On Characterizing Information Dissemination During City-Wide Cascading Failures in Smart Grid

Mingkui Wei, Zhuo Lu, *Member, IEEE*, and Wenye Wang

***Abstract*—Although the smart gird is expected to eliminate cascading failures with the help of real-time system monitoring and control, it is yet unknown whether its underlying communication network is fast and reliable enough to achieve this goal. In this paper, we take an in-depth study on this issue by addressing three specific questions: 1) what is the evolution process of information dissemination and fault propagation in the smart grid?; 2) how to quantify the impact of cascading failures?; and 3) what are the conditions that information dissemination becomes either a booster or an adversary in mitigating cascading failures? To answer these questions, we build an innovative framework, the cascading failure with communications framework, to consolidate both communication networks and power grids, and provide quantitative evaluation on the impact of cascading failures. By studying and observing the progress of cascading failures in two city-wide power grids, we find that information dissemination is not always the winner in the race against fault propagation. Particularly, while fast and reliable communications can help in mitigating the consequences of cascading failures, anomalies such as massage delays may weaken its capability. Moreover, severely under-achieved communications, counter-intuitively, can even exacerbate the consequence of cascading failures.**

***Index Terms*—Communication networks, computer networks, cyber-physical systems, power system faults, power system protection, smart grids.**

I. Introduction

CASCADING failure is one of the most critical issues in the power system because it can lead to large area blackouts. The initiation of a cascading failure can be sporadic and accidental, such as a lighting striking and disconnecting a transmission line. If not being handled properly and timely, the disconnected transmission line may cause power flow redistribution, and potential overload and failure on other remaining lines.

Seemingly intuitive, cascading failures are extremely difficult to prevent [1] and can cause huge loss [2]. A closer inspection of recent large-scale blackouts reveals that one major cause of such events is the lack of real-time information exchange on a fast and reliable communication network [2]. This fact indicates that there is an urgent need to revamp the legacy monitoring and control systems to scale down, if not to eliminate, similar incidents in the future, and such a demand in essence calls for integrating advanced communications into next-generation power systems, i.e., the communication-assisted *smart grid*. With an advanced communication network, a system fault can be quickly located and isolated, thus, the further fault propagation can be eliminated. In this regard, there is no doubt that the concept of the smart grid is promising in mitigating or even preventing cascading failures.

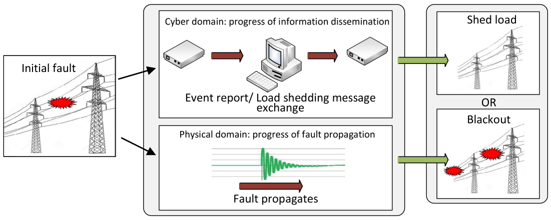

Initial fault
Cyber domain: progress of information dissemination
Event report/ Load shedding message exchange
Physical domain: progress of fault propagation
Fault propagates
Shed load
OR
Blackout

Fig. 1. Example of information dissemination and fault propagation.

This promising welfare, however, is based on an underlying assumption that communication networks can *always* achieve beneficial objectives in the smart grid, and this very assumption serves as the basis of most studies on cascading failures [3], [4]. Particularly, the communication network is first assumed to be ideal, i.e., with zero-delay and zero-packet loss, and advanced power system management schemes are then developed. In practice, however, messages in a communication network are subject to random delay, loss, or even cyber-attacks [5], which makes the ideal communication assumption doubtable. In this paper, we take an in-depth study on *information dissemination* against *fault propagation* in smart grids.

An example of this study is demonstrated in Fig. 1. The occurrence of the initial failure triggers a series of reactions in two domains. In the cyber domain, a failure is detected by the nearest monitoring devices and reported to a control center. The control center makes load shedding decisions and sends commands back to the controllers, trying to stop the fault propagation by shedding loads on buses. In the physical domain, the initial failure causes the power flow on remaining lines to gradually change (either increase or decrease), and potentially causes more failures due to overload. And the eventual consequence depends on whether information dissemination is fast enough to halt fault propagation in the power grid.

Our approach is to study the following three specific questions.

Manuscript received January 6, 2017; revised May 21, 2017; accepted September 30, 2017. Date of publication November 8, 2017; date of current version November 22, 2018. The work of Z. Lu was supported in part by the Office of Naval Research (ONR) under Grant N000141712109 and in part under the FC2 2017 seed grant. *(Corresponding author: Mingkui Wei.)*

M. Wei is with the Department of Computer Science, Sam Houston State University, Huntsville, TX 77341 USA (e-mail: mwei2@ncsu.edu).

Z. Lu is with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: zhuo.lu.us@ieee.org).

W. Wang is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606 USA (e-mail: wwang@ncsu.edu).

Digital Object Identifier 10.1109/JSYST.2017.2763462

1) What is the evolution process of information dissemination in communication networks and fault propagation in power grids?
2) How to quantify the impact of cascading failures?
3) What are the conditions that information dissemination becomes either a *booster* or an *adversary* in mitigating cascading failures?

To quantitatively evaluate the impact of cascading failures, we build an innovative framework, i.e., the cascading failure with communications (CFC) framework, to integrate both the communication network and the power grid for the system-wide simulation study with realistic settings. Based on the CFC framework, two city-wide power grids, IEEE 14-bus [6] and FREEDM 18-bus [7] systems are investigated. We observe that information dissemination does not always outrun fault propagation in the smart grid. In particular, we find that although fast and reliable communications can help in mitigating the consequence of cascading failures, anomalies such as massage delays can weaken its capability. Furthermore, severely under-achieved communications with extra long delays, ironically, may even exacerbate the consequence. To the best of our knowledge, it is for the first time that the race between the communication network and the power grid is systematically investigated with a wide range of settings.

The rest of this paper is organized as follows. In Section II, we review the background of cascading failure in smart grid, and related work in cascading failure study. In Section III, we introduce in detail the development of the *CFC* framework. In Section IV, we implement the *CFC* framework, and use it to evaluate the impact of cascading failures on two city-scale smart grids, and present and analysis the simulation result in Section V. Finally, we conclude our work in Section VI.

II. Backgrounds and Related Work

Topological modeling is the most widely used approach in recent cascading failure studies. In such modeling, the power system is modeled as a graph, where the substations or buses, are denoted by vertices, and the transmission lines are denoted by edges.

To start with, the power system is assumed to be running under steady state, in particular, power generation is sufficient to accommodate the consumption, and the power flow on transmission lines are blow their capacity. The numerical value of power flow on each transmission line under the steady state, i.e., the "predisturbance" power flow, can be calculated following classic power flow models, such as the direct current (dc) power flow model [4], [8], or the alternating current (ac) power flow model [9], [10].

The massive cascading failure usually starts with a small scale, or even single-point failure, which serves as the "trigger event." In the real-world power system, such triggering failure can be the result of various reasons, including inclement weather (e.g., lighting strikes), hostile natural environment (e.g., tree limbs touches or falls on transmission lines), and system malfunction (e.g., inadvertant device failure). For the purpose of cascading failure studies, however, it is not a big concern of the actual cause of a triggering event. Rather, we are more interested in understanding how such initial failure propagates across the system and causes more failure. Thus, in most existing simulation-based cascading failure studies, the initial failure is realized by selecting and removing a transmission line from the topological power system [4], [8], [10].

