



Reliability Assessment and Modeling of Cyber Enabled Power Systems with Renewable Sources and Energy Storage

Final Project Report

T-53

Power Systems Engineering Research Center

*Empowering Minds to Engineer
the Future Electric Energy System*



Reliability Assessment and Modeling of Cyber Enabled Power Systems with Renewable Sources and Energy Storage (T-53)

Final Project Report

Project Team

Chanan Singh, Project Leader
Alex Sprintson
Texas A&M University

Visvakumar Aravinthan
Wichita State University

PSERC Publication #16-07

November 2016

For information about this project, contact

Chanan Singh
Regents Professor and Irma Runyon Chair Professor
Department of Electrical and Computer Engineering
Texas A&M University
3128 TAMU
College Station, Texas 77843-3128
Email: singh@ece.tamu.edu
Phone: 979-845-7589

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: <http://www.pserc.org>.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

PSERC members are given permission to copy without fee all or part of this publication for internal use if appropriate attribution is given to this document as the source material. This report is available for downloading from the PSERC website.

© 2016 Texas A&M University. All rights reserved.

Acknowledgements

This is the final report for the Power Systems Engineering Research Center (PSERC) research project titled “Reliability Assessment and Modeling of Cyber Enabled Power Systems with Renewable Sources and Energy Storage” (project T-53). We express our appreciation for the support provided by PSERC’s industry members and by the National Science Foundation under the Industry / University Cooperative Research Center program.

We wish to thank the following Industry Team Members whose advice we acknowledge: M. J. Mousavi, ABB Inc.; J. Fleeman, AEP; K. D. Phillips, AEP; M. Shao, GE Energy Management; W. Li, BC Hydro; X. Luo, ISO New England; L. Min, Lawrence Livermore National Lab; D. Arjona, Idaho Power; and M. Papic, Idaho Power; Adam Wigington, EPRI.

Executive Summary

Quantitative reliability indices are important to utility companies, vendors, and regulators for planning, operation, maintenance, and regulatory purposes. System reliability evaluation methodologies have been mostly focusing on the current-carrying part which is called physical part in this report and the cyber part is assumed to be perfectly reliable in these evaluations. As a part of the efforts to make the grid smarter, the cyber part has been rapidly expanding. The term “cyber” refers to the devices and activities residing in the secondary side of the power system, associated with the functionalities of measurement, control, monitoring, and protection. The overall power system is also referred to as a “cyber-physical power system”, in which the communication networks and power components are interdependent. The cyber-physical interdependencies exist extensively in power systems at all levels and associate with various aspects. The main focus of this work is to investigate this cyber-physical interdependence and develop methods for evaluation of reliability considering this mutual relationship. There are three sections in this report. Part I is concerned with modeling of interdependence between cyber and physical at the composite system level and Part II is focused at the distribution level. During the course of these investigations we also developed a new efficient algorithm for approximating failure frequency with provable guarantees that can be used both for the physical and cyber parts and this is described in Part III.

Part I. Reliability Modeling and Analysis of Cyber Enabled Power Transmission Systems

Information and Communication Technologies (ICTs) are becoming more pervasive in electric power systems to improve system control, protection, monitoring, and data processing capabilities. Generally the ICT technologies are assumed to be perfectly reliable in the process of composite power system reliability evaluation. The failure of these technologies, however, can widen the scope and impact of the failures in the current carrying part. This assumption may thus have significant effect on the reliability indices calculated and result in too optimistic reliability evaluation. For realistic reliability evaluation, it is necessary to consider ICT failures and their impact on composite power systems.

We have extended the scope of bulk power system reliability modeling and analysis with the consideration of cyber elements. Analysis of composite power system together with the cyber part can become computationally challenging. A novel computationally tractable methodology with the use of *Cyber-Physical Interface Matrix (CPIM)* is proposed and demonstrated. The CPIM decouples the analysis of cyber system from the evaluation of the physical system and provides the means of performing the overall analysis in a manageable fashion.

Using the concept of Cyber-Physical Interface Matrix (CPIM), we perform reliability modeling and analysis at the substation level. We have enhanced the substation model with consideration of cyber-link failures. In an attempt to use a non-sequential MCS for dependent failures induced by the cyber failures, we investigate the major difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in the presence of dependent failures and propose a method to overcome these difficulties.

Part II. Reliability Assessment and Modeling of Cyber Enabled Power Distribution Systems

The distribution system reliability evaluation in presence of cyber system is modeled and analyzed in this work. The primary focus of this work has two directions:

a) Distribution System Reliability Evaluation

This part focuses on developing a reliability evaluation technique that is scalable and could be utilized in multi system framework. The proposed model utilizes failure modes and effects analysis and reduces computational complexity using branch and node information of the network. Load point based model is used to preserve topology information and include cyber network in the next part.

b) Cyber-Power System based Reliability Modeling and Analysis

This part focuses on incorporating the properties of cyber – physical systems to develop a reliability evaluation model. Two types of cyber failures (i) cyber unavailability and (ii) cyber-attacks are considered in this work to determine system reliability. Fault detector and automated switch placement as considered as example applications in this work.

Part III. An Efficient Algorithm for Approximating Failure Frequency with Provable Guarantees

In this work, we consider the problem of approximating the failure frequency of large-scale composite systems whose terminals are connected through components that experience random failure and repair processes over time. At any given time, a system failure occurs if the surviving system fails to have all-terminal connectivity. We assume that each component's up-times and down-times are modeled by statistically independent stationary random processes, and these processes are statistically independent across the components. In this setting, the exact computation of failure frequency is known to be computationally intractable (NP-hard). This work, for the first time, provides a polynomial-time algorithm to approximate the failure frequency with high probability within an arbitrary multiplicative error factor using near-minimum cut sets. Moreover, our numerical results show that not only is the proposed method computationally more efficient than the commonly-used bounding technique, but it also has a superior performance in terms of the accuracy of the approximation.

Principal Outcomes

1. Developed a reliability evaluation methodology for composite power systems using Cyber-Physical Interface Matrix (CPIM) and Consequent Events Matrix (CEM) that decouple the analysis of cyber from the physical part.
2. Illustration of the methodology on an extended standard reliability test system for system-wide reliability analysis.
3. A non-sequential Monte Carlo approach for systems having dependent failures induced by cyber failures.
4. A new approach to calculation of frequency of failure using cut sets.
5. A new algorithm for reliability evaluation of radial distribution networks. New analytical model based on physical characteristics of distribution system paves the way for new reliability model developments in the presence of cyber enabled devices

6. Identification of ways emerging cyber enabled devices and logics put power system at risk. These vulnerabilities are categorized into unavailability of data and cyber security threats.
7. Investigation of the necessity of updating the traditional reliability models to incorporate cyber enabled logic. The developed model leads to more accurate and realistic reliability indices at distribution feeder level.
8. Common mode failures are introduced as a potential vulnerability in cyber enabled power distribution network as multiple devices can fail due to a common cause.
9. The proposed probabilistic model is incorporated into a traditional power distribution network planning problem to illustrate the effectiveness of the developed model.

Project Publications

1. H. Lei, C. Singh, and A. Sprintson, "Reliability modeling and analysis of IEC 61850 based substation protection systems," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2194–2202, September 2014.
2. H. Lei and C. Singh, "Power system reliability evaluation considering cyber-malfunctions in substations," *Electric Power Systems Research*, vol. 129, pp. 160-169, December 2015.
3. H. Lei and C. Singh, "Non-Sequential Monte Carlo Simulation for Cyber-Induced Dependent Failures in Composite Power System Reliability Evaluation," *IEEE Transactions on Power Systems*, (accepted for publication).
4. M. Heidari, M. Sepehry, and V. Aravinthan, "Fault Detector and Switch Placement in Cyber-Enabled Power Distribution Network," *IEEE Trans. Smart Grid*, (accepted for publication).
5. M. Heidari, T. Balachandran, V. Aravinthan, V. Namboodiri, and G. Chen, "ALARM: Average Low-Latency Medium Access Control Communication Protocol for Smart Feeders," *IET Generation, Transmission & Distribution*, (accepted for publication).
6. H. Lei and C. Singh, "Incorporating protection systems into composite power system reliability assessment," *IEEE Power and Energy Society 2015 General Meeting*, July 26-30, 2015, Denver, Colorado.
7. H. Lei, C. Singh, and A. Sprintson, "Reliability analysis of modern substations considering cyber link failures," *IEEE Power and Energy Society Innovative Smart Grid Technologies 2015 Asian Conference*, November 4-6, 2015, Bangkok, Thailand.
8. M. Sepehry, M. Heidari, and V. Aravinthan, "Modeling of Uncertainty in Distribution Network Reconfiguration Using Gaussian Quadrature Based Approximation Method," *2016 IEEE PES General Meeting*.
9. M. Sepehry, M. Heidari, V. Aravinthan, "A Stochastic Modeling of Aggregated Vehicle-to-Grid for Reliability Assessment of Distribution Network," in *Proc. North American Power Symposium 2015*.
10. M. Heidari, M. Sepehry, L. Zhao, V. Aravinthan, "Reliability Analysis of Cyber-Enabled Power Distribution System Using Sequential Monte-Carlo," in *Proc. North American Power Symposium 2015*.
11. A. Banajiger and V. Aravinthan, "Radial Feeder Reliability Evaluation in the Presence of Battery Storage using Modified Sequential Monte Carlo Simulation," in *Proc. 7th International Conference on Industrial Automation for Sustainability*, Dec. 2014.

12. M. Sepehry, M. Heidari, A. Banajigar, and V. Aravinthan “A New Algorithm for Reliability Evaluation of Radial Distribution Networks,” in Proc. North American Power Symposium 2014, Sep. 2014. Received first place in the student paper competition.

Part I

Reliability Modeling and Analysis of Cyber-Enabled Power Transmission Systems

Chanan Singh

Alex Sprintson

Hangtian Lei, Graduate Student

Texas A&M University

For information about this project, contact

Chanan Singh, Regents Professor and Irma Runyon, Chair Professor
Department of Electrical and Computer Engineering
Texas A&M University
3128 TAMU
College Station, Texas 77843-3128
Phone: (979) 845-7589
Email: singh@ece.tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: <http://www.pserc.org>.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
551 E. Tyler Mall
Engineering Research Center #527
Tempe, Arizona 85287-5706
Phone: (480) 965-1643
Fax: (480) 965-0745

Notice Concerning Copyright Material

PSERC members are given permission to copy without fee all or part of this publication for internal use if appropriate attribution is given to this document as the source material. This report is available for downloading from the PSERC website.

2016 Texas A&M University.

All rights reserved.

