure, each worker-node acts as a new master-node for its sub-branches.

- *Reduce*: The master-node hierarchically collects the results from the entire set of worker-nodes, combines them in an output format, and releases the results.

Sometimes, there is a "Shuffle" step mid-stage between the Map and Reduce steps, wherein the output of Map operation is sorted in parallel or exchanged between the nodes, in order to prepare the data for a Reduce step.

### C. Implementation of the DOPF on a Distributed Computing Platform

Figure 3 shows the flowchart for the implementation of the DOPF model using a GA-based solution approach on the proposed distributed computing platform. After defining the objective function $F$, the parameter values for $G$, $PS$, $EN$, $SV$, $TL$, $CR$, and $MR$ are chosen. For a 24-hour timeframe, the individuals, which are determined by the GA procedure, are appropriately distributed from the master-node among the $Nd$ worker-nodes of the SGCM system. Each worker-node runs the DLF model, and the master-node then combines the results, and ranks them by their value of $F$. Based on the ranks, a pool of individuals is selected, and the cross-over and mutation operators are applied to generate an off-spring; the best $PS$ fits among parents and off-springs are then selected for the new generation. The stopping criteria of the GA algorithm are the following:

1) *The run-time is longer than time $TL$*, which is a criterion checked at the end of each generation. In order to feed the system with the new optimized controllable values in real-time applications, $TL$ should be less than the time interval of the application (e.g., 1 hour for the DOPF model).
2) *The solution does not change over $SV$ number of iterations*. Here, $SV$ = $G$ in order to continue the GA up to the last generation, but has been found that $SV$ = 10 is adequate for the DOPF to reach a solution.
3) *$G_m$ reaches $G$*. If the first two stopping criteria are not satisfied before the $G$ number of generations, the solution for the last generation is used.

The output of this model provides the optimal 24-hour LTC tap positions, number of capacitor blocks switched on, and peak demand caps for the smart loads represented by (3).

In the distributed computing approach, there are different overheads, such as communication overhead time, fault overhead time, and bandwidth overhead. The communication overhead time is the time taken to transfer data between the master-node and worker-nodes or amongst worker-nodes. The larger the size of data needed to be transferred, the longer the communication overhead time of the program. Also, when the number of the worker-nodes increases, the limit on the data transfer bandwidth should be considered. For example, if a series of data from different worker-nodes arrived simultaneously at the communication channel towards the master-node, a sequential data transfer, which requires longer run-times, is configured in the SGCM system to satisfy the bandwidth constraint. In addition, different faults may take

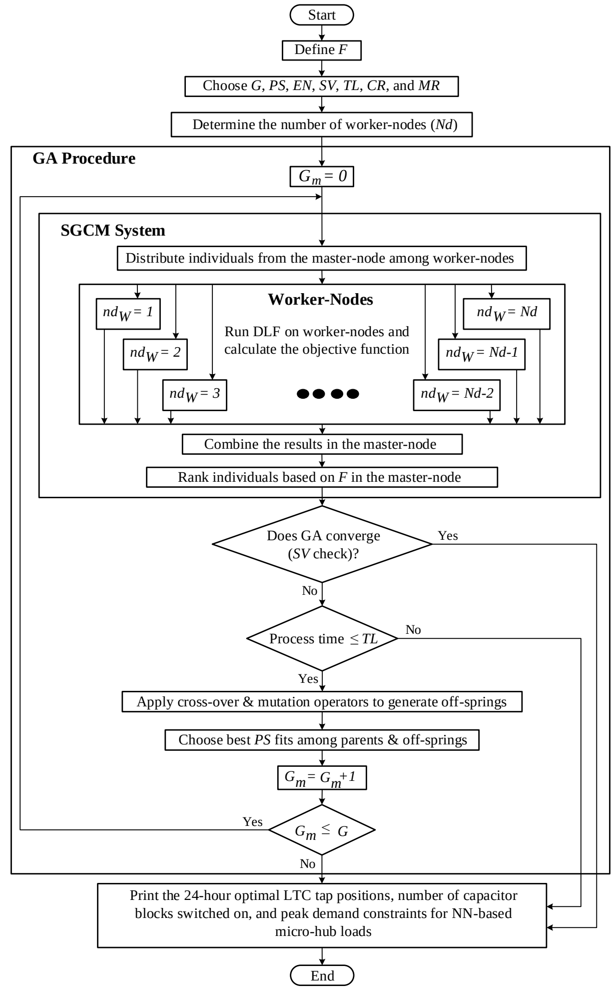



Fig. 3. Flowchart of proposed implementation of the DOPF model with a GA-based solution.

place, such as failure of SGCM nodes, which are addressed by assigning the task to the alternative nodes or re-running the task on the same nodes; this improves the reliability of the system, but increases run-times.

