

Received June 4, 2017, accepted August 20, 2017, date of publication August 29, 2017, date of current version September 19, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2746093

Delay Aware Intelligent Transient Stability Assessment System

**JAMES J. Q. YU, (Member, IEEE), ALBERT Y. S. LAM, (Senior Member, IEEE),
DAVID J. HILL, (Life Fellow, IEEE), AND VICTOR O. K. LI, (Fellow, IEEE)**

Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

Corresponding author: James J. Q. Yu (jqyu@eee.hku.hk)

This work was supported by the Theme-based Research Scheme of the Research Grants Council of Hong Kong under Grant T23-701/14-N.

ABSTRACT Transient stability assessment is a critical tool for power system design and operation. With the emerging advanced synchrophasor measurement techniques, machine learning methods are playing an increasingly important role in power system stability assessment. However, most existing research makes a strong assumption that the measurement data transmission delay is negligible. In this paper, we focus on investigating the influence of communication delay on synchrophasor-based transient stability assessment. In particular, we develop a delay aware intelligent system to address this issue. By utilizing an ensemble of multiple long short-term memory networks, the proposed system can make early assessments to achieve a much shorter response time by utilizing incomplete system variable measurements. Compared with existing work, our system is able to make accurate assessments with a significantly improved efficiency. We perform numerous case studies to demonstrate the superiority of the proposed intelligent system, in which accurate assessments can be developed with time one third less than state-of-the-art methodologies. Moreover, the simulations indicate that noise in the measurements has trivial impact on the assessment performance, demonstrating the robustness of the proposed system.

INDEX TERMS Transient stability assessment, communication delay, long short-term memory, phasor measurement units, voltage phasor, intelligent system.

I. INTRODUCTION

Transient stability refers to the capability of a power system to maintain its synchronism subject to large disturbances [1]. The transient stability issues caused by large disturbances are considered more serious than before as the power systems are being operated close to their stability limits to satisfy the increasing power demand [2]. Consequently, critical contingencies may lead to significant system failures or power blackouts. In order to prevent such situations, system operators need to assess the stability condition of the grid and, when necessary, plan a collection of remedial control actions to retain the stability. Therefore, transient stability assessment (TSA) in real-time is regaining interest from the community [3].

Many previous studies on TSA were conducted by using offline dynamic simulations for a collection of credible contingencies [4]. This methodology is widely adopted in designing the protective and control systems for secure operations. Meanwhile, online TSA techniques are employed to evaluate the progress of transient dynamics of a power system in real time.

With the gradual adoption of synchrophasor measurement facilities, e.g., phasor measurement units (PMUs), a significant amount of effort has been devoted to utilizing real-time system variables for TSA decision making [4]–[6]. On top of that, post-contingency remedial actions can be taken in real-time to give guaranteed TSA results [3]. With post-contingency system dynamics, techniques such as the piecewise constant current load equivalent method [7] and emergency single machine equivalent method [8] were proposed for online dynamic power system security assessment. Machine learning techniques for TSA, on the other hand, received a lot of attention in recent years due to their relatively low assessment computational complexity. Approaches like pattern recognition [9], decision tree [10], artificial neural networks (ANN) [11], support vector machine [12], and fuzzy knowledge-based systems [13] were employed to realize fast-response online TSA. For instance, our previous work [14] handles TSA with a modern variant of ANN, and assessment results can be generated in an online manner. These techniques extract the relationship between system variable measurements and their respective stability indices. With this

relationship, new transient stability dynamics can be assessed with minimal computational efforts.

It appears that all existing research on synchrophasor-based online TSA implicitly assumes that the wide area monitoring systems can provide reliable, accurate, and synchronized system variable measurements and there is no data transmission delay between PMUs and the central controller, see [14], [15] for examples. However, although being measured in a synchronous manner thanks to PMUs, synchrophasors cannot reach the central controller in perfect synchrony due to unpredictable communication link congestion and routing delay [16]. Therefore, it is necessary to design a data transmission delay-aware TSA system that can provide robust and reliable assessment results given delayed or missing system variable measurements [17].

One feasible approach to address the delayed synchrophasor problem is to recover those missing data by learning the system model and other measurements when the delayed data packets are being transmitted [18], [19]. However, this approach suffers from a major drawback that the high computational complexity of these nonlinear state estimators results in a large delay for estimating the delayed synchrophasor, far from being commensurate with the PMU sampling rate [17]. To overcome this drawback in the state estimation approaches, a data-mining based neural network ensemble prediction technique is utilized as an alternative in this paper.

This work focuses on establishing an intelligent system to address the transient stability assessment problem with time-delayed synchrophasors. We construct the system using multiple advanced machine learning techniques and heuristics, namely Long Short-term Memory (LSTM) [20], ensemble learning [21], rule-based decision machine heuristic, and two optimizers (Adam [22] for neural network training, and Social Spider Algorithm [23] for non-convex optimizations). Utilizing the advantages of involved techniques, the proposed system is able to make preliminary assessments at the earliest possible time and revise the predictions when more information is available. Compared with previous TSA systems, the proposed one can achieve a faster response time to make accurate assessments. In addition, the combination of multiple techniques contributes to the superior assessment performance as will be illustrated.

The remainder of this paper is organized as follows. Section II introduces the delay aware TSA problem and its difference from tradition TSA. Section III elaborates on the formulation and implementation of the proposed intelligent system. Section IV demonstrates numerical results on a modified New England 10-machine test system. Finally we conclude in Section V with a discussion of potential future research.

II. DELAY AWARE TRANSIENT STABILITY ASSESSMENT

Typical stability assessment methodologies introduce the idea of observation window to facilitate their data collection process (See [15, Sec. 2.1] for a detailed introduction). After the clearance of a fault, the power system dynamic behavior is

observed for a certain period of time by PMUs, and this information is later utilized to make a stability prediction for the future. While some recent work has improved the TSA system by replacing the block observation with continuous observation [15], it still relies on an implicit but strong assumption: the central controller needs to receive measurements with the same timestamp from different data sources simultaneously.

