# Summary of Changes to Submission RESS_2019_1308: Identification of Interdependencies and Prediction of Fault Propagation for Cyber-Physical Systems

We thank the reviewers and the editor-in-chief for the time and expertise they have invested in reviewing our submission. We have carefully revised the paper in response to this constructive feedback.

To facilitate the review process for the second submission of the manuscript, we have included a verbatim copy of the full decision letter including reviewers' comments. For each concern raised by the reviewers, we have provided our responses under respective comments. Where applicable, we have also explained how we have addressed the concerns and provided verbatim copies of the revision(s) made in the new submission.

---

# 1 Decision Letter

Dear Dr. Sedigh Sarvestani,

Thank you for submitting your manuscript to Reliability Engineering and System Safety.

I have completed my evaluation of your manuscript. The reviewers recommend **reconsideration of your manuscript following minor revision and modification**. I invite you to **resubmit your manuscript after addressing the comments below**. Please resubmit your revised manuscript by **Oct 11, 2020**.

When revising your manuscript, please consider all issues mentioned in the reviewers' comments carefully: please outline every change made in response to their comments and provide suitable rebuttals for any comments not addressed. Please note that your revised submission may need to be re-reviewed.

To submit your revised manuscript, please log in as an author at `https://www.editorialmanager.com/jress/`, and navigate to the "Submissions Needing Revision" folder under the Author Main Menu.

Reliability Engineering and System Safety values your contribution and I look forward to receiving your revised manuscript.

Please remember to sign in to receive the list of contents of next issues of RESS in: `https://www.sciencedirect.com/science/serial-alerts/save/09518320`

Kind regards
Prof. Carlos Guedes Soares
Editor-in-Chief
Reliability Engineering and System Safety

# 2 Reviewers' Comments

## 2.1 Reviewer 1

In this paper, identification of interdependencies and prediction of fault propagation are studied for cyber-physical systems. However, the innovation is weak, and there are some issues in this paper.

The comments are as follows:

1. In Introduction, the research background is not clear, and the research motivation is not specified.

2. The main contribution and the novelty of the paper should be described in detail.

3. What is the relation between identification of interdependencies and prediction of fault propagation? Are the two independent?

4. How are the interdependencies identified in the simulations? Please explain it clearly.

5. The prediction of fault propagation is a general neural network prediction. What is the innovation?

6. What is the basis for selecting the number of hidden layers and hidden layer nodes of the ANN?

7. On page 19, "while the validation dataset is employed to minimize overfitting during the training cycles" is mentioned. Please explain it clearly.

8. This paper lacks comparisons with other related papers in simulation. It's essential to embody the superiority of the control method in this paper.

9. The conclusion should be concise.

To address the problem mentioned in item 1, we included the following sections that explain the background and motivation of the work, respectively.

Section 1

The tight coupling of the physical and cyber networks are through links and various means of interactions, which makes them intertwined systems with intense interdependencies. *Interdependence* is a general term that accounts for a relationship among components of a system, where state of each component influences or is correlated with the state of others [1]. Interdependence can be categorized into different types and may exist at different levels, i.e., between systems, subsystems, or components. For example, a malfunctioning fuel pressure sensor in a car affecting its fuel economy manifests a functional interdependency in the engine subsystem. In an interdependent CPS, impairments originating in cyber and physical components may propagate through unprotected channels and escalate into a system-level failure. In general, interdependencies in a system can severely complicate the development of accurate models. This paper seeks to present a model of interdependency for CPSs, and develop a tool for predicting the components that are in risk of failure by reason of an earlier impairment.

```
Section 1


There are numerous examples of large-scale systems being affected by

impaired components or due to impairments in other correlated systems.

The Arizona-Southern California blackout in September 2011, which occurred

as a result of an 11-minute outage of a 500 kV transmission line is an

example of cascading failures with profound consequences due to internal

interdependencies within the boundaries of the Southwestern US power grid.

The outage triggered a cascade that left approximately 2.7 million customers

without power for up to 12 hours [2].  The Italy blackout, occurred in

September 2003, is another example in which interdependencies among different

infrastructures caused further outages [3].  This failure was exacerbated

by the loss of Internet communication nodes left without power, which in

turn caused further breakdown of communication and control at multiple

power stations.  Such incidents motivate our work on identification of

cyber-physical interdependencies and prediction of imminent failures across

the physical and cyber domains.
```

To address the concerns specified in items 2, 3, and 5, we modified a section
of the paper as shown below.