Table of Contents

1. Introduction.....	1
2. Substation Level Reliability Modeling and Analysis.....	3
2.1 Introduction	3
2.2 Protection Systems Using IEC 61850	4
2.3 Reliability Analysis of the Integrated System.....	8
2.4 Summary	26
3. Reliability Analysis Of Modern Substations Considering Cyber-Link Failures.....	27
3.1 Introduction	27
3.2 System Configuration and Parameters	28
3.3 Reliability Analysis	34
3.4 Results and Discussions	35
3.5 Summary	39
4. Composite Power System Reliability Evaluation Considering Cyber-Malfunctions In Substations	40
4.1 Introduction	40
4.2 Methodology Outline and Objectives	41
4.3 Test System Configuration.....	44
4.4 Reliability Analysis	52
4.5 The Scalability of the Overall Methodology.....	62
4.6 Considerations in Software Implementation for Large Power Systems	63
4.7 Summary	65
5. Non-Sequential Monte Carlo Simulation For Power System Reliability Analysis Considering Dependent Failures	66
5.1 Introduction	66
5.2 Origination of Dependent Failures.....	67
5.3 Problem of Applying Non-sequential Sampling for Dependent Failures	69
5.4 Proposed Method.....	75
5.5 Case Studies	78
5.6 Summary	80
6. CONCLUSIONS	82
REFERENCES.....	84

List of Figures

Figure 1. Typical architecture of an IEC 61850-based substation automation system.	4
Figure 2. An IEC 61850 based protection system for a 230-69 kV substation.	6
Figure 3. The reliability block diagram of the line protection unit.	8
Figure 4. The states diagram for an individual component.	9
Figure 5. The states diagram of the process bus.	11
Figure 6. A composite system consisting of a substation and other components.	21
Figure 7. An example of random number mapping.	25
Figure 8. The physical part of the test system.	29
Figure 9. The integrated test system.	30
Figure 10. The cyber part of substation 1.	32
Figure 11. Single line diagram of the RBTS.	44
Figure 12. The protection system for bus 3.	48
Figure 13. The protection system for bus 4.	49
Figure 14. The protection system for bus 5.	49
Figure 15. State transition diagram of individual element.	50
Figure 16. State transition diagram of the process bus.	51
Figure 17. EENS comparison at each bus.	60
Figure 18. Relationship between switching time and system EENS.	62
Figure 19. An example of sampling.	70
Figure 20. A three-component system.	70
Figure 21. The system state space diagram for completely independent scenario.	71

Figure 22. The state transition diagrams for individual components.	72
Figure 23. The state transition diagrams for partially independent scenario.	73
Figure 24. The system state space diagram for partially independent scenario.	73
Figure 25. The system state space diagram for fully dependent scenario.	75
Figure 26. The system state space diagram for an actual power system.	76
Figure 27. Expression of the proposed method.	78

List of Tables

Table 1 Substation protection zone division	7
Table 2 Reliability data for individual components	7
Table 3 Reliability data for protection units.....	8
Table 4 Probability data for individual components	10
Table 5 Summary of scenarios of the line fault clearance at A.....	13
Table 6 Summary of scenarios of the line fault clearance at B.....	13
Table 7 Summary of scenarios of the line fault clearance at I	13
Table 8 Summary of scenarios of the line fault clearance at J	14
Table 9 Summary of scenarios of the transformer fault clearance at E	15
Table 10 Summary of scenarios of the transformer fault clearance at F.....	16
Table 11 Summary of scenarios of the bus fault clearance at C	18
Table 12 Summary of scenarios of the bus fault clearance at D	19
Table 13 Summary of scenarios of the bus fault clearance at G	19
Table 14 Summary of scenarios of the bus fault clearance at H	20
Table 15 Elements of matrix M.....	20
Table 16 Components in the composite system	22
Table 17 Generation and load capacities.....	29
Table 18 Cyber component names and meanings	31
Table 19 Reliability data for components	31
Table 20 Communication path failure probabilities.....	33
Table 21 The consequent event matrix	36

Table 22 The cyber-physical interface matrix.....	37
Table 23 Effect of main path failure on successful operation for line 1	38
Table 24 Bus data	45
Table 25 Generating unit data	45
Table 26 Transmission line physical parameters	47
Table 27 Transmission line outage data	47
Table 28 Reliability data for protection system elements	50
Table 29 The cyber-physical interface matrix for bus 3	54
Table 30 The consequent event matrix for bus 3	55
Table 31 The cyber-physical interface matrix for bus 4	55
Table 32 The consequent event matrix for bus 4	55
Table 33 The cyber-physical interface matrix for bus 5	56
Table 34 The consequent event matrix for bus 5	56
Table 35 Reliability indices for buses	59
Table 36 Simulated transmission line failure rates	59
Table 37 EENS comparison	60
Table 38 Effect of switching time on system EENS	61
Table 39 The format of a cyber-physical interface matrix	68
Table 40 Estimated reliability indices	80

1. Introduction

The quantitative reliability indices of bulk power systems are important to utility companies, vendors, and regulators for planning, operation, maintenance, and regulatory purposes. Studies of bulk power system reliability evaluation have been mostly focusing on the current-carrying part. The pertinent theories and methodologies are well established and documented [1]-[3].

Information and Communication Technologies (ICTs) are widely deployed in electric power systems to improve system control, protection, monitoring, and data processing capabilities. ICT functionalities are generally assumed to be perfectly reliable in the process of composite power system reliability evaluation. This assumption may have significant effect on the reliability indices calculated and result in too optimistic reliability evaluation. For realistic reliability evaluation, it is necessary to consider ICT failures and their impact on composite power systems.

This research aims at extending the scope of bulk power system reliability modeling and analysis with the consideration of cyber elements. In particular, a novel methodology with the use of *Cyber-Physical Interface Matrix (CPIM)* is proposed and demonstrated. The CPIM decouples the analysis of the cyber system from the evaluation of the physical system and provides means of performing the overall analysis in a tractable fashion.

In this part of the report, the term “cyber” refers to the devices and activities residing in the secondary side of the power system, associated with the functionalities of measurement, control, monitoring, and protection. The term “physical” refers to the equipment and activities in the primary side of the power system, associated with the generation, transmission, and distribution of electric power and energy. A whole power system is also referred to as a “cyber-physical power system”, in which the communication networks and power components are interdependent [4]-[6]. The cyber-physical interdependencies exist extensively in power systems and associate with various aspects. This research focuses on the aspect of protection as a facet to study such interdependencies since protection system hidden failures are recognized as common causes of multiple or cascading outages [7]-[10].

A protection system consists of circuit breakers, current and voltage transformers, communication cables, protective relays, and possibly some auxiliary devices [11]-[13]. With the advent of microprocessor-based relays and the rapid progress of communication technologies, modern protection panels are equipped with multifunctional Intelligent Electronic Devices (IEDs) that are connected to communication networks [14]-[17].

Some studies [8], [10], [18]-[22] have been done to consider protection system failures in composite power system reliability evaluation. In most of the previous work, protection system failures were either concentrated on circuit breaker trip mechanisms [10], [18] or represented abstractly by multistate models [8], [19]-[22], in which the protection system was treated as a compact object. Some important technical details inside the protection system, such as the placement of cyber elements (e.g., CT/PTs, MUs, and IEDs) and their wire connections, were absent in those publications. Due to the absence of such details, the interdependencies between cyber elements and physical components were not sufficiently covered. It is necessary to develop a novel reliability evaluation methodology which not only covers such technical details, but is also scalable for applications in large cyber-physical power systems. Furthermore, non-sequential Monte Carlo methods are typically easier to implement and require much less CPU time and memory as compared to sequential methods [3], [23], [24]. It would be beneficial for the application of developed methodology in large systems if the efficiency is further improved with the use of non-sequential techniques.

The remainder of this part of the report is organized as follows: Section 2 proposes a novel methodology with the use of Cyber-Physical Interface Matrix (CPIM) and performs reliability modeling and analysis at the substation level. Section 3 enhances the substation model with consideration of cyber-link failures. Section 4 enhances and implements the proposed methodology on an extended standard reliability test system to obtain system-wide reliability indices. Section 5 researches the major difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in the presence of dependent failures and proposes a method to overcome these difficulties. The conclusions and outlook are given in Section 6. References are attached at the end.

2. Substation Level Reliability Modeling and Analysis*

2.1 Introduction

The electric power substations are vital nodes among electric power generation, transmission, and distribution systems. In recent years, utilities, vendors, and research institutions have paid close attention to the reliability of substation automation systems (SASs) since the failure or malfunctioning of a SAS may have a huge impact on a power system.

The electric power SAS was designed in the past using control and protection schemes with electro-mechanical and hard-wired relay logic [15]. This architecture has undergone significant changes with the advent of multi-functional microprocessor-based Intelligent Electronic Devices (IEDs). Traditional panels with dedicated stand-alone relays, control switches, meters, and status indicators have been replaced by multi-functional, smart, and communicative IEDs [16].

IEC 61850, the international standard for substation automation, provides interoperability between IEDs from different vendors by standardizing the aspects of the information exchange between them [25]. The vision of IEC 61850 series is to define an interoperable communication system, which facilitates the implementation of functions that are distributed among various IEDs from different vendors [15].

Some progress has been made recently in the reliability evaluation of substation automation systems. In [26], reliability indices have been investigated and selected to identify the critical components in an all-digital protection system. The challenges for implementing IEC 61850 based new communication architecture were discussed and some possible solutions to several major implementation issues were suggested in [27]. In [28], the reliability block diagram approach was used to evaluate the reliability and availability of practical Ethernet switch architectures for the SASs.

However, most of the previous publications focused on either the physical components (e.g., transformers, circuit breakers, and transmission lines) or the cyber part separately. In [29], the cyber and physical parts of a SAS were integratively analyzed, but different bays were analyzed separately at a conceptual level, which may not be always true in

* Part of this section is reprinted from copyrighted material with permission from IEEE. © 2014 IEEE. Reprinted, with permission, from Hangtian Lei, Chanan Singh, and Alex Sprintson, "Reliability Modeling and Analysis of IEC 61850-based Substation Protection Systems," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2194-2202, Sep. 2014.

practice. To overcome this shortcoming, a specific cyber-physical system of a typical substation is designed and analyzed using the methodology proposed in this section. Furthermore, the concept of Cyber-Physical Interface Matrix (CPIM) is introduced and its utility is illustrated. The CPIM depicts the inter-dependencies among the failures of different physical components due to various cyber failure modes. It decouples the analysis of the cyber part from the physical part and provides the means of performing the overall analysis in a computationally tractable way.

The remainder of this section is organized as follows: In Section 2.2, the overall architecture and main components of the IEC 61850-based SAS are described, a typical IEC 61850 based substation layout, including both physical and cyber parts, is presented. In Section 2.3, a general technique for cyber-physical system reliability analysis is presented, the detailed analysis for a substation is performed, and the CPIM is obtained. The utilization of the interface matrix is then illustrated by incorporating this substation into a composite power system. Section 2.4 is the summary of Section 2.

2.2 Protection Systems Using IEC 61850

2.2.1 IEC 61850 Hierarchy and Architecture

A typical architecture of the IEC 61850 based SAS, which consists of three levels, is shown in Figure 1.

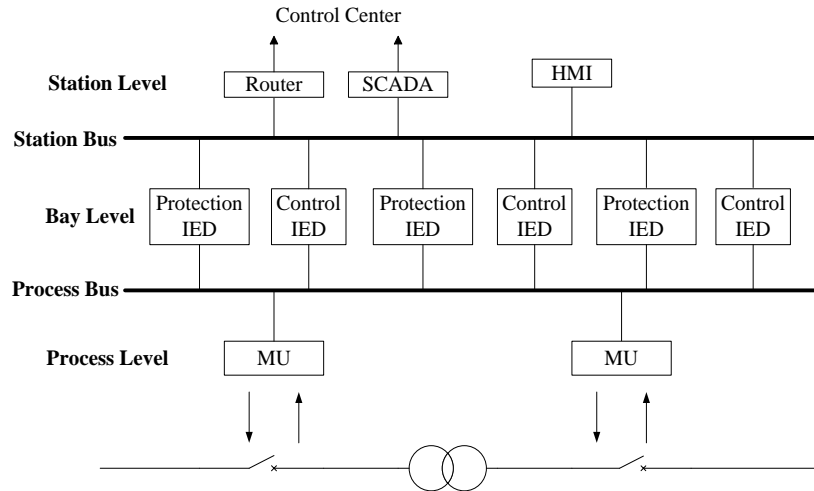


Figure 1. Typical architecture of an IEC 61850 based substation automation system.

Process level: This level includes Current Transformers (CTs) / Potential Transformers (PTs), Merging Units (MUs), actuators, etc. The voltage and currents signals acquired by CTs / PTs are digitized by MUs and sent over the Ethernet network to the bay level.

Bay level: This level includes microprocessor-based relays (also known as protection IEDs) and bay controllers (control IEDs). Protection IEDs receive information coming from the process level, conduct elaborate calculations and send decision signals over the Ethernet network.

Station level: Station level includes the Human Machine Interface (HMI) and Supervisory Control and Data Acquisition (SCADA) system. At this level, the status data of various components in the substation are available to operators for monitoring and operation purposes. Operators can also issue signals at this level to perform certain kinds of manual control.

Process bus: The process bus enables the time critical communication between the process level and the bay level, which builds a bridge for voltage and currents information going from MUs to protection IEDs, and for trip signals going the opposite direction.