From the communication perspective, however, it is not realistic to postulate that no data transmission delay is incurred over such communication links [16]. Suppose the system measurements from PMU p reach the central controller at time $t + \Delta t_{p,t}$, where t is the time index after the fault clearance, and $\Delta t_{p,t}$ is the data transmission delay induced by the message communication from p at t . The previous TSA work generally assumes $\Delta t_{p,t} \equiv 0$. However, investigations on data transmission protocols demonstrate that modeling such a delay is a complicated task and cannot be simply replaced with a constant [24]. Moreover, the measurements actually follow an asynchronous and disordered pattern: $\Delta t_{p,t} < \Delta t_{q,i}$ has no direct implication on $\Delta t_{p,t+1} < \Delta t_{q,t+1}$, where p and q are arbitrary PMUs.

According to the IEEE Standard for Synchrophasor Data Transfer for Power Systems (C37.118.2-2011) [25], typical values for the transmission delay between PMU and Phasor Data Concentrator (PDC) are between 20 ms to 50 ms, and this delay may be further increased due to temporary data congestion over the communication links. The delay should be added to the system response time the TSA calculation time to give, which obstructs the early adoption of subsequent control actions. Therefore, research needs to be carried out on proposing mechanisms that can make transient assessments without waiting for the arrival of all PMU measurements. Such mechanisms should develop a similar TSA accuracy compared with the state-of-the-art TSA algorithms, while making assessments far earlier than those made with full communication delays. In this paper we aim to devise an intelligent delay aware TSA system to achieve this objective. The proposed system copes with the asynchronous arrival of synchrophasor data, and minimizes their influence on the system response time.

III. PROPOSED DELAY AWARE ASSESSMENT SYSTEM

In this section, we elaborate on how we design an intelligent system to perform TSA considering delayed synchrophasor data packets arriving in an asynchronous manner. We will first introduce Long-short Term Memory (LSTM) networks, and the reason why we use such networks to construct the system. Then we overview the structure of the system followed by the illustrations of other building blocks and the decision machine. Finally we will explain the overall work flow of the proposed delay aware TSA system.

A. AN INTRODUCTION TO LONG SHORT-TERM MEMORY

ANN is one of the automatic machine learning techniques and it has been employed in a variety of disciplines in the last few decades [26]. One major merit of this technique that makes it

suitable for online data processing is its excellent responsiveness [26]. ANN functions like a black box. By training with supervised datasets, ANN learns the mathematical relationship of inputs X and outputs Y , or their distributions [27].

LSTM is a variant of ANN [20]. From the functional point of view, LSTM differs from the conventional ANN in that it considers the temporal data correlation in X . In addition, LSTM generally outperforms other time-dependent variants of ANN, e.g., recurrent neural network, as it avoids the occurrence of the “vanishing gradient problem”, which deteriorates the temporal correlation extraction efficiency [20].

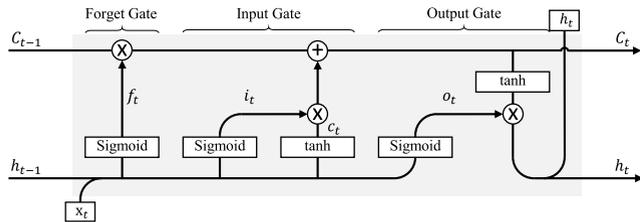


FIGURE 1. The LSTM memory block with the memory cell C_t .

Fig. 1 illustrates one memory block of a typical LSTM network. This block utilizes the input data with a specific timestamp as well as the memory from the previous timestamp for feature extraction. The processed information, or memory, is stored in the memory cell C_t and passed in the next time slot [20]. The output data h_t is also generated based on C_t . This process is depicted in Fig. 2.

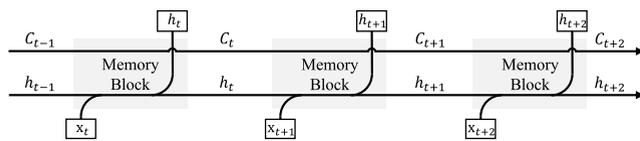


FIGURE 2. Unrolled form of a typical LSTM network.

As illustrated, the LSTM memory block comprises three gates, namely, the forget, input, and output gates; the forget and input gates manage the existing network memory and the new input information, while the output gate controls the output information. To compute C_t , these gates manipulate the temporal data correlation stored in the LSTM network as follows:

$$\mathbf{f}_t = \text{Sigmoid}(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f) \quad (1a)$$

$$\mathbf{i}_t = \text{Sigmoid}(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i) \quad (1b)$$

$$\mathbf{c}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c) \quad (1c)$$

$$\mathbf{C}_t = \mathbf{f}_t * \mathbf{C}_{t-1} + \mathbf{i}_t * \mathbf{c}_t. \quad (1d)$$

Consequently, the output is generated at the output gate, given by

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o) \quad (2a)$$

$$\mathbf{h}_t = \mathbf{o}_t * \tanh(\mathbf{C}_t). \quad (2b)$$

In (1) and (2), $*$ is the element-wise product, $\text{Sigmoid}(x) = (1 + e^{-x})^{-1}$ is the sigmoid function, and \mathbf{W} , \mathbf{U} , \mathbf{b} are matrices corresponding to the LSTM learning parameters.

With the LSTM memory blocks, neural networks can accept a sequence of inputs $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T$ and develop timestamped outputs $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_t, \dots, \mathbf{h}_T$, where T is the length of input data. Output \mathbf{h}_t are generated with the complete existing knowledge from time 1 to t . For instance, \mathbf{h}_1 is calculated solely using \mathbf{x}_1 , and \mathbf{h}_3 is generated by $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 . Utilizing this characteristic, LSTM can develop preliminary results once the first set of input \mathbf{x}_1 is available. Therefore, the proposed LSTM-based system has the capability to generate very early TSA results, which can be later revised when more data is available. It meets the delay aware TSA system requirements as stated in Section II.

B. STRUCTURE OF PROPOSED ENSEMBLE-BASED INTELLIGENT SYSTEM

Ensemble of neural networks is a learning paradigm in which multiple neural networks are employed to solve a problem. In statistical learning, it is widely recognized that such ensembles demonstrate improved generalization capabilities compared to standalone networks [21]. In an ensemble, classification errors of one single network can be compensated by others, to provide an enhanced robustness of the complete system.