By distributing the individuals from the master-node to the worker-nodes, the run-time can be significantly reduced. If the overhead times, specifically communication overheads, are considerably longer than computing node run-times, they might affect the speed of the distributed approach. In the SGCM system, by reducing the transferred data size, the communication overheads are negligible and thus insignificant with respect to computing nodes run-times. Therefore, the SGCM system, with enough processors and worker-nodes, and short overhead times is expected to arrive at the required DOPF solutions in real-time applications.

## IV. Results and Analysis

### A. DOPF Results

The real unbalanced distribution feeder with 41 nodes shown in Fig. 4 is used to test and demonstrate the proposed approach. The system has three three-phase transformers

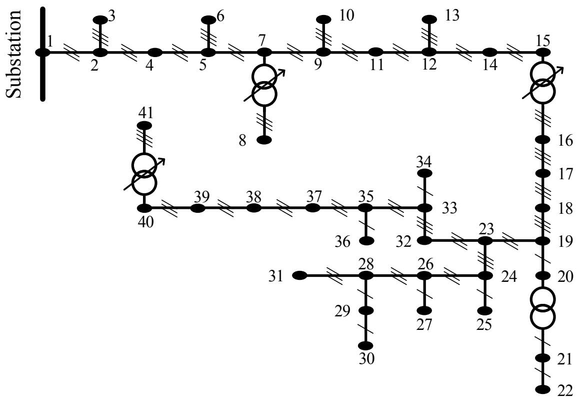



Fig. 4. Practical distribution test feeder [29].

TABLE I
COMPARISON OF $F_1$, $J_1$, AND $R$ FOR DIFFERENT NUMBER OF WORKER-NODES

| | | | Generation | | |
|---|---|---|---|---|---|
| | | | 25 | 50 | 100 |
| ***Nd*** | **1** | $J_1$ [kWh] | 7892.04 | 8071.10 | 7757.61 |
| | | $R$ [kWh] | 1851.99 | 1956.57 | 2046.32 |
| | | $F_1$ [kWh] | -47667.63 | -50626.03 | -53632.05 |
| | **2** | $J_1$ [kWh] | 8133.57 | 8054.96 | 7905.68 |
| | | $R$ [kWh] | 1958.79 | 1971.66 | 2008.46 |
| | | $F_1$ [kWh] | -50630.16 | -51094.78 | -52348.06 |
| | **5** | $J_1$ [kWh] | 8419.75 | 8011.12 | 8331.62 |
| | | $R$ [kWh] | 1887.80 | 1990.12 | 2069.62 |
| | | $F_1$ [kWh] | -48214.25 | -51692.36 | -53757.04 |
| | **10** | $J_1$ [kWh] | 8172.41 | 7912.79 | 8017.24 |
| | | $R$ [kWh] | 1890.12 | 1938.29 | 2069.41 |
| | | $F_1$ [kWh] | -48531.19 | -50235.79 | -54065.09 |
| | **20** | $J_1$ [kWh] | 7882.74 | 8103.66 | 7881.85 |
| | | $R$ [kWh] | 1915.08 | 1990.83 | 2029.93 |
| | | $F_1$ [kWh] | -49569.72 | -51621.09 | -53015.99 |

TABLE II
COMPARISON OF $F_2$, $J_2$, AND $R$ FOR DIFFERENT NUMBER OF WORKER-NODES

| | | | Generation | | |
|---|---|---|---|---|---|
| | | | 25 | 50 | 100 |
| ***Nd*** | **1** | $J_2$ [MWh] | 301.76 | 301.16 | 300.29 |
| | | $R$ [kWh] | 1778.10 | 1771.74 | 1771.33 |
| | | $F_2$ [MWh] | 248.42 | 248.00 | 247.15 |
| | **2** | $J_2$ [MWh] | 305.69 | 303.84 | 304.04 |
| | | $R$ [kWh] | 1645.12 | 1676.95 | 1741.88 |
| | | $F_2$ [MWh] | 256.34 | 253.53 | 251.79 |
| | **5** | $J_2$ [MWh] | 303.04 | 301.60 | 301.85 |
| | | $R$ [kWh] | 1868.16 | 1858.93 | 1942.44 |
| | | $F_2$ [MWh] | 247.00 | 245.84 | 243.58 |
| | **10** | $J_2$ [MWh] | 304.17 | 302.55 | 302.49 |
| | | $R$ [kWh] | 1684.36 | 1785.52 | 1862.03 |
| | | $F_2$ [MWh] | 253.64 | 248.99 | 246.63 |
| | **20** | $J_2$ [MWh] | 302.65 | 301.00 | 301.26 |
| | | $R$ [kWh] | 1657.41 | 1664.87 | 1790.17 |
| | | $F_2$ [MWh] | 252.93 | 251.06 | 247.55 |

