A Distributed Algorithm for Economic Dispatch over Time-Varying Directed Networks with Delays

Tao Yang, *Member*, *IEEE*, Jie Lu, *Member*, *IEEE*, Di Wu, *Member*, *IEEE*, Junfeng Wu, Guodong Shi, *Member*, *IEEE*, Ziyang Meng, *Member*, *IEEE*, and Karl H. Johansson, *Fellow*, *IEEE*

***Abstract*—In power system operation, the economic dispatch problem (EDP) aims to minimize the total generation cost while meeting the demand and satisfying generator capacity limits. This paper proposes an algorithm based on the gradient push-sum method to solve the EDP in a distributed manner over communication networks potentially with time-varying topologies and communication delays. This paper shows that the proposed algorithm is guaranteed to solve the EDP if the time-varying directed communication network is uniformly jointly strongly connected. Moreover, the proposed algorithm is also able to handle arbitrarily large but bounded time-varying delays on communication links. Numerical simulations are used to illustrate and validate the proposed algorithm.**

***Index Terms*—Distributed algorithm, economic dispatch, gradient push-sum method, time-varying delays, time-varying networks.**

I. INTRODUCTION

THE economic dispatch problem (EDP) is one of important problems in power system operation. It is essentially an optimization problem where the objective is to minimize the total generation cost while meeting total demand and satisfying individual generator output limits. There exist many centralized approaches for solving the EDP, such as the lambda-iteration method and the gradient search method [1]. The centralized methods require a single control center that accesses the entire network's information, and therefore may be subject to performance limitations, such as high communication requirement and cost, substantial computational burden, and limited flexibility and scalability, and disrespect of privacy. It is thus desirable to develop distributed approaches to overcome these limitations and accommodate various resources in the future smart grid.

During the past few years, due to the rising of distributed control and multi-agent systems research [2]–[6], various distributed algorithms have been developed for power system applications [7]–[10]. As for the EDP, various consensus-based distributed algorithms have been proposed by choosing the generation incremental cost as the consensus variable. In these algorithms, each agent maintains a few variables and updates them through the information exchange with its neighboring agents. For instance, the authors of [11] propose a leader-follower consensus-based algorithm where the leader collects the mismatch between demand and generation, and then leads the updates of marginal cost in the system. To avoid the requirement of a leader, a two-level consensus-based algorithm is proposed in [12], where the upper level is the consensus and gradient algorithm, and the lower level executes the classical consensus by choosing the local mismatch as the consensus variable. In the algorithm proposed in [13], in addition to consensus part, an innovation term is introduced to ensure the balance between system generation and demand. All these three algorithms are only applicable to undirected communication networks, i.e., the information must be exchanged bidirectionally. Because directed communication networks cost less than their undirected counterparts [2], it is desirable to develop control and coordination algorithms that only require directed information flow. The capability of utilizing low-cost communication networks is favorable in the future smart grid. Realizing such a need, distributed algorithms have been proposed in [14], [15] to solve the EDP over both undirected and directed communication networks. In [14], the authors propose a ratio consensus based algorithm which relies on two linear iterations. The one in [15] estimates the mismatch with all the agents being participated. The authors of [16] propose minimum-time consensus-based algorithm to solve the EDP in a minimum number of time steps. As for generation cost functions, most of existing studies assume

Manuscript received February 6, 2016; revised May 11, 2016 and July 23, 2016; accepted October 3, 2016. This work was supported in part by the Laboratory Directed Research and Development (LDRD) program at the Pacific Northwest National Laboratory, the Knut and Alice Wallenberg Foundation, the Swedish Research Council, National Natural Science Foundation of China under Grant 61503249, Natural Science Foundation of Shanghai under Grant 16ZR1422500, and Shanghai Pujiang Program under Grant 16PJ1406400.

T. Yang is with the Department of Electrical Engineering, University of North Texas, Denton, TX, 76203 (e-mail: Tao.Yang@unt.edu).

J. Lu is with the School of Information Science and Technology, ShanghaiTech University, Shanghai 200031, China (e-mail: lujie@shanghaitech.edu.cn).

D. Wu is with the Pacific Northwest National Laboratory, Richland, WA 99352 (corresponding author; phone: +1 509-375-3975; fax: +1 509-372-4353; e-mail: di.wu@pnnl.gov).

J. Wu and K. H. Johansson are with the ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden (e-mail: {junfengw, kallej}@kth.se).

G. Shi is with the Research School of Engineering, The Australian National University, Canberra, ACT 0200 Australia (e-mail: guodong.shi@anu.edu.au).

Z. Meng is with the Department of Precision Instrument and the State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China (e-mail: ziyangmeng@mail.tsinghua.edu.cn).

quadratic functions, whereas [17] considers general convex functions.

One common assumption on communication networks in the literature is that the communication links are time-invariant and are not subject to time delays. However, in practice, communication network topology may vary due to unexpected loss of communication links. In addition, time delays are ubiquitous in communication networks [18]. Therefore, it is desirable to investigate the potential impacts of imperfect communications on the existing distributed EDP algorithms, and develop algorithms that are robust to imperfect communications in the future smart grid. In [19], the authors study impacts of communication delays on the two-level consensus algorithm, and find that the algorithm may fail to converge due to time delays. In [20], the authors investigate the impacts of communication time delays on the algorithm proposed in [15]. Several potential negative impacts of time delays have been found, such as slower convergence rate, convergence to incorrect value, and divergence. The authors of [21] propose a nonnegative-surplus based distributed algorithm to solve the EDP over time-varying directed communication networks but without time delays.

The existing literature is inadequate to solve the EDP in imperfect communication networks that are subject to time-varying topologies and time delays. In order to better handle these practical restrictions, this paper proposes a distributed algorithm based on the gradient push-sum method. Compared with existing EDP studies, the main contributions of this paper are summarized as follows.

1) We propose a distributed algorithm to solve the EDP over time-varying directed communication networks that are uniformly jointly strongly connected. This is a mild condition on the connectivity of communication topologies, since the network can be disconnected at any time instance as long as the joint graph over a period of time is strongly connected. Therefore, the requirement on network topologies is more general compared to the fixed strongly connected topologies considered in the existing studies [11]–[17], [22].
2) While time-varying communication networks are also considered in [21], the proposed algorithm in this paper can handle the EDP with general convex cost functions, which are more general compared to the quadratic cost functions considered in [21].
3) The proposed algorithm can also handle arbitrarily large but bounded time-varying delays in addition to time-varying directed topologies, which is a distinguishing feature compared with many existing studies, such as [11], [12], [15], [16], [21]. To the best of our knowledge, this is the first algorithm that is capable to solve the EDP over time-varying directed communication networks with communication delays.

The remainder of the paper is organized as follows: In Section II, some preliminaries on graph theory and notations are introduced. Section III presents the EDP formulation and the centralized Lagrangian-based approach. In Section IV, a distributed algorithm based on the gradient push-sum method is proposed to solve the EDP over communication networks with imperfections, such as time-varying topologies and time delays. Case studies are presented in Section V to illustrate and validate the proposed algorithm. Finally, concluding remarks are offered in Section VI.

II. Preliminaries and Notations

This section first presents some background on graph theory [23], which is needed to describe the communication network. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed graph (digraph) with the set of nodes (agents) $\mathcal{V} = \{1, \ldots, N\}$ and the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. A directed edge from node i to node j is denoted by $(i, j) \in \mathcal{E}$. For notational simplification, we assume that the digraph does not have any self loop, i.e., $(i, i) \notin \mathcal{E}$ for all $i \in \mathcal{V}$ although each node i has an access to its own information. A directed *path* from node i_1 to node i_k is a sequence of nodes $\{i_1, \ldots, i_k\}$ such that $(i_j, i_{j+1}) \in \mathcal{E}$ for $j = 1, \ldots, k-1$. If there exists a directed path from node i to node j, then node j is said to be reachable from node i. A digraph $\mathcal{G}$ is said to be *strongly connected* if every node is reachable from every other node.

In this paper, an agent is assigned to each bus in the power system. The topology of communication network could be different from the physical network, and is modeled as a time-varying directed graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$, where the edge set changes over time due to unexpected loss of communication links. All agents that can transmit information to node i directly at time t are said to be its in-neighbors and belong to the set $\mathcal{N}_i^{\text{in}}(t) = \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}(t)\}$. The nodes which receive information from agent i at time t belong to the set of its out-neighbors, denoted by $\mathcal{N}_i^{\text{out}}(t) = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}(t)\}$. The cardinality of $\mathcal{N}_i^{\text{out}}(t)$ is called its out-degree at time t and is denoted by $d_i(t) = |\mathcal{N}_i^{\text{out}}(t)|$. The joint graph of $\mathcal{G}(t)$ in the time interval $[t_1, t_2)$ with $t_1 < t_2 \leq \infty$ is denoted as $\mathcal{G}([t_1, t_2)) = \cup_{t\in[t_1,t_2)}\mathcal{G}(t) = (\mathcal{V}, \cup_{t\in[t_1,t_2)}\mathcal{E}(t))$. A time-varying directed network $\mathcal{G}(t)$ is said to be *uniformly jointly strongly connected* if there exists a constant $T > 0$ such that $\mathcal{G}([t_0, t_0 + T))$ is strongly connected for any $t_0 \geq 0$.