Due to the outstanding performance of neural network ensembles for classification tasks, we develop a novel LSTM ensemble-based intelligent system for efficient TSA. The structure of such a system is shown in Fig. 3. In the system, one *Main Block* LSTM network (the striped block in Fig. 3) is employed to develop a *Primary Result* using post-contingency bus voltage magnitude and angle measurements of the power system. Meanwhile, N *Ensemble Block* networks generate N *Secondary results* in parallel. While primary and secondary results share the same characteristics and purposes, we consider the primary result more reliable, as the main block has a larger network structure, which will be elaborated on in the following sub-sections, than individual secondary results. These $N + 1$ results are jointly considered in a rule-based *Decision Machine* to produce the final TSA result. The purpose of introducing this machine is to gather enough system stability assessment information from multiple sources before making a final conclusion.

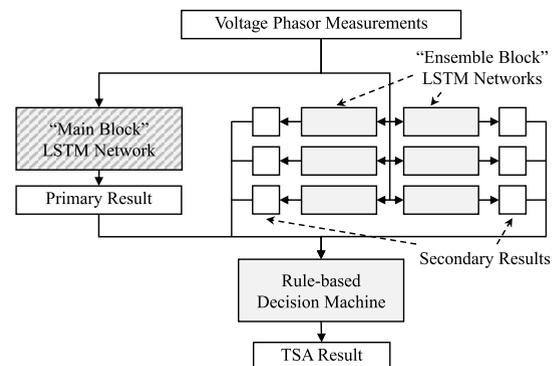


FIGURE 3. Structure of the proposed ensemble-based intelligent system for TSA.

Besides the different sizes of the main and ensemble blocks, their ways of addressing communication delays are also different. The main block can develop TSA results despite the incompleteness of input data due to latency, but the primary results may sometimes be inaccurate. The ensemble blocks will only develop their results when all required input measurements are known, and the secondary results are utilized to correct the primary result. Consequently, when most measurements are not received, the main block can already give preliminary assessments. With the increase of received data, ensemble blocks start to correct the primary result if necessary.

C. MAIN BLOCK

As illustrated in Fig. 3, the main block is employed to develop primary results for generating the final TSA result in the proposed system. As this result is considered essential - this part will be further elaborated in Section III-E - we pay more attention to model its input-output dependency by employing a Deep Neural Network (DNN) with four layers of LSTM blocks and two fully-connected hidden neuron layers. The structure of this DNN is depicted in Fig. 4. While the LSTM blocks are capable of extracting the temporal data dependency from the input data, the neuron layers translate the extracted features to human-readable assessment results. A final Sigmoid function is utilized to cast the results into (0, 1). Owing to its superior capability of modeling highly non-linear relationships, this main block network is expected to output a more reliable result than the secondary results produced by the ensemble blocks, each of which contains significantly fewer number of layers.

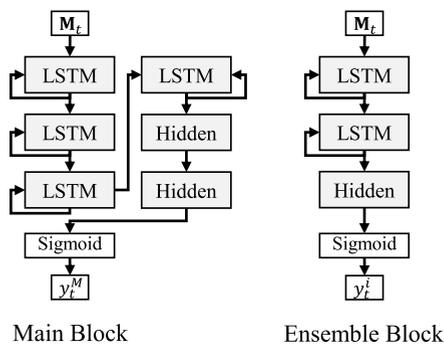


FIGURE 4. Structure of the main block and ensemble blocks.

1) OFFLINE TRAINING PROCESS

Training the main block is conducted offline with all bus voltage phasor data obtained by contingency simulations. Using 50/60-Hz sampling, the normalized measured voltage phasors are presented in the form of $\mathbf{M}^{B \times T}$, where B is the number of buses.¹ The input data \mathbf{M} can be sliced along the

¹In practice, it can be more economically efficient to install PMUs on a subset of buses in the power grid. In such cases, $\mathbf{M}^{B \times T}$ may also include “measurements” which are actually computed using power flow model.

time axis, resulting in vectors $\mathbf{M}_t = [M_{1,t}, M_{2,t}, \dots, M_{B,t}]'$, where $M_{b,t}$ is the system variable measurements of bus b at timestamp t after the fault clearance. Each \mathbf{M}_t comprises measurements from different buses with an identical timestamp. The output stability assessment y is obtained by observing the generation angle derivation and presented in 0-1 binary form.²

Given a collection of C training cases $\{\mathbf{M}_{(c)}, y_{(c)}\}_{c=1}^C$ where $y_{(c)}$ is the assessment y of case c , the training process aims to obtain the parameters of the LSTM [20] and fully-connected neuron layers [26]. In this paper, we employ the Adam optimizer [22] to find the optimal values for the above-mentioned parameters, and the binary cross entropy error function is selected as the objective function:

$$\text{minimize } - \sum_{c=1}^C [y_{(c)} \log y_{(c)}^M + (1 - y_{(c)}) \log(1 - y_{(c)}^M)], \quad (3)$$

where $y_{(c)}^M$ is the actual assessment result of $\mathbf{M}_{(c)}$ with the main block.

2) ONLINE ASSESSMENT PROCESS

Online assessment cannot be conducted in the same way as training. As introduced in Section III-A, the proposed main block accepts one or multiple \mathbf{M}_t vectors instead of an integral \mathbf{M} block as input and outputs y_t . However, some values in \mathbf{M}_t can be unknown when the first several measurements in \mathbf{M}_t reach the central controller. In all previous work, e.g., [3]–[7], [15], the system may hang on until all synchrophasors in \mathbf{M}_t become ready, but the waiting time ($\max\{\Delta t_{p,t}\}$) is significantly increased compared with premature assessments cases.

In this work, we employ a “zero-padding” scheme to pad all unknown values in \mathbf{M}_t with zeros. In addition, a user-defined parameter ϕ is introduced to determine whether a specific \mathbf{M}_t is included in the input. It is included only when 1) the percentage of known values in \mathbf{M}_t is greater than ϕ , and 2) \mathbf{M}_{t-1} is also included when $t > 1$. In this design, the value zero in the input is considered “unknown”, and the main block can start developing primary results much earlier than waiting for all synchrophasors.

During the online assessment process, one stability assessment result is developed whenever a new synchrophasor reaches the central control and $\mathbf{M} \neq \emptyset$. This primary result is output to the decision machine to develop the TSA result.