Station bus: The station bus enables information exchange between the bay level and station level, which makes the status data of the entire substation available to the control center for monitoring and operation purposes.

2.2.2 IEC 61850 Based Protection System Layout and Configuration

An IEC 61850-based protection system for a typical 230-69 kV substation is designed for the reliability analysis. The integrated system, including physical components (e.g., transformers, transmission lines, and circuit breakers) and cyber components (e.g., merging units, Ethernet switches, and Prot. IEDs), is shown in Figure 2. The physical part has been used in [30] to illustrate the protection zones and schemes. The cyber part is designed according to IEC 61850 standards [31].

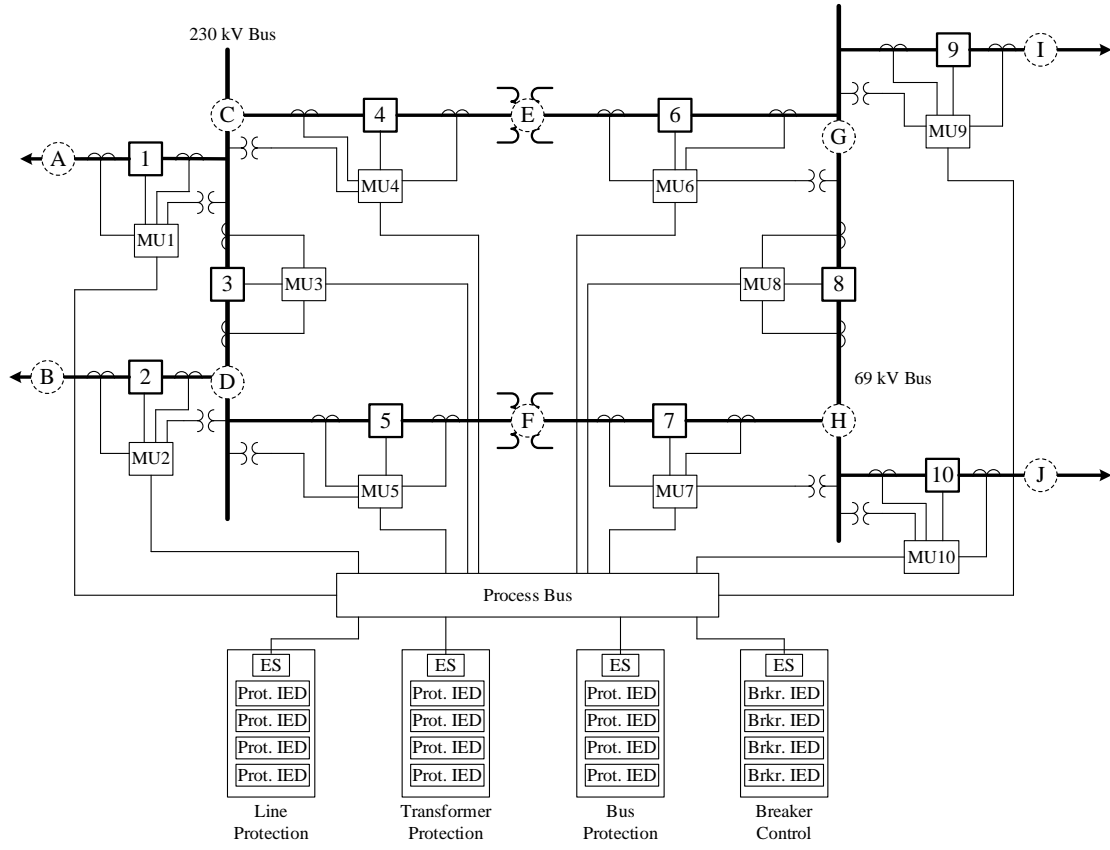


Figure 2. An IEC 61850 based protection system for a 230-69 kV substation.

The primary protection zones associated with various fault locations and the corresponding circuit breakers needed to trip for fault clearances are listed in Table 1.

The MTTF [32] values of individual components for reliability calculations are obtained from [15], [28], [33]-[35] and tabulated in Table 2.

Table 1 Substation protection zone division

Type	Fault Location	Associated Circuit Breakers
Line	A	Breaker 1
	B	Breaker 2
	I	Breaker 9
	J	Breaker 10
Transformer	E	Breakers 4, 6
	F	Breakers 5, 7
Bus	C	Breakers 1, 3, 4
	D	Breakers 2, 3, 5
	G	Breakers 6, 8, 9
	H	Breakers 7, 8, 10

The MTTF varies for CBs at different voltage levels, or serving different functions in the system [35], for the study in this section, a typical value of 100 years is chosen. Using MRT of 8 hours from [15] and [28], the failure and repair rates of individual components are tabulated in Table 2.

Table 2 Reliability data for individual components

	MTTF (year)	Failure Rate λ (/year)	MRT (h)	Repair Rate μ (/year)
CB	100	0.01	8	1095
MU	150	0.00667	8	1095
PB	100	0.01	8	1095
ES	50	0.02	8	1095
Prot. IED	150	0.00667	8	1095

For convenience of analysis, the ES and protection IEDs located at the same protection panel are combined into one line protection unit. Normally, for each protection unit, redundant protection IEDs are equipped, hence, in the reliability block diagram shown in Figure 3, both the protection IEDs are shown in parallel, and then in series with the ES. The reliability data after combination is shown in Table 3.

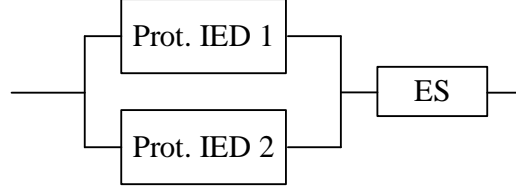


Figure 3. The reliability block diagram of the line protection unit.

Table 3 Reliability data for protection units

	Failure Rate λ (/year)	Mean Repair Time (h)	Repair Rate μ (/year)
Line Protection Unit	0.02000008	7.99998	1095.002
Transformer Protection Unit	0.02000008	7.99998	1095.002
Bus Protection Unit	0.02000008	7.99998	1095.002

2.3 Reliability Analysis of the Integrated System

2.3.1 A General Technique for Reliability Analysis of Cyber-Physical Systems

The complexity and dimensionality of substation automation systems make it difficult, if not impossible, to conduct the reliability analysis of the whole system, physical and cyber, in a single step. Even for the current carrying part alone, it is not computationally efficient to model all the components distinctly and simultaneously. Therefore, it is necessary to perform the analysis sequentially.

Our proposed approach is formulated with the following steps:

- 1) Develop an interface matrix between the cyber and physical subsystems. This matrix is called *Cyber-Physical Interface Matrix (CPIM)*. It defines the relationship between the cyber subsystems and physical subsystems in terms

of failure modes and effects. In developing this matrix, the major interaction points between the cyber and physical part need to be identified.

- 2) Analyze the cyber part to determine parameters of the interface matrix.
- 3) Determine probabilities of interface events.
- 4) Analyze the physical system using the interface matrix.

An example of the Cyber-Physical Interface Matrix M is shown in (2.1). The elements of this matrix are the probabilities of interface events. In the following sections, the reliability analysis for the system shown in Figure 2 will be presented to illustrate the procedures of obtaining these probabilities.

$$M = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{bmatrix} \quad (2.1)$$

2.3.2 Individual Component Analysis

For each component (except the process bus) listed in Table 2, only two states, UP and DOWN, are considered in our study as shown in Figure 4.

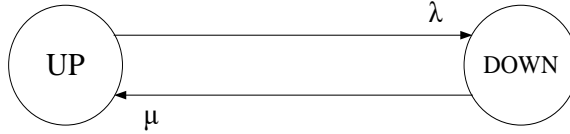


Figure 4. The states diagram for an individual component.

The probabilities of being in the UP and DOWN states can be calculated using equations (2.2) and (2.3), respectively.

$$p_{up} = \frac{\mu}{\lambda + \mu} \quad (2.2)$$

$$p_{down} = \frac{\lambda}{\lambda + \mu} \quad (2.3)$$

The calculated probabilities for Circuit Breaker (CB), Merging Unit (MU), Line Protection Unit, Transformer Protection Unit, and Bus Protection Unit are tabulated in Table 4.

In an ideal environment that utilizes non-blocking switches and has prioritization mechanisms, the time of delay is negligibly small. However, in practice, delays might occur due to the use of legacy technology, such as bus Ethernet, which is still used in some substations, or due to the use of wireless technology.

Table 4 Probability data for individual components

Component	p_{up}	p_{down}
Circuit Breaker (CB)	0.999990867	0.000009132
Merging Unit (MU)	0.999993912	0.000006088
Line Protection Unit	0.999981735	0.000018265
Transformer Protection Unit	0.999981735	0.000018265
Bus Protection Unit	0.999981735	0.000018265

Compared with fiber-optic communication, wireless communication technologies are more economically feasible for small-scale automation systems in rural areas. In wireless environments, the delays can be quite large due to the electromagnetic interference in high voltage environments. Meanwhile, the radio frequency interference from wireless equipment can also affect the functioning of equipment [36], and Generic Object Oriented Substation Event (GOOSE) packets might be occasionally dropped by the network due to errors. Therefore, there is some finite probability that delay may happen in the network. Taking into account this probability, delay is modeled as a state of the process bus. Of course, the delay probability can be set to zero, if needed.

The state DELAY means that due to the temporarily heavy traffic, the process bus does not physically fail, but the message transfer is delayed and the delay time is over the threshold value that causes the breakers associated with the primary protection zone fail to trip in time. The probability of delay given that the PB is not in the DOWN state is denoted by p_d (=0.003). The state transition diagram of the PB is shown in Figure 5.

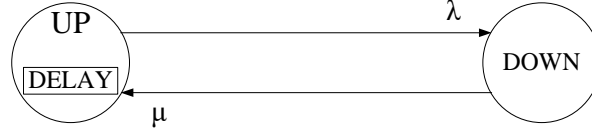


Figure 5. The states diagram of the process bus.

Thus, for the Process Bus (PB), the probabilities of being in the UP, DELAY, and DOWN states can be calculated using equations (2.4)-(2.6).

$$p_{up} = \frac{\mu}{\lambda + \mu} (1 - p_d) \quad (2.4)$$

$$p_{delay} = \frac{\mu}{\lambda + \mu} p_d \quad (2.5)$$

$$p_{down} = \frac{\lambda}{\lambda + \mu} \quad (2.6)$$

The reliability analysis of line, transformer, and bus fault clearances for the substation shown in Figure 2 are discussed in sections 2.3.3, 2.3.4, and 2.3.5, respectively.

Several assumptions are made:

- 1) The cable links between various devices and all the CT/PTs are assumed not to fail.
- 2) If the breaker(s) for the primary protection fail to trip correctly due to the message delay or due to the failure of the components other than the process bus, the trip signal can be transferred to an adjacent protection zone. However, if the process bus fails, the entire system is assumed to fail.
- 3) If the primary protection fails to trip correctly and the trip signal is transferred to an adjacent protection zone, the subsequent failure of the trip, is not considered.
- 4) The fault events happening in different areas (e.g., areas A, E, and C) are analyzed independently. The possibility that faults occur simultaneously at different locations is not considered.
- 5) If the message delay at the process bus is beyond a threshold value, the breaker(s) for the primary protection will fail to trip and the trip signal will be transferred to an adjacent protection zone.

2.3.3 Reliability Analysis of Line Fault Clearance

As shown in Figure 2, line faults can happen in areas A, B, I, or J. Here, the line fault at area A is taken as an example for illustration. The same techniques can be applied to faults at B, I, or J.

When a line fault happens at A, the voltage/current information will be sensed by PT/CTs and will be sent to MU1. The information will be digitized at MU1 and then be sent to Line Protection Unit via the PB. Based on the information received, relay algorithms will be performed at the Line Protection Unit and a trip signal will be sent to Circuit Breaker 1 via the PB. There are 4 components associated with this procedure, namely, MU1, PB, Line Protection Unit, and CB1. One or more components' failure will result in malfunction. The detailed descriptions of different scenarios are listed as below.