Notations: In this paper, variables in boldface represent vectors or matrices. For a matrix $\mathbf{A}$, we use $\mathbf{A}_{ij}$ or $[\mathbf{A}]_{ij}$ to denote its (i, j)-th entry and $\mathbf{A}^{\mathsf{T}}$ to denote its transpose. A matrix is nonnegative if all its entries are equal to or greater than zero. A vector is a stochastic vector if all entries are nonnegative and sum up to 1. For a vector $\mathbf{x}$, we use $\mathbf{x}_i$ to denote its i-th entry. $\mathbf{0}$ and $\mathbf{1}$ denote the column vectors with all entries being 0 and 1, respectively. $\mathbf{0}$ and $\mathbf{I}$ denote the matrix with all entries being 0 and the identity matrix, respectively. The set of real (integer) numbers is denoted by $\mathbb{R}$ ($\mathbb{Z}$) and the set of nonnegative real (integer) numbers is denoted by $\mathbb{R}_+$ ($\mathbb{Z}_+$).

III. Problem Formulation and Lagrangian-based Approach

This section first presents the mathematical formulation and then the centralized Lagrangian-based approach for the EDP.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics

A. Formulation of EDP

The goal of EDP is to minimize the total generation cost while meeting total demand and satisfying individual generator output limits, as formulated in (1):

$$\min_{x_i} \quad \sum_{i=1}^{N} C_i(x_i) \tag{1a}$$

$$\text{subject to} \quad \sum_{i=1}^{N} x_i = D, \tag{1b}$$

$$x_i \in X_i := [x_i^{\min}, x_i^{\max}],\ i = 1, \ldots, N, \tag{1c}$$

where N is the number of generators, x_i is the power generation of generator i, $C_i(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is the cost function of generator i, $x_i^{\min}$ and $x_i^{\max}$ are respectively the lower and upper bounds of the power generation of generator i, and D is the total demand satisfying $\sum_{i=1}^{N} x_i^{\min} \le D \le \sum_{i=1}^{N} x_i^{\max}$ in order to ensure the feasibility of problem (1).

Compared to most studies [11]–[16], [21], [24] in the EDP literature where cost functions are assumed to be quadratic, this paper considers general convex cost functions that satisfy Assumption 1.

Assumption 1: For each $i \in \{1, \ldots, N\}$, the cost function $C_i(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly convex and continuously differentiable.

B. Centralized Lagrangian-based Approach

Since i) each cost function $C_i(\cdot)$ is convex, ii) the constraint (1b) is affine, and iii) the set $X_1 \times \cdots \times X_N$ is a polyhedral set, if we dualize problem (1) with respect to the constraint (1b), there is zero duality gap. Moreover, the dual optimal set is nonempty [25]. Consequently, solutions of the EDP can be obtained by solving its dual problem.

For convenience, let $\mathbf{x} = [x_1, \ldots, x_N]^{\mathrm{T}} \in \mathbb{R}_+^N$. Then, define the Lagrangian function

$$\mathcal{L}(\mathbf{x}, \lambda) = \sum_{i=1}^{N} C_i(x_i) - \lambda \left(\sum_{i=1}^{N} x_i - D \right).$$

The corresponding Lagrange dual problem is

$$\max_{\lambda \in \mathbb{R}_+} \sum_{i=1}^{N} \Psi_i(\lambda) + \lambda D, \tag{2}$$

where

$$\Psi_i(\lambda) = \min_{x_i \in X_i} C_i(x_i) - \lambda x_i. \tag{3}$$

Under Assumption 1, for any given $\lambda \in \mathbb{R}_+$, the right-hand side of (3) has a unique minimizer given by

$$x_i(\lambda) = \min\{\max\{\nabla C_i^{-1}(\lambda), x_i^{\min}\}, x_i^{\max}\}, \tag{4}$$

where ∇C_i^{-1} denotes the inverse function of ∇C_i, which exists over $[\nabla C_i(x_i^{\min}), \nabla C_i(x_i^{\max})]$ since ∇C_i is continuous and strictly increasing due to Assumption 1. Furthermore, there is at least one optimal solution to dual problem (2), and the unique optimal solution of the primal EDP is given by

$$x_i^* = x_i(\lambda^*), \quad \forall i = 1, 2, \ldots, N, \tag{5}$$

where λ^* is any dual optimal solution.

For any given $\lambda \in \mathbb{R}_+$, because of the uniqueness of $x_i(\lambda)$, the dual function $\sum_{i=1}^{N} \Psi_i(\lambda) + \lambda D$ is differentiable at λ and its gradient is given by $-(\sum_{i=1}^{N} x_i(\lambda) - D)$ [26]. We can then apply the gradient method to solve the dual problem in (2):

$$\lambda(t+1) = \lambda(t) - \gamma(t) \left(\sum_{i=1}^{N} x_i(\lambda(t)) - D \right), \tag{6}$$

where $\lambda(0) \in \mathbb{R}$ can be arbitrarily assigned and $\gamma(t)$ is the step-size at time step t. When designing a distributed algorithm based on (6), the main challenge is how to obtain the global quantity $\sum_{i=1}^{N} x_i(\lambda(t)) - D$ in a distributed manner. In this paper, we will propose a distributed algorithm to avoid the need of the global quantity.

IV. Main Results

This section proposes an algorithm that is capable to solve the EDP in a distributed fashion over time-varying communication networks potentially with arbitrarily large but bounded time delays. In Section IV-A, the dual problem in (2) is first converted to an agent-based distributed convex optimization problem. Then a distributed algorithm is proposed for the EDP based on the gradient push-sum method [27]. Section IV-B shows that the proposed algorithm is able to solve the EDP over time-varying directed communication networks. Finally, Section IV-C shows that the proposed algorithm is also robust to communication time delays.

A. Distributed Gradient Push-Sum Algorithm

The dual problem in (2) can be converted into

$$\max_{\lambda \in \mathbb{R}} \sum_{i=1}^{N} \Phi_i(\lambda), \tag{7}$$

where

$$\Phi_i(\lambda) = \min_{x_i \in X_i} C_i(x_i) - \lambda(x_i - D_i), \tag{8}$$

and D_i is a virtual local demand at each bus such that $\sum_{i=1}^{N} D_i = D$. The gradient of $\Phi_i(\lambda)$ is

$$\nabla \Phi_i(\lambda) = -\left(x_i(\lambda) - D_i\right). \tag{9}$$

Motivated by the gradient push-sum method [27], a distributed algorithm is proposed in Algorithm 1 to solve the equivalent dual problem (7). In the proposed algorithm, each agent i maintains scalar variables $v_i(t)$, $w_i(t)$, $y_i(t)$, $\lambda_i(t)$, $x_i(t)$, where $x_i(t)$ and $\lambda_i(t)$ are estimations of the primal and dual optimal solution, respectively. At each time step t, each agent $i \in \mathcal{V}$ updates its variables according to (10).

$$w_i(t+1) = \frac{v_i(t)}{d_i(t)+1} + \sum_{j \in \mathcal{N}_i^{\text{in}}(t)} \frac{v_j(t)}{d_j(t)+1}, \tag{10a}$$

$$y_i(t+1) = \frac{y_i(t)}{d_i(t)+1} + \sum_{j \in \mathcal{N}_i^{\text{in}}(t)} \frac{y_j(t)}{d_j(t)+1}, \tag{10b}$$

$$\lambda_i(t+1) = \frac{w_i(t+1)}{y_i(t+1)}, \tag{10c}$$

$$x_i(t+1) = \min\{\max\{\nabla C_i^{-1}(\lambda_i(t+1)), x_i^{\min}\}, x_i^{\max}\}, \tag{10d}$$

$$v_i(t+1) = w_i(t+1) - \gamma(t+1)(x_i(t+1) - D_i). \tag{10e}$$

Algorithm 1 Distributed algorithm for the EDP

Input: The time-varying graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$, the step-size $\gamma(t)$, an arbitrarily assigned $v_i(0)$, and $y_i(0) = 1$ for all $i \in \mathcal{V}$.
Output: The optimal incremental cost λ^* and the optimal generation x_i^*.
repeat
  for $i = 1$ **to** N **do**
    Run the update rule (10).
    Return $\lambda_i(t+1)$ and $x_i(t+1)$.
  end for
  Update t as $t := t + 1$.
until $|\lambda_i(t) - \lambda_i(t-1)| < \epsilon_1$ **and** $\max_{i,j \in \mathcal{V}} |\lambda_i(t) - \lambda_j(t)| < \epsilon_2$.