We use a 4-bus power system to demonstrate the commencement and progress of a cascading failure in Fig. 2. As shown by step $t = t_0$ in Fig. 2, we assume that line l_1 failed and is removed from the graph at time t_0, which necessitates the load on l_1 to be carried by remaining lines, and the actual value of which can be recalculated according to the power flow model as mentioned above. As a result of the load redistribution, it is likely some lines are added with more load than that they can carry. Such lines are considered overload and will be removed from the system again, which causes even more load to be added to even less remaining lines. Consequently, more overload and failure may occur to the system. As shown by step $t = t_i$ in Fig. 2, the failure of line l_1 causes failure on line l_2 at time t_i.

The failure of line l_2 causes generator G_2 to be isolated and consequently a power supply deficiency, i.e., power generation is less than power consumption. In this case, *load shedding* [11], [12] needs to be applied to regain a new balance of the system. In conventional power grids, load shedding is conducted with a *preset* manner, in which all loads have preconfigured priority, and it is always the load with the lowest priority be shed first, regardless of how much such actions contribute in stopping the failure propagation. Since the objective of the conventional load shedding is to re-establish power balance rather than dismiss transmission line overload, it is not guaranteed that overload, and thus failures, can be eliminated from the system. As shown by step $t = t_{ii}$ in Fig. 2, although power balance is achieved by shedding load on buses b_3 and b_4, line l_3 still becomes overloaded and fails.

On the contrary to conventional load shedding, smart grid affords *intelligent load shedding* [11], [13] by enabling critical information exchange among power devices. As shown by steps $t = t_1$ and $t = t_2$ in Fig. 2, the initial failure on line l_1 is detected and immediately reported by adjacent intelligent electronic devices (IEDs) to the control center, which then calculates an optimal solution that achieve multiple objectives including minimizing the cost of load shedding, regaining generation and consumption balance, and eliminating transmission line overload. This solution is then disseminated to all IEDs in the form of commanding messages. On receiving these messages, buses shed load as required and the cascading failure will be completely stopped beyond time t_2. This whole procedure seems plausible at the first glance, and as a matter of fact, it serves as as one fundamental assumption to most existing works in cascading failure study [4], [8], [10]. However, a critical factor that has been neglected in this procedure is the timing comparison between the message dissemination in the communication network, and the fault propagation in the power grid. As shown in Fig. 2, since the load shedding calculation is based on the information of failure on line l_1 and aims at stopping the failure that *will* happen on line l_2, if all messages can be correctly delivered *on time*, i.e., if $t_2 \leq t_i$, the cascading failure can be dismissed. If it is not the case, however, the cascading failure may not be completely stopped, and the fault will keep propagating.

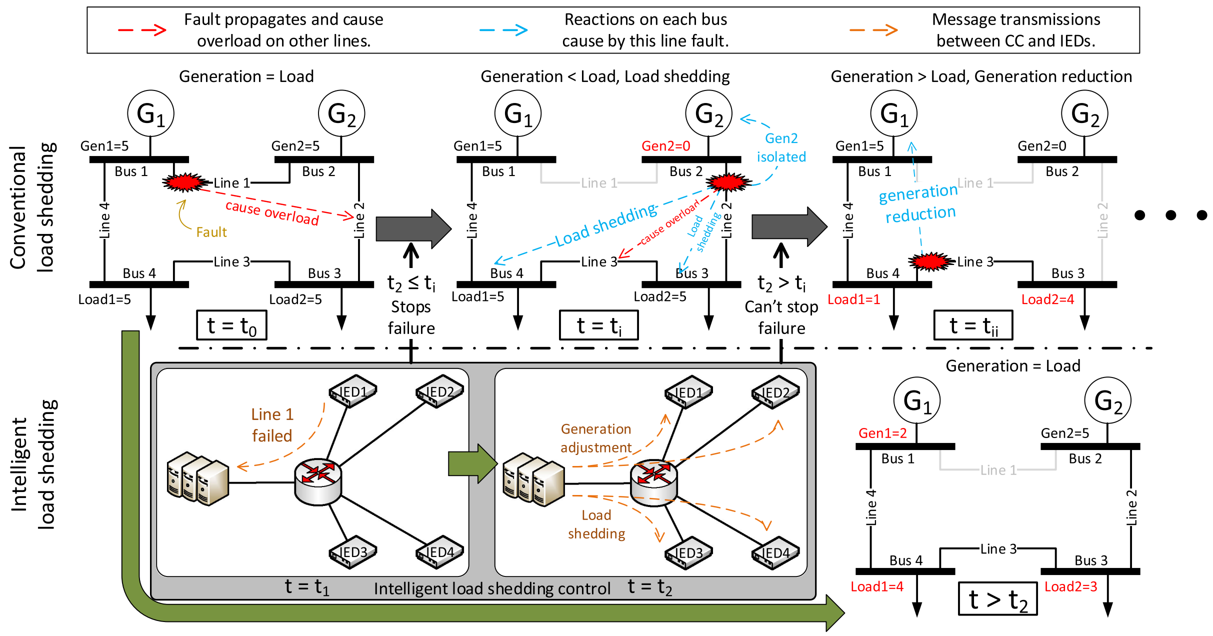


Fig. 2. Cascading failures in conventional power grid, and communication-assisted smart power grid.

Based on above demonstration, we argue that to understand the practical impact of cascading failures in the smart grid, it is critical to effectively evaluate and study the competition between the message dissemination and fault propagation with realistic settings. In the following, we showcase our approach in tackling this problem, and the conclusions drawn from our study.

III. CFCs Framework

In this section, we introduce the *CFC* framework, which includes the power grid system part, the communication network part and key modules to integrate the two systems together.

A. *Modeling of Power Grids*

Our objective is to consolidate the process of fault propagation in the power grids and the message dissemination in the communication networks, such that their real time interaction can be characterized. This objective essentially necessitates a new approach to inspect the power system from a time-wise perspective. During the progress of a cascading failure, the power system changes both topologically (e.g., a failed line will cause the removal of an edge from the graph) and numerically (e.g., load shedding will change the values of power injection on buses, and power flow on lines). To capture this dynamic system change, we model the power system with a graph whose status varies over time.

Denote a power system with n buses and m transmission lines by a directed graph $\mathcal{G} = (\mathcal{B}, \mathcal{L})$, where $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ denotes the set of buses, and $\mathcal{L} = \{l_1, l_2, \ldots, l_m\}$ denotes the set of lines. The topological evolution of this graph, i.e., the power grid, can be tracked by its *incidence matrix* at any given time t. More specifically, the *incidence matrix* of a directed graph is an $n \times m$ matrix, in which each column i has only a “1” and a “-1,” indicating the origin and ending nodes of the ith line. Therefore, if a line fails at time t, the corresponding column in the incidence matrix will be set to all 0. In our framework, the incidence matrix is a time-varying process that denotes the change of the power system over time.