equipped with 32-step LTCs, one single-phase transformer, and it is assumed that a portion of the load is controllable through a demand cap. Feeder current limits are not available for this system; hence, (4b) is not used here.

The proposed distributed computing approach to solve the DOPF model is executed on three physical servers. Two of the servers are based on the Windows 64-bit operating system with two E5-2650 v2 Intel Xeon 2.60 GHz processors and 48 GB RAM with 32 cores, and the third is based on Windows 64-bit operating system with two E5-2670 v3 Intel Xeon 2.30 GHz processors and 128 GB RAM with 48 cores. In all the three servers, the download/upload speed is 95 Mbps.

In order to compare the GA results and solution times, a series of simulations are run, and based on the results, the GA parameters $PS$ = 24, $EN$ = 4, $SV$ = $G$, $CR$ = 0.85, and $MR$ = 0.005 for $\alpha_1$ = $\alpha_2$ = 30 are chosen for this case study. It should be mentioned that even with a different set of GA parameters, the results are close for different GA parameters, which demonstrates the robustness and consistency of the DOPF model. The size of data, which is processed in each worker-node is about 2.46 MB, and the size of the results that should be transferred to the master-node is 8 kB in text format. The size of data processed in the master-node is about 2.56 MB, which is slightly larger than the processed files of each worker-node, and the size of the new set of individuals, which should be transferred to each worker-node is 4 kB in text format. Note that changing the GA parameters such as $PS$ does affect the size of the text files, but not significantly. For the DOPF problem with an MINLP model, the size of data is reasonable, i.e., neither small nor too large; in this study, the DOPF is a 24-hour problem with different control variables, including three LTCs (i.e., 3×24 = 72 control variables) and peak demand caps for 48 load nodes, for a total of 120 control variables. The discrete values of LTC taps can vary between -16 to +16, and the peak demand cap varies between 2.55 and 7.331 kW with a resolution of 1 W.

A comparison of the results with different number of worker-nodes and generations, for $r$ = 1 and $r$ = 2, are presented in Tables I and II. The closeness of $F_r$ results for different generations with different number of worker-nodes shows that 25 generations are sufficient to arrive at an acceptable solution for the DOPF model. When the number of worker-nodes is low (i.e., 1, 2, or 5), all the worker-nodes are installed and run on one server, whereas for 10 worker-nodes, two servers are used (each server has five worker-nodes), and for 20 worker-nodes three servers are used (two servers with five and one server with ten worker-nodes). It is worth mentioning that there is no limit on the number of worker-nodes on the servers. These servers are identified with their IP addresses, and since each server can handle more than one worker-node, these nodes are identified based on their port numbers; hence, each worker-node has identical node addresses (i.e., a mixture of an IP address and a port number) in the SGCM system. There is access to five open ports on two of the servers and ten unblocked ports on the other one. It is possible to administer the machines in order to have more open ports, and therefore, more worker-nodes on one machine. If there are more than one worker-nodes installed in a server, the multi-tasking feature of the platforms (i.e., Windows, Linux, and Mac) is responsible for assigning the cores and required physical memories to the worker-nodes in the SGCM system.