The step-size $\gamma(t+1)$ satisfies the following decay conditions:

$$\sum_{t=1}^{\infty} \gamma(t) = \infty, \quad \sum_{t=1}^{\infty} \gamma^2(t) < \infty,$$
$$\gamma(t) \leq \gamma(s) \text{ for all } t > s \geq 1. \tag{11}$$

One typical selection is $\gamma(t) = \frac{a}{t+b}$, where $a > 0$ and $b \geq 0$. In this algorithm, each agent i needs to know its out-degree $d_i(t)$ and sends the quantities $\frac{v_i(t)}{d_i(t)+1}$ and $\frac{y_i(t)}{d_i(t)+1}$ to all the agents j in its out-neighbors set. In initialization, $v_i(0)$ is assigned with an arbitrary value and $y_i(0) = 1$ for all $i \in \mathcal{V}$.

According to (9), $-(x_i(t+1) - D_i)$ in (10e) is the gradient of the function $\Phi_i(\lambda)$ at $\lambda = \lambda_i(t+1)$. Without (10d) and the gradient term in (10e), the algorithm would be reduced to a particular version of push-sum algorithm [28], or ratio consensus algorithm [29], [30] for computing the average of initial values in directed graphs. In this case, all $\lambda_i(t+1)$ converge to a common value. The inclusion of the gradient term in the update of $v_i(t+1)$ is to ensure that all $\lambda_i(t+1)$ converge to the optimal incremental cost λ^*.

Remark 1: Note that at each step t, agent i runs the update rule (10):

- $d_i^{\text{in}}(t) + 1$ multiplications are performed in each of (10a) and (10b), where $d_i^{\text{in}}(t)$ is the in-degree of agent i at time t. Note that $d_i^{\text{in}}(t) \leq N - 1$, where N is the size of the communication network.
- Only one division is performed in (10c).
- For quadratic cost functions $C_i(x_i) = a_i x_i^2 + b_i x_i + c_i$ where $a_i > 0$, b_i and c_i are cost parameters, the update equation (10d) has a closed form expression
$$x_i(t+1) = \min\{\max\{\frac{\lambda_i(t+1) - b_i}{2a_i}, x_i^{\min}\}, x_i^{\max}\},$$
therefore, only one multiplication is performed in (10d). For general convex cost functions, the update equation (10d) may not have a closed form expression. Nevertheless, its numerical solution can be obtained in a finite number of time steps by using the bisection method due to the continuity and strict monotonicity of $\nabla C_i^{-1}(\cdot)$.
- Only one multiplication is performed in (10e).

Moreover, it requires a finite number of time steps for the algorithm to converge based on the stopping criteria. Therefore, the computational complexity of the proposed algorithm is $\mathcal{O}(N)$.

B. Robustness to Time-Varying Communication Networks

In this subsection, we will show that the proposed distributed Algorithm 1 is capable to solve the EDP over time-varying directed communication networks which satisfy Assumption 2, as stated in Theorem 1.

Assumption 2: The time-varying directed communication network $\mathcal{G}(t)$ is uniformly jointly strongly connected, i.e., the jointly communication network $\mathcal{G}([t_0, t_0 + T))$ is strongly connected for any $t_0 \geq 0$ with some constant $T > 0$.

Theorem 1: Under Assumptions 1 and 2, distributed Algorithm 1 with the step-size $\gamma(t)$ satisfying conditions in (11) solves the EDP, i.e., $\lambda_i(t) \to \lambda^*$, and $x_i(t) \to x_i^*$ as $t \to \infty$ for all $i \in \mathcal{V}$.

Proof: Note that the equivalent dual problem (7) has the same form as the optimization problem considered in [27]. The only difference is that the dual problem is a maximization problem while the problem in [27] is a minimization problem. In order to apply [27, Theorem 1] to show Theorem 1, we need to verify that all the conditions are satisfied.

- The condition (a) is that the network is uniformly jointly strongly connected, which is satisfied in our case as assumed in Assumption 2.
- The condition (b) is that each function in the minimization problem is convex and the optimal set is nonempty. This is also satisfied in our case since each function $\Phi_i(\lambda)$ in the maximization problem (7) is concave and the optimal set is nonempty, which is guaranteed by Assumption 1.
- The condition (c) is that the (sub)gradient of each function in the problem is uniformly bounded. This is indeed satisfied in our case since it follows from (9) that the gradient of each function $\Phi_i(\lambda)$ is uniformly bounded, i.e.,

$$\Big|\nabla \Phi_i(\lambda_i(t+1))\Big| = \Big|-(x_i(\lambda_i(t+1)) - D_i)\Big| \leq \max_{i \in \mathcal{V}} x_i^{\max} + \max_{i \in \mathcal{V}} D_i. \tag{12}$$

Therefore, all the conditions are satisfied and the result follows. ■

Remark 2: The key result used in the proof of [27, Theorem 1] is [27, Lemma 3]. It shows that the matrix $\mathbf{A}(t)$ which captures the weights used in the update equations (10a) and (10b), defined as

$$\mathbf{A}_{ij}(t) = \begin{cases} \frac{1}{d_j(t)+1}, & \text{if } j \in \mathcal{N}_i^{\text{in}}(t) \cup \{i\}, \\ 0 & \text{otherwise}, \end{cases} \tag{13}$$

has the special property under Assumption 2, that is, for each $s \in \mathbb{Z}_+$, there is a stochastic vector $\boldsymbol{\phi}(s)$ such that for all i, j and $t \geq s$,

$$\Big|[\mathbf{A}^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t-1) \cdots \mathbf{A}^{\mathrm{T}}(s+1)\mathbf{A}^{\mathrm{T}}(s)]_{ij} - \boldsymbol{\phi}_j(s)\Big| \leq \alpha \beta^{t-s},$$

with

$$\alpha = 2, \quad \beta = \Big(1 - \frac{1}{N^{NT}}\Big)^{\frac{1}{NT}}, \tag{14}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics

where T is the bound on the intercommunication interval given in Assumption 2.

Remark 3: In the proof of Theorem 1, we have built upon the recent developed result in [27]. However, our work is substantially different.

- We consider the EDP, which is a constrained optimization problem as shown in (1). We have converted the problem to an equivalent dual problem (7), which has the same form as the unconstrained problem considered in [27].
- In order to apply the result in [27], there are substantial details need to be verified as shown in Theorem 1.
- We will show in Theorem 2 that the proposed algorithm is also robust to arbitrarily large but bounded time-varying communication delays, which are not considered in [27].

Remark 4: Theorem 1 shows that the proposed distributed Algorithm 1 solves the EDP over time-varying directed communication networks that are uniformly jointly strongly connected. This is a mild condition on the connectivity of communication topologies, since the network can be disconnected at any time instance as long as the joint graph over a period of time is strongly connected. Therefore, the requirement on network topologies is more general than the existing studies [11]–[17], [22], where the fixed strongly connected topologies are required.

C. Robustness to Communication Time Delays

This subsection studies the impact of time delays on the proposed Algorithm 1. We first model time delays in the directed communication networks. In particular, the communication link (j, i) at time step t undergoes a priori unknown delay $\tau_{ji}(t) \in \mathbb{Z}_+$. We impose the following assumption on time-varying delays.

Assumption 3: The time-varying delays are uniformly bounded at all times, i.e., $0 \le \tau_{ji}(t) \le \bar{\tau}$ for all $t \in \mathbb{Z}_+$ with some finite $\bar{\tau} \in \mathbb{Z}_+$. Moreover, each agent has access to its own value without any time delay, i.e., $\tau_{ii}(t) = 0$ for all $i \in \mathcal{V}$ and for all time steps $t \in \mathbb{Z}_+$.

Note that in the proposed algorithm (10), only the updates of $w_i(t+1)$ and $y_j(t+1)$ rely on communications among the agents, while the updates of $\lambda_i(t+1)$, $x_i(t+1)$, and $v_i(t+1)$ are executed locally without the need of further communications. When communications are subject to time delays, each agent i updates the values of $w_i(t+1)$ and $y_i(t+1)$ by combining its own values and the delayed information received from its in-neighbors. More specifically, under time-delays, executing the update rule (10) results in:

$$w_i(t+1) = \frac{v_i(t)}{d_i(t)+1} + \sum_{j \in \mathcal{N}_i^{\text{in}}(t)} \frac{v_j\left(t-\tau_{ji}(t)\right)}{d_j(t)+1}, \tag{15a}$$

$$y_i(t+1) = \frac{y_i(t)}{d_i(t)+1} + \sum_{j \in \mathcal{N}_i^{\text{in}}(t)} \frac{y_j\left(t-\tau_{ji}(t)\right)}{d_j(t)+1}, \tag{15b}$$

$$\lambda_i(t+1) = \frac{w_i(t+1)}{y_i(t+1)}, \tag{15c}$$

$$x_i(t+1) = \min\{\max\{\nabla C_i^{-1}(\lambda_i(t+1)), x_i^{\min}\}, x_i^{\max}\}, \tag{15d}$$

$$v_i(t+1) = w_i(t+1) - \gamma(t+1)(x_i(t+1) - D_i). \tag{15e}$$

The following theorem shows that the proposed Algorithm 1 is able to solve the EDP over time-varying directed communication networks even when communication links are subject to arbitrarily large but bounded delays.