B. *Modeling of Communication Networks*

We model the communication network as another graph that can be either dependent or independent from the power grid, and associate each bus in the power system with one IED in the communication network. Although in practice a substation consists of multiple sensors, controllers, and communication hosts, we hereby assume one IED on each bus as a symbolic aggregation for all monitoring and control components for simulation purpose. The IED equipped on a bus is in charge of monitoring bus operation, communicating with and executing instructions from the control center. The architecture of such a network can be either centralized, in which one control center governs all IEDs, or distributed where multiple control centers cooperate and control the overall system.

C. *Comprising Modules of the CFC Framework*

We demonstrate the *CFC* framework in Fig. 3. The implementation of this framework is based on iterative simulation. Each iteration takes six steps to complete, while each step is comprised with one or few modules, which are summarized in Fig. 4, and described in the following.

Step 1: Failure Initiation: In the *failure initiation* step, the *failure initiation module* is used to decide how, when, and where a cascading failure is to be triggered, as has been discussed in Section II.

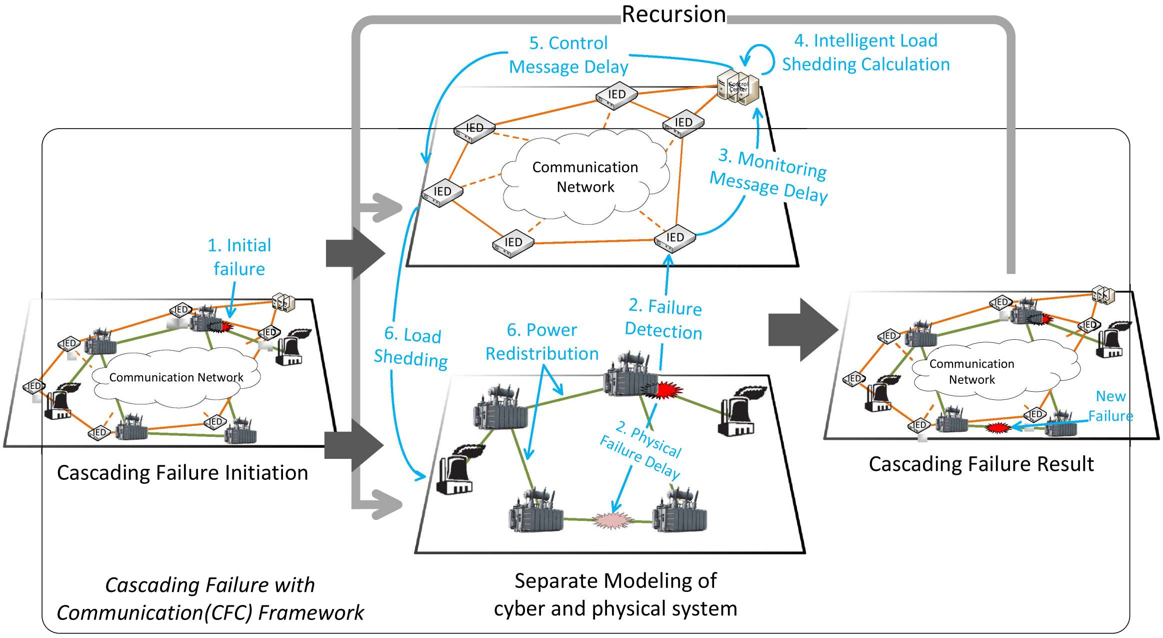
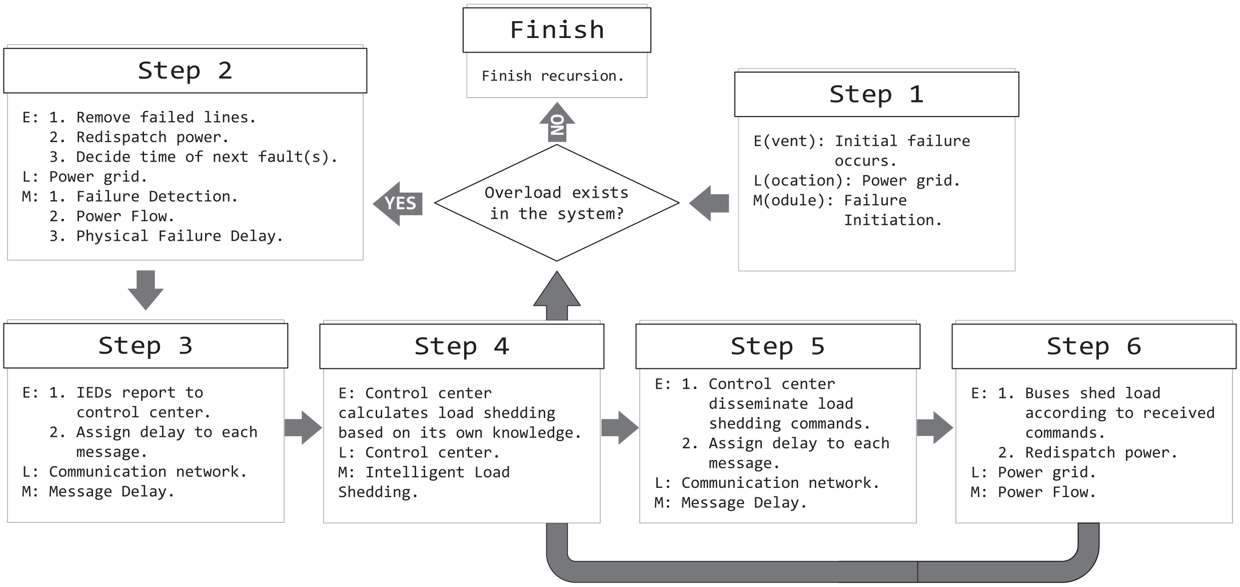

Recursion
5. Control Message Delay
4. Intelligent Load Shedding Calculation
IED
Communication Network
3. Monitoring Message Delay
1. Initial failure
6. Load Shedding
6. Power Redistribution
2. Failure Detection
2. Physical Failure Delay
New Failure
Cascading Failure Initiation
Cascading Failure Result
Cascading Failure with Communication(CFC) Framework
Separate Modeling of cyber and physical system

Fig. 3. CFC framework.


Finish

Finish recursion.

Step 1

E(vent): Initial failure occurs.
L(ocation): Power grid.
M(odule): Failure Initiation.

Overload exists in the system?

Step 2

E: 1. Remove failed lines.
2. Redispatch power.
3. Decide time of next fault(s).
L: Power grid.
M: 1. Failure Detection.
2. Power Flow.
3. Physical Failure Delay.

Step 3

E: 1. IEDs report to control center.
2. Assign delay to each message.
L: Communication network.
M: Message Delay.

Step 4

E: Control center calculates load shedding based on its own knowledge.
L: Control center.
M: Intelligent Load Shedding.

Step 5

E: 1. Control center disseminate load shedding commands.
2. Assign delay to each message.
L: Communication network.
M: Message Delay.