In the objective function (1), the component $R$ varies

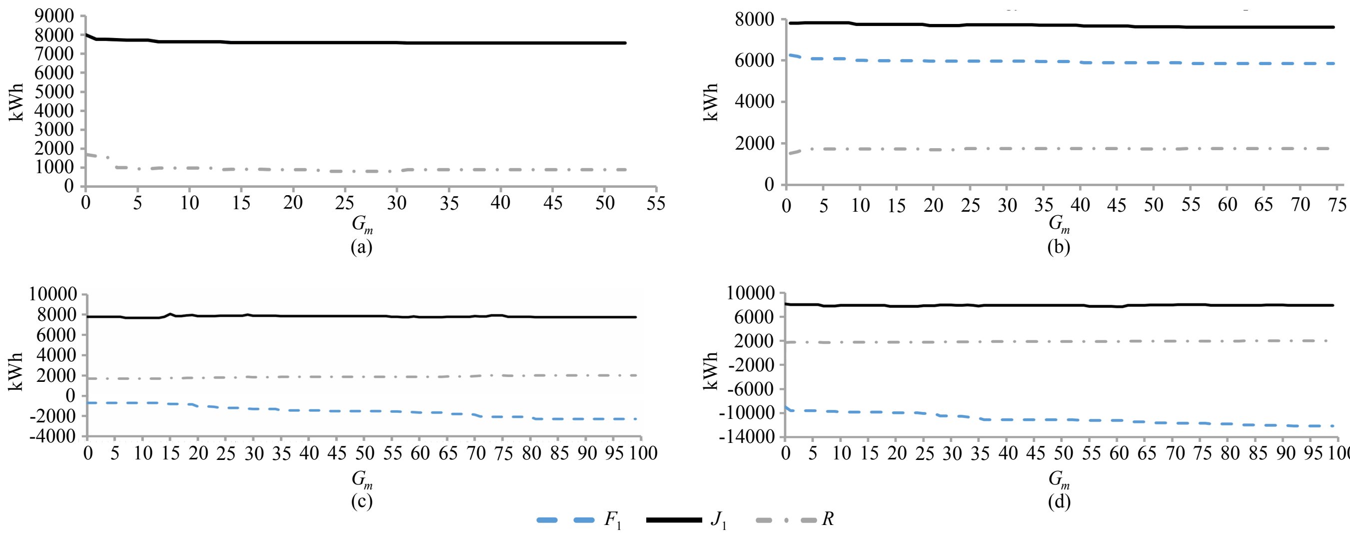



Fig. 5. GA convergence for minimization of $F_1$ with: (a) $\alpha_1$ = 0; (b) $\alpha_1$ = 1; (c) $\alpha_1$ = 5; (d) $\alpha_1$ = 10.

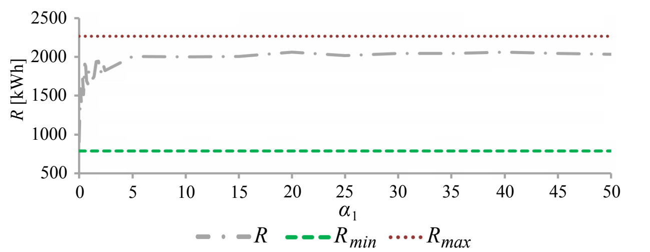



Fig. 6. Effect of $\alpha_1$ on $R$.

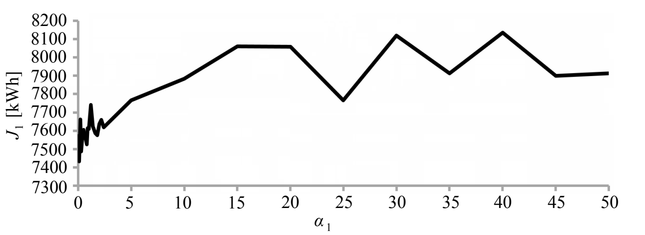



Fig. 7. Effect of $\alpha_1$ on $J_1$.

between $R_{min}$ and $R_{max}$, where:

$$R_{min} = \frac{1}{24} \sum_{h} \sum_{n} \sum_{p} \gamma_{min_{p,n,h}} \tag{5a}$$

$$R_{max} = \frac{1}{24} \sum_{h} \sum_{n} \sum_{p} \gamma_{max_{p,n,h}} \tag{5b}$$

When $\alpha_r$ = 0, the objective is essentially to minimize energy loss ($J_1$) or the energy drawn from the substation ($J_2$); in either case, the optimal solution yields a low value of $\gamma_r$, and the $R$ profile drifts towards $R_{min}$. Figure 5 shows the GA convergence considering energy loss minimization ($r$ = 1) for different values of $\alpha_1$. It is noted that for $\alpha_1$ = 0, the value of $R$ drifts towards $R_{min}$ (Fig. 5(a)), while as $\alpha_1$ is increased, $R$ increases, as shown in Fig. 5(b)-5(d), with the value of $\gamma_{n,p,h}$ increasing, which means that the peak demand caps for controllable smart loads increase. In Fig. 6, observe that for $\alpha_1 > 5$, the $R$ profile is close to $R_{max}$, which indicates a high peak demand cap for controllable smart loads. On the other hand, in Fig. 7, the effect of $\alpha_1$ on $J_1$ is shown, demonstrating that as $\alpha_1$ increases, $J_1$ generally increases; the variations of $J_1$ are due to the GA convergence characteristics, which yield relatively small changes in the $J_1$ values.