Theorem 2: Under Assumptions 1, 2, and 3, distributed Algorithm 1 with the step-size $\gamma(t)$ satisfying conditions in (11) solves the EDP even when communication links are subject to arbitrarily large but bounded delays, i.e., $\lambda_i(t) \to \lambda^*$, and $x_i(t) \to x_i^*$ as $t \to \infty$ for all $i \in \mathcal{V}$.

Proof: We first note that under our modeling on communication time delays, executing the update rule (10) results in (15). The proof is based on an augmented digraph representation which allows us to reduce the original system with bounded delays (15) to a system without delays. More specifically, for each agent i in the original graph, we introduce $\bar{\tau}$ virtual agents $i^{(1)}$, $i^{(2)}$, …, $i^{(\bar{\tau})}$, where at each time step t, virtual agent $i^{(r)}$ holds information that is destined to arrive to node i in r steps. Since time delays are bounded by $\bar{\tau}$, there are in total $N(\bar{\tau}+1)$ agents in the augmented digraph. In the augmented digraph, we enumerate the agents in the original digraph first and then the virtual agents. Moreover, the virtual agents are indexed so that the first N agents model the delay of 1 time step, the next N agents model the delay of 2 time steps, and so on.

We now describe how these agents communicate in the augmented digraph. In particular, at time step t, for each edge (j, i) in the original network, that edge also exists in the augmented digraph along with edges $(j, i^{(1)})$, $(j, i^{(2)})$, …, $(j, i^{(\bar{\tau})})$, and edges $(i^{(1)}, i)$, $(i^{(2)}, i^{(1)})$, …, $(i^{(\bar{\tau})}, i^{(\bar{\tau}-1)})$.

For each virtual agent $i^{(r)}$, where $r = 1, \ldots, \bar{\tau}$, we associate it with the states $v_i^{(r)}$, $y_i^{(r)}$ and $w_i^{(r)}$. We then define $\mathbf{v}^{(r)}(t)$, $\mathbf{y}^{(r)}(t)$ and $\mathbf{w}^{(r)}(t)$ as the column stack vectors of $v_i^{(r)}$, $y_i^{(r)}$ and $w_i^{(r)}$ where $i \in \mathcal{V}$, respectively. For example, $\mathbf{v}^{(r)}(t) = [v_1^{(r)}(t), \ldots, v_N^{(r)}(t)]^{\mathrm{T}}$. Finally, we define $\tilde{\mathbf{v}}(t) = [\mathbf{v}(t)^{\mathrm{T}}, \mathbf{v}^{(1)}(t)^{\mathrm{T}}, \ldots, \mathbf{v}^{(\bar{\tau})}(t)^{\mathrm{T}}]^{\mathrm{T}}$. Similarly, we define $\tilde{\mathbf{w}}(t)$ and $\tilde{\mathbf{y}}(t)$. For the agent in the original network $i \in \mathcal{V}$, the initial states are given by $v_i(0)$ and $y_i(0) = 1$, while for all the virtual agents, the initial states are given by $\mathbf{v}^{(r)}(0) = \mathbf{0}$ and $\mathbf{y}^{(r)}(0) = \mathbf{0}$ for all $r = 1, 2, \ldots \bar{\tau}$.

In the augmented digraph, the original system with delays (15) can be rewritten in a more compact form without delays as

$$\tilde{\mathbf{w}}(t+1) = \tilde{\mathbf{A}}(t)\tilde{\mathbf{v}}(t), \tag{16a}$$

$$\tilde{\mathbf{y}}(t+1) = \tilde{\mathbf{A}}(t)\tilde{\mathbf{y}}(t), \tag{16b}$$

$$\lambda_i(t+1) = \frac{w_i(t+1)}{y_i(t+1)}, \quad i \in \mathcal{V}, \tag{16c}$$

$$x_i(t+1) = \min\{\max\{\nabla C_i^{-1}(\lambda_i(t+1)), x_i^{\min}\}, x_i^{\max}\}, \tag{16d}$$

$$\tilde{\mathbf{v}}(t+1) = \tilde{\mathbf{w}}(t+1) - \gamma(t+1)[\mathbf{x}^{\mathrm{T}}(t) - \tilde{\mathbf{D}}^{\mathrm{T}}, \mathbf{0}_{N\bar{\tau}}^{\mathrm{T}}]^{\mathrm{T}}, \tag{16e}$$

where $\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^{\mathrm{T}}$, $\tilde{\mathbf{D}} = [D_1, \ldots, D_N]^{\mathrm{T}}$, and

$$\tilde{\mathbf{A}}(t) = \begin{bmatrix} \mathbf{A}^{(0)}(t) & \mathbf{I}_{N\times N} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{A}^{(1)}(t) & \mathbf{0} & \mathbf{I}_{N\times N} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}^{(\bar{\tau}-1)}(t) & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{I}_{N\times N} \\ \mathbf{A}^{(\bar{\tau})}(t) & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \end{bmatrix}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics

Note that in (16), the update equations (16c) and (16d) are only for the agents in the original graph, which are identical to the update equations (10c) and (10d), respectively, for the case without communication delays. Also note that $\tilde{y}_i(t+1) > 0$ for all $t \in \mathbb{Z}_+$ and for all original agents $i \in \mathcal{V}$. This is not necessarily true for the virtual agents. Therefore, $\lambda_i(t+1)$ for $i \in \mathcal{V}$ in (16c) is a finite quality for all $t \in \mathbb{Z}_+$. $x_i(t+1)$ for $i \in \mathcal{V}$ is then updated according to (16d). These resulting values are finally used in the update equations (16e) for the agents in the original graph, while for the virtual agents, update equations (16e) reduces to $\tilde{v}_i(t+1) = \tilde{w}_i(t+1)$.

Herein, $\mathbf{A}^{(0)}(t)$, $\mathbf{A}^{(1)}(t)$, ..., $\mathbf{A}^{(\bar{\tau})}(t)$ are appropriately defined nonnegative matrices that depend on the communication link delays which are experienced by messages sent at time step t. Specifically, $\mathbf{A}^{(r)}(t)$ for $r = 0, \ldots, \bar{\tau}$ is a matrix associated only with the communication links for which the message was delayed by r steps at time step t, and satisfies

$$\mathbf{A}_{ij}^{(r)}(t) = \begin{cases} \mathbf{A}_{ij}(t), & \text{if } \tau_{ij}(t) = r, \ \ (i,j) \in \mathcal{E}(t), \\ 0, & \text{otherwise}, \end{cases}$$

with $\mathbf{A}_{ij}(t)$ given by (13).

Notice that at time step t, for each edge (j, i), only one of $\mathbf{A}_{ij}^{(0)}(t)$, ..., $\mathbf{A}_{ij}^{(\bar{\tau})}(t)$ is nonzero and is equal to $\mathbf{A}_{ij}(t)$. Thus, the special structure of the matrix $\tilde{\mathbf{A}}(t)$ allows us to analyze their products. More specially, it follows from [31, Lemma 5] that $\tilde{\mathbf{A}}(t)$ has the special property under Assumptions 2 and 3, that is, for each $s \in \mathbb{Z}_+$, there is a stochastic vector $\tilde{\boldsymbol{\phi}}(s)$ such that for all $i, j \in \{1, \ldots, N(\bar{\tau}+1)\}$ and $t \geq s$,

$$\left| [\tilde{\mathbf{A}}^{\mathrm{T}}(t)\tilde{\mathbf{A}}^{\mathrm{T}}(t-1)\cdots\tilde{\mathbf{A}}^{\mathrm{T}}(s+1)\tilde{\mathbf{A}}^{\mathrm{T}}(s)]_{ij} - \tilde{\boldsymbol{\phi}}_j(s) \right| \leq \alpha\tilde{\beta}^{t-s},$$

with

$$\alpha = 2, \quad \tilde{\beta} = \left(1 - \frac{1}{N^{NB}}\right)^{\frac{1}{NB}},$$

where $B = T + \bar{\tau}$, T is the bound on the intercommunication interval given in Assumption 2 and $\bar{\tau}$ is the upper bound of the communication time delays in Assumption 3. Therefore, the special property given in Remark 2 is satisfied with the matrix $\tilde{\mathbf{A}}$.