D. ENSEMBLE BLOCK

While the main block is expected to develop considerably reliable assessment results, it is still inevitable that the DNN may suffer from the problem of overfitting which can result in undermined accuracy on unknown test cases. What is worse, missing synchrophasors introduce noise to the input data, which further compromise the prediction performance.

²In this paper, assessment result 1 means that the system will remain stable and 0 is unstable.

In order to provide more information for the assessment system, multiple ensemble blocks are employed, each of which comprises two layers of LSTM and a fully-connected hidden neuron layer. Similarly, the results are post-processed by a Sigmoid function. Fig. 4 presents a comparison on the rolled structure of a main block and an ensemble block.

As the ensemble blocks have a much simpler network structure, they are less aware to outliers in the input characteristics at the expense of limited abilities to model complex data with scarce neurons. Thus only a small portion of all the available synchrophasors are utilized as the inputs of an ensemble block. Meanwhile, potential overfitting to the chosen PMU data can bias the results. Multiple ensembles with voltage phasors generated by different PMUs are required to give sufficient secondary results for the decision machine.

In such a system, how to choose voltage phasor collections by PMUs as network input for different ensemble blocks greatly influences the overall prediction accuracy. To optimally allocate PMU collections for each ensemble block as data input, a PMU input optimization problem is formulated. The objective function is constructed by considering the system voltage phasor observabilities of multiple sets of PMU. The overall formulation of the problem is given as follows:

$$\text{maximize}_{S_1, \dots, S_N} \frac{1}{N} (\sum O_i + \sum (O_i - \frac{1}{N} \sum O_i)), \quad (4a)$$

$$\text{subject to } |S_i| = \lfloor P/N \rfloor, \quad \forall i \in \{1, 2, \dots, N-1\}, \quad (4b)$$

$$\bigcup S_i = \{1, 2, \dots, P\}, \quad (4c)$$

$$S_i \cap S_j = \emptyset, \quad \forall i, j \in \{1, 2, \dots, N\}, i \neq j, \quad (4d)$$

where S_i is the set of PMUs for ensemble block i , N and P are the total numbers of ensemble blocks and PMUs, respectively, and O_i is the voltage phasor observability of PMU set S_i . The constraints in (4b) limit the total number of PMUs for each ensemble block. The constraints in (4c) dictate that all PMUs in the system are considered in one ensemble block, and (4d) guarantee that each PMU can only be included by one ensemble block. This combinatorial optimization problem can be solved by using a suitable metaheuristic and we adopt a recently proposed Social Spider Algorithm (SSA) [23] as the problem solver in this paper. Note that the problem solver is among the possible techniques to tackle (4a). Alternative methods may be further investigated in future research.

After determining the inputs for each ensemble block, the same training method used in the main block can also be adopted to train these blocks offline. Meanwhile, online assessments are made only when the input vectors \mathbf{M}_t are available as a whole for ensemble blocks. The computational expense can be reduced when compared with the main block, and the assessed secondary results are passed to the decision machine for further processing.

E. DECISION MACHINE

Recall that both the primary and secondary results made by the main and ensemble blocks are values between zero

and one. However, practitioners would prefer a meaningful conclusion on stability instead of statements like “the system will be 35% stable”. Thus it is essential to develop a transformation scheme to map the fuzzy real-valued results into the affirmative 0-1 binary form.

Here we adopt a series of threshold values $\theta_t \in (0, 0.5)$ for each block to map the network result y_t to binary values. Recall that both the main block and the ensemble blocks can generate a sequence of y_t values at different timestamps, each value has a corresponding binary mapping, denoted z_t as follows:

$$z_t = \begin{cases} 1 \text{ (Stable)} & \text{for } y_t > 1 - \theta_t, \\ 0 \text{ (Unstable)} & \text{for } y_t < \theta_t, \\ ? \text{ (Unknown)} & \text{otherwise.} \end{cases} \quad (5)$$

The result is considered reliable when we have either $z_t = 1$ or $z_t = 0$. The threshold θ_t of each block in the system is determined offline by solving the threshold optimization problem:

$$\min_{\theta_1, \dots, \theta_T} (1 - A) \times \omega + D - 1. \quad (6)$$

Here $\omega \in (0, \infty)$ is a weight coefficient, A and D are the testing accuracy and average cycles required to generate a reliable result. When solving (6), the training cases $\{\mathbf{M}_{(c)}, y_{(c)}\}$ are employed to generate the value of A using the binary cross entropy error function. Different values of θ can result in different 0-1 assessment results z_t from the same y_t , thus A is developed from the control variables. Variable D is calculated with the method introduced in [15], where proper θ values help the system make assessments early. Due to space limitations, the relationship among θ 's, A , and D is not provided here. Interested readers may refer to [15] for more details. This optimization problem can also be solved by an appropriate metaheuristic (e.g., SSA), and the determined θ_t values remain constant during the online stability assessment process.

Besides the above transformation process, another main objective achieved by this decision machine is to combine multiple results and to develop a final assessment result. The basic idea is to make assessments largely based on the primary result, and use the secondary results for correction. As the primary and secondary results are generated at different time instants, the decision machine follows the charts given in Fig. 5 to consolidate these results. If the machine receives the primary result, it requires one secondary result for confirmation. Meanwhile, if one or multiple secondary results are available, the system still need to wait for the arrival of the primary one. Therefore, we have the following rules:

- The algorithm starts when the first reliable result is generated by either the main block or any of the ensemble blocks.
- If the main block makes a reliable assessment first, the algorithm will wait for the first secondary result from any ensemble block.

- If the secondary result is identical to the primary one, the algorithm outputs the result. Otherwise, another secondary result is requested, and considered as the final assessment.
- If the first result is from an ensemble block, the algorithm will wait until the main block generates its assessment.
- If only one secondary result is made before the primary one, the algorithm compares them and makes a final assessment if they are identical. Otherwise, another secondary result is requested.
- If multiple secondary results are made before the primary one, the algorithm uses the more popular result as the final assessment.

They are summarized in Fig. 5. During this whole process, each block can only have one assessment at a time. New assessments from the same block overwrites the existing one.