Step 6

E: 1. Buses shed load according to received commands.
2. Redispatch power.
L: Power grid.
M: Power Flow.

Fig. 4. Simulation sequence in each iteration.

Step 2: Failure Detection: The *failure detection module* is used to decide when and how an overloaded line will be removed. After the removal of overloaded lines, the power redispatch will be recalculated by the *power flow module*, which can be either the dc or the ac power flow model.

Meanwhile, the *physical failure delay module* will be used to determine the time interval (denoted as τ_p) after which the next failure(s) will happen. Note this time also defines the time span of current simulation iteration.

Step 3: Failure Report Delivery: Failure events will be reported from IEDs to the control center, and the *message delay module* is used to determine the properties of the communication channel such as message delay and packet loss.

Step 4: Intelligent Load Shedding Calculation: The control center is aware of the system topology at the beginning, and will update its knowledge each time it receives a failure report. At this step, the control center will calculate load shedding, not based on what is really happening in the system, but on its own knowledge. Remind that because of the delay of event report messages, the control center may not be able to know the exact system topology at this time.

The *Intelligent Load Shedding Module* chooses the load shedding algorithm. For instance, the load shedding can be performed as solving an optimization problem, which retains the most possible load in the system while eliminating all possible overload on existing lines [8], [14].

Step 5: Control Message Delivery: The result of the intelligent load shedding calculation is a vector that contains the value of load shedding on each bus, and this message is then disseminated by the control center to all IEDs. The *message delivery module* is applied again on these control messages.

Step 6: Load Shedding and Power Redispatch: Load will be shed immediately when it is received at corresponding buses, and *Power Flow Module* will be used again for power redispatch. This power redispatch may cause overload and failures on more lines (that will happen at time τ_p), which serve as the first step of the next iteration.

At this time, it is possible that some messages, including both failure reports and control messages, may have not been delivered, and all undelivered messages will be carried on to the next iteration.

The iteration of the simulation will continue until at least one of the following three criteria is met:

1) no more overload exist on any lines;
2) all lines have been removed;
3) all generators have been isolated.

D. Summary

In this section, we demonstrate the modeling of the *CFC* framework. The modular design of the framework provides flexibility such that different modules can be chosen to meet various objectives. For instance, message delays in the *message delay module* can be set as simple as a constant value, and the *power flow module* can be either dc or ac power flow model to achieve different accuracy. This framework can also be adapted for more complex system interactions, such as the *interdependent networks* [15], in which the power grid and the communication network depend on each other and the failure in one domain impacts the other.

IV. Framework Implementation

In the previous section, we have described the modeling of the *CFC* framework, and demonstrated in Fig. 3 that how this framework can capture the real-time interaction between the power system and the communication network. In this section, we implement this framework in order to observe and evaluate this interaction.

A. Power System Prototypes

A practical question of cascading failure study lies in the scale of the power systems. The power system is one of the largest and most complicated infrastructures, which is deployed nation-wide and contains hundreds of thousands of interconnected power devices. On the other hand, however, in the microscopic perspective the power grid consists of numerous autonomous small grids that are run by local utility companies. Therefore, it is critical to find suitable scale of the smart grid to carry out cascading failure studies.

In this paper, we choose to study the cascading failure based on power systems with the *city-scale*, i.e., systems with the size comparable to a city area, for the following reasons.

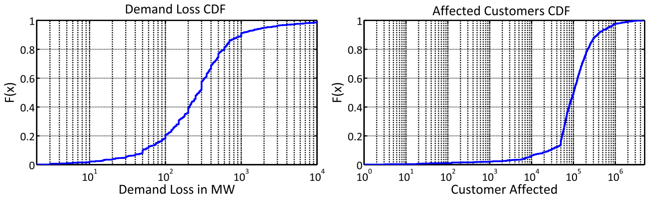


Fig. 5. Empirical CDF of “demand loss” and “customer affected” for blackout events happened during 2002 and 2015 in the U.S.

First, by analyzing historical data of electric disturbance events in the recent decades [1], we observed that most blackout events happen in the city scale. In Fig. 5, we present the empirical CDF of the demand loss and affected customers of blackout events during 2003 and 2015 in the US. We observe that more than 90% events have their demand loss less than 1000 MW and the number of affected customers less than half million, which typically fits a medium-sized city.

Second, large scale blackouts are usually results of multiple causes, e.g., human errors and control software bugs played important roles in the 2003 Southeast Blackout [2], which are difficult to be accurately modeled.

Based on above two reasons, in this paper, we choose to study two city-scale systems, namely the the FREEDM 18-bus system [7] and the IEEE 14-bus system [6]. The FREEDM 18-bus system (also known as the *GreenHub*) is built at the NC State University for smart grid research, which contains 18 buses and 24 transmission lines. Meanwhile, the IEEE 14-bus system is a generic power system model that has been widely used for power system study. The detailed information for both systems can be easily found in recent literatures such as [7] and [6], and, therefore, is omitted here.

B. Communication Use Cases

We adopted a centralized control scheme since we are studying a relatively small scale power system. In particular, we assume that there is one control center in both systems, and each IED has a direct link to the control center to achieve the best communication performance. We develop four communication use cases to study the reaction of cascading failures to communications with different delay performances.

1) Ethernet Communication (EthBase): In this case, all IEDs and control center are directly interconnected with high-speed Ethernet links. The data rate for each link is set to be 10 Mbps. Each IED sets up a TCP session with the control center and periodically sends system updates. The sample rate (i.e., message sending rate) is set as 50 samples per second according to practical power devices [16], [17], and the size of a message uniformly distributes between 256 and 2560 bytes, i.e., 1 and 10 Modbus packets [18].

2) Ethernet Communication Under Distributed Denial-of-Service (DDoS) Attack (EthDos): Cyber security is one of the biggest concerns in the smart grid [19], and in this case we consider the communication network that is under cyber-attacks. In particular, based on the Ethernet network stated above, we assume an attacker launches a DDoS attack, i.e., the attacker breaches the communication network, and intensively sends useless data to cause congestion on the communication channel. In this case, we assume there are four attackers who send messages with length uniformly distributed between 5K and 10K bytes, and the time interval between two messages is uniformly distributed between 0.01 and 0.001 s. Note that this setting result a rather "mild" attack, nevertheless, it is intensive enough to slow down the legit message exchange, and make significant impact to the procedure of stopping a cascading failure.

3) Wireless Communication (WifiBase): Because of its ubiquitous availability, the wireless technology such as Wifi or 3G is desirable to be integrated in smart grid operation [20]. Nevertheless, concerns also rise that it may not be suitable for critical infrastructures because it is slower and less reliable compared to wired communication. In this case, we assume that IEDs communicate the the control center via wireless communication instead of wired link. We set the Wifi standard to be 802.11 g, which provides 54-M bandwidth.