The rest of the proof follows from the proof of Theorem 1 by noticing that in the update equation (16e), the gradient of each function $\Phi_i(\lambda)$ for the original agent is uniformly bounded as shown in (12) and there is no perturbation for the virtual agents. ■

Remark 5: Theorem 2 shows that the proposed distributed Algorithm 1 also solves the EDP even when time-varying directed communication links are subject to arbitrarily large but bounded time-varying delays. In the EDP literature, various distributed algorithms have been proposed for fixed networks without time delays, e.g., [11], [13], [15], [17]. As shown in [19], [20], these algorithms fail to converge when the estimate of the mismatch is subject to time delays and when the time delays on the estimate of the optimal incremental cost are large. The authors of [21] propose a nonnegative-surplus based distributed algorithm to solve the EDP over time-varying communication networks but without time delays. To the best of our knowledge, our proposed algorithm is the first algorithm that is capable to solve the EDP over switching networks with arbitrarily large but bounded time-varying delays.

Remark 6: Note that the convergence rate of the existing algorithms for the EDP over fixed topologies in [11]–[13], [15], [17] depends on the real part of the dominant eigenvalue of the Laplacian matrix associated with network topology and the step-size. In our work, we consider switching topologies with time-varying communication delays, therefore the convergence rate of our proposed algorithm also depends on the nature of switching sequence and the nature of time delays. The explicit relationship is thus more complicated and is left as an interesting future direction.

Remark 7: Although the distributed Algorithm 1 is proposed to solve the EDP over switching communication networks with time-varying delays, it can also be applied to solve a particular type of the optimal resource allocation problem [32], [33], which can be formulated as (1).

V. Case Studies

In this section, various case studies are presented in order to illustrate and validate the proposed algorithm. Test systems have been developed and studied for distributed EDP algorithms in existing works, e.g., the IEEE 14-bus system and the IEEE-118 bus system for fixed communication networks [13], [15], [17] and a 4-bus system for time-varying communication networks [21]. These test systems are adopted to study the proposed algorithm with the corresponding type of communication networks, considering both without and with time delay scenarios.

A. Fixed Communication Networks

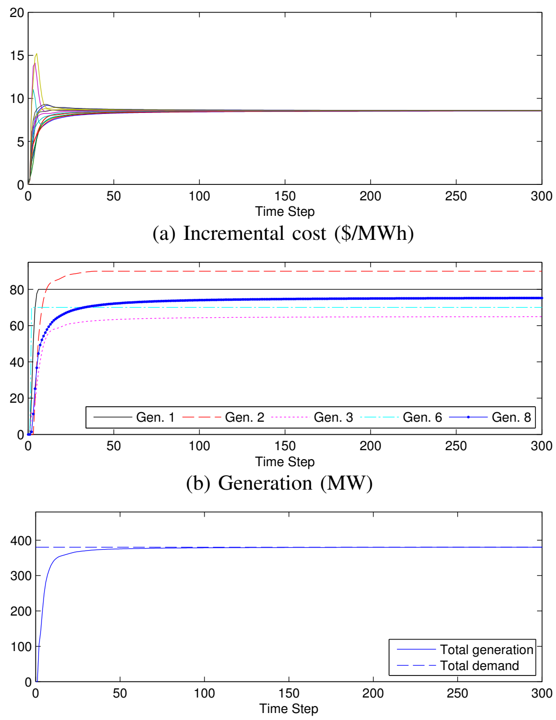
First, the IEEE 14-bus system is used to demonstrate the implementation of the proposed algorithm for a fixed directed communication network, which is modeled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with the edge set $\mathcal{E} = \{(i, i+1), (i, i+2) | 1 \leq i \leq 12\} \cup \{(13, 14), (13, 1), (14, 1), (1, 7), (2, 8), (3, 2), (3, 9), (4, 10), (5, 2), (5, 11), (6, 12)\}$. Note that generator buses are $\{1, 2, 3, 6, 8\}$, and load buses are $\{2, 3, 4, 5, 6, 9, 10, 11, 13, 14\}$. The generator parameters including the parameters of the quadratic cost functions are adopted from [13], [15], [17], which are given in Table I. When a bus does not contain generators, the power generation at that bus is set to zero. Thus, the update in (10d) simply becomes $x_i(t+1) = 0$ for $i \notin \{1, 2, 3, 6, 8\}$. The virtual local demands at each bus are given as $D_1 = 0$ MW, $D_2 = 9$ MW, $D_3 = 56$ MW, $D_4 = 55$ MW, $D_5 = 27$ MW, $D_6 = 27$ MW, $D_7 = 0$ MW, $D_8 = 0$ MW, $D_9 = 8$ MW, $D_{10} = 24$ MW, $D_{11} = 53$ MW, $D_{12} = 46$ MW, $D_{13} = 16$ MW, and $D_{14} = 40$ MW. The total demand is $D = \sum_{i=1}^{14} D_i = 380$ MW, which is unknown to the agent at each bus.

1) Without delay: This case represents ideal communication network, where the communication links are time-invariant and not subject to time delays. This most basic case has been used in many existing studies, and therefore is selected as a starting point for testing the proposed algorithm. The results with a step-size of $\gamma(t) = \frac{0.15}{t}$ are plotted in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics

TABLE I
IEEE 14-BUS SYSTEM GENERATOR PARAMETERS

Bus	a_i(MW^{2}h)	b_i ($/MWh)	Range (MW)
1	0.04	2.0	[0,80]
2	0.03	3.0	[0,90]
3	0.035	4.0	[0,70]
6	0.03	4.0	[0,70]
8	0.04	2.5	[0,80]



(a) Incremental cost ($/MWh)

(b) Generation (MW)

(c) Total generation vs. demand (MW)

Fig. 1. Simulation results for the IEEE-14 bus over fixed communication networks without time delays.

Fig. 1, where Fig. 1(a) shows the evolution of incremental cost $\lambda_i(t)$, Fig. 1(b) shows the evolution of power generation $x_i(t)$, and Fig. 1(c) shows the evolution of the total generation in comparison of total demand. As can be seen, the incremental cost $\lambda_i(t)$ computed at each agent converges to the optimal value $\lambda^* = 8.52$ $/MWh. In particular, at time step $t = 300$, the maximum difference between all the λ_i is 0.0045 $/MWh, which is quite small. The generation at each generator bus also converges to their optimal values, which are $x_1^* = 80$ MW, $x_2^* = 90$ MW, $x_3^* = 64.65$ MW, $x_6^* = 70$ MW, and $x_8^* = 75.31$ MW. As $\lambda_i(t)$ and $x_i(t)$ converge to their optima, the total generation meets total demand $D = 380$ MW. The proposed algorithm is able to handle the power output constraints of individual generator and find the correct optimal solutions. For example, generators 1, 2 and 6 are at the upper bounds of their power output because they are cheaper than the other two generators and therefore provide generation as much as possible.

For this communication network, the proposed algorithm has also been studied for generators with non-quadratic cost functions. In particular, we adopt the following case from [17], where the generator at bus 6 is replaced with a fixed generation $x_6 = 100$ MW and the cost function of generators at bus 1 and 3 become:

$$C_1(x_1) = \frac{(x_1 + 25)^2}{25} + 50\exp(\frac{x_1 + 40}{100}),$$
$$C_3(x_3) = \frac{(x_3 + 57.14)^2}{28.58} + 7 \times 10^{-6} x_3^4.$$

Using the proposed algorithm, the incremental cost at each bus converges to 8.85 $/MWh, the generation at each generator bus respectively converges to $x_1 = 67.85$ MW, $x_2 = 90$ MW, $x_3 = 41.47$ MW, $x_6 = 100$ MW, and $x_8 = 79.96$ MW, which agrees with the centralized solution and the one found in [17].

2) With time-varying delays: In this case, the proposed algorithm is studied using the same communication network topology but with time delays. While there exist works that study the impacts of uniform fixed time delays on distributed EDP algorithm [19], [20], this work considers more general and challenging cases, where delays could be arbitrary time-varying with an upper bound. In this case study, we assume that the upper bound is $\bar{\tau} = 20$. In particular, at each time step, the time delay of each link has a probability of $1/21$ to be any integer in $\{0, 1, \ldots, 20\}$. Since time delays are stochastic, the iteration results at each agent vary from one simulation to another. Nevertheless, the proposed algorithm always converges to the optimal solutions. As an example, Fig. 2 plots the simulation results for a particular run. As can be seen, even in the presence of communication time delays, each variable eventually converges to the same value as the case without time delays. In particular, at time step $t = 5000$, the maximum difference between all the λ_i is 0.0412 $/MWh, which is only 0.48% of the optimal incremental cost λ^*. Compared with the results in Fig. 1, the optimal solutions are obtained with a slower rate.