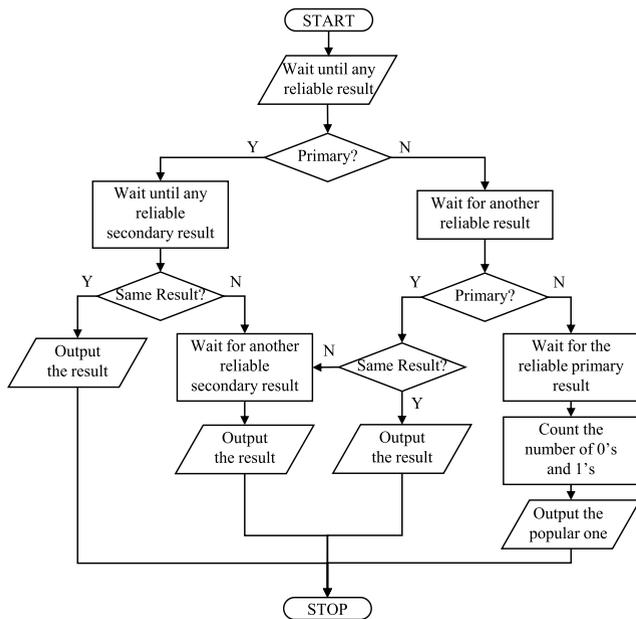


FIGURE 5. Flow chart of generating final assessment result with the rule-based decision machine.

F. WORK FLOW OF THE INTELLIGENT SYSTEM

The complete work flow of our proposed system can be divided into two phases as shown in Fig. 6, namely, an offline training phase and an online assessment phase. Utilizing the power system model as well as pre-defined contingency cases, post-contingency system dynamics are calculated in an offline manner. The dynamics are later employed to train the main block and ensemble blocks of the assessment system. Meanwhile, the power system topology is considered to develop network inputs of each ensemble block. After training all blocks in the system, their respective threshold values are optimized using the same input data. The optimized values as well as the blocks are regarded as the trained ensemble-based intelligent system for the online TSA process.

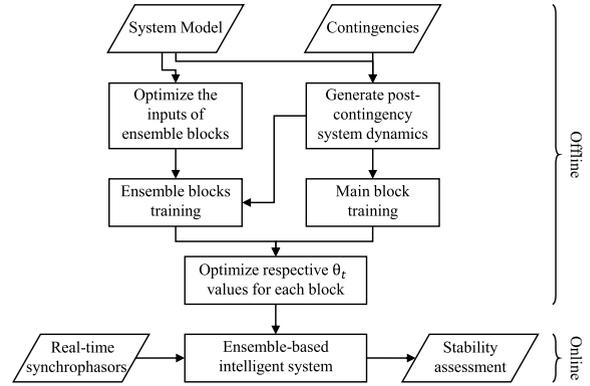


FIGURE 6. Work flow of the proposed intelligent system.

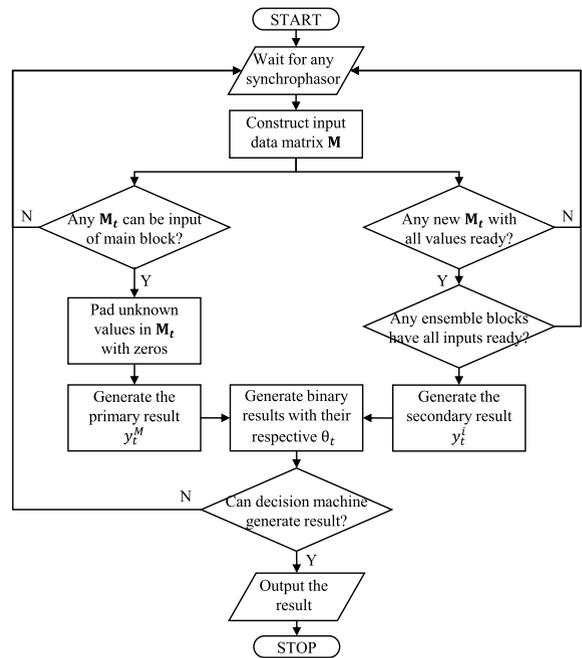


FIGURE 7. Online assessment of the proposed intelligent system.

Fig. 7 depicts the online work flow of our proposed ensemble-based intelligent system for TSA. The assessment is triggered when any new synchrophasor reaches the central controller. The synchrophasor is placed in the respective position in the input matrix M . The system then checks if either the main block or any ensemble blocks can generate new results y_t with the incomplete input matrix M . If so, the generated results are mapped to binary values using pre-optimized θ_t values. The calculated z_t results are placed in the decision machine for final assessment generation. The system stops when the machine outputs an assessment stating either the system will or will not be stable in the future.

IV. NUMERICAL STUDIES

We evaluate the performance of the proposed intelligent system through dynamic simulation on the New England 10-machine system benchmark [28] comprising 39 buses,

10 synchronous generators, 34 transmission lines, and 12 transformers. It has 19 constant loads consuming a total 6097.1MW and 1408.9 MVar. This system is an abstraction of an actual power system in New England [2]. Among all generators, G10 represents the aggregated generation from the rest of the eastern interconnection. All other generators are equipped with an IEEE Type-1 (IEEET1) exciter [29] and a WSCC Type G (BPA_GG) governor with parameters taken from [2]. PMUs are assumed installed on all buses.

A. TRAINING AND TESTING PREPARATION

The collection of training and testing cases are generated by time-domain simulation of post-contingency system dynamics. These cases comprise the PMU measurements of voltage phasors of all buses in the system and their transient stability 0-1 classifications for selected $N - 2$ contingencies. The loss of any transmission lines or transformers is considered as an $N - 1$ contingency. Then the loss of any remaining component is considered as an $N - 2$ contingency.

Furthermore, we consider four operating conditions in which the consumed power is set to 80%, 90%, 100%, and 110% of the nominal load level, respectively. A three-phase short-circuit fault is applied at either of the terminal buses of the removed components in $N - 2$ contingencies with a fault clearance time of 0.2 seconds. As a result, 4058 transient contingency cases are created. All cases are simulated for ten seconds after the fault clearance, and are considered unstable if any generator is out of step. The time-domain post-contingency system simulation is performed with DIGSILENT PowerFactory [30], and all simulations are conducted on an Intel Core i7 CPU at 3.40 GHz clock speed.