Figure 8 shows the convergence of the GA considering the energy drawn from the substation ($r$ = 2) for different values of $\alpha_2$. Note that the $R$ profile converges towards $R_{min}$ for $\alpha_2$ = 0, while as $\alpha_2$ is increased, the peak demand cap increases and effectively allowing more flexibility of the controllable smart loads. In Fig. 9, observe that for $\alpha_2 > 200$, the $R$ profile is close to $R_{max}$, and when $\alpha_2$ is increased beyond 200, the contribution of $R$ to $F_2$ increases, and the one from $J_2$ decreases, which leads to a general increase of $J_2$ value (Fig. 10).

### *B. Impact of the Distributed Computing Platform on Solution Time*

Figures 11 and 12 present the impact of $Nd$ on the solution times with different generation numbers for the GA. Observe that the distributed computing approach has a significant impact on the reduction of solution time, taking about 15 minutes, whereas the centralized approach requires a much larger computation time. Furthermore, these studies also reveal that increasing the number of parallel computers enhances the parallelization of the computation process and decreases the run-time. However, increasing $Nd$ cannot always reduce the run-time; thus, $Nd$ needs to be optimally determined for each case study, which corresponds to a value of $Nd$ = 10 for the present study.

The closeness of the objective function values $F_1$ and $F_2$ for different generation numbers demonstrates the robustness of the model, although the solution time varies. Fast solution times help with shorter feeder dispatch intervals, which is

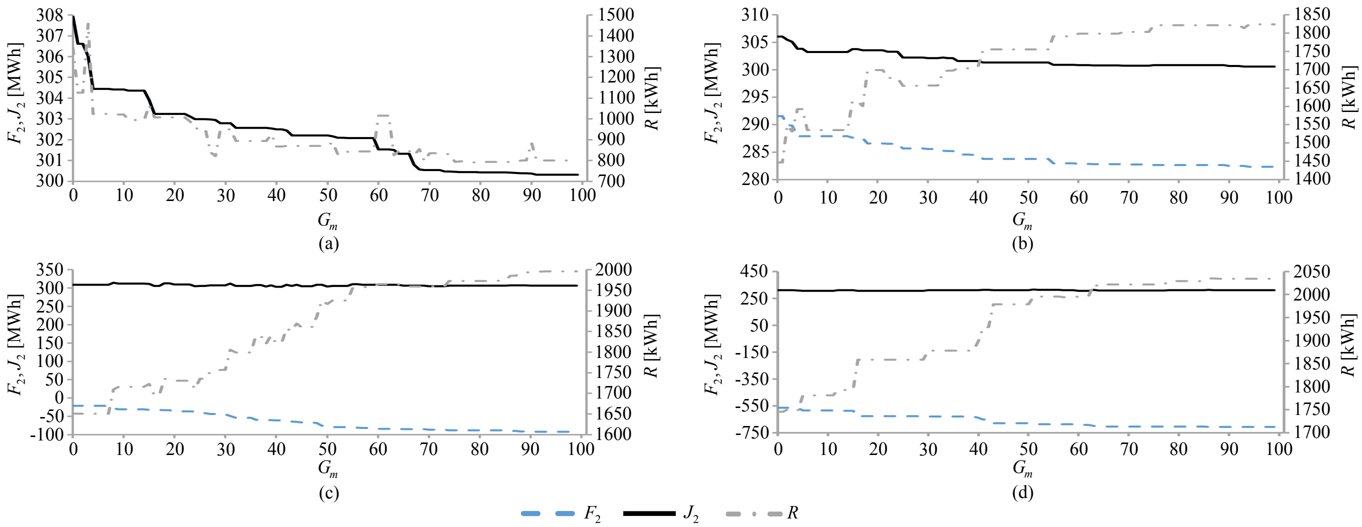



Fig. 8. GA convergence for minimization of $F_2$ with: (a) $\alpha_2$ = 0; (b) $\alpha_2$ = 10; (c) $\alpha_2$ = 200; (d) $\alpha_2$ = 500.