Section 1

** OLD **

In this paper, our objective is to present a model of interdependency by observing sequence of failures for a set of failure cases, and subsequently, providing a tool for prediction of failure sequences. We identify interdependencies among components of a CPS using correlation and causation analyses and quantify them with dependency metrics presented in our previous work [15]. Using the knowledge of the interdependency, we develop an artificial neural network for prediction of fault propagation paths and cascading failures. To illustrate our approach, we have applied it to two smart power grids based on the well-studied IEEE 14-bus and 57-bus test systems.

** NEW **

In this paper, our objective is to present a model of interdependency by observing sequence of failures for a set of failure cases and capturing the extent of interdependence with quantitative metrics as well as to provide a tool for prediction of failure sequences. We identify interdependencies among components of a CPS using correlation and causation analyses and quantify them with dependency metrics presented in our previous work [15]. Using the knowledge of the interdependency, we identify areas of the system where faults can propagate and create larger failures. We then create a set of failure cases and use the data from the resulting failure sequences to develop and configure an artificial neural network for prediction of fault propagation paths and cascading failures. Our research paves the road to a better understanding of interdependencies in a system and prediction of its future failures using historical failure records or data from simulation of failure scenarios. To illustrate our approach, we have applied it to two smart power grids based on the well-studied IEEE 14-bus and 57-bus test systems.

In response to the concern in item 5, the novelty of the work lies in the link from the ''identification of interdependencies'' to the ''prediction of failures''. The neural network method is only used to demonstrate that the underlying knowledge of interdependency is beneficial in determining the optimal configuration for achieving the best prediction performance.

Regarding the concern in item 4, we explain our approach in the following portion of Section 2.2.

---

Section 2.2

Our approach to identification of dependencies relies on observation of the system's behavior in response to each of a set of failure cases. The disruption associated with each failure case triggers a failure sequence, from which we infer dependency links between the components. The extent of dependency can be quantified with statistical methods such as correlation or causation analysis. The greater the number of failure cases, and the longer the duration of observation for each, the more accurate this inference will be. For collecting the failure sequences, it is possible to utilize data from simulation, laboratory and/or field observation, and/or historical data about failures of a system. In this paper, we present our results using data from a simulation environment.

---

Additionally, please refer to Sections 3.1 through 3.3 where we explain all the steps that we take for selection of the failure cases, simulating the failure scenarios, and extracting quantitative interdependency metrics from failure sequences observed in the simulation results.

Regarding the concern mentioned in item 6, please note that this study is not focused on the technical aspects of the machine learning tools. We only demonstrate applicability of such techniques in the field of CPS interdependency. Therefore, we consciously restricted the extent of technical information regarding the neural network architecture, configuration, and tuning to prevent distraction from the main focus of the study. However, to address the concern of the respected reviewer, we have added the following section.

> Section 2.4
>
>
> The choice of ANN architecture is generally based on heuristic rules and
> is only for the sake of demonstrating applicability of the method.  The
> choice of number and size of the hidden layers is according to the previous
> studies on the performance of different ANNs in classification and time
> series prediction problems [21].  The bottleneck hidden layer structure
> used in this work is known to create a compressed representation of the
> information, and hence, acts as a nonlinear transformation and dimensionality
> reduction stage for the inputs.  The architecture presented here has shown an
> excellent performance on the test cases investigated in Section 3; however,
> depending on the type and size of the system under test, reconfiguration and
> adjustments may be necessary.