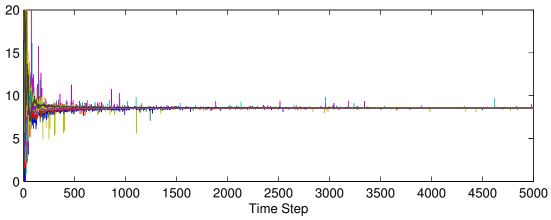
B. Time-Varying Communication Networks

The test system and the communication topology are adopted from [21] for comparison purpose. Four generators are selected from three types, and the power output ranges and parameters of quadratic cost functions for each generator type are given in Table II. The communication network is modeled as a time-varying directed graph $\mathcal{G}(t)$ switching among three fixed topologies $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ shown in Fig. 3 at each time step. In particular,

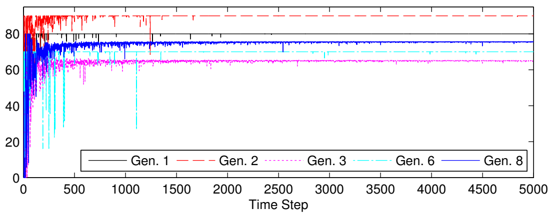
$$\mathcal{G}(t) = \begin{cases} \mathcal{G}_1, & \text{if } t \in [0,1) \cup [3,4) \cdots \cup [3s, 3s+1) \cdots, \\ \mathcal{G}_2, & \text{if } t \in [1,2) \cup [4,5) \cdots \cup [3s+1, 3s+2) \cdots, \\ \mathcal{G}_3, & \text{if } t \in [2,3) \cup [5,6) \cdots \cup [3s+2, 3s+3) \cdots, \end{cases}$$

where $s \in \mathbb{Z}_+$. It is easy to check that each of the fixed topologies $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ is not strongly connected. For example, there is no directed path from agent 2 to agent 4 in $\mathcal{G}_1$. However, the time-varying directed graph $\mathcal{G}(t)$ is uniformly jointly strongly connected since the joint graph $\mathcal{G}([t_0, t_0 + T))$ is strongly connected for any $t_0 \geq 0$ with $T = 3$. Thus, Assumption 2 is satisfied with $T = 3$. According to Theorems 1 and 2, the proposed algorithm solves the EDP over the time-varying communication network without delays and with delays, respectively. To implement the proposed

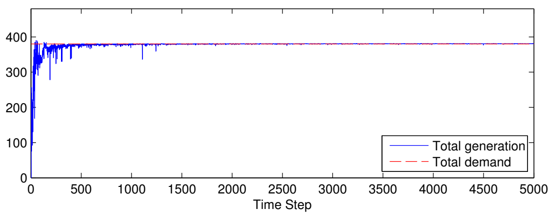
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics



(a) Incremental cost (\$/MWh)



(b) Generation (MW)



(c) Total generation vs. demand (MW)

Fig. 2. Simulation results for the IEEE-14 bus over fixed communication networks with time-varying delays.

TABLE II
GENERATOR PARAMETERS

Type	A (Gen. 1&2)	B (Gen. 3)	C (Gen. 4)
Range (MW)	[150,600]	[100,400]	[50,200]
a_i(MW^{2}h)	0.00142	0.00194	0.00482
b_i (\$/MWh)	7.2	7.85	7.97
c_i (\$/h)	510	310	78

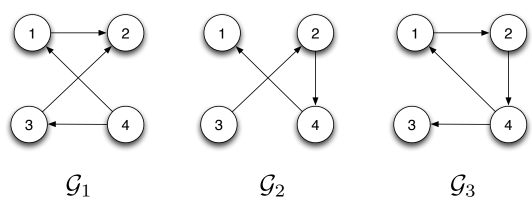
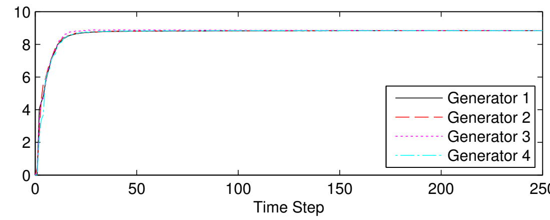


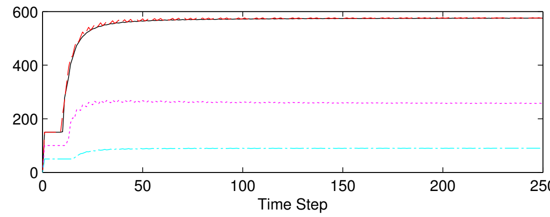
Fig. 3. Time-varying directed communication network.

algorithm, we first choose the virtual local demands at each bus as $D_1 = 500$ MW, $D_2 = 500$ MW, $D_3 = 350$ MW, and $D_4 = 150$ MW. The total demand is $D = \sum_{i=1}^{4} D_i = 1500$ MW, which is unknown to the agent at each bus.

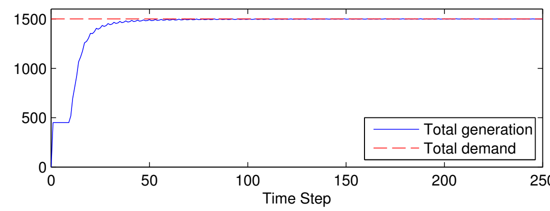
1) Without delay: First, we consider the case where communication links are not subject to delays. With a step-size of $\gamma(t) = \frac{0.01}{t}$, the simulation results are shown in Fig. 4. As can be seen, $\lambda_i(t)$ converge to the optimal incremental cost $\lambda^* = 8.84$ \$/MWh, and x_i converges to the optimal generation $x_1^* = 577.46$ MW, $x_2^* = 577.46$ MW, $x_3^* = 255.16$ MW, and $x_4^* = 90.25$ MW as shown in Fig. 4(b), which agree with the centralized solution and the one obtained in [21]. In particular, at time step $t = 250$, the maximum difference between all the λ_i is 0.0126 \$/MWh, which is quite small. The total generation



(a) Incremental cost (\$/MWh)



(b) Generation (MW)



(c) Total generation vs. demand (MW)

Fig. 4. Simulation results for the system of four generators over time-varying communication networks without time delays.

gradually meets the total demand 1500 MW.

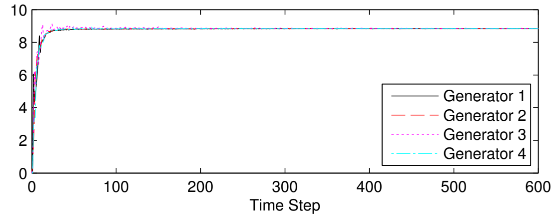
2) With time-varying delays: Note that the nonnegative-surplus based algorithm proposed in [21] cannot handle time delays since it is build upon the algorithm in [15] which is not robust to even uniformly constant delays as shown in [20]. On the other hand, the proposed algorithm in this paper is robust to time-varying delays in addition to time-varying topologies as shown in Theorem 2. In order to demonstrate this distinguishing feature, we herein consider the same time-varying communication topology but with time delays. In this case study, we consider the case where time-varying delays are upper bounded by $\bar{\tau} = 3$ and the probability mass function of time delay on any communication link is given by $P_\tau(\tau) = 0.5$ for $\tau = 0$, $P_\tau(\tau) = 0.35$ for $\tau = 1$, $P_\tau(\tau) = 0.1$ for $\tau = 2$, and $P_\tau(\tau) = 0.05$ for $\tau = 3$.

Since time delays are stochastic, the dynamics at each agent vary from one simulation to another. Nevertheless, the proposed algorithm always converges to the optimal solutions. Simulation results for a particular run are plotted in Fig. 5, As can been seen, even in the presence of communication time delays, each variable still converges to the same value as the case without time delays. In particular, at time step $t = 600$, the maximum difference between all the λ_i is 0.0126 \$/MWh, which is quite small. Compared with the results in Fig. 4, the optimal solutions are obtained with a slower rate.

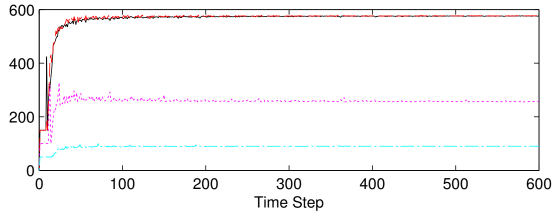
C. IEEE-118 Bus System

In this case we test our algorithm to the IEEE 118-bus system to further show the effectiveness of the proposed algorithm.