The generated transient contingencies are randomly divided into a training set and a testing set. To comply with the 3:1 training/testing ratio [15], 3044 cases are used for training the system, while the remaining 1014 are employed to test the system performance. By adopting this configuration, one can easily tell if the system has over-fit, in which the performance on the training set is superior but that on the testing set is unsatisfactory.

Based on the observation result of [31], the transmission delay $\Delta t_{p,t}$ is formulated as a shifted gamma distribution with $k = 20$ and $\theta = 2.0$. Although this delay model can only approximate the general one-way delay over the transmission network, real world delay data can be applied to our proposed system and the performance gain shall not be influenced significantly.

All system parameters are listed in Table 1. Among these defined parameters, the number of ensemble blocks N are

TABLE 1. Simulation Parameters.

System Parameters				
ϕ	0.5	N	3 to 7	ω 100
SSA Parameters				
$ pop $	30	max_iter	2000	r_a 1.0
p_c	0.7	p_m	0.1	

set to five values, i.e., 3 to 7. This means that in the test, Problem (4a) is optimized five times and each optimization generates N blocks. As a result, $3 + 4 + 5 + 6 + 7 = 25$ ensemble blocks are created to facilitate the generation of secondary results.

B. IMPACT OF DELAYED MEASUREMENTS

We first compare the proposed delay aware TSA mechanism, labeled by “Delay aware TSA” in the sequel, with conventional techniques in which the assessments are generated after receiving all measurements. For fair comparison, we employ the existing fastest TSA mechanism proposed in [15], labeled by “Synchronous TSA (STSA)”, for performance assessment. The communication latency values for both mechanisms are identical, and STSA is carried out when the control center can calculate the complete system state. In addition, we further assume that STSA can generate assessments with information of the first post-contingency cycle, and the calculation is instantaneous.

The simulation results are presented in Table 2, where the better performing results are in bold. In this table, the assessment accuracy and response times for both mechanisms are presented. The response times for synchronous TSA in practice are always greater than the listed value, thus prepended with “>” signs. It can be concluded that delayed measurements have a significant impact on the TSA response time. On average, the proposed delay aware TSA can achieve around 1.7x speedup than state-of-the-art conventional TSA mechanism. In addition, thanks to all the block networks in delay aware TSA, the proposed mechanism generates correct assessments in all test cases, and can provide almost perfect accuracy in the training cases.

TABLE 2. Comparison of assessment accuracy and response time.

Mechanism	Response Time (ms)			Accuracy (%)	
	Average	Best	Worst	Training	Testing
Delay aware TSA	48.0	35.4	79.9	99.8	100.0
Synchronous TSA	>82.6	>58.9	>141.3	N/A	99.4

For completeness of performance assessment, we also measure the average training time for the proposed delay aware TSA. On average, the training time for the main block is 1341 seconds, and that for each of the ensemble blocks is 176.2 seconds. Therefore, the whole system can be trained within 5746 seconds sequentially. Moreover, as the training of different blocks are independently, it is simple and effective to train the networks paralleled. In such a case, the proposed system can adapt to significant changes in operating conditions in little time. Note that the training data already contains different operating conditions. So insignificant changes can be addressed using the same network without re-training.

C. ASSESSMENT ACCURACY AND RESPONSE TIME

As presented in Section IV-B, the proposed delay aware TSA can achieve a superior performance compared with conventional mechanisms. It is of interest to determine which

component(s) of the mechanism contribute to the performance. In this test, we separate the delay aware TSA mechanisms into two sub-systems, composed of the main block and all ensemble blocks, respectively. For the ensemble blocks sub-system, final assessment is made when at least two ensembles generate identical secondary assessment results. In addition, we also investigate the impact of measurement inclusion parameter ϕ .

The simulation results are compared in Table 3. It can be seen that neither main block nor ensemble blocks can generate perfect assessment results, but combining them together with the decision machine yields superior performance. In addition, the proposed mechanism can have a slightly shorter average response time than either of the sub-systems. So we can conclude that the system performance is contributed to by all major components of the delay aware TSA, namely the main block, ensemble blocks, and the decision machine.

TABLE 3. Comparison of different variants of proposed delay aware TSA mechanism.

Mechanism	Response Time (ms)			Accuracy (%)	
	Average	Best	Worst	Training	Testing
Delay aware TSA	48.0	35.4	79.9	99.8	100.0
Main Block, $\phi = 0.5$	48.9	31.2	93.7	94.7	92.9
Ensemble Blocks	49.5	36.6	79.3	99.5	98.4
Main Block, $\phi = 1.0$	85.4	46.3	144.7	99.7	99.4
Main Block, $\phi = 0.8$	70.9	41.5	104.3	98.0	97.5
Main Block, $\phi = 0.6$	56.1	35.0	89.3	95.1	93.8
Main Block, $\phi = 0.4$	45.7	27.9	65.7	91.9	92.4

In addition, it can be summarized that the parameter ϕ plays an important role in balancing the trade-off between assessment accuracy and response time. While a larger ϕ leads to more accurate assessments, a small ϕ makes the system wait for less measurements, rendering a faster response speed.

The rates of assessments with respect to response time are depicted in Fig. 8 for a better understanding of the results. The figure depicts that the proposed delay aware TSA can generally develop most assessment in the shortest average response time. It can also be observed that while main block sub-system with $\phi = 0.5$ can start making assessment earliest among all mechanisms, it suffers from random latency spikes

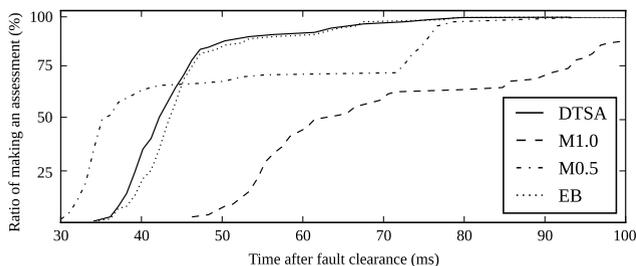


FIGURE 8. Rate of assessments with respect to response time. “DTSA” is the proposed delay aware TSA mechanism, “M1.0” is the main block sub-system with $\phi = 1.0$, “M0.5” is the main block sub-system with $\phi = 0.5$, and “EB” is the ensemble blocks sub-system.

in some test cases and has to wait for the delay measurements in these cases. Both conclusions accord with the results in Table 3, and demonstrate the superiority of the proposed mechanism.