For item 7, we modified the following section and better explained the reasons
for splitting failure data to training, validation, and test.

```
Section 3.4


** OLD **

The training dataset is used for adjusting the weights of the ANN shown in
Section 2.4 by backpropagation technique, while the validation dataset is
employed to minimize overfitting during the training cycles.  Upon completion
of the training, the test dataset is used for measuring the predictive
performance of the ANN.

** NEW **

The training dataset is used for adjusting the weights of the ANN shown in
Section 2.4 by backpropagation, while the validation dataset is employed as
a recurrent feedback loop to minimize prediction error during the training
cycles.  This separation of training and validation datasets is done to
prevent overfitting of the neural network.  An overfitted model is one that
corresponds very closely to the training dataset, and hence, may fail to
make accurate predictions on future observations.  Upon completion of the
training, the test dataset is used for measuring the predictive performance
of the ANN.
```

In response to the concern mentioned in item 8, our intent in this paper has been to only focus on a method for capturing interdependency between components of a CPS and exploring the possibility of predicting unforeseen failures using a model that has been trained with a small data from past failures. Although we have briefly described our simulation environment as well as the reasons for choosing specific cyber control schemes, we believe that comparing our simulation and control method with those presented in other studies will distract readers from the main topic. Since have discussed our simulation environment and control strategy in our previous studies, we added references to these publications [26, 27] in Section 4.

Regarding the comment in item 9, we removed the first sentence as it only presented a background for the work. All other discussions are either related to

```
key research findings or future directions of the research and can not be further

summarized.
```

## 2.2 Reviewer 2

In this work, the authors identify the inter-dependencies among the components of a cyber physical system using correlation metrics. They have presented a model of inter-dependency for cyber physical systems and developed a tool for predicting the components that are in risk of failure. Authors talked about physical-physical, physical-cyber, cyber-cyber, cyber-physical dependencies.

The format and literary presentation of the paper is satisfactory. The structure of the paper is deliberate and transparent. The work is presented in 4 sections. Inter-dependency concepts and fault prediction concepts are well explained. Reference are written in an adequate manner and results are properly demonstrated with tabular results and graphs. The list of symbols and abbreviations is also given for the better understanding of the mathematical terms.

Some minor suggestions are mentioned here.

1. Section 2 is very lengthy. It may be divided in two parts for better readability.

2. The causation dependency should be more clearly defined with supportive example.

3. The difference between cyber-physical and physical-cyber dependency should be explained with some clear example.

4. Authors did not mentioned that how the causation dependencies can be identified.

5. In equation 2, alpha value is set to 0.9. Its reason should be mentioned.

```
    For item 1, we restructured the paper and divided this section to

''Interdependency Analysis'' and ''Prediction of Failure Sequences''.

    For item 2, we modified the following section and added an example.
```

Section 2.2

** OLD **

Interdependency between components can be due to causality or simply
a correlation.  In a causation relationship, state of a component is
responsible for that of another.  On the other hand, state of two components
are correlated when they have a statistical relationship, whether causal
or not.  Depending on the purpose of interdependency analysis, either
correlation or causation may be of interest.

** NEW **

Interdependency between components can be due to causality or simply
correlation.  In a causation relationship, the state of a component is
responsible for that of another.  On the other hand, the state of two
components are correlated when they have a statistical relationship, whether
causal or not.  In a simple system with a processor that controls two
actuators, it may be observed that the two actuators fail together, however,
no cause-and-effect relationship is established.  In fact, further analysis
may reveal that the failure of the processor unit causes failure of both
actuators, hence indicates simultaneity (a type of correlation) between
failure of the two actuators without one being the actual reason for failure
of the other.  Depending on the purpose of interdependency analysis, either
correlation or causation may be of interest.

In order to address the concern mentioned in item 3, we modified the following
section and added an example.