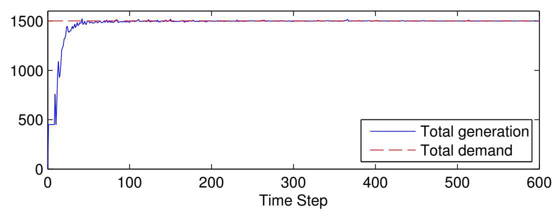
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics



(a) Incremental cost (\$/MWh)

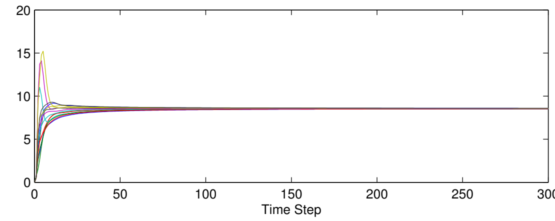


(b) Generation (MW)

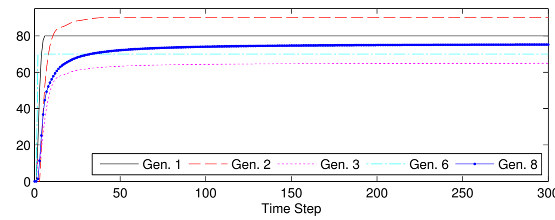


(c) Total generation vs. demand (MW)

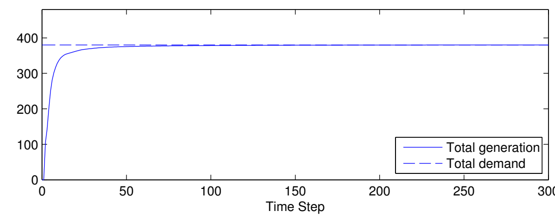
Fig. 5. Simulation results for the system of four generators over time-varying communication networks with time delays.



(a) Incremental cost (\$/MWh)



(b) Generation (MW)



(c) Total generation vs. demand (MW)

Fig. 6. Simulation results for the IEEE 118-bus over fixed networks.

1) Fixed topologies without delays: We first assume that the communication topology is the same as the physical network. The results with a step size of $\gamma(t) = \frac{0.6}{t}$ are plotted in Fig. 6, where Fig. 6(a) shows the evolution of incremental cost $\lambda_i(t)$, Fig. 6(b) shows the evolution of power generation $x_i(t)$, and Fig. 6(c) shows the evolution of the total generation in comparison of total demand. As can be seen, the incremental cost $\lambda_i(t)$ computed at each agent converges to the optimal value $\lambda^* = 39.38$ \$/MWh, which agrees with the result obtained by running the MATPOWER [34]. The generation at each generator bus also converges to the same optimal values as obtained by the MATPOWER. As $\lambda_i(t)$ and $x_i(t)$ converge to their optima, the total generation meets total demand $D = 4242$ MW.

2) Time-varying topologies with time-varying delays: In this case, the proposed algorithm is evaluated using a time-varying communication network with time delays. The communication network is modeled as a time-varying graph switching among two fixed topologies. In particular,

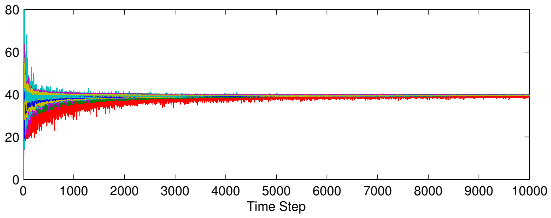
$$\mathcal{G}(t) = \begin{cases} \mathcal{G}_1, & \text{if } t \in [0,1) \cup [2,3) \cdots \cup [2s, 2s+1) \cdots, \\ \mathcal{G}_2, & \text{if } t \in [1,2) \cup [3,4) \cdots \cup [2s+1, 2s+2) \cdots, \end{cases}$$

where $s \in \mathbb{Z}_+$, $\mathcal{G}_1$ is the graph obtained by disconnecting Zone 1 and Zone 2 in [35, Figure 6], where IEEE 118-bus system have been partitioned into three different zones, and $\mathcal{G}_2$ is the graph obtained by disconnecting Zone 2 and Zone 3 in [35, Figure 6]. It is easy to check that each of the fixed topologies $\mathcal{G}_1$ and $\mathcal{G}_2$ is not connected. However, the time-varying graph $\mathcal{G}(t)$ is uniformly jointly connected since the joint graph $\mathcal{G}([t_0, t_0 + T))$ is connected for any $t_0 \geq 0$ with $T = 2$. Thus, Assumption 2 is satisfied with $T = 2$. According to Theorem 2, the proposed algorithm solves the EDP over the time-varying communication network with arbitrary time-varying communication delays with an upper bound. In this case study, we assume that the upper bound is $\bar{\tau} = 3$. In particular, at each time step, the time delay of each link has a probability of $1/4$ to be any integer in $\{0, 1, 2, 3\}$. Since time delays are stochastic, the dynamics at each agent vary from one simulation to another. Nevertheless, the proposed algorithm always converges to the optimal solutions. Simulation results for a particular run are plotted in Fig. 7. As can been seen, even in the presence of both time-varying communication links and communication time delays, each variable still converges to the same value as the case for the fixed communication network without time delays.

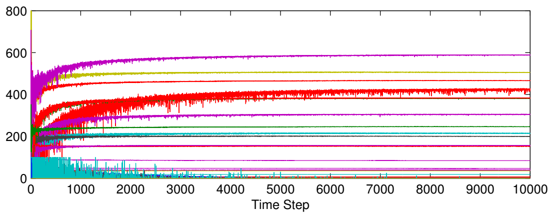
VI. Conclusions

This paper proposes a distributed algorithm based on the gradient push-sum method to solve the EDP over time-varying directed communication networks with time-varying delays. The cost functions are assumed to be convex rather than quadratic as in most existing studies. The proposed algorithm is fully distributed, without requiring global information of the system. Both theoretical proofs and simulation results showed that the proposed algorithm can solve the EDP over time-varying directed communication networks provided that the network is uniformly jointly strongly connected. Moreover, the proposed algorithm is also robust to arbitrarily large but bounded time-varying communication delays. One future work will focus on the robustness of the proposed distributed algorithm against unreliable communication links that may drop packets. Another interesting direction is to extend the

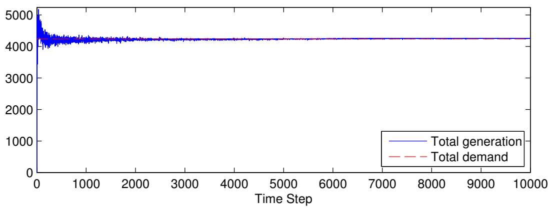
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics



(a) Incremental cost ($/MWh)



(b) Generation (MW)



(c) Total generation vs. demand (MW)

Fig. 7. Simulation results for the IEEE 118-bus over time-varying networks with delays.

proposed distributed algorithm to accommodate additional physical constraints, such as transmission line loss and power flow and transmission line flow constraints.

REFERENCES

[1] A. Wood and B. Wollenberg, *Power Generation, Operation and Control*, 2nd ed. New York: Wiley, 1996.

[2] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[3] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[4] R. Gupta and M.-Y. Chow, "Networked control system: Overview and research trends," *IEEE Trans. Ind. Electron.*, vol. 57, no. 7, pp. 2527–2535, Jul. 2010.

[5] J. Qin, C. Yu, and H. Gao, "Coordination for linear multiagent systems with dynamic interaction topology in the leader-following framework," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2412–2422, May 2014.

[6] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 427–438, Feb. 2013.

[7] C. Zhao, U. Topcu, and S. Low, "Optimal load control via frequency measurement and neighborhood area communication," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3576–3587, Nov. 2013.

[8] P. Yi, Y. Hong, and F. Liu, "Distributed gradient algorithm for constrained optimization with application to load sharing in power systems," *Systems & Control Letters*, vol. 83, pp. 45 – 52, 2015.

[9] M. Liu, Y. Shi, and X. Liu, "Distributed MPC of aggregated heterogeneous thermostatically controlled loads in smart grid," *IEEE Trans. Ind. Electron.*, vol. 63, no. 2, pp. 1120–1129, Feb. 2016.

[10] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control & economic optimality in microgrids," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 3, pp. 241–253, Sep. 2016.

[11] Z. Zhang and M.-Y. Chow, "Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1761–1768, Nov. 2012.

[12] Z. Zhang, X. Ying, and M.-Y. Chow, "Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment," in *Proc. North Amer. Power Symp. (NAPS)*, Aug. 2011.

[13] S. Kar and G. Hug, "Distributed robust economic dispatch in power systems: A consensus+innovations approach," in *Proc. IEEE Power Energy Soc. Gene. Meet.*, Jul. 2012.

[14] A. Dominguez-Garcia, S. Cady, and C. N. Hadjicostis, "Decentralized optimal dispatch of distributed energy resources," in *Proc. 51st IEEE. Conf. Decision and Control (CDC)*, Dec. 2012, pp. 3688–3693.

[15] S. Yang, S. Tan, and J.-X. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.

[16] T. Yang, D. Wu, Y. Sun, and J. Lian, "Minimum-time consensus based approach for power system applications," *IEEE Trans. Ind. Electron.*, vol. 63, no. 2, pp. 1318–1328, Feb. 2016.