1) All components operate as intended

If all aforementioned components operate as intended, CB1 will trip in time, and only the faulted line will be isolated. The rest of the substation will stay in service.

2) PB fails to work

Since PB is the hub of all the cyber links inside the substation, the failure of PB will cause all the relays to be unable to receive or send information. All breakers will fail to trip and the entire system will be affected by the fault.

3) One or more components of MU1, Line Protection Unit, CB1 fail to operate

CB1 will not trip at first. When the fault comes to affect area C, bus protection will be triggered and CB1, CB3, CB4 will trip. Areas A and C will be out of service.

4) PB is in UP state, but message delay happens due to temporarily heavy information traffic

In this case, the tripping signal will not arrive at CB1 in time and thus CB1 will not trip before the breaker failure timer expires, bus protection for area C will then be triggered and CB1, CB3, CB4 will trip.

The probabilities and effects corresponding to each case are shown in Table 5. These probabilities are conditional and are calculated given that a fault has already happened at location A. To obtain the actual probabilities, they need to be multiplied by the probability of this fault occurrence.

Table 5 Summary of scenarios of the line fault clearance at A

Scenario No.	Probability	Areas Affected
1	0.996957511	A
2	0.000009132	Entire Substation
3	0.000033384	A, C
4	0.002999973	A, C

Table 6 Summary of scenarios of the line fault clearance at B

Areas Affected	Probability
B	0.996957511
Entire Substation	0.000009132
B, D	0.003033357

Table 7 Summary of scenarios of the line fault clearance at I

Areas Affected	Probability
I	0.996957511
Entire Substation	0.000009132
G, I	0.003033357

Table 8 Summary of scenarios of the line fault clearance at J

Areas Affected	Probability
J	0.996957511
Entire Substation	0.000009132
H, J	0.003033357

Similarly, the probabilities and effects of line fault clearances at locations B, I, and J are shown in Tables 6-8, respectively.

2.3.4 Reliability Analysis of Transformer Fault Clearance

As shown in Figure 2, a transformer winding/ground fault can happen at areas E or F. Here, the transformer fault at E is taken as an example for illustration. The same techniques can be applied to faults at F.

When a transformer fault happens at E, the voltage/current information will be sensed by corresponding PT/CTs and will be sent to MU4 and MU6. The information will be digitized at these merging units and then will be sent to Transformer Protection Unit via the PB. Based on the information received, relay algorithms will be performed at Transformer Protection Unit and trip signals will be sent to CB4 and CB6 via the PB. There are 6 components associated with this procedure, namely, MU4, MU6, PB, Transformer Protection Unit, CB4, and CB6. One or more components' failure will result in malfunction. The detailed descriptions of different scenarios are listed as below.

1) All components operate as intended

If all aforementioned components operate as intended, CB4 and CB6 will trip as intended, only the faulted part E will be isolated and the rest of this substation will stay in service.

2) PB fails to work

Since PB is the hub of all the cyber links inside the substation, the failure of PB will cause all the relays to be unable to receive or send information, all breakers will fail to trip, and the entire system will be affected by this fault.

- 3) *One or more components of MU4, CB4 fail, while all other components work as intended*

CB6 will trip as intended, but CB4 will not. When the fault affects area C, the bus protection will be triggered and CB1 and CB3 will trip. Areas C and E will be out of service.

- 4) *One or more components of MU6, CB6 fail, while all other components work as intended*

CB4 will trip as intended, but CB6 will not. When the fault affects area G, the bus protection will be triggered, and CB8 and CB9 will trip. Areas E, G and I will be out of service.

- 5) *Both CB4 and CB6 fail to trip due to breakers failure or PB delay, or failure of Transformer Protection Unit, but PB is not down*

When the fault comes to affect areas C and G, protection devices of these zones will be triggered. Areas C, E, G, and I will be isolated from the system.

The probabilities and effects corresponding to each scenario are shown in Table 9. These probabilities are calculated given that a fault has already happened at location E. The actual probabilities need to be multiplied by the probability of this fault occurrence.

Table 9 Summary of scenarios of the transformer fault clearance at E

Scenario No.	Probability	Areas Affected
1	0.996942336	E
2	0.000009132	Entire Substation
3	0.000015174	C, E
4	0.000015174	E, G, I
5	0.003018182	C, E, G, I

Similarly, the probabilities and effects of a transformer fault clearance at location F are shown in Table 10.

Table 10 Summary of scenarios of the transformer fault clearance at F

Areas Affected	Probability
F	0.996942336
Entire Substation	0.000009132
D, F	0.000015174
F, H, J	0.000015174
D, F, H, J	0.003018182

2.3.5 Reliability Analysis of Bus Fault Clearance

A bus fault can happen at locations C, D, G, or H. Here, the bus fault at C is taken as an example for illustration. The same techniques can be applied to faults at D, G, and H.

When a bus fault happens at C, the voltage/current information will be sensed by the corresponding PT/CTs, and will be sent to MU1, MU3, and MU4. The information will be digitized at these merging units, and then will be sent to Bus Protection Unit via the Process Bus. Based on the information received, relay algorithms will be performed at the Bus Protection Unit and trip signals will be sent to CB1, CB3, and CB4 via the PB. There are 8 components associated with this procedure, namely, MU1, MU3, MU4, PB, Bus Protection Unit, CB1, CB3, and CB4. One or more components' failure will result in malfunction. The detailed descriptions of different scenarios are listed as below.

1) *All components operate as intended*

If all components operate as intended, CB1, CB3, and CB4 will trip as intended, only the faulted bus will be cut off, and the rest of this substation will stay in service.

2) *PB fails to work*

Since PB is the hub of all the cyber links inside the substation, the failure of PB will cause all the relays to be unable to receive or send information, all breakers will fail to trip, and the entire system will be affected by this fault.

3) *One or more components of MU1, CB1 fail, while all other components work as intended*

CB3 and CB4 will trip as intended, but CB1 will not. When the fault comes to affect area A, the breaker located at the substation on the other side of this line will trip. Areas A and C will go out of service.

- 4) *One or more components of MU3, CB3 fail, while all other components work as intended*

CB1, CB4 will trip as intended, but CB3 will not. When the fault comes to affect area D, the bus protection for D will be triggered and CB2 and CB5 will trip. Buses C and D will be out of service.

- 5) *One or more components of MU4, CB4 fail, while all other components work as intended*

CB1 and CB3 will trip as intended, but CB4 will not. When the fault affects area E, the transformer protection will be triggered and CB6 will trip. Bus C and transformer E will be out of service.

- 6) *One or more components of MU1, CB1 fail and one or more components of MU3, CB3 fail, while all other components work as intended*

Both CB1 and CB3 will not trip. When the fault comes to affect areas A and D, the protection IEDs for these zones will be triggered, CB2, CB5 as well as the breaker on the other side of line A will trip. Areas A, C, and D will go out of service.

- 7) *One or more components of MU1, CB1 fail and one or more components of MU4, CB4 fail, while all other components work as intended*

Both CB1 and CB4 will not trip. When the fault comes to affect areas A and E, the protection IEDs for these zones will be triggered, CB6 and the breaker on the other side of line A will trip. Areas A, C, and E will be cut off from the system.

- 8) *One or more components of MU3, CB3 fail and one or more components of MU4, CB4 fail, while all other components work as intended*

Both CB3 and CB4 will not trip as intended. When the fault comes to affect areas D and E, the protection IEDs for these zones will be triggered, CB2, CB5, and CB6 will trip. Areas C, D, and E will be cut off from the system.

- 9) *All of the CB1, CB3, and CB4 fail to trip due to breakers failure or PB delay or failure of Bus Protection Unit, but PB is not down*

When the fault comes to affect areas A, D and E, the protection IEDs for these zones will be triggered. Areas A, C, D, and E will be isolated from the system.

Table 11 Summary of scenarios of the bus fault clearance at C

Scenario No.	Probability	Affected Areas
1	0.996927163	C
2	0.000009132	Entire Substation
3	0.000015174	A, C
4	0.000015174	C, D
5	0.000015174	C, E
6	2.31×10^{-10}	A, C, D
7	2.31×10^{-10}	A, C, E
8	2.31×10^{-10}	C, D, E
9	0.003018182	A, C, D, E

The probabilities and effects corresponding to all the scenarios are shown in Table 11. These probabilities are calculated given that a fault has already happened at location C. The actual probabilities need to be multiplied by the probability of this fault occurrence.

Similarly, the probabilities and effects of bus fault clearances at locations D, G, and H are shown in Tables 12-14, respectively.

Table 12 Summary of scenarios of the bus fault clearance at D

Affected Areas	Probability
D	0.996927163
Entire Substation	0.000009132
B, D	0.000015174
C, D	0.000015174
D, F	0.000015174
B, C, D	2.31×10^{-10}
B, D, F	2.31×10^{-10}
C, D, F	2.31×10^{-10}
B, C, D, F	0.003018182

Table 13 Summary of scenarios of the bus fault clearance at G

Affected Areas	Probability
G	0.996927163
Entire Substation	0.000009132
G, I	0.000015174
G, H	0.000015174
E, G	0.000015174
G, H, I	2.31×10^{-10}
E, G, I	2.31×10^{-10}
E, G, H	2.31×10^{-10}
E, G, H, I	0.003018182

Table 14 Summary of scenarios of the bus fault clearance at H

Affected Areas	Probability
H	0.996927163
Entire Substation	0.000009132
H, J	0.000015174
G, H	0.000015174
F, H	0.000015174
G, H, J	2.31×10^{-10}
F, H, J	2.31×10^{-10}
F, G, H	2.31×10^{-10}
F, G, H, J	0.003018182

2.3.6 Construction and Utilization of Cyber-Physical Interface Matrix

The Cyber-Physical Interface Matrix (CPIM) M can be obtained by synthesizing the data from Table 5 to Table 14. The elements of this matrix are the probabilities of interface events. For the Cyber-Physical Interface Matrix (CPIM) M corresponding to the substation illustrated in previous sections, some elements are tabulated in Table 15. The CPIM shown in Table 15 can be improved by eliminating the off-diagonal zeros to make it more compact. Some examples of better developed CPIMs are presented in Section 4.4.1 of this part of the report.

Table 15 Elements of matrix M

	Column 1	Column 2	Column 3	...	Column 58
Row 1	0.996958	0	0	...	0
Row 2	0	9.1×10^{-6}	0	...	0
Row 3	0	0	0.003033357	...	0
...
Row 58	0	0	0	...	0.00301818

Once the CPIM is obtained, its results can be utilized for the reliability analysis of a wider area by incorporating this substation into a larger system without considering the details of the cyber part.

To illustrate the utility of the interface matrix, the Monte Carlo simulation process for a composite system shown in Figure 6 is explained in the following example. The simulation process is for illustration only. Its specific implementation will be presented in section 4 of this part of the report.

Since the CPIM, which depicts the inter-dependencies among the failures of different physical components due to various cyber failure modes, is already obtained, the simulation process can be performed without considering the details of the cyber part. Therefore, only the physical components are shown in Figure 6. All these components are numbered in Table 16.

The next event sequential simulation [37], in which the time is advanced to the occurrence of the next event, is used. For each individual component in this composite system, two states, UP and DOWN, are considered.

The simulation process for the composite system can be formulated in the following steps:

1) *Step 1*

Set the initial state of all components as UP and set the simulation time t to 0.