4) Wireless Communication Under DDoS Attack (WifiDos): In this case, we study the performance of wireless communication that is under DDoS attack as well. We assume the same intensity of DDoS attack as in case 2. In particular, each legitimate host sends messages with size ranging from 256 to 1024 bytes (uniformly distributed), with the frequency of 50 messages per second using 802.11-g standard. And there are four malicious hosts who send messages with the size uniformly distributed between 5K and 10K bytes, and sending interval uniformly distributed between 0.01 and 0.001s.

C. Implementation of CFC Modules

1) Failure Initiation Module: At the beginning of each simulation, we randomly choose one line and disconnect it to serve as the initial failure.

2) Power Flow Module: We choose to use the dc power flow model in this implementation, which provides a good balance between accuracy and complexity, and is also used in most existing cascading failure studies [4], [8], [14]. To facilitate further illustration, we hereby denote the power flow as a vector $F = [f_1\ f_2 \dots f_m]$, where m is the number of transmission lines in the power system before any failure happens, and $|f_i|, i \in [1, m]$ is the value of power flow on line l_i, while the polarity of f_i indicates its direction.

3) Failure Detection Module: Assume a power system is running at stable state, i.e., the power flows have reached equilibrium and do not change on all lines, denote the vector of *stable power flow* as

$$\hat{F}_s = [\hat{f}_1\ \hat{f}_2\ \dots\ \hat{f}_m] \tag{1}$$

in which $\hat{f}_i, i \in [1, m]$ is the power flow on the ith line.

Accordingly, denote the vector of *power flow threshold* as

$$\hat{F}_{\text{th}} = [\epsilon_1 \hat{f}_1\ \epsilon_2 \hat{f}_2\ \dots\ \epsilon_m \hat{f}_m] \tag{2}$$

where $\epsilon_i, i \in [1, m]$ is called *safety factor*, which determines line l_i's tolerance to overload. Intuitively, a larger value of ϵ_i means the system is more robust against failure, however, in the meanwhile more system capacity will be wasted. We choose $\epsilon_i = 1.1, \forall i$ in this study for better demonstration purpose, which also complies with existing study [4].

4) Load Shedding Module: The load shedding [21] can be formulated as an optimization problem, whose objective is to minimize the cost of load shedding while maintaining the balance between power generation and consumption.

The objective of optimal load shedding can be written as

$$\min \Sigma_{\forall b_j \in \mathcal{B}} (w_j \cdot |\Delta p_j|) \tag{3}$$

where Δp_j, for $j \in [1, n]$, is the load to be shed on the jth bus, and w_j is the *unit cost* for load shedding (e.g., dollars per kilowatt), which is used to differentiate load priorities and is set as $w_j = 1, \forall j$ in our study. The optimization in (3) is subject to constraints of power balance, generation, and load capacity, as well as overload requirements [4], [8]. It is then solved with classic linear programming solvers.

5) Message Delay Module: The delay of each message in such a homogeneous network is assumed to be an identically and independently distributed (i.i.d) random variable, and we obtain its distribution using real-time simulations. In particular, we model and simulate the communication network within OMNeT++ [22] according to the topology of the power system and the communication use cases. We then collect message delays during the simulation as samples and use curve fitting to find the empirical distribution.

6) Physical Fault Delay Module: It is critical yet challenging to find the accurate value of τ_p, i.e., the time interval between two consecutive failures. Because of the complexity of power systems, it is computationally unfeasible to exhaustively identify all possible combinations of the line failure sequences, and record the exact time for each combination. In this implementation, we assume τ_p between any two consecutive failures is an *i.i.d* random variable, and use real-time simulation to find its distribution.

In particular, we build the power system in PSCAD [23], a high-fidelity power system simulator. To start with, we begin the simulation and allow the power system to operate until it reaches steady state. Then, we choose one transmission line from the system and trip it to serve as the initial failure. We monitor all other transmission lines concurrently, and record the time when the power flow on these lines exceed their capacity (1.1 times of power flow in steady state), if the overload ever happens. We run this simulation multiple times with each time choosing different lines to trigger the cascading failure, and use the collected sample time to fit for an empirical distribution.

It is worth noting that in this setup we do not consider the reaction time of relays and circuit breakers. In a real-power system, different types of relays are deployed at different location to serve different purposes. And the reaction time of relays can ranging from less than 1 ms, e.g., with the solid state relay, to 10 ms or even longer for conventional mechanical relays. Thus, assuming a relay reaction time on each bus not only causes excessive complication to the simulation procedure, but also makes the study overly scenario dependent. Furthermore, including the relay reaction time does not change the fact that possibility still exists where a control message comes later than physical fault, therefore, it won't affect the nature that cascading failure cannot be stopped under imperfect communication.

D. Simulation Implementation

The simulation is conducted in MATLAB. The implementation of the framework includes the follow steps.

1) Matrix-Representation of the Power System: The power system is abstracted into an *incidence matrix*. Particularly, a power system with n buses and m edges formulates a $n \times m$ matrix. Each of the m columns of the matrix represents the beginning and ending of a transmission line, e.g., if line l_1 begins from bus b_1 and ends at b_2 (the direction of a line can be arbitrary chosen and won't affect the result), then the first column of the matrix will have the first element as 1, the second element as -1, and all other elements as 0.

2) Power Flow Calculation: The power flow on each transmission lines can be calculated with knowledge of the incidence matrix, and the power injection on each bus. The detailed procedure is omitted, which can be easily found in the literature, such as [4].

3) Overload Detection and Load Shedding Calculation: When a line is tripped, it is removed from the graph, which changes the topology. Accordingly, the column representing the tripped line will be erased from the incidence matrix (in case a bus is completely isolated, the row that representing the bus will be erased).

Then, power flow on remaining lines will be recalculated, based on the principle in the previous step. The new power flow will be compared with a line's capacity, and overloaded lines will be marked.

In the meanwhile, load shedding algorithm will be running to calculate which bus will shed load, and how much load need to be shed, such that all overload can be eliminated.

4) New Failure Occurrence: A physical fault delay is assigned to each overloaded line to represent the time when the line will be tripped.

On the other hand, the result of the load shedding calculation will be sent to buses whose load will be shed, and each such message is assigned with a message delay.

Whenever a message arrives at a bus, load will be shed at that bus, and power flow will be recalculated accordingly. When this procedure hits the time instance when a line ought to fail, this line will be tested again for overload. If this line is still under overloaded condition, the line will be tripped and triggers a new round of failure.

V. Simulation Result

We evaluate the result of cascading failures with two metrics, i.e., the percentage of overloaded lines, and the percentage of lost load.

Definition 1 (Percentage of overloaded lines): The *percentage of overloaded lines at time t*, denoted as $N(t)$, is defined as

$$N(t) = \frac{\Sigma_{i \leq k} N_i(t)}{k \cdot m} \tag{4}$$

in which m is the number of lines in the power system at the beginning, $N_i(t)$ is the number of overloaded lines at time t in the ith simulation run, and k is the total number of simulation runs. Denote $N = N(\infty)$, i.e., the percentage of overloaded lines at the end of a cascading failure.

Similarly, we define the *percentage of lost load* as follows.