D. TRANSIENT ASSESSMENT ON NOISY-DELAYED MEASUREMENTS

All previous simulation assumes that the PMU measurements, whenever they arrive at the central controller, can represent the system dynamic state accurately. However, in practice these measurements may be noisy. Therefore, additional numerical simulations are carried out to study the influence of noisy and delayed measurements on the proposed system for TSA.

According to IEEE Standard for Synchrophasor Data Transfer for Power Systems (C37.118.2-2011) [25], all PMUs complying with the standard shall generate system variable measurements with a total vector error less than 1%. Thus in this paper we follow the approach introduced in [17] to generate noisy test cases:

$$\tilde{V} \angle \tilde{\theta}_V = V \angle \theta_V + \Delta V \angle \Delta \theta_V, \tag{7}$$

where $\tilde{V} \angle \tilde{\theta}_V$ is the measured voltage phasor, $V \angle \theta_V$ is the actual voltage phasor, and $\Delta V \angle \Delta \theta_V$ is the noise phasor imposed, which satisfies a truncated complex normal distribution [17]. The newly generated test cases are employed to test the assessment performance of the proposed system trained using noiseless training cases.

The assessment result is summarized in Table 4 and Fig. 9. It can be observed from the comparison that data noise makes trivial influence on the assessment accuracy and response time. This is contributed by the outstanding classification ability of neural networks on noisy data [22], [26], and the introduction of multiple networks to form an ensemble [21].

TABLE 4. Comparison of noisy and noiseless data.

Mechanism	Response Time (ms)			Accuracy (%)	
	Average	Best	Worst	Training	Testing
Delay aware TSA	48.0	35.4	79.9	99.8	100.0
Noisy Data	48.2	31.7	83.5	99.8	99.9

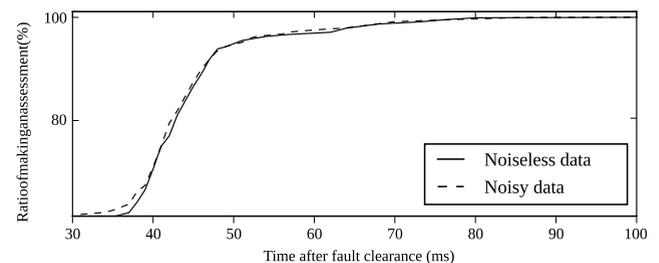


FIGURE 9. Performance of proposed intelligent system on making early assessments with noiseless and noisy data.

V. CONCLUSION

In this paper, an intelligent system is proposed to address the data transmission delay in an online TSA process. This system is based on Long-short Term Memory (LSTM) ensemble neural network with strategically designed decision machine. In particular, one large LSTM network called main block and multiple small LSTM networks called ensemble blocks cooperate to provide a collection of transient assessments considering different input data for the decision machine to make a final TSA conclusion. Different from existing TSA work, the proposed system can adapt to the delayed PMU measurements and make reliable assessments at the earliest possible time to facilitate later control actions. Simulation results show that the developed system outperforms conventional TSA methodologies in terms of the average response time and maintains a perfect assessment accuracy. In addition, the simulation demonstrates that both the main block and ensemble blocks contribute to superior TSA performance. The proposed system is also tested with noisy system measurement data, and the result indicates that the system is robust with noisy data complying with the related IEEE standard.

The contributions of this paper are summarized as follows:

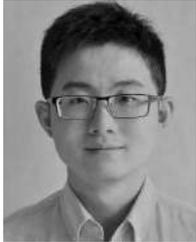
- We propose a delay aware transient stability assessment to address the impact of communication delay in the process of transient stability assessment. The delay is critical for fast response time, but has not been considered in the previous literature.
- We develop an LSTM ensemble-based intelligent system to handle the assessment problem with delayed and missing data caused by communication delay. The system can be further extended to more power system data-driven applications.
- We assess the system on a widely adopted testbed, and provide configuration guidelines to fully utilize its assessment capability. The simulation results demonstrate a significant improvement in system response time while maintaining a perfect accuracy.

Future efforts will focus on the availability of PMUs in the system. It is assumed that PMUs are installed on all available buses in the system, which may be practical due to the decreasing PMU prices. However it is still of interest to find the minimal number and locations of PMUs needed to fulfill the assessment task. Besides, different assessment predictors may contribute to a better average response time.