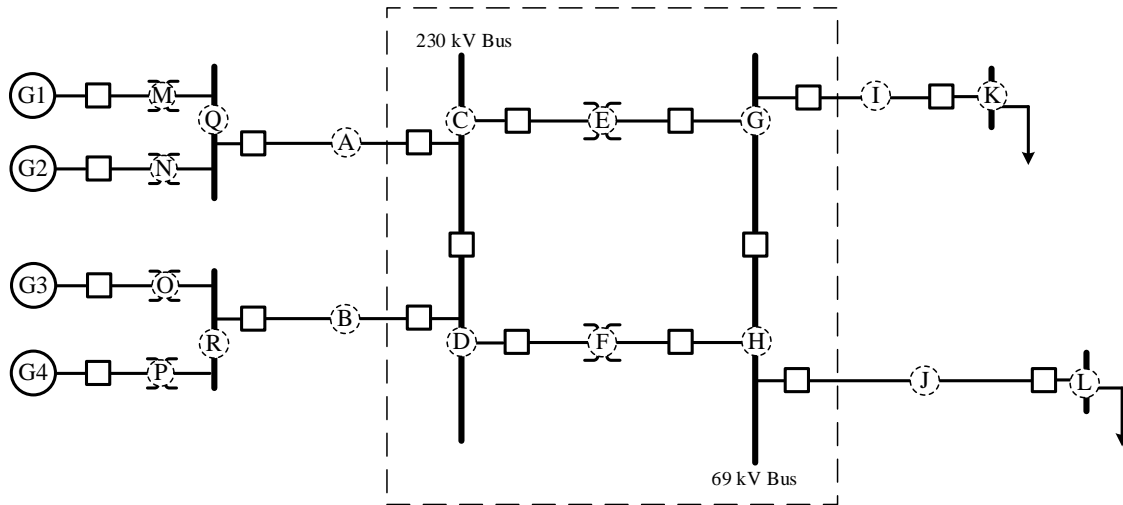


Figure 6. A composite system consisting of a substation and other components.

Table 16 Components in the composite system

Number	Component Name
1	Generator G1
2	Generator G2
3	Generator G3
4	Generator G4
5	Transformer M
6	Transformer N
7	Transformer O
8	Transformer P
9	Bus Q
10	Bus R
11	Line A
12	Line B
13	Bus C
14	Bus D
15	Transformer E
16	Transformer F
17	Bus G
18	Bus H
19	Line I
20	Line J
21	Bus K
22	Bus L

2) *Step 2*

For each individual component, draw a random decimal number between 0 and 1 to compute the time to the next event.

In this section, the distributions of UP and DOWN times for all components are assumed to be exponential. Let N_c be the total number of components, z_i ($0 < z_i < 1$, $1 \leq i \leq N_c$) be the random number drawn for the i^{th} component. The time to the next transition of this component is given by:

$$T_i = -\frac{\ln(z_i)}{\rho_i} \quad (2.7)$$

In (2.7), depending on whether the i^{th} component is UP or DOWN, λ_i or μ_i is used in place of ρ_i .

3) *Step 3*

Find the minimum time, change the state of the corresponding component, and update the total time.

The time to the next system transition is given by:

$$T = \min\{T_i\}, 1 \leq i \leq N_c \quad (2.8)$$

If this T corresponds to T_q , that is, the q^{th} component, then the next transition takes place by the change of state of this component. The total simulation time t is increased by T .

4) *Step 4*

Change the q^{th} component's state accordingly. For each component i , $1 \leq i \leq N_c$, subtract T from T_i

$$T_{i,res} = T_i - T \quad (2.9)$$

where $T_{i,res}$ is the residual time to transition of component i . The time T_i is updated to:

$$T_i = T_{i,res} \quad (2.10)$$

Since the residual time for component q causing transition becomes 0, therefore, the time to its next transition T_q is determined by drawing a new random number and using (2.7).

5) *Step 5*

If the state of the q^{th} component transits from UP to DOWN, which means a primary fault happens to this component, then the CPIM is used to determine if there are some subsequent failures causing more components out of service due to the cyber part's malfunction.

Let n_q be the number of possible scenarios if a primary fault happens to the q^{th} component, and $p_{q,j}$ ($1 \leq j \leq n_q$) be the probability of the j^{th} scenario given that a primary fault already happened at the q^{th} component. These probabilities are directly available from the CPIM. According to the analysis in sections 2.3.3-2.3.5, the following relationship exists:

$$\sum_{j=1}^{n_q} p_{q,j} = 1 \quad (2.11)$$

For the convenience of further illustration, a zero probability $p_{q,0}$ is added to the left side of (2.11), which yields:

$$\sum_{j=0}^{n_q} p_{q,j} = 1 \quad (2.12)$$

Draw a random decimal number y ($0 < y \leq 1$). Let s ($1 \leq s \leq n_q$) be an integer which satisfies (2.13).

$$\sum_{j=0}^{s-1} p_{q,j} < y \leq \sum_{j=0}^s p_{q,j} \quad (2.13)$$

Then the s^{th} scenario is determined to happen. Let S be the set of components that would go out of service if the s^{th} scenario happens, and S_2 be the set such that:

$$S_2 = \{k | k \in S, k \neq q\} \quad (2.14)$$

For every component k whose state is currently UP and in S_2 , change its state to DOWN, draw a new random number, and calculate the time to its next transition T_k using (2.7). Since the failure of component k is caused by the cyber failure rather than a primary fault, therefore, an expedited repair rate $\mu_{k, exp}$ instead of μ_k is used in (2.7). The value of $\mu_{k, exp}$ is normally available from engineering practice and is called a switching rate.

The following specific case is shown as an example to illustrate the details from step 3 to step 5. For the system shown in Figure 6, in the first iteration of Monte Carlo simulation, if $q = 11$ is obtained in step 3, which means a primary fault happens at Line A, then T_{11} is updated in step 4. In step 5, from the CPIM shown in Table 15, $n_{11} = 3$, $p_{11,1} = 0.996958$, $p_{11,2} = 9.1 \cdot 10^{-6}$, and $p_{11,3} = 0.003033357$ can be obtained.

If the random number y is generated to be 0.9177, as shown in Figure 7, then the 1st scenario is determined to happen, which means only Line A is going out of service. Thereby, $S = \{11\}$ and $S_2 = \{\}$ can be obtained.

If the random number y is generated to be 0.9987, also as shown in Figure 7, then the 3rd scenario is determined to happen, which means both Line A and Bus C would go out of service. $S = \{11, 13\}$ and $S_2 = \{13\}$ can be obtained, a new random number z_{13} is generated and $\mu_{13, exp}$ is used in (2.7) to calculate the time to its next transition T_{13} .

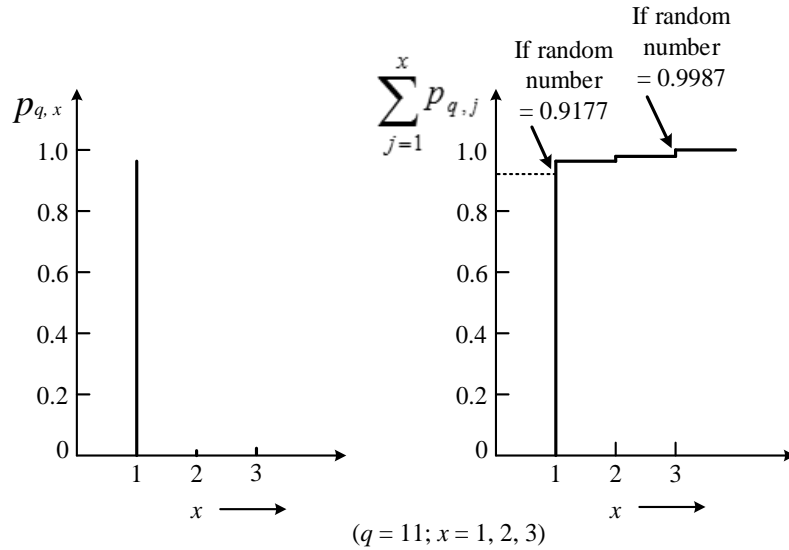


Figure 7. An example of random number mapping.

6) Step 6

Perform the network power flow analysis to assess system operation states. Update reliability indices.

7) *Convergence*

Steps 3–6 are iteratively continued until a convergence criterion is satisfied. The simulation is said to have converged when the reliability indices attain stable values. For any index i , the convergence is measured by its standard error, defined as:

$$\eta = \frac{\sigma_i}{\sqrt{N_y}} \quad (2.15)$$

where σ_i is the standard deviation of the index i and N_y is the number of years simulated.

Convergence is said to occur when the standard error in (2.15) drops below a preselected value, ε_i , as shown in (2.16).

$$\eta < \varepsilon_i \quad (2.16)$$

When the simulation finishes, reliability indices, such as the Loss of Load Expectation, can be finally obtained.

2.4 Summary

A novel methodology for modeling and analysis of cyber enabled substation protection systems is presented. A typical protection system based on the IEC 61850 concepts, incorporating both physical and cyber components, is designed. The probabilities of various faults and tripping scenarios are calculated. General techniques for reliability analysis of cyber-physical systems are presented. The concept of Cyber-Physical Interface Matrix (CPIM) is introduced and its utility is illustrated. The CPIM decouples the analysis of the cyber part from the physical part and provides the means of performing the overall analysis of a composite system in a more tractable fashion.

This methodology of finding the CPIM also applies to the reliability analysis of substation automation systems with more complex configuration and larger scale. In such systems, more effort is needed in detailed analysis of various cyber failure modes as well as effects on the physical side.

3. Reliability Analysis of Modern Substations Considering Cyber-Link failures*

3.1 Introduction

This section enhances the substation protection system reliability model with the consideration of cyber-link failures.

A substation protection system is a typical cyber-physical system. It consists of circuit breakers, current/potential transformers, merging units, and protection panels with intelligent electronic devices. These components are connected in an Ethernet-based environment [38]-[41]. In recent years, numerous research efforts have been devoted to study of the reliability considerations and implementation issues of modern substation automation systems [28], [38]-[44].

Due to the complexity of monitoring, control, and communication functions as well as the variety of cyber-physical interdependencies, it is challenging to model and analyze the complete cyber-physical system with explicit technical details. Therefore, most research work focuses either on the cyber part or on the physical part. To cover the whole cyber-physical system, it is necessary to divide the overall analysis into subsections and proceed sequentially. A tractable methodology of performing the overall analysis by decoupling the cyber part from the physical part has been proposed in Section 2 [42] of this part of the report by introducing the concept, *Cyber-Physical Interface Matrix (CPIM)*.

The example provided in Section 2 is for the purpose of illustration and some technical details have been simplified in modeling the cyber network. For example, the traffic delay is modeled as a state with a predefined probability value and the links in the communication network are assumed to never fail.

This section applies the methodology proposed in Section 2 to a 4-bus power system with the consideration of more technical details in the cyber part. Unlike some previous publications [38], [39] performing simulations to study the issue of packet delay, this section mathematically models delay as the unavailability of communication links. The remainder of this section is organized as follows:

* Part of this section is reprinted from copyrighted material with permission from IEEE. © 2015 IEEE. Reprinted, with permission, from Hangtian Lei, Chanan Singh, and Alex Sprintson, "Reliability Analysis of Modern Substations Considering Cyber link Failures," in *Proc. IEEE Power and Energy Society Innovative Smart Grid Technologies 2015 Asian Conference*, Bangkok, Thailand, Nov. 2015.

Section 3.2 presents the test system configuration and parameters. The issue of link unavailability due to packet delay is also discussed and modeled in Section 3.2. Section 3.3 outlines the overall procedures. In Section 3.4, the results and discussions are provided. Section 3.5 is the summary of this section.

3.2 System Configuration and Parameters

To illustrate the interactions between cyber and physical components, the reliability of a 4-bus power system with Ethernet-based protection configurations is analyzed in this section as an example. The physical part of the system is taken from [45]. The cyber part is designed according to the typical configurations of modern substation protection systems.

3.2.1 Configuration and Parameters of the Physical Part

The physical part of the system shown in Figure 8 is as described in [45]. The load and generation capacities are tabulated in Table 17, of which the loads are directly obtained from [45]. The generation capacities at substations (buses) 1 and 4 are assumed to be 250 MW and 300 MW, respectively. This assumption is based on the consideration of a 10% capacity reserve for the whole system. Compared to line faults, bus faults are relatively rare, and thus are not considered in the reliability analysis of this section.