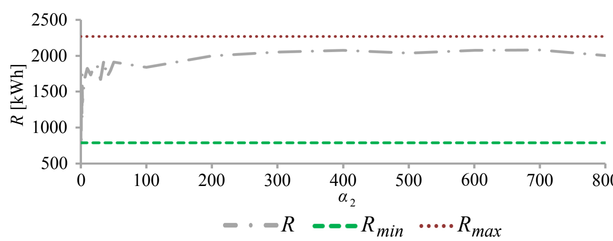



Fig. 9. Effect of $\alpha_2$ on $R$.

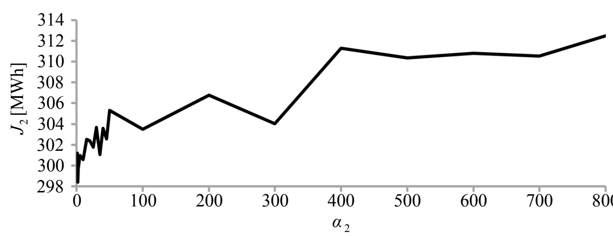



Fig. 10. Effect of $\alpha_2$ on $J_2$.

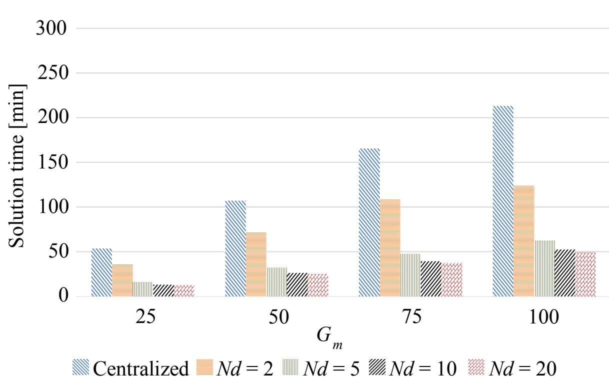



Fig. 11. Computation times for centralized versus proposed distributed computing approach with $F_1$, for different $Nd$ values.

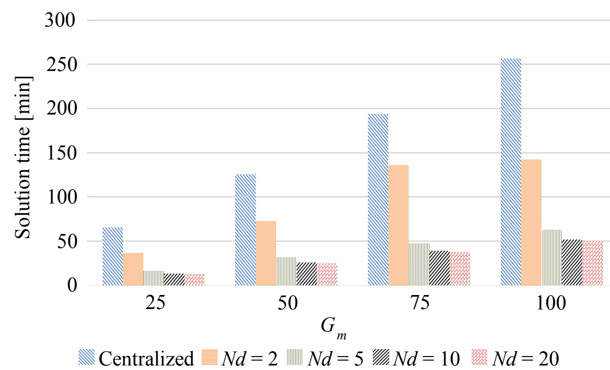



Fig. 12. Computation times for centralized versus proposed distributed computing approach with $F_2$, for different $Nd$ values.

necessary with controllable loads. In distribution systems, real-time intervals for the application being considered would be about 15 min; hence, with the run-times shown for the proposed approach, it should be possible to yield reasonably optimal solutions of the DOPF model for real-time applications.

## V. Conclusions

This paper proposed a novel distributed computing architecture based on an SGCM system for solving a DOPF model using a GA-based solution approach to determine the optimal solutions for controllable distribution feeder devices, in particular LTCs, SCs, and peak demand caps for controllable smart loads. An unbalanced three-phase distribution system, considering uncontrollable loads (i.e., fixed ZIP load models) and controllable smart loads was modeled considering a practical feeder constraints. The decentralized SGCM system allowed to reduce run-times, thus, rendering the DOPF model suitable for real-time applications. The studies carried out on a practical feeder demonstrated that compared to the centralized approach, the proposed distributed computing architecture yielded significantly faster solutions by increasing the number of SGCM worker-nodes, with realistic peak demand caps for controllable smart loads, considering both LDC operator and customer interests.

## References


[1] N. Singh, E. Kliokys, H. Feldmann, R. Kussel, R. Chrustowski, and C. Joborowicz, "Power System Modelling and Analysis in a Mixed Energy Management and Distribution Management System," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1143 – 1149, Aug. 1998.

[2] F. Goodman and M. McGranaghan, "EPRI Research Plan for Advanced Distribution Automation," in *Proc. 2005 IEEE Power Engineering Society General Meeting*, vol. 3, June 2005, p. 2620.

[3] A. Ipakchi and F. Albuyeh, "Grid of the Future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52 – 62, Apr. 2009.

[4] H. Farhangi, "The Path of the Smart Grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18 – 28, Feb. 2010.