Definition 2 (Percentage of lost load): The *percentage of lost load at time t*, denoted as $P(t)$, is defined as

$$P(t) = \frac{\Sigma_{i \leq k} P_i(t)}{k \cdot P_{\text{total}}} \tag{5}$$

in which $P(t)$ is the amount of load that has been lost (shed due to load shedding, or disconnected due to line failure) at time t in the ith simulation run, and k is the total number of simulation runs. And P_{total} is the summation of total load in the original power grid. Denote $P = P(\infty)$, i.e., the percentage of lost load at the end of a cascading failure.

The simulation was carried on MATLAB, and for each setup in the following, the results were averaged over 25 000 simulation runs. Since most results for the 14-bus system and GreenHub follow the same trend but differ only in the value, in the following we only discuss the results for GreenHub. The result of the 14-Bus system is provided as a reference in the Appendix.

A. Does Communication Help? How?

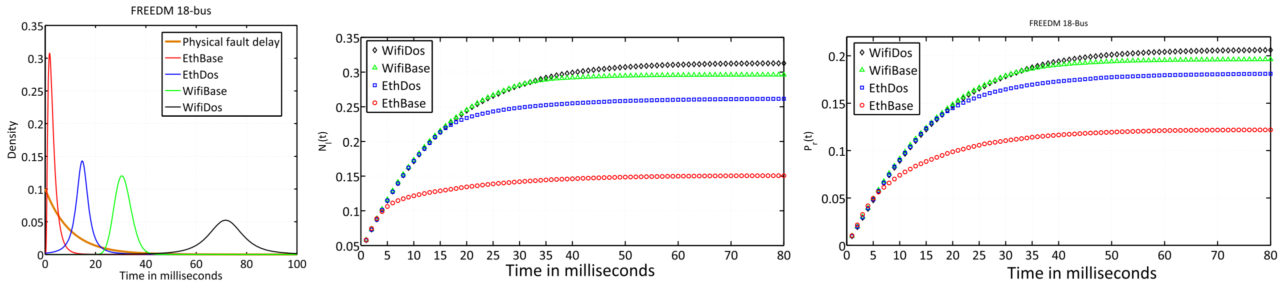
In prior to answer these questions, we first present the comparison of the message delays in four communication cases and the physical delay in the power grid, such that the readers can gain an intuitive sense on their performances. We plot the curve-fitted distribution for all four cases in Fig. 6(a), from which we can observe that among four cases, *EthBase* has the best delay performance, and the case *WifiDos* has the worst.

1) Aggregated Results: In Fig. 6(b) and (c), we plot $N(t)$ and $P(t)$ for the GreenHub under 4 use cases.

By looking at Fig. 6, our first intuition is that communications with shorter delays do help on mitigating the consequence of cascading failures. As shown in all figures, the *EthBase* case achieves much better performance, i.e., smaller value in both $N(t)$ and $P(t)$, than other three cases.

We also notice that, for the first few milliseconds, the differences between the four cases are almost indistinguishable. Taking Fig. 6(c) as an example, for the first 6 ms, all four lines have almost the same value, which means during this period of time, physical fault is the dominant force, and the improvement on communications can hardly provide any benefit. Therefore, we can conclude that *excessive upgrade on communication networks may not be able to provide increasingly better results*.

For the third observation, we find that the differences between all cases tend to diminish as the delay performance becomes worse. Therefore, *the contribution of communication networks does not have a linear relationship with their performances*. This can be explained by observing the message delay distribution compared with physical fault delay distribution in Fig. 6(a). Since the value of $N(t)$ or $P(t)$ essentially depends on the probability that the message delay is shorter than the physical fault delay, as the messages delay getting larger, this probability



(a) (b) (c)

Fig. 6. $N(t)$ and $P(t)$ for FREEDM 18-bus system. (a) PDF of physical delay and message delay. (b) $N(t)$ for GreebHub. (c) $P(t)$ for GreenHub.

TABLE I
MEAN, MIN, MAX VALUE FOR N AND P OF GREENHUB

	Case	Mean	Max	Min	≥ NoComm Max
$N_i, i \in k$	NoComm	7.50	15	4	-
	EthBase	3.29	20	1	291/25 000 (1.2%)
	EthDos	3.62	20	1	459/25 000 (1.8%)
	WifiBase	7.11	19	1	107/25 000 (0.4%)
	WifiDos	7.51	20	1	39/25 000(0.2%)
$P_i, i \in k$	NoComm	3041	10 893	638	-
	EthBase	1810	14 823	0	382/25 000 (1.5%)
	EthDos	2692	14 823	0	506/25 000 (2.0%)
	WifiBase	2910	14 823	0	312/25 000 (1.2%)
	WifiDos	3058	10 893	0	75/25 000 (0.3%)

gradually approaches to 0 with slower changing speed. Practically, this observation suggests the necessity and significance of a fast and reliable communication network.

2) Statistical Results: Although we have shown above with aggregated results that communications do help in alleviating the consequences of cascading failures, it is still not safe to conclude that communications can *always* benefit the smart grid. In the following, we present statistical results that are recorded during all 25 000 simulation runs.

In Table I, we present the mean, min, and max value of $N_i(\infty)$ and $P_i(\infty)$ for $i \in k$ (denoted as N_i and P_i), i.e., *number* of overloaded lines and *amount* of lost load, that are recorded during the 25 000 runs. As a comparison, in these tables we also present the result of "NoComm," which denotes the case where communications do not exist in the power grid. In this case, if load imbalance happens, we simply increase the generation at all generators or decrease the load at all buses by the same percentage, as a special case of conventional load shedding where all loads have the same priority.

It is shown in Table I that in all the four cases, the value ranges for both N_i and P_i have been enlarged. Taking N_i as an example, in *NoComm* case, if a cascading failure happens, for the best case, there will be only 4 overloaded lines as the result, whereas in the worst case this number will be 15. Remind that the total number of lines in this system is 24. Therefore, a cascading failure in this system can only cause impacts with moderate significance. The existence of the communications, however, makes the minimum value smaller and the maximum value larger. For instance, for *EthBase*, in the best case, there will be only 1 line fault, however, the maximum value of N_i is enlarged to 20, meaning more severe events can happen in this *EthBase* case.

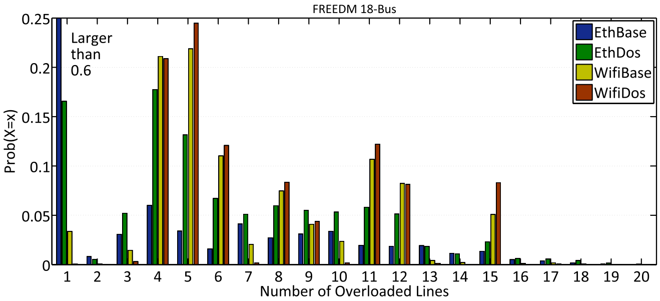


Fig. 7. Histogram of N for GreenHub of 25 000 simulation runs.