[17] H. Xing, Y. Mou, M. Fu, and Z. Lin, "Distributed bisection method for economic power dispatch in smart grid," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3024–3035, Nov. 2015.

[18] K. Aström and P. Kumar, "Control: A perspective," *Automatica*, vol. 50, no. 1, pp. 3–43, 2014.

[19] Z. Zhang and M.-Y. Chow, "The influence of time delays on decentralized economic dispatch by using incremental cost consensus algorithm," in *Control and Optimization Methods for Electric Smart Grids*. London: Springer Verlag, 2012, vol. 371, pp. 313–326.

[20] T. Yang, D. Wu, Y. Sun, and J. Lian, "Impacts of time delays on distributed algorithms for economic dispatch," in *Proc. IEEE Power and Energy Society General Meeting*, Jul. 2015.

[21] Y. Xu, K. Cai, T. Han, and Z. Lin, "A fully distributed approach to resource allocation problem under directed and switching topologies," in *Proc. 10th Asian Control Conference (ASCC)*, May 2015.

[22] W. Zhang, W. Liu, X. Wang, L. Liu, and F. Ferrese, "Online optimal generation control based on constrained distributed gradient algorithm," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 35–45, Jan. 2015.

[23] C. Godsi and G. Royle, *Algebraic Graph Theory*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 2001, vol. 207.

[24] C. Li, X. Yu, W. Yu, T. Huang, and Z.-W. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Informat.*, vol. PP, no. 99, pp. 1–1, 2015.

[25] D. P. Bertsekas, A. Nedić, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.

[26] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[27] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, Mar. 2015.

[28] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *in Procs. 44th Annu. IEEE Symp. Found. Comp. Sci.*, Oct. 2003, pp. 482–491.

[29] A. Dominguez-Garcia and C. N. Hadjicostis, "Distributed strategies for average consensus in directed graphs," in *Proc. 50th IEEE Conf. Decision and Control (CDC)*, Dec. 2011, pp. 2124–2129.

[30] T. Charalambous, Y. Yuan, T. Yang, W. Pan, C. N. Hadjicostis, and M. Johansson, "Distributed finite-time average consensus in digraphs in the presence of time-delays," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 4, pp. 370–381, Dec. 2015.

[31] A. Nedić and A. Ozdaglar, "Convergence rate for consensus with delays," *Journal of Global Optimization*, vol. 47, no. 3, pp. 437–456, Jul. 2010.

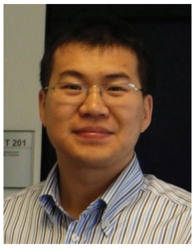
[32] L. Xiao and S. Boyd, "Optimal scaling of a gradient method for distributed resource allocation," *Journal of Optimization Theory and Applications*, vol. 129, no. 3, pp. 469–488, Jun. 2006.

[33] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," in *Proc. 47th IEEE Conf. Decision and Control (CDC)*, Dec. 2008, pp. 4185–4190.

[34] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[35] D. Zhang and S. Li, "Optimal dispatch of competitive power markets by using PowerWorld simulator," *International Journal of Emerging Electric Power Systems*, vol. 14, no. 6, pp. 535-547, Oct. 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2016.2617832, IEEE Transactions on Industrial Electronics



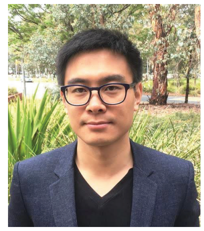
Tao Yang (M'12) received the B.S. degree in Computer Science from Harbin University of Science and Technology in 2003, the M.S. degree with distinction in control engineering from City University, London in 2004, and the Ph.D. degree in electrical engineering from Washington State University in 2012. Between August 2012 and August 2014, he was an ACCESS Post-Doctoral Researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Sweden. He is currently an Assistant Professor at the Department of Electrical Engineering, University of North Texas (UNT). Prior to joining the UNT, he was a Scientist/Engineer II with Energy & Environmental Directorate, Pacific Northwest National Laboratory. His research interests include distributed control and optimization in power systems, Cyber Physical Systems, networked control systems, and multi-agent systems.

Jie Lu (M'13) received the B.S. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2007 and the Ph.D.degree in electrical and computer engineering from the University of Oklahoma, Norman, in 2011.

From 2012 to 2015, she was a postdoctoral researcher with KTH Royal Institute of Technology, Stockholm, Sweden, and with Chalmers University of Technology, Gothenburg, Sweden. Since 2015, she has been an Assistant Professor in the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. Her research interests include distributed optimization, optimization theory and algorithms, and networked dynamical systems.

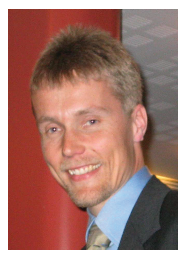
Di Wu (M'12) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University in 2003 and 2006, respectively, and the Ph.D. in electrical and computer engineering from Iowa State University, Ames, in 2012. Dr. Wu is currently a senior power system engineer in the Electricity Infrastructure group at Pacific Northwest National Laboratory (PNNL). He serves as an Associate Editor for the *IEEE Power and Energy Technology Systems Journal*. His research interests include distributed control and coordination in power systems, operational planning and optimization of energy storage, production cost models, and power system dynamic simulation. Before joining PNNL, his research mainly focuses on the impacts of plug-in electric vehicles on power systems and planning of national energy and transportation infrastructure.

Junfeng Wu received the B.Eng. from the Department of Automatic Control, Zhejiang University, Hangzhou, China, in 2009 and the Ph.D. degree in Electrical and Computer Engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2013. From September to December 2013, he was a Research Associate in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. He is currently a Postdoctoral Researcher at ACCESS (Autonomic Complex Communication nEtworks, Signals and Systems) Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, Sweden. His research interests include networked control systems, state estimation, and wirelesssensor networks, multi-agent systems. He received the Guan Zhao-Zhi Best Paper Award at the 34th Chinese Control Conference in 2015.



Guodong Shi (M'15) received his B.Sc. degree in Mathematics and Applied Mathematics from School of Mathematics, Shandong University, Jinan, China, in July 2005, and his Ph.D. in Systems Theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in July 2010, respectively. From Aug. 2010 to Apr. 2014 he was a postdoctoral researcher at the ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. He held a visiting position from Oct. 2013 to Dec. 2013 at the School of Information and Engineering Technology, University of New South Wales, Canberra, Australia. Since May 2014 he has been with the Research School of Engineering, The Australian National University, Canberra, Australia, as a Lecturer and Future Engineering Research Leadership Fellow. His current research interests include distributed control systems, quantum networking and decisions, and social opinion dynamics.

Ziyang Meng (M'14) received his Bachelor degree with honors from Huazhong University of Science & Technology, Wuhan, China, in 2006, and Ph.D. degree from Tsinghua University, Beijing, China, in 2010. He was an exchange Ph.D. student at Utah State University, Logan, USA from Sept. 2008 to Sept. 2009. From 2010 to 2015, he held postdoc, researcher, and Humboldt research fellow positions at, respectively, Shanghai Jiao Tong University, Shanghai, China, KTH Royal Institute of Technology, Stockholm, Sweden, and Technical University of Munich, Munich, Germany. He joined Department of Precision Instrument, Tsinghua University, China as an Associate Professor since Sept. 2015. His research interests include multiagent systems, small satellite systems, distributed control and optimization, nonlinear systems and information fusion techniques. He was selected to the national 1000-Youth Talent Program of China in 2015.



Karl Henrik Johansson (SM'08/F'13) is the director of the ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, KTH Royal Institute of Technology, Sweden. He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. He also heads the Stockholm Strategic Research Area ICT The Next Generation. He received MSc and PhD degrees in Electrical Engineering from Lund University. He has held visiting positions at UC Berkeley, California Institute of Technology, Nanyang Technological University, and Institute of Advanced Studies, Hong Kong University of Science and Technology. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation systems. He has been a member of the IEEE Control Systems Society Board of Governors and the Chair of the IFAC Technical Committee on Networked Systems. He has been on the Editorial Boards of several journals, including Automatica, IEEE Transactions on Automatic Control, and IET Control Theory and Applications. He is currently a Senior Editor of IEEE Transactions on Control of Network Systems and Associate Editor of European Journal of Control. He has been Guest Editor for a special issue of IEEE Transactions on Automatic Control on cyber-physical systems and one of IEEE Control Systems Magazine on cyber-physical security. He was the General Chair of the ACM/IEEE Cyber-Physical Systems Week 2010 in Stockholm and IPC Chair of many conferences. He has served on the Executive Committees of several European research projects in the area of networked embedded systems. He received the Best Paper Award of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems in 2009 and the Best Theory Paper Award of the World Congress on Intelligent Control and Automation in 2014. In 2009 he was awarded Wallenberg Scholar, as one of the first ten scholars from all sciences, by the Knut and Alice Wallenberg Foundation. He was awarded Future Research Leader from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999. He is a Fellow of the IEEE.