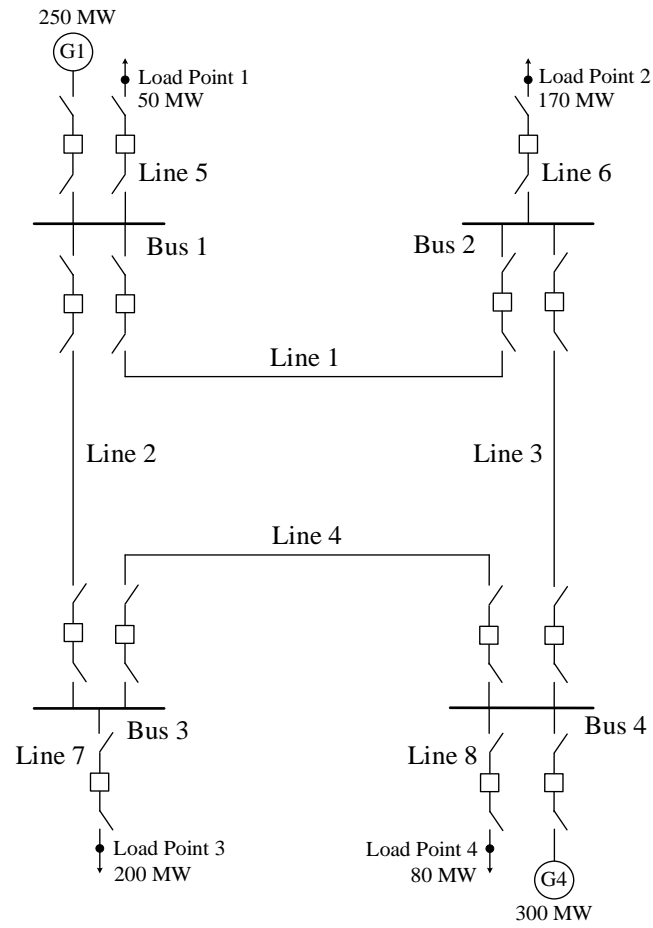


Figure 8. The physical part of the test system.

Table 17 Generation and load capacities

Bus No.	Generation Capacity (MW)	Load Capacity (MW)
1	250	50
2	0	170
3	0	200
4	300	80
Total	550	500

3.2.2 Configuration and Parameters of the Cyber Park

The Ethernet-based protection system is designed for each substation (bus), as shown in Figure 9. For the protection system at each substation, three Ethernet switches are used and they are connected in a ring topology. Take substation 1 as an example, the cyber component names and their meanings are tabulated in Table 18.

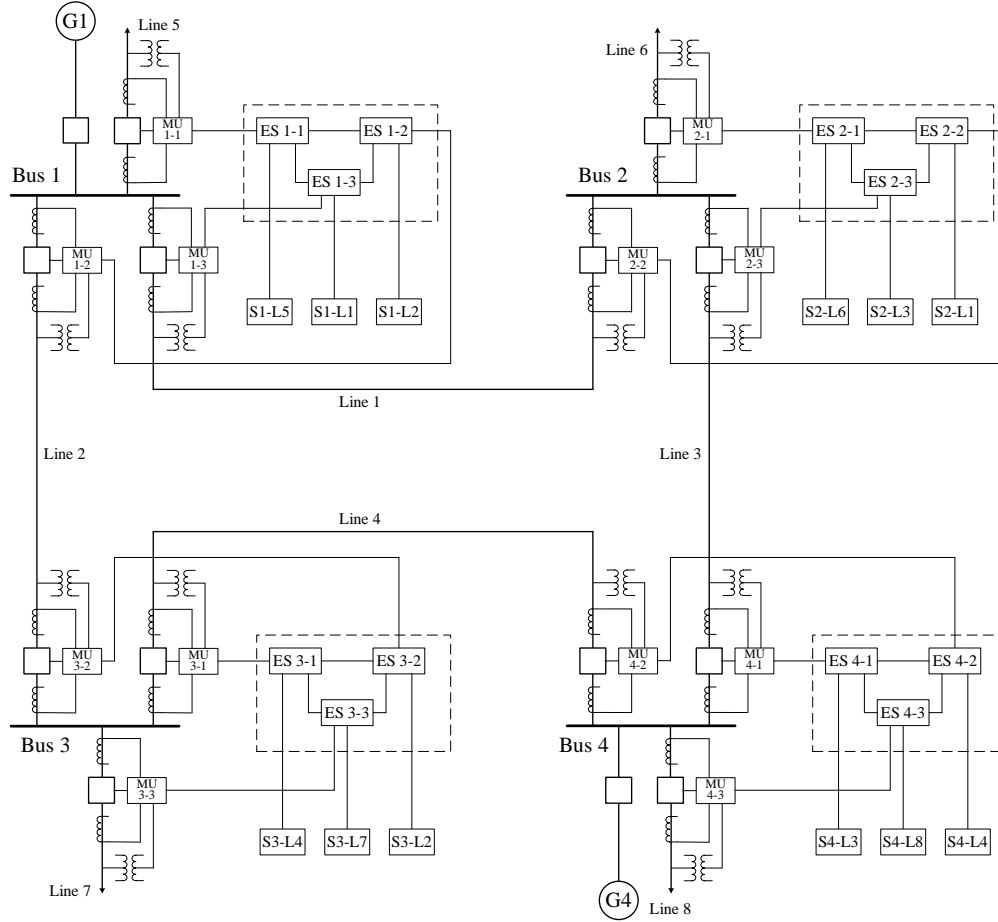


Figure 9. The integrated test system.

Table 18 Cyber component names and meanings

Component Name	Meaning
MU 1-1	Merging Unit 1 at Substation 1
MU 1-2	Merging Unit 2 at Substation 1
MU 1-3	Merging Unit 3 at Substation 1
ES 1-1	Ethernet Switch 1 at Substation 1
ES 1-2	Ethernet Switch 2 at Substation 1
ES 1-3	Ethernet Switch 3 at Substation 1
S1-L5	Line 5 Protection Panel at Substation 1
S1-L1	Line 1 Protection Panel at Substation 1
S1-L2	Line 2 Protection Panel at Substation 1

The circuit breaker reliability data for this section are based on the data from [28], [35]. The reliability data for merging units, Ethernet switches, and line protection panels are not widely available. Based on [28], [33], [34], the failure rates and Mean Repair Times are tabulated in Table 19. Components of the same category are assumed identical and therefore have the same reliability data. In this section, the current and potential transformers are assumed to never fail.

Table 19 Reliability data for components

Component	Failure Rate (/year)	Mean Repair Time (h)
Circuit Breaker	0.01	8
Merging Unit	0.02	8
Ethernet Switch	0.01	8
Line Protection Panel	0.02	8

3.2.3 Link Failure in the Cyber Network

Generally, there are two types of cyber link failures: (a) A link is unavailable due to packet delay resulting from traffic congestion or queue failure; (b) A link is physically damaged. Failure type (b) is relatively rare and thus only failure type (a) is considered in this section.

To illustrate how to model the cyber link unavailability, the cyber part of substation 1 is separated from the physical part and the cyber links are numbered from 1 to 21, as shown in Figure 10.

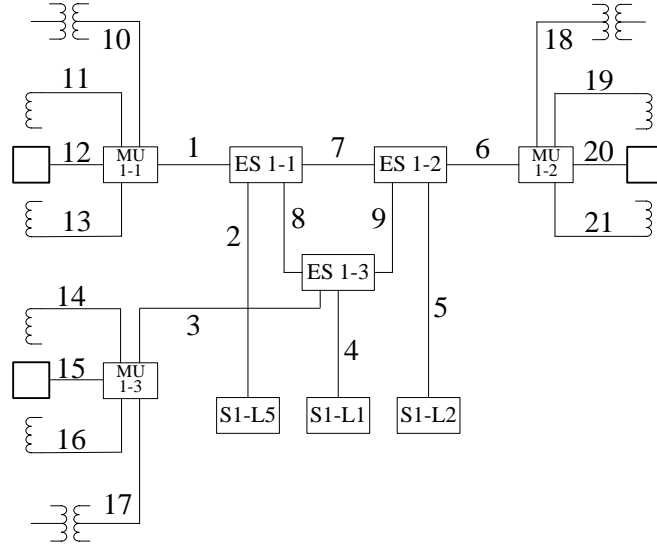


Figure 10. The cyber part of substation 1.

In Figure 10, the traffic on links 10-21 is relatively light compared with other links. Therefore, links 10-21 are considered congestion free.

Each link is bidirectional and each direction has a queue. Consider a link i connecting components a and b . The time it takes for a packet to travel from component a to b is a random variable denoted by $t_{i,1}$. For the reverse direction (from b to a), the random time is denoted by $t_{i,2}$. Depending on the arrival and departure stochastic processes associated with each queue on link i , the values of $t_{i,1}$ and $t_{i,2}$ follow some probability distribution functions.

Consider the communication from component a to x . We say this communication is unavailable if the time it takes on every possible path from a to x is greater than a predefined threshold delay value.

For example, consider the communication from MU 1-1 to S1-L1. There are two possible paths, 1-8-4 and 1-7-9-4. Assuming that all associated links are in forward directions, the probability of communication path failure is:

$$p_{fail} = \Pr[(t_{1.1} + t_{8.1} + t_{4.1} > T_{tsd}) \text{ and } (t_{1.1} + t_{7.1} + t_{9.1} + t_{4.1} > T_{tsd})] \quad (3.1)$$

where T_{tsd} is a predefined threshold delay value for the two paths.

Therefore, the effects of link failures can be modeled as the probabilities of “communication path failure” between any two components. These probabilities can be obtained by modeling and analyzing the queuing process at both the link and the path levels. The detailed procedures are based on queuing theory and are beyond the scope of this part of the report. These probabilities are assumed directly at the path level. Only the paths between merging units and protection panels are of our interest and the corresponding probabilities are tabulated in Table 20. The probabilities given are an example. The methodology of modeling link failures proposed in this section is general, and also applicable for other probability values. For each cyber link, we assume its forward queue and reverse queue are independent. Therefore, for each path, it is assumed that the forward failure and reverse failure are independent.

Table 20 Communication path failure probabilities

From	To	Forward Path Failure Probability	Reverse Path Failure Probability
MU 1-1	S1-L5	0.002	0.002
MU 1-1	S1-L1	0.001	0.001
MU 1-1	S1-L2	0.001	0.001
MU 1-2	S1-L5	0.001	0.001
MU 1-2	S1-L1	0.001	0.001
MU 1-2	S1-L2	0.002	0.002
MU 1-3	S1-L5	0.001	0.001
MU 1-3	S1-L1	0.002	0.002
MU 1-3	S1-L2	0.001	0.001

3.3 Reliability Analysis

The overall procedures can be divided into two stages: (a) Reliability analysis of the cyber part; (b) Reliability evaluation for the entire power system.

In stage (a), the failure modes of individual cyber components and their combinations are examined and analyzed. For each cyber component, only two states, UP and DOWN, are considered. Besides, there are some assumptions:

- 1) If the primary protection fails to isolate the fault and the trip signal is transferred to adjacent protection zones, this fault can always be isolated by adjacent protection zones.
- 2) The current and potential transformers are assumed to always work.
- 3) Only the first-order primary faults are considered. The situations in which multiple primary faults occur concurrently on different lines are not considered.
- 4) All the cyber components in this system are considered statistically independent, and thus, common-mode failures are not considered.
- 5) A successful operation requires that neither the forward path nor the reverse path is in failure mode.

The objective of stage (a) is to obtain the Consequent Event Matrix (CEM) and the Cyber-Physical Interface Matrix (CPIM), in which the consequent events and their probabilities are summarized. To obtain these probabilities, the consequent events after a primary fault occurs at each line are analyzed. Two states, UP and DOWN, are considered for each component. The probability of a consequent event can be obtained by multiplying the UP/DOWN probabilities of the components associated with this event. To illustrate the analysis in stage (a), the procedures of obtaining the probabilities in the first row (corresponding to the primary fault occurring at line 1) are provided as follows. The procedures are similar for primary faults occurring at other lines.

When a primary fault occurs at line 1, there are four possible scenarios.

- 1) Both of the two terminal circuit breakers operate as intended.