[5] K. R. C. Mamandur and R. D. Chenoweth, "Optimal Control of Reactive Power Flow for Improvements in Voltage Profiles and for Real Power Loss Minimization," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 7, pp. 3185 – 3194, July 1981.

[6] Y. Y. Hsu and F. C. Lu, "A Combined Artificial Neural Network-Fuzzy Dynamic Programming Approach to Reactive Power/Voltage Control in a Distribution Substation," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1265 – 1271, Nov. 1998.

[7] J. Kearly, A. Y. Chikhani, R. Hackam, M. M. A. Salama, and V. H. Quintana, "Microprocessor Controlled Reactive Power Compensator for Loss Reduction in Radial Distribution Feeders," *IEEE Trans. Power Del.*, vol. 6, no. 4, pp. 1848 – 1855, Oct. 1991.

[8] R. Baldick and F. F. Wu, "Efficient Integer Optimization Algorithms for Optimal Coordination of Capacitors and Regulators," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 805 – 812, Aug. 1990.

[9] I. Roytelman and V. Ganesan, "Coordinated Local and Centralized Control in Distribution Management Systems," *IEEE Trans. Power Del.*, vol. 15, no. 2, pp. 718 – 724, Apr. 2000.

[10] J. J. Grainger and S. Civanlar, "Volt/Var Control on Distribution Systems with Lateral Branches Using Shunt Capacitors and Voltage Regulators Part I: The Overall Problem," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 11, pp. 3278 – 3283, Nov. 1985.

[11] I. Roytelman, B. K. Wee, and R. L. Lugtu, "Volt/Var Control Algorithm for Modern Distribution Management System," *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1454 – 1460, Aug. 1995.

[12] S. Bruno, S. Lamonaca, G. Rotondo, U. Stecchi, and M. L. Scala, "Unbalanced Three-Phase Optimal Power Flow for Smart Grids," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4504 – 4513, Oct. 2011.

[13] M. B. Liu, C. A. Cañizares, and W. Huang, "Reactive Power and Voltage Control in Distribution Systems with Limited Switching Operations," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 889 – 899, May 2009.

[14] S. Paudyal, C. A. Cañizares, and K. Bhattacharya, "Optimal Operation of Distribution Feeders in Smart Grids," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4495 – 4503, Oct. 2011.

[15] S. Paudyal, C. A. Cañizares, and K. Bhattacharya, "Three-Phase Distribution OPF in Smart Grids: Optimality versus Computational Burden," in *Proc. 2011 2$^{nd}$ IEEE PES Int. Conf. and Exhib. Innovative Smart Grid Technologies (ISGT Europe)*, Dec. 2011, pp. 1 – 7.

[16] H. Fan and Y. Chen, "System and Method for Providing P2P Based Reconfigurable Computing and Structured Data Distribution," Patent, July 2013, WO Patent App. PCT/CA2012/000,912, [online]. Available: http://www.google.com/patents/WO2013102253A1?cl=en.

[17] M. Geidl, G. Koeppel, P. F. Perrod, B. Klockl, G. Andersson, and K. Frohlich, "Energy Hubs for the Future," *IEEE Power Energy Mag.*, vol. 5, no. 1, pp. 24 – 30, Feb. 2007.

[18] S. Pazouki, M. R. Haghifam, and S. Pazouki, "Short Term Economical Scheduling in an Energy Hub by Renewable and Demand Response," in *Proc. 2013 3$^{rd}$ Int. Conf. Electric Power and Energy Conversion Systems (EPECS2013)*, Oct. 2013, pp. 1 – 6.

[19] L. Carradore and F. Bignucolo, "Distributed Multi-Generation and Application of the Energy Hub Concept in Future Networks," in *Proc. 2008 UPEC 43$^{rd}$ Int. Universities Power Engineering Conf.*, Sept. 2008, pp. 1 – 5.

[20] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The Hadoop Distributed File System," in *Proc. 2010 IEEE 26$^{th}$ Symp. Mass Storage Systems and Technologies (MSST)*, May 2010, pp. 1 – 10.

[21] M. Khan, Y. Jin, M. Li, Y. Xiang, and C. Jiang, "Hadoop Performance Modeling for Job Estimation and Resource Provisioning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 2, pp. 441 – 454, Feb. 2016.