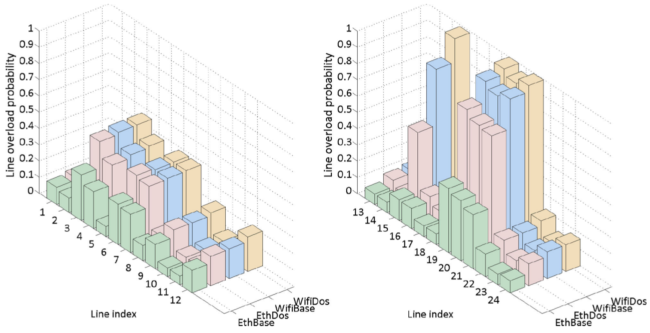


Fig. 8. Line failure probability for FREEDM 18-bus system.

In the last column of Table I, we list the number of events whose N_i is larger than 15, i.e., the maximum value in *NoComm*, and calculate its percentage during the 25 000 simulation runs. Ironically, we observe that this percentage does not increase monotonically as message delays getting worse. For the two *Wifi* cases, while their mean values grow to be much larger than those of the two *Eth* cases, their "worse than NoComm max" percentages become smaller. To investigate the cause, in Fig. 7, we plot the histogram of N_i, from which we observe that *longer delay only increases probability of medium-sized blackouts*. In particular, we see that for the two *Wifi* cases, in most cascading failure events, the number of overloaded lines falls between 4 and 15, while the probability, it is smaller than 4 or larger than 15, has been depressed significantly compared with the two *E*th cases.

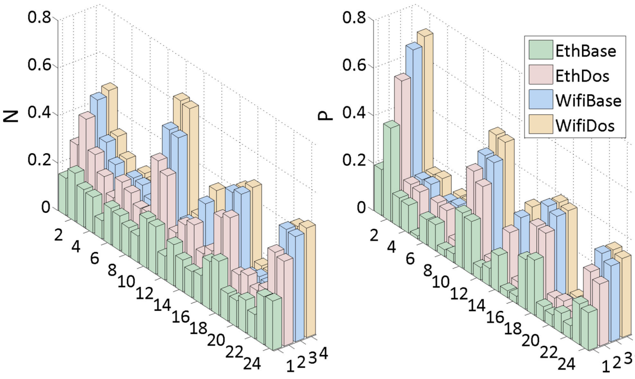


Fig. 9. Line contribution for FREEDM 18-bus system.

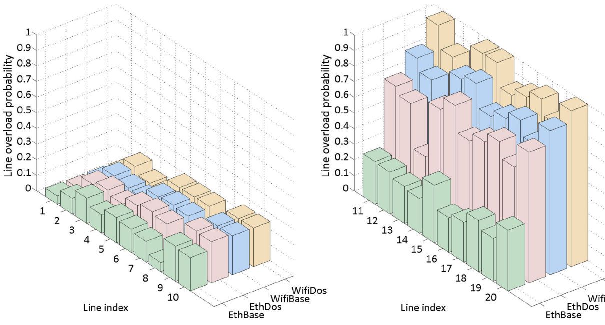


Fig. 10. Line failure probability for IEEE 14-bus system.

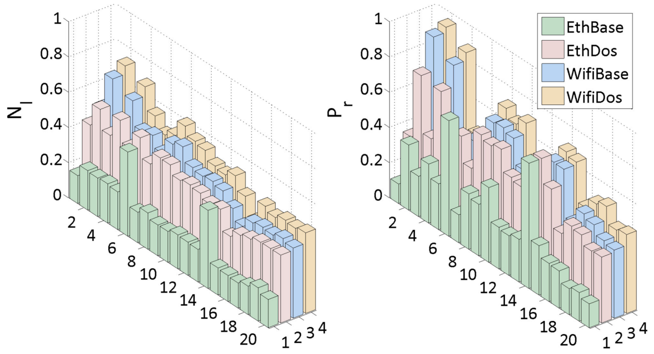


Fig. 11. Line contribution for IEEE 14-bus system.

TABLE II
MEAN, MIN, MAX VALUE FOR N AND P OF IEEE 14-BUS

	Case	Mean	Max	Min	Worse Than NoComm
N	NoComm	10.34	15	7	-
	EthBase	4.20	17	1	92/25 000 (0.36%)
	EthDos	8.64	18	1	117/25 000 (0.47%)
	WifiBase	9.47	17	1	34/25 000 (0.14%)
	WifiDos	10.33	16	1	6/25 000(0.02%)
P	NoComm	97 427	163 880	74360	-
	EthBase	48 242	168 860	0	377/25 000 (1.5%)
	EthDos	85 116	168 860	0	766/25 000 (3.1%)
	WifiBase	91 692	168 860	0	677/25 000 (2.7%)
	WifiDos	97 177	163 880	0	113/25 000 (0.5%)

We also observe that for the *WifiDos* case, the values of both N_i and P_i are slightly larger than those of *NoComm*, which indicates, counter-intuitively, that *severely under-achieved communications can be an adversary in mitigating cascading failures.*

B. How Does Individual Line React to Cascading Failures?

1) Line Failure Probabilities: We define the *failure probability* of a line as the percentage of the event in which this line fails, compared to the 25 000 simulation runs. The failure probability for GreenHub is plotted in Fig. 8 (IEEE 14-bus result is plotted in Fig. 10). From this figure we can make a number of observations.

We first observe that there are some lines that are particularly vulnerable to cascading failure. Such lines include line l_{15}, l_{19}, l_{20}, and l_{21}. Except the *EthBase* case, we find the failure probability of these three lines are significantly higher than that of others. In reality, we may consider either to increase the capacity of these lines, or deploy extra lines to offload them.

Our second observation is that the improvement made on communication does not generate equal benefit on these lines. In particular, the most vulnerable lines gains the most benefit. Take line l_{15} as an example. Its failure probability in the *WifiDos* case is almost 1, but it drops substantially to be lower than many of the others in the *EthBase* case. On the contrary, benefits on lines that are already robust can be negligible, such as on lines l_9 and l_{12}. This phenomenon in essential suggests us to consider the tradeoff between improving the communication network and upgrading the power gird. For instance, improving the communication and bring its performance from *WifiDos* to *EthBase* may becomes less desirable compared with upgrading the few vulnerable lines.

2) Line Contribution to Cascading Failures: Having seen how a line is susceptible to cascading failures, in this subsection we explore from the reverse direction, that is, how a cascading failure is susceptible to the failure of a particular line.

In Fig. 9 we plot the average values of N and P for the cascading failure events which are triggered by the failure of a particular line in the GreenHub (IEEE 14-bus result is plotted in Fig. 11). From this figure, it is interesting to observe that there exist some lines whose failure can cause devastating impact to the power grid. For instance, the failure of line l_2 under the $WifiDos$ case will cause more than 70% of the total load to be lost, and more than 50% of lines to be tripped. On the other hand, lines such as l_{15} and l_{16} has only minimum impact to the overall system.

These two figures suggest the weakest points of a smart grid power system. Lines such as l_2, l_{10}, and l_{11} should be actively monitored to prevent them from any inadvertent failures. From the communication side, these lines should be effectively protected from cyber-attacks, since such weak points are usually the target whose failure can leverage the impact of an attack.