Section 2.1


** OLD **

Depending on the source and destination of an edge, it can represent one of four types of dependency, namely, physical-physical, physical-cyber, cyber-physical, and cyber-cyber. Note that in this notation, $s_1 \quad - \quad s_2$ dependence represents a relation in which components of subsystem 1 ($s_1$) influence components of subsystem 2 ($s_2$). As an example, Figure 1 illustrates the dependency graph for a hypothetical CPS.


** NEW **

Depending on the source and destination of an edge, it can represent one of four types of dependency, namely, physical-physical, physical-cyber, cyber-physical, and cyber-cyber. Note that in this notation, $s_1 \quad - \quad s_2$ represents a relation in which components of subsystem 1 ($s_1$) influence components of subsystem 2 ($s_2$). Figure 1 illustrates a dependency graph for a hypothetical CPS and specifies these four types of dependency. Note that in this context, ''physical-cyber'' and ''cyber-physical'' represent different types of dependencies. For example in the 2003 Italy blackout [3], the loss of telecommunication services that occurred following the initial power outage manifests a physical-cyber dependency, while the additional outages in the grid that took place because of the loss of control in the power stations indicates a cyber-physical dependency.

For item 4, please note that Section 2.2.2 is devoted to identification of causation dependencies. We also added a sentence to this section to specify that the validation of the proposed technique (whether a cause-and-effect relationship truly exists) is shown in the study that inspired our work.

> ```
> Section 2.2.2
>
>
> For validation of the basic proposed approach readers are referred to the
>
> original study [17].
> ```

To address the concern mentioned in item 5, we added the following section and a new figure (Figure 2) to support our claim.

> ```
> Section 2.2.2
>
>
> In Equation (2), α controls the threshold in detecting the causative
>
> relationships.  Selection of α depends on the distribution of w_{ij} among
>
> the members of F_k set.  In the histogram shown in Figure 2, we observe a
>
> significant separation between the members of F_k at about w_{ij}  =   0.9, which
>
> allows us to easily distinguish correlative and causative relationships.  For
>
> the systems studied in this work, the w_{ij} histograms follow a similar pattern
>
> to what is shown in Figure 2.  Therefore, in all following analyses, we set α
>
> to 0.9 (shown as a red dotted line on Figure 2), i.e., H_{k,j} captures members
>
> of F_k whose w_{ij} value is larger than 0.9 (on the right of the red dotted
>
> line).
> ```

## 2.3   Reviewer 3

This paper identify the critical interdependencies and predict the fault propagation for power CPS. The authors should address the following issues.

1. In the interdependency identification, the authors employ the method of reference [17] to analyze the causation. Compared with [17], the author considers the impacts of FACT and PMUs on the system. However, this paper still focuses on identifying the critical interdependencies among branches.

2. The contribution is not clear. The contribution is to propose a new method or to apply the old method to the failure propagation analysis of CPS?

3. The authors should explain the differentiation of correlation and causation, for example, the role of correlation and causation in fault propagation. The reviewer suggests that the authors give an example to explain it.

4. In section 3, the authors claim "PMU device is disabled as soon as a voltage violation occurs at the bus on which it is installed". What is the basis of this assumption?

5. In section 3, the authors don't consider the communication failures in the paper. If the authors do not consider the communication failures, this paper still focuses on studying the fault propagation of pure power system. In this paper, what is the interdependency of cyber-cyber.

6. The authors employ the state variable to calculate the RDC. The state variable is power flow?

7. The authors should explain how to propagate a fault after the PMU or FACTS fails.

8. The authors should employ a large-scale system to verify the effectiveness of proposed method.

9. Please give some diagrams to explain figs 2-4 although reference [17] has been explained clearly.

Overall, we recommend the major revision.