For the convenience of further illustration, the circuit breaker on line 1 at bus 1 side is referred to as “breaker 1-1” and the circuit breaker on line 1 at bus 2 side is referred to as “breaker 1-2”. The operation of breaker 1-1 associates with these components: breaker 1-1, MU 1-3, ES 1-3, Line Protection Panel S1-L1, forward communication path, and reverse communication path. The UP and DOWN probabilities of each component can be calculated from the data in Tables 19 and 20. Therefore, the successful operation probability of breaker 1-1, denoted by p_1 , can be obtained by multiplying the UP

probabilities of all associated components. Similarly, the successful operation probability of breaker 1-2, denoted by p_2 , can be obtained by multiplying the UP probabilities of the associated components in bus 2. The product of p_1 and p_2 would be the probability of this consequent event, in which only line 1 will be isolated after the primary fault.

2) Breaker 1-2 operates as intended while breaker 1-1 does not.

In this case, at least one associated component in bus 1 is in DOWN state while all associated components in bus 2 are in UP state. As a result, lines 1, 2, and 5 will be isolated. The probability of this consequent event is therefore $(1 - p_1)p_2$.

3) Breaker 1-1 operates as intended while breaker 1-2 does not.

In this case, at least one associated component in bus 2 is in DOWN state while all associated components in bus 1 are in UP state. As a result, lines 1, 3, and 6 will be isolated. The probability of this consequent event is therefore $(1 - p_2)p_1$.

4) Neither breaker 1-1 nor breaker 1-2 operates as intended.

This consequent event is a result from the failure of at least one associated component in bus 1 and at least one associated component in bus 2. The probability of this event is therefore $(1 - p_1)(1 - p_2)$. In this event, lines 1, 2, 3, 5, and 6 will be isolated.

Following similar procedures as performed above, the results of a primary fault occurring at each line are obtained and tabulated in Tables 21 and 22.

After the probabilities of all cyber induced consequent events are obtained in stage (a), the analysis proceeds to stage (b), in which the reliability evaluation is performed at the power system level. Using the two matrices obtained in stage (a), the effects of protection malfunctions can be taken into account without considering the details of the cyber part. The procedures of a Monte Carlo simulation for the reliability evaluation in stage (b) have been formulated in Section 2 [42] of this part of the report. The detailed implementation will be presented in Section 4 of this part of the report. The reliability indices at load points can be obtained after the two stages of analysis.

3.4 Results and Discussions

3.4.1 Results

After the analysis of the cyber part, the Consequent Event Matrix (CEM) and the Cyber-Physical Interface Matrix (CPIM) are obtained, as shown in Tables 21 and 22, respectively.

In the CEM, each entry is an 8-digit binary code in which each digit corresponds to the status of a line. A “1” means that the corresponding line is going out of service after a primary fault whereas a “0” means that the corresponding line is not affected. For example, an entry “11001000” in the first row means that lines 1, 2, and 5 are going out of service after a primary fault occurs on line 1. It should be noted that the CEM is a result of the failure of cyber part. If the cyber part worked as it is meant to, then a primary fault on a line will not result in the isolation of other lines.

In the CPIM, each entry gives the probability of a consequent event given that a primary fault occurs on a particular line. Each row corresponds to the location of a primary fault. For example, the probability corresponding to the event “11001000” in the first row is the probability of lines 1, 2, and 5 going out of service given that a primary fault already occurred on line 1. If the cyber part had perfect reliability, then column 1 would have probabilities 1 and other columns zero.

Table 21 The consequent event matrix

Primary Fault Location	Consequent Events			
Line 1	10000000	11001000	10100100	11101100
Line 2	01000000	11001000	01010010	11011010
Line 3	00100000	10100100	00110001	10110101
Line 4	00010000	01010010	00110001	01110011
Line 5	00001000	11001000	00000000	00000000
Line 6	00000100	10100100	00000000	00000000
Line 7	00000010	01010010	00000000	00000000
Line 8	00000001	00110001	00000000	00000000

Table 22 The cyber-physical interface matrix

Primary Fault Location	Probabilities of Consequent Events			
Line 1	0.9919152	0.0040342	0.0040342	0.0000164
Line 2	0.9919152	0.0040342	0.0040342	0.0000164
Line 3	0.9919152	0.0040342	0.0040342	0.0000164
Line 4	0.9919152	0.0040342	0.0040342	0.0000164
Line 5	0.9959494	0.0040506	0	0
Line 6	0.9959494	0.0040506	0	0
Line 7	0.9959494	0.0040506	0	0
Line 8	0.9959494	0.0040506	0	0

3.4.2 The Effect of Path Failure Probability

In section 3.2.3, a probability with value 0.002 is assumed, as both the forward and the reverse path failure probabilities for the communication path between a merging unit and the corresponding line protection panel. For the convenience of illustration, this probability is called the probability of main path failure. In practice, its value may vary due to data traffic or some other factors. Its value also has an important impact on the probability of successful operations (such as the event “10000000”).

Table 23 shows the relationship between the probability of main path failure and the probability of only line 1 being isolated given that a primary fault already occurred on line 1. Consider the symmetry of the system configuration, similar relationships exist for the primary faults on other lines.

Table 23 Effect of main path failure on successful operation for line 1

Main Path Failure Probability	Probability of Successful Operation	Compared to 0.9919152
0.002	0.9919152	0.000%
0.004	0.9839879	- 0.799%
0.006	0.9761082	- 1.594%
0.008	0.9682758	- 2.383%
0.01	0.9604907	- 3.168%
0.02	0.9222671	- 7.022%
0.04	0.8492535	- 14.382%

3.4.3 Discussions

From the results shown in Tables 21 and 22, it can be seen that the probabilities of undesired trips due to cyber failures are relatively small. However, such events have significant impact on system reliability. For example, when a primary fault occurs on line 1, the probability of lines 1, 2, 3, 5, and 6 being isolated concurrently is only 0.0000164. But if this consequent event happens, buses 1 and 2 will be isolated from the system with significant amount of load affected. Furthermore, such events may possibly cause severe stability issues. Therefore, the impact of cyber failures on power system reliability is significant. More detailed evaluation of such impact will be performed in Section 4 of this part of the report.

The results shown in Table 23 indicate a close relationship between the link failure and successful operation probability. The probability of successful operation decreases drastically with increasing communication path failure probability. In some severe cases, for instance, when the main path failure probability increases to 0.02 due to heavy data traffic or queue failure, the probability of successful operation drops below 0.95, which is not acceptable.

3.5 Summary

In this section, a power system configuration is extended to include the Ethernet-based protection architectures with the consideration of cyber link failures. A systematic methodology is implemented to evaluate the reliability of this extended system. The analysis performed in this section is mostly at the substation level. In the following section, reliability evaluation at the composite system level will be performed.

4. Composite Power System Reliability Evaluation Considering Cyber-Malfunctions in Substations*

4.1 Introduction

In composite power system reliability evaluation, due to the variety of protection system architectures as well as the diversity of control and communication mechanisms, it is hard to explicitly model protection systems with detailed configurations. As a result, in most of the previous work, protection system failures were either concentrated on circuit breaker trip mechanisms [10] or represented abstractly by multistate models [8], [19]-[21], in which the protection system was treated as a compact object. Some important technical details inside the protection system, such as the placement of cyber elements (e.g., CT/PTs, MUs, and IEDs) and their wire connections, were absent in those publications. Due to the absence of such details, the interdependencies between cyber elements and physical components were not sufficiently covered. In [4] and [5], to study the direct and indirect cyber-physical interdependencies, some mathematical terms and operations were defined and proposed with applications on small test systems including monitoring, control, and protection features. The results in [4] and [5] provide valuable information that indicates the impact of cyber element failures on physical system reliability indices. However, excessive self-defined reliability terms and tedious mathematical operations were introduced in [4] and [5]. These terms are hardly available from engineering practice, making it difficult to implement the overall methodology in practical applications.

A more systematic and scalable methodology was proposed in Section 2 [42] of this part of the report. This methodology performs the overall analysis in a tractable fashion with the use of *Cyber-Physical Interface Matrix (CPIM)*. In Section 2, a typical substation protection system with detailed architecture was designed and analyzed as an example to illustrate the procedures of obtaining a CPIM. The steps on how to use a CPIM in composite power system reliability evaluation were also formulated. The substation model was further enhanced in Section 3 with the consideration of cyber-link failures.

The composite power system displayed in Section 2 is simple and is used for illustration only. The overall methodology with the use of CPIM needs to be further demonstrated

* Part of this section is reprinted from copyrighted material with permission from Elsevier.

© 2015 Elsevier. Reprinted, with permission, from Hangtian Lei and Chanan Singh, "Power System Reliability Evaluation Considering Cyber-malfunctions in Substations," *Electric Power Systems Research*, vol. 129, pp. 160-169, Dec. 2015.

with its implementation on a standard test system so that the impact of protection failures on system-wide reliability indices can be numerically validated. Also, the scalability of the overall methodology needs further illustration as this is very important to its application for large power systems. Moreover, the unavailability of a standard reliability test system containing practical protection features is an obstacle for validation of the impact of protection failures on system-wide reliability indices. Extending a standard reliability test system with detailed descriptions on the cyber part would be beneficial for future studies in this area. With these objectives, this section continues and enhances the work that has been performed in Section 2. The remainder of this section is organized as follows: Section 4.2 outlines the overall methodology. Section 4.3 presents the test system configuration and parameters. In Section 4.4, the overall analysis, including the reliability analysis at the substation level and the reliability evaluation at the composite system level, is performed. Also, the results are presented and summarized. The scalability of the overall methodology is illustrated in Section 4.5. Some major considerations in software implementation for large power systems are discussed in Section 4.6. Section 4.7 is the summary of this section.

4.2 Methodology Outline and Objectives

The cyber-physical interdependencies exist in many aspects of power systems, including, but not limited to supervisory control, protection, monitoring, and metering. This section focuses on the aspect of protection, since protection hidden failures are recognized as common causes of expanded outages and have significant impact on power system reliability [7]-[10], [19]-[21].

In this section, reliability evaluation is performed in a composite power system consisting of current-carrying components and protection systems. The Roy Billinton Test System (RBTS) [46] is used as the test system with extensions at load buses to include detailed configuration in terms of protection system elements.

The size of this system is small to permit reasonable time for extension of cyber part and development of interface matrices, but the configuration of this system is sufficiently detailed to reflect the actual features of a practical system [47]. The methodology performed in this section also applies for large systems. For large systems, in spite of more efforts needed in detailed analysis of cyber failure modes, as well as effects on the physical side, the main procedures are identical to those performed in this section. In short, the selected system is adequate to illustrate the methodology and extension to larger systems is more mechanical effort rather than illustrating the validity of the technique.

The overall analysis mainly consists of two stages: 1) Reliability analysis of protection systems at the substation level; 2) Reliability evaluation from the system-wide perspective.

4.2.1 Reliability Analysis at the Substation Level

The failure modes of protection systems in terms of basic cyber elements and their relationships to transmission line tripping scenarios are analyzed in this stage. The CPIMs, which depict the interdependencies among the failures of physical components due to various cyber failure modes, are obtained at the end of this stage.

4.2.2 Reliability Evaluation from the System-wide Perspective

In this stage, a sequential Monte Carlo simulation is performed on the composite system to obtain system-wide reliability indices. The results of CPIMs obtained in the previous stage are directly utilized in this stage without the necessity of considering protection system configuration details. At the end of this stage, system-wide reliability indices, such as Loss of Load Probability (LOLP), Loss of Load Expectation (LOLE), Expected Energy Not Supplied (EENS), and Expected Frequency of Load Curtailment (EFLC), for each bus and for the overall system, can be obtained.

4.2.3 System-wide Reliability Indices

The following system-wide reliability indices [8], [20], [47] are defined and used in this section.

4.2.3.1 Loss of Load Probability (LOLP)

$$LOLP = \sum_{i=1}^{N_s} \frac{H_i t_i}{t_{total}} \quad (4.1)$$

where,

N_s	Total number of iterations simulated;
H_i	Equals 1 if load curtailment occurs in the i^{th} iteration; otherwise it equals 0;
t_i	Simulated time in the i^{th} iteration, with the unit of year;
t_{total}	Total simulated time, with the unit of year.