Comparing Fig. 11 with Fig. 10, we are able to find a weak negative correlation between the vulnerability and the influence of a line. Particularly, lines with less failure probability tend to have larger influence. This observation essentially suggests different protection strategies for two types of lines: for lines who are vulnerable to cascading failures, the focus should be put on protect them from being impacted by other lines' failure, such as increasing their threshold to tolerant larger overload;

while for lines whose failures trigger severe consequences, the focus should be put on preventing them from being self-failure and triggering a cascading failure.

VI. Conclusion

In this paper, we studied the interaction between information dissemination in communication networks and fault propagation in power grids. We built an innovative framework, the *CFC* framework, to capture the dynamics between the two domains, implemented it with two smart grid prototypes, and evaluated the impact of communications in alleviating the consequence of cascading failures. We found that information dissemination does not always outrun fault propagation in power grids, and severely under-achieved communication can be an adversary and exacerbate the cascading failure instead. Our study bears significance in that it demonstrates the importance of fast and reliable communication in the smart grid.

Appendix

Simulation result for the IEEE 14-Bus system, which shows similar result as those of the GreenHub, is shown in Table II.

References

[1] Electric Disturbance Events (OE-417) Annual Summaries. [Online]. Available: https://www.oe.netl.doe.gov/OE417_annual_summary.aspx

[2] B. Liscouski and W. Elliot, "Final report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendations," U.S. Dept. Energy, Tech. Rep., vol. 40, no. 4, 2004.

[3] M. Rahnamay-Naeini and M. M. Hayat, "On the role of power-grid and communication-system interdependencies on cascading failures," in *Proc. IEEE Global Conf. Signal Inf. Process.*, 2013, pp. 527–530.

[4] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures—Analysis and control implications," in *Proc. IEEE INFOCOM*, 2012, pp. 2634–2642.

[5] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Comput. Netw.*, vol. 57, pp. 1344–1371, 2013.

[6] "Power systems test case archive." [Online]. Available: http://www.ee.washington.edu/research/pstca/

[7] A. Q. Huang, M. L. Crow, G. T. Heydt, J. P. Zheng, and S. J. Dale, "The future renewable electric energy delivery and management (FREEDM) system: The energy internet," *Proc. IEEE*, vol. 99, no. 1, pp. 133–148, Jan. 2011.

[8] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "An initial model for complex dynamics in electric power system blackouts," in *Proc. IEEE 34th Annu. Hawaii Int. Conf. Syst. Sci.*, 2001, pp. 710–718.

[9] J. Zhu, *Optimization of Power System Operation*. Hoboken, NJ, USA: Wiley.

[10] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with dc power flow model and transient stability analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 285–297, Jan. 2015.

[11] D. Xu and A. A. Girgis, "Optimal load shedding strategy in power systems with distributed generation," in *Proc. IEEE PES Winter Meeting*, 2001, vol. 2, pp. 788–793.

[12] J. Tang, J. Liu, F. Ponci, and A. Monti, "Adaptive load shedding based on combined frequency and voltage stability assessment using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 2035–2047, May 2013.

[13] A. Arulampalam and T. K. Saha, "Fast and adaptive under frequency load shedding and restoration technique using rate of change of frequency to prevent blackouts," in *Proc. IEEE PES Gen. Meeting*, 2010, pp. 1–8.

[14] I. Dobson, J. Chen, J. Thorp, B. A. Carreras, and D. E. Newman, "Examining criticality of blackouts in power system models with cascading events," in *Proc. IEEE 35th Annu. Hawaii Int. Conf. Syst. Sci.*, 2002, 10 pp.

[15] X. Lu, W. Wang, J. Ma, and L. Sun, "Domino of the smart grid: An empirical study of system behaviors in the interdependent network architecture," in *Proc. IEEE SmartGridComm*, 2013, pp. 612–617.

[16] Wally RTU. [Online]. Available: http://www.teamware.it/

[17] Varec 8130 Remote Terminal Unit, [Online]. Available: http://www.varec.com/docs/

[18] T. Morris, R. Vaughn, and Y. Dandass, "A retrofit network intrusion detection system for MODBUS RTU and ASCII industrial control systems," in *Proc. IEEE 45th Hawaii Int. Conf. Syst. Sci.*, 2012, pp. 2338–2345.

[19] M. Wei and W. Wang, "Greenbench: A benchmark for observing power grid vulnerability under data-centric threats," in *Proc. IEEE INFOCOM*, 2014, pp. 2625–2633.

[20] A. R. Devidas and M. V. Ramesh, "Wireless smart grid design for monitoring and optimizing electric transmission in India," in *Proc. IEEE SENSORCOMM*, 2010, pp. 637–640.

[21] X. Lu, Z. Lu, W. Wang, and J. Ma, "On network performance evaluation toward the smart grid: A case study of DNP3 over TCP/IP," in *Proc. IEEE GLOBECOM*, 2011, pp. 1–6.

[22] A. Varga *et al.*, "The omnet++ discrete event simulation system," in *Proc. EUROSIS ESM*, 2001, vol. 9, 11 pp.

[23] O. Anaya-Lara and E. Acha, "Modeling and analysis of custom power systems by PSCAD/EMTDC," *IEEE Trans. Power Del.*, vol. 17, no. 1, 2002, p. 266–272.

Mingkui Wei received the M.S. degree in computer engineering from Southeast University, Nanjing, China, in 2008, and the Ph.D. degree in computer engineering from North Carolina State University, Raleigh, NC, USA, in 2016.

From 2008 to 2011, he was with Bell-labs, Alcatel-Lucent, as an Engineer. He is currently an Assistant Professor with the Department of Computer Science, Sam Houston State University, Huntsville, TX, USA. His research interests include cyber security in cyber physical systems, especially the smart gird. He is also interested in software and network security and digital forensics.

Zhuo Lu (GS'10–M'14) received the Ph.D. degree from North Carolina State University, Raleigh, NC, USA, in 2013.

He is an Assistant Professor with the Department of Electrical Engineering, University of South Florida, Tampa, FL, USA. He is also affiliated with the Florida Center for Cybersecurity and with the Department of Computer Science and Engineering. He currently leads the Communications, Security, and Analytics (CSA) Laboratory, University of South Florida. His research interests include modeling and analytical perspectives on communication, networks, and security. His recent research is equally focused on practical and system perspectives on networking and security.

Dr. Lu is a member of the ACM.

Wenye Wang received the M.S.E.E and Ph.D degrees in computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1999 and 2002, respectively.

She is an Professor with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. Her research interests include mobile and secure computing, modeling and analysis of wireless networks, network topology, and architecture design.

Dr. Wang has been a member of the Association for Computing Machinery since 1998 and a member of the Eta Kappa Nu and Gamma Beta Phi honorary societies since 2001. She was the recipient of the NSF CAREER Award 2006. She was a co-recipient of the 2006 IEEE GLOBECOM Best Student Paper Award—Communication Networks and the 2004 IEEE Conference on Computer Communications and Networks Best Student Paper Award.