In response to item 1, we argue that our work has addressed the needs for a CPS interdependency analysis method. The framework presented in reference [17] is only limited to the conventional power systems, while our improvements add a layer of abstraction and allow modeling of cyber components with discrete-time behaviors. Therefore, our method can be applied to different CPSs with minimal changes. To better explain the differences between reference [17] and our work, we added some explanations to Section 2.2.2 as shown below.

Section 2.2.2

In general, a causal relationship is harder to establish than correlation, and hence, fewer interdependency studies have investigated causality. We use a method inspired by the interaction model introduced in [17] to identify causation relationships and estimate $\mathbf{D}$. The work presented in [17] determines the interactions among components of a power grid, finds key dependency links, and provides strategies for mitigating cascading failures using a heuristic method. We present a similar method that is generalized to be applicable for cyber-physical systems. Specifically, we have extended the method to incorporate heterogeneous components, control the sensitivity in detecting causality, and account for dependency relationships between degraded states rather than binary states. We utilize a cyber-physical simulation environment and account for discrete-time behaviors of the cyber infrastructure, which is not possible in the OPA simulation environment employed in [17]. For validation of the basic proposed approach readers are referred to the original study [17].

To address the concern specified in items 2, we modified a section of the paper as shown below and better explained our objectives and contributions.

Section 1


** OLD **

In this paper, our objective is to present a model of interdependency by
observing sequence of failures for a set of failure cases, and subsequently,
providing a tool for prediction of failure sequences.  We identify
interdependencies among components of a CPS using correlation and causation
analyses and quantify them with dependency metrics presented in our previous
work [15].  Using the knowledge of the interdependency, we develop an
artificial neural network for prediction of fault propagation paths and
cascading failures.  To illustrate our approach, we have applied it to two
smart power grids based on the well-studied IEEE 14-bus and 57-bus test
systems.

** NEW **

In this paper, our objective is to present a model of interdependency by
observing sequence of failures for a set of failure cases and capturing the
extent of interdependence with quantitative metrics as well as to provide
a tool for prediction of failure sequences.  We identify interdependencies
among components of a CPS using correlation and causation analyses and
quantify them with dependency metrics presented in our previous work [15].
Using the knowledge of the interdependency, we identify areas of the system
where faults can propagate and create larger failures.  We then create a set
of failure cases and use the data from the resulting failure sequences to
develop and configure an artificial neural network for prediction of fault
propagation paths and cascading failures.  Our research paves the road to
a better understanding of interdependencies in a system and prediction of
its future failures using historical failure records or data from simulation
of failure scenarios.  To illustrate our approach, we have applied it to
two smart power grids based on the well-studied IEEE 14-bus and 57-bus test
systems.

For item 3, we modified the following section and added an example.

Section 2.2

** OLD **

Interdependency between components can be due to causality or simply
a correlation.  In a causation relationship, state of a component is
responsible for that of another.  On the other hand, state of two components
are correlated when they have a statistical relationship, whether causal
or not.  Depending on the purpose of interdependency analysis, either
correlation or causation may be of interest.

** NEW **

Interdependency between components can be due to causality or simply
correlation.  In a causation relationship, the state of a component is
responsible for that of another.  On the other hand, the state of two
components are correlated when they have a statistical relationship, whether
causal or not.  In a simple system with a processor that controls two
actuators, it may be observed that the two actuators fail together, however,
no cause-and-effect relationship is established.  In fact, further analysis
may reveal that the failure of the processor unit causes failure of both
actuators, hence indicates simultaneity (a type of correlation) between
failure of the two actuators without one being the actual reason for failure
of the other.  Depending on the purpose of interdependency analysis, either
correlation or causation may be of interest.

In response to the concern mentioned in item 4, we added a brief explanation in
the manuscript as shown below.

> Section 4
>
> In addition to the intrinsic functional dependencies between components of
> smart grids, we assumed that the operation of PMU devices depends on the
> underlying power grid, i.e., a PMU device is disabled as soon as a voltage
> violation occurs at the bus on which it is installed.  Voltage violation
> is defined to be outside of 0.9 to 1.1 per-unit range, according to the
> EN-50160 standard [28].  Note that we consciously simplify this dependency
> (no backup power or fallback mechanism) as it allows us to better understand
> the consequences of the dependency of the cyber network on the physical
> system.