[22] Z. Ren, J. Wan, W. Shi, X. Xu, and M. Zhou, "Workload Analysis, Implications, and Optimization on a Production Hadoop Cluster: A Case Study on Taobao," *IEEE Trans. Serv. Comput.*, vol. 7, no. 2, pp. 307 – 321, Apr. 2014.

[23] J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107 – 113, Jan. 2008, [online]. Available: = http://doi.acm.org/10.1145/1327452.1327492.

[24] "Smart Grid Resource Center," *EPRI: Electric Power Research Institute*, [online]. Available: http://smartgrid.epri.com/SimulationTool.aspx.

[25] A. Mosaddegh, C. A. Cañizares, and K. Bhattacharya, "Distributed Computing Approach to Solve Unbalanced Three-Phase DOPFs," in *Proc. 2015 IEEE Electrical Power and Energy Conf. (EPEC2015)*, Oct. 2015, pp. 408 – 413.

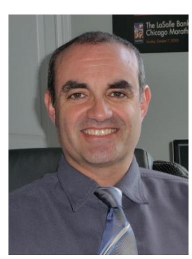
[26] M. C. Bozchalui, S. A. Hashmi, H. Hassen, C. A. Cañizares, and K. Bhattacharya, "Optimal Operation of Residential Energy Hubs in Smart Grids," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1755 – 1766, Dec. 2012.

[27] I. Sharma, "Operation of Distribution Systems with PEVs and Smart Loads," Ph.D. dissertation, Department of Electrical and Computer Engineering, University of Waterloo, 2014.

[28] I. Sharma, K. Bhattacharya, and C. A. Cañizares, "Smart Distribution System Operations with Price-Responsive and Controllable Loads," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 795 – 807, Mar. 2015.

[29] M. Graovac, X. Wang, and R. Iravani, "Integration of Storage in Electrical Distribution Systems and Its Impact on the Depth of Penetration of DG," Clean Energy Technology Centre (CTEC), Tech. Rep. 2009-174, May 2008, [online]. Available: https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/canmetenergy/files/pubs/2009-174_RP-TEC_411-IMPACT_Graovac_Wang_Iravani_e.pdf.


**Abolfazl Mosaddegh** (S'13) received the B.Sc. and M.Sc. degrees in electrical engineering from Iran University of Science and Technology, Tehran, Tehran, Iran in 2008 and 2011, respectively. He is currently working toward the Ph.D. degree in electrical and computer engineering at the University of Waterloo, Waterloo, ON, Canada. His research interests are in distributed computing approaches and distribution system modeling and analysis in the context of smart grids.



**Claudio A. Cañizares** (S'85-M'91-SM'00-F'07) received the Diploma degree in electrical engineering from Escuela Politecnica Nacional, Quito, Ecuador, and the M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin-Madison, WI, USA, in 1984, 1988, and 1991, respectively. He has held various academic and administrative at the University of Waterloo, ON, Canada, since 1993, where he is currently a Full Professor and the Hydro One Endowed Chair in the Department of Electrical and Computer Engineering. His research and consulting work concentrates on stability, operating, control, modeling, simulation, and computational issues in sustainable power and energy systems in the context of competitive markets, smart grids, and microgrids.

Dr. Cañizares was awarded the IEEE Canada Electric Power Medal in 2016, has received several IEEE Power and Energy Society (PES) Working Group Awards, and has held several leadership positions in PES Technical Committees and Subcommittees. He is a Fellow of the Royal Society of Canada and the Canadian Academy of Engineering, and is a Registered Professional Engineer in the province of Ontario.

**Kankar Bhattacharya** (M'95-SM'01) received the Ph.D. degree in electrical engineering from the Indian Institute of Technology, New Delhi, India, in 1993. He was with the Faculty of Indira Gandhi Institute of Development Research, Mumbai, India, from 1993 to 1998, and with the Department of Electric Power Engineering, Chalmers University of Technology, Gothenburg, Sweden, from 1998 to 2002. In 2003, he joined the Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON, Canada, where he is currently a Full Professor. His current research interests include power system economics and operational aspects.

**Hongbing Fan** (S'01-M'04) received the PhD degree in computer science from the University of Victoria in 2003 and joined Wilfrid Laurier University in 2004. Previously, he received the BS degree in mathematics and the PhD degree in operational research and control theory from Shandong University in 1982 and 1990, respectively. He has worked at Shandong University, the Chinese University of Hong Kong, and the University of Lethbridge. His current research interests are interconnection networks, algorithms for CAD of VLSI, network on-chip-based System-on-Chip applications, combinatorial algorithm, and complexity.