REFERENCES

- [1] P. Kundur et al., "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.
- [2] P. Demetriou, M. Asprou, J. Quiros-Tortos, and E. Kyriakides, "Dynamic IEEE test systems for transient analysis," *IEEE Syst. J.*, to be published. [Online]. Available: <http://ieeexplore.ieee.org/document/7151773>
- [3] J. C. Cepeda, J. L. Rueda, D. G. Colomé, and D. E. Echeverría, "Real-time transient stability assessment based on centre-of-inertia estimation from phasor measurement unit records," *IET Generat., Transmiss. Distrib.*, vol. 8, no. 8, pp. 1363–1376, 2014.
- [4] R. Diao, V. Vittal, and N. Logic, "Design of a real-time security assessment tool for situational awareness enhancement in modern power systems," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 957–965, May 2010.
- [5] K. Sun, S. Likhate, V. Vittal, V. S. Kolluri, and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and decision trees," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1935–1943, Nov. 2007.
- [6] Q. Gao and S. Rovnyak, "Decision trees using synchronized phasor measurements for wide-area response-based control," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 855–861, May 2011.
- [7] C. W. Liu and J. Thorp, "Application of synchronised phasor measurements to real-time transient stability prediction," *IEE Proc.-Generat., Transmiss. Distrib.*, vol. 142, no. 4, pp. 355–360, Jul. 1995.
- [8] M. Pavella, D. Ernst, and D. Ruiz-Vega, *Transient Stability of Power Systems: A Unified Approach to Assessment and Control*. Boston, MA, USA: Springer, Oct. 2000.
- [9] L. Zhu, C. Lu, and Y. Sun, "Time series shapelet classification based online short-term voltage stability assessment," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1430–1439, Mar. 2016.
- [10] T. Guo and J. V. Milanovic, "Probabilistic framework for assessing the accuracy of data mining tool for online prediction of transient stability," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 377–385, Jan. 2014.
- [11] F. Hashiesh, H. E. Mostafa, A.-R. Khatib, I. Helal, and M. M. Mansour, "An intelligent wide area synchrophasor based system for predicting and mitigating transient instabilities," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 645–652, Jun. 2012.
- [12] T. Guo and J. V. Milanović, "Online identification of power system dynamic signature using PMU measurements and data mining," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1760–1768, May 2016.
- [13] I. Kamwa, S. R. Samantaray, and G. Joos, "On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 152–161, Mar. 2012.
- [14] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Trans. Power Syst.*, to be published. [Online]. Available: <http://ieeexplore.ieee.org/document/7932885>
- [15] R. Zhang, Y. Xu, Z. Y. Dong, and K. P. Wong, "Post-disturbance transient stability assessment of power systems by a self-adaptive intelligent system," *IET Generat., Transmiss. Distrib.*, vol. 9, no. 3, pp. 296–305, 2015.
- [16] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proc. IEEE*, vol. 83, no. 10, pp. 1374–1396, Oct. 1995.
- [17] M. He, V. Vittal, and J. Zhang, "Online dynamic security assessment with missing pmu measurements: A data mining approach," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1969–1977, May 2013.
- [18] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stofopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.
- [19] F. Bai et al., "Measurement-based correlation approach for power system dynamic response estimation," *IET Generat., Transmiss. Distrib.*, vol. 9, no. 12, pp. 1474–1484, 2015.
- [20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
- [21] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 10, pp. 993–1001, Oct. 1990.
- [22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, Jul. 2015, pp. 1–9.
- [23] J. J. Q. Yu and V. O. K. Li, "A social spider algorithm for global optimization," *Appl. Soft Comput.*, vol. 30, pp. 614–627, May 2015.
- [24] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
- [25] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Standard C37.118.2-2011, Dec. 2011.
- [26] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, no. 2, pp. 4–22, Apr. 1987.
- [27] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Netw.*, vol. 1, no. 1, pp. 4–27, Mar. 1990.
- [28] A. Pai, *Energy Function Analysis for Power System Stability*. New York, NY, USA: Springer, 1989.

- [29] *IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*, IEEE Standard 421.5-2005, 2006.
- [30] *PowerFactory—DigSILENT Germany*. Accessed: Dec. 1, 2016. [Online]. Available: <http://www.digsilent.de/index.php/products-powerfactory.html>
- [31] A. Hernandez and E. Magana, "One-way delay measurement and characterization," in *Proc. 3rd Int. Conf. Netw. Services (ICNS)*, Jun. 2007, p. 114.



JAMES J. Q. YU (S'11–M'15) received the B.Eng. and Ph.D. degrees in electrical and electronic engineering from The University of Hong Kong, Hong Kong, in 2011 and 2015, respectively. He is currently a Post-Doctoral Fellow with the Department of Electrical and Electronic Engineering, The University of Hong Kong. His research interests include smart city technologies, power stability analysis, evolutionary algorithm design and analysis, and deep neural network applications.



the Department of Electrical and Electronic Engineering, HKU. His research interests include optimization theory and algorithms, evolutionary computation, smart grid, and smart city. He is a Croucher Research Fellow.

ALBERT Y. S. LAM (S'03–M'10–SM'16) received the B.Eng. (Hons.) degree in information engineering and the Ph.D. degree in electrical and electronic engineering from The University of Hong Kong (HKU), Hong Kong, in 2005 and 2010, respectively. He was a Post-Doctoral Scholar with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA, from 2010 to 2012. He is currently a Research Assistant Professor with



DAVID J. HILL (S'72–M'76–SM'91–F'93–LF'14) received the Ph.D. degree in electrical engineering from The University of Newcastle, Australia, in 1976. From 2005 to 2010, he was an Australian Research Council Federation Fellow with The Australian National University. He has also held academic and substantial visiting positions with the universities of Melbourne, California (Berkeley), Newcastle (Australia), Lund (Sweden), Munich, and Hong Kong (City and Polytechnic). He currently holds the Chair of electrical engineering with the Department of Electrical and Electronic Engineering, The University of Hong Kong. He is also a Part-Time Professor of electrical engineering with The University of Sydney, Australia. Since 1994, he has held various positions with The University of Sydney, including the Chair of electrical engineering until 2002 and again from 2010 to 2013 along with an ARC Professorial Fellowship.

His general research interests include control systems, complex networks, power systems, and stability analysis. His work is now mainly on control and planning of future energy networks and basic stability and control questions for dynamic networks. He is a fellow of the Society for Industrial and Applied Mathematics, USA, the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Hong Kong Academy of Engineering Sciences. He is also a Foreign Member of the Royal Swedish Academy of Engineering Sciences.



VICTOR O. K. LI (S'80–M'81–SM'86–F'92) received the B.S., M.S., E.E., and D.Sc. degrees in electrical engineering and computer science from MIT in 1977, 1979, 1980, and 1981, respectively. He has also served as the Associate Dean of engineering and the Managing Director of Versitech Ltd., the technology transfer and commercial arm of HKU. He served on the board of China.com Ltd., and serves on the board of Sunevision Holdings Ltd., listed on the Hong Kong Stock Exchange. He was a Professor of electrical engineering with the University of Southern California (USC), Los Angeles, CA, USA, and the Director of the USC Communication Sciences Institute. He is currently the Chair of information engineering, the Cheng Yu-Tung Professor in sustainable development, and the Head of the Department of Electrical and Electronic Engineering with The University of Hong Kong (HKU). His research interests include the technologies and applications of information technology, including clean energy and environment, social networks, wireless networks, and optimization techniques. He is a Registered Professional Engineer and a fellow of the Hong Kong Academy of Engineering Sciences, the IEEE, the IAE, and the HKIE. He has received numerous awards, including the PRC Ministry of Education Changjiang Chair Professorship at Tsinghua University, the U.K. Royal Academy of Engineering Senior Visiting Fellowship in Communications, the Croucher Foundation Senior Research Fellowship, and the Order of the Bronze Bauhinia Star, Government of the Hong Kong Special Administrative Region, China. Sought by government, industry, and academic organizations, he has lectured and consulted extensively around the world.

...