Regarding the concern mentioned in item 5, we clarified our approach and
reasoning by including some additional explanations to Section 4.1 as shown below.

> Section 4.1
>
> It is worth mentioning that upon availability of simulation environments
> capable of modeling the communication infrastructure with high resolution,
> considering the effects of respective impairments will improve the quality of
> the model.  Unless a sophisticated model for channel impairments is utilized,
> inclusion of communication failures simply adds redundant failure cases and
> complicates representation of the results without actually capturing the
> behavior of data transport mechanisms.  Therefore in this work, we have
> aggregated the failures of each communication link with those of the cyber
> component that receives the respective data.  For example, a corrupted
> data package transmitted to a FACTS device is captured as a FACTS failure.
> Incorporating the manifestations of communication impairments guarantees
> that the model, despite the simplification, maintains all cyber-physical
> interdependencies.  For instance, influence of a faulty PMU data on the
> operation of the decision support algorithm and corresponding FACTS device
> represents cyber-cyber dependencies.

For the question asked in item 6, please note that Equation (9) defines state variables for transmission lines, FACTS devices, PMU devices, and the decision support platform. For the case of transmission lines, we have used the magnitude of active power flow in per unit.

To address the concern in item 7, we explained the propagation of faults in the simulation environment in Section 4.2 as shown below.

Section 4.2

The simulation environment is used to determine power flows and voltages in
the system during the failure cases. For each failure case, specific faults
are injected to predetermined components of the system by disabling FACTS
devices, PMU devices, and decision support as well as tripping transmission
lines. At each time step, PSAT performs power flow analysis and determines
active power flow on each line and voltage at each bus. Active power flow of
the lines are compared to their capacity, and if any line is overloaded,
it is considered failed and the topology is updated accordingly. Since
the cyber-physical dependencies are incorporated in the data models and
enforced by the simulation environment, the propagation of the injected
faults takes place automatically. For example, when the decision support
platform receives a faulty data from a malfunctioning PMU, it is likely to
send incorrect commands to respective FACTS devices. The FACTS devices will
in turn apply wrong compensations and decrease power transfer capability of
the system. The resulting load imbalance can overload and trip transmission
lines and cause further cyber and/or physical failures. The simulation
continues this automatic propagation of faults until no further failures
are detected.

In response to item 8, we argue that the usage of IEEE-14 and IEEE-57 with 27 and 100 components, respectively, allows us to verify the scalability of our approach. In our proposed approach, the main concern in terms of applicability to large-scale systems is related to the prediction of failures using the neural network tool. We have shown in Section 4.4 that our proposed method is very

16

scalable as it can provide very similar performance metrics (Table 4) by only
doubling the size of the training dataset (from 8,000 to 16,000) while the size
of the system has increased from $n = 27$ (IEEE-14) to $n = 100$ (IEEE-57). In order
to better clarify this claim and present our reasoning, we revised parts of Section
4.4 as shown below.

> Section 4.4
>
> As mentioned earlier, we have performed the case study on a larger test
> system based on the IEEE-57 in order to evaluate the scalability of our
> approach. From Table 4, we can see that the ANN has a similar performance
> for IEEE-14 and IEEE-57 systems. Considering the fact that IEEE-14 is
> composed of 27 components (20 transmission lines, 3 FACTS devices, 3 PMU
> devices, and 1 decision support platform) while IEEE-57 has a total of 100
> components (80 transmission lines, 7 FACTS devices, 12 PMU devices, and
> 1 decision support platform), the $100/27 \approx 3.7\times$ increase in the size of
> the system only requires a $2\times$ increase in the size of the training failure
> dataset from 8,000 to 16,000 (Table 3). This verifies the scalability of our
> prediction approach and its applicability to large-scale systems.