Since a sequential Monte Carlo simulation is performed, “iteration” here means a time period between two instants of system state change.

4.2.3.2 Loss of Load Expectation (LOLE)

$$LOLE = LOLP \cdot 8760 \quad (4.2)$$

with the unit of hours/year.

4.2.3.3 Expected Energy not Supplied (EENS)

$$EENS = \sum_{i=1}^{N_s} \frac{8760 R_i t_i}{t_{total}} \quad (4.3)$$

with the unit of MWh/year,

where,

N_s	Total number of iterations simulated;
R_i	Load curtailment during the i^{th} iteration, with the unit of MW;
t_i	Simulated time in the i^{th} iteration, with the unit of year;
t_{total}	Total simulated time, with the unit of year.

4.2.3.4 Expected Frequency of Load Curtailment (EFLC)

$$EFLC = \sum_{i=2}^{N_s} \frac{Z_i}{t_{total}} \quad (4.4)$$

with the unit of (/year),

where,

N_s	Total number of iterations simulated;
Z_i	Equals 1 if load curtailment does not happen in the $(i-1)^{\text{th}}$ iteration AND load curtailment happens at the i^{th} iteration; otherwise it equals 0;
t_{total}	Total simulated time, with the unit of year.

4.3 Test System Configuration

The Roy Billinton Test System (RBTS) [46] is used as the test system in this section. The single line diagram of the RBTS is shown in Figure 11. The bus, generation, load, and transmission line data are also provided in this section. 100 MVA and 230 kV are used as the base values of power and voltage throughout this section.

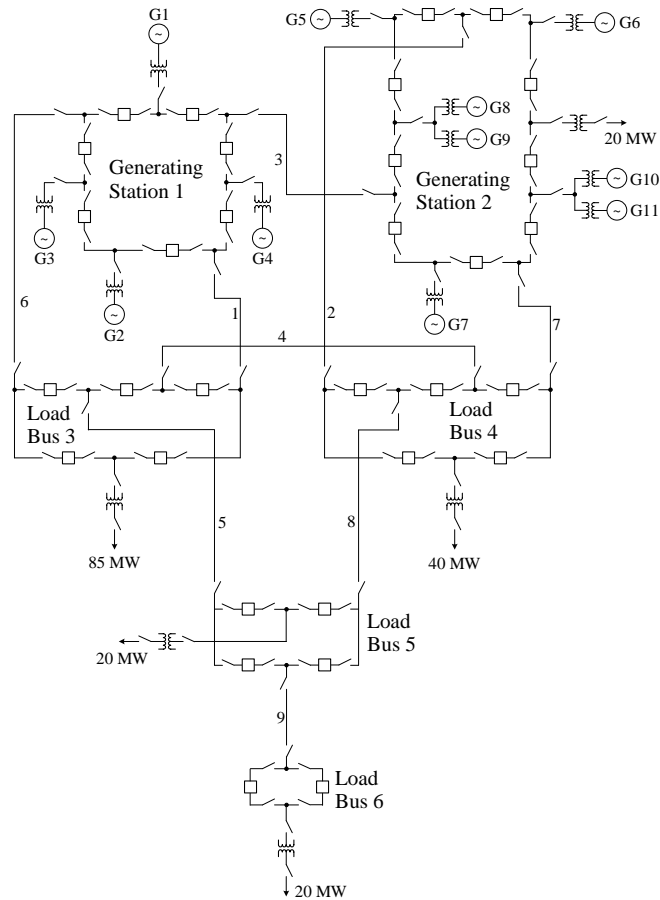


Figure 11. Single line diagram of the RBTS.

4.3.1 Bus, Generation, and Load Data

The data for all the buses and generating units are obtained from [46] and are tabulated in Tables 24 and 25, respectively. A DC optimal power flow model is used in case of load curtailment. Therefore, only the real power data are considered.

Table 24 Bus data

Bus No.	Name in Figure 11	Peak Load (p.u.)	Generation Capacity (p.u.)
1	Generating Station 1	0.00	1.10
2	Generating Station 2	0.20	1.30
3	Load Bus 3	0.85	0
4	Load Bus 4	0.40	0
5	Load Bus 5	0.20	0
6	Load Bus 6	0.20	0

Table 25 Generating unit data

Unit No.	Bus	Rating (MW)	Failure Rate (per year)	MRT (hours)
1	1	40	6.0	45
2	1	40	6.0	45
3	1	10	4.0	45
4	1	20	5.0	45
5	2	5	2.0	45
6	2	5	2.0	45
7	2	40	3.0	60
8	2	20	2.4	55
9	2	20	2.4	55
10	2	20	2.4	55
11	2	20	2.4	55

4.3.1.1 Generation Variation

The generators are represented by reliability models with two states, UP and DOWN. The corresponding failure rate and Mean Repair Time (MRT) are obtained from [46] and are tabulated in Table 25.

4.3.1.2 Load Variation

The annual peak load data for each bus are obtained from [46] and are shown in Table 24. The hourly load profile is created based on the information in Tables 1, 2, and 3 of the IEEE Reliability Test System [48].

4.3.2 Transmission Line Data

The transmission line physical parameters and outage data are obtained from [46] and are tabulated in Tables 26 and 27, respectively.

A DC optimal power flow model with simplified line parameters is used in case of load curtailment. Therefore, the line resistance (R) as well as the charging susceptance (B) are not considered in the transmission line model and only the line reactance (X) is provided in Table 26. Furthermore, in the DC optimal power flow model, since the voltage magnitude at each bus is assumed to be 1.0 p.u., the current rating for each line shown in Table 26 is numerically equal to the power rating.

For the transmission line outage data, compared with [46], the transient outage (normally with duration of less than one minute) is not considered in this section. Instead, a new term *switching time* is defined. The switching time for each transmission line, which is tabulated in Table 27, defines the time needed to switch a line back to service when this line is tripped due to a protection failure rather than resulting from a primary fault occurs at this line. The reciprocal of a switching time is called a *switching rate* and has been illustrated in Section 2 of this part of the report.

Table 26 Transmission line physical parameters

Line No.	Buses		Reactance X (p.u.)	Current Rating (p.u.)
	From	To		
1	1	3	0.180	0.85
2	2	4	0.600	0.71
3	1	2	0.480	0.71
4	3	4	0.120	0.71
5	3	5	0.120	0.71
6	1	3	0.180	0.85
7	2	4	0.600	0.71
8	4	5	0.120	0.71
9	5	6	0.120	0.71

Table 27 Transmission line outage data

Line No.	Buses		Permanent Outage Rate (per year)	Outage Duration (hours)	Switching Time (hours)
	From	To			
1	1	3	1.5	10.0	4.0
2	2	4	5.0	10.0	4.0
3	1	2	4.0	10.0	4.0
4	3	4	1.0	10.0	4.0
5	3	5	1.0	10.0	4.0
6	1	3	1.5	10.0	4.0
7	2	4	5.0	10.0	4.0
8	4	5	1.0	10.0	4.0
9	5	6	1.0	10.0	4.0

4.3.3 Protection System Architecture and Reliability Data

For bus 6, since it is connected with only one transmission line (line 9), even if its own protection system fails, line 9 will always be de-energized by opening the breakers at bus 5 without isolating any other lines. Therefore, the protection system configuration at bus 6 is not considered and only buses 3, 4, and 5 in the RBTS are extended to include detailed protection system configurations, as shown in Figures 12, 13, and 14, respectively.

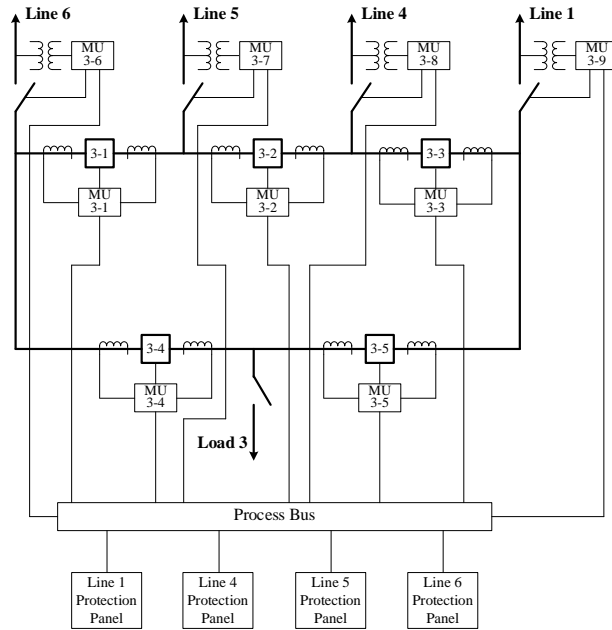


Figure 12. The protection system for bus 3.

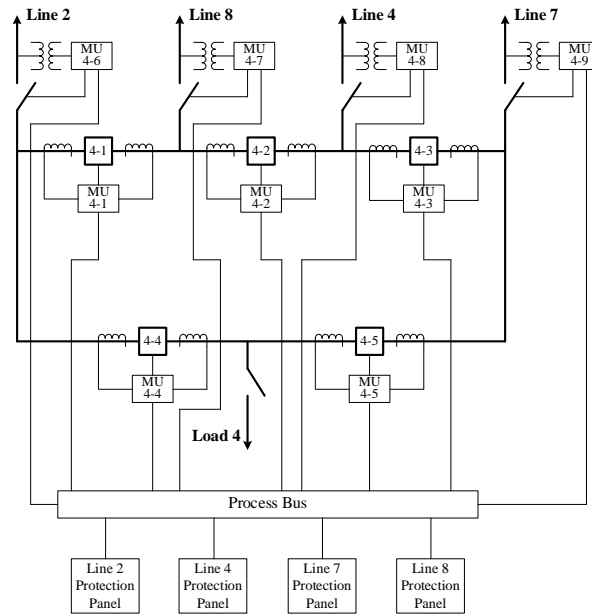


Figure 13. The protection system for bus 4.

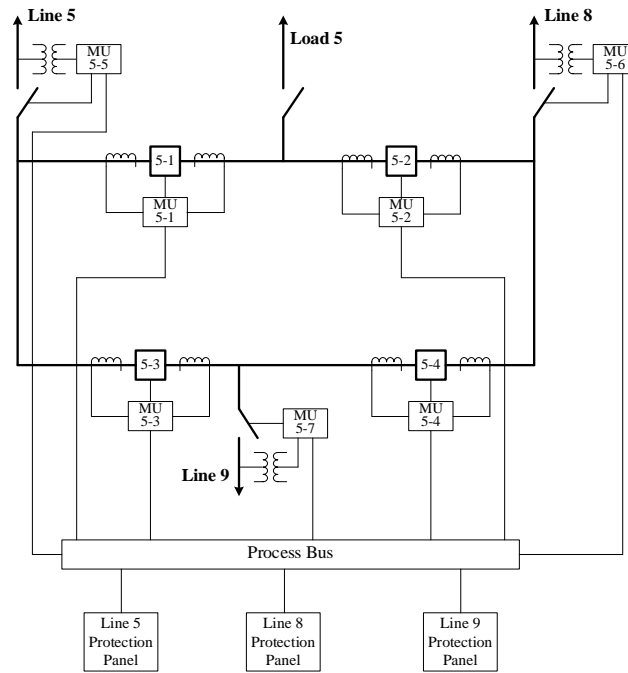


Figure 14. The protection system for bus 5.

The reliability data for Circuit Breakers (CBs), Merging Units (MUs), Process Buses (PBs), and Line Protection Panels are tabulated in Table 28. We assume that the same type of elements at different substations are identical and thereby have the same reliability data.

According to engineering practice, the Mean Time to Failure (MTTF) varies for Circuit Breakers at different voltage levels, or serving different functions in the system [35]. For the study in this section, a typical value of 100 years is chosen for the MTTF and a value of 8 hours is used for the Mean Repair Time (MRT).

The reliability data for MUs, PBs, and Line Protection Panels are reasonably chosen based on the information from [15], [28], [33], and [34].

Table 28 Reliability data for protection system elements

Element Name	MTTF (year)	