Regarding the comment mentioned in item 9, unfortunately we do not understand
the change requested by the respected reviewer. Figures 2 (the ANN architecture),
3 (IEEE-14 and IEEE-57 test systems), and 4 (the failure sequence represented in
terms of state variables) are accompanied by discussions and explanations. Also,
reference [17] is the study that inspired our work on causation analysis, but
is not directly related to any of these three figures. If there are any other
concerns that we have not taken into account, please let us know and we will do
our best to address them.

## 2.4 Reviewer 4

In the paper, the authors studied the intrinsic interdependencies in cyber-physical power systems and predicted the failure propagation with neural networks. In general, the manuscript is clearly presented and the topic is up to date. The reviewer has the following comments.

1. What is the meaning of $c_{ij}$ in Equation (2)?

2. How does the IEEE test network correspond to four types of dependency in section 2.1? In other words, which components have dependencies in Figure 3, such as P-P, P-C, C-P and C-C? Please give a more detailed illustration about simulated networks

3. Matrix $W$ in subsection 2.2.2 has not been used since it was proposed, and what is the difference between $w_{ij}$ and $e_{ij}$?

4. Subsection 2.2.2 should be better reorganized for clearer narrative purposes.

5. Is the decision support algorithm in section 3 proposed by the authors? Please give more details about the algorithm.

6. In section 3, the authors use the IEEE 57 bus system to demonstrate the scalability of the proposed method. However, the IEEE 57 bus system is still relatively a small-scale system, and the method may be infeasible to large-scale system due to the high computational complexity. As shown in Table 1, the number of total buses of the system increased by about four times (from 14 to 57), but the total number of cases required to be simulated increased by about 100 times (from 6,720 to 673,920), the complexity of the proposed should be discussed by the authors.

7. As mentioned by the authors in the last paragraph of section 1 "Using the knowledge of the interdependency, we develop an artificial neural network for prediction of fault propagation paths and cascading failure". However, it seems that the interdependency knowledge have not been used to develop the ANN. Thus, two parts of research (interdependency identification and fault prediction) seem to be separate.

8. As is mentioned in comment (7), how to validate the results of correlation and causation analysis, rather than just calculating the numerical value. Thus, the authors are suggested to give more experiments or illustrations about this concern.

For item 1, we fixed the typo and replaced $c_{ij}$ with the correct notation of $w_{ij}$. We appreciate your attention and apologize for the error in the manuscript.

Regarding the concern in item 2, we understand that the illustrations may not be able to present all identified interdependency links. We considered multiple different ways of presenting the interdependency data and found the graph model (shown in Figures 6 and 7) most suitable. In order to better understand the levels of dependency between cyber and physical networks and compare the four types of interdependency, we also used a bar chart shown in Figure 8.

Regarding the comment in item 3, please note that $w_{ij}$ is indeed used in the calculation of $\mathcal{H}$ (see our response to item 1). The difference between $w_{ij}$ and $e_{ij}$ is that $w_{ij}$ captures all components that fail consecutively (hence $\mathbf{W}$ is named succession frequency matrix), while $e_{ij}$ is intended to capture only failure successions that entail causative relationships.

For item 4, we have reorganized Section 2.2.2 and also added detailed explanations and reasoning regarding the proposed method.

Regarding the concern mentioned in item 5, since we have discussed our control

strategy in our previous studies, we added a reference to these publications [26, 27] in Section 4.

In response to item 6, we argue that the usage of IEEE-14 and IEEE-57 with 27 and 100 components, respectively, allows us to verify the scalability of our approach. In our proposed approach, the main concern in terms of applicability to large-scale systems is related to the prediction of failures using the neural network tool. We have shown in Section 4.4 that our proposed method is very scalable as it can provide very similar performance metrics (Table 4) by only doubling the size of the training dataset (from 8,000 to 16,000) while the size of the system has increased from n = 27 (IEEE-14) to n = 100 (IEEE-57). Please note that for IEEE-14 from the 6,720 cases simulated, we extracted 17,968 training data, but only used 8,000 data for training of the ANN. Similarly for IEEE-57 from the 673,920 cases simulated, we extracted 1,181,871 training data, but only used 16,000 data for training of the ANN. In order to better clarify this claim and present our reasoning, we revised parts of Section 4.4 as shown below.

---

Section 4.4

As mentioned earlier, we have performed the case study on a larger test system based on the IEEE-57 in order to evaluate the scalability of our approach. From Table 4, we can see that the ANN has a similar performance for IEEE-14 and IEEE-57 systems. Considering the fact that IEEE-14 is composed of 27 components (20 transmission lines, 3 FACTS devices, 3 PMU devices, and 1 decision support platform) while IEEE-57 has a total of 100 components (80 transmission lines, 7 FACTS devices, 12 PMU devices, and 1 decision support platform), the $100/27 \approx 3.7\times$ increase in the size of the system only requires a $2\times$ increase in the size of the training failure dataset from 8,000 to 16,000 (Table 3). This verifies the scalability of our prediction approach and its applicability to large-scale systems.

---

To address the concern mentioned in item 7, we explained how we have used the knowledge of interdependency in the prediction of fault propagation in Section 1. In fact, the neural network method is only used to demonstrate that the underlying

19

knowledge of interdependency is beneficial in determining the optimal configuration for achieving the best prediction performance.

> Section 1
>
> In this paper, our objective is to present a model of interdependency by observing sequence of failures for a set of failure cases and capturing the extent of interdependence with quantitative metrics as well as to provide a tool for prediction of failure sequences. We identify interdependencies among components of a CPS using correlation and causation analyses and quantify them with dependency metrics presented in our previous work [15]. Using the knowledge of the interdependency, we identify areas of the system where faults can propagate and create larger failures. We then create a set of failure cases and use the data from the resulting failure sequences to develop and configure an artificial neural network for prediction of fault propagation paths and cascading failures. Our research paves the road to a better understanding of interdependencies in a system and prediction of its future failures using historical failure records or data from simulation of failure scenarios. To illustrate our approach, we have applied it to two smart power grids based on the well-studied IEEE 14-bus and 57-bus test systems.

The concern mentioned in item 8 warrants further research study that we plan to perform in the future, however, we have provided a brief summary of our preliminary verification of the results on causation analyses in Section 4.3 as shown below.

20

Section 4.3


To verify correctness of the causation links identified, further analytical study is required. For the sake of brevity, we present our analysis on only the top five causation links shown in Figure 6b for IEEE-14. The most fragile section of the IEEE-14 system is the combination of $L_{1-2}$ and $L_{1-5}$. In normal operation, the load on $L_{1-5}$ is very small and the majority of the power is transmitted to the rest of the grid through $L_{1-2}$. In research studies on standard power test systems the capacity of each transmission line is usually assumed to be 120% or 125% of its power flow in normal operation. With this assumption, the capacity of $L_{1-5}$ is significantly lower than that of $L_{1-2}$, and hence, upon tripping of $L_{1-2}$, the additional power flow that will have to be transmitted through $L_{1-5}$ causes an overload. This confirms the causation link of $L_{1-2} - L_{1-5}$. In all cases of $L_{2-3} - F_{2-3}$, $L_{2-4} - F_{2-4}$, and $L_{1-5} - F_{1-5}$ a tripped transmission line causes its FACTS device to fail. These causation link are also valid as the FACTS devices are rendered inoperative upon outage of the respective transmission lines. Finally, the causation link of $F_{1-5}$ — $L_{1-2}$ is valid as the only means of controlling the balance between $L_{1-2}$ and $L_{1-5}$ is through $F_{1-5}$, and hence, failure of $F_{1-5}$ creates and imbalance that leads to an overload on $L_{1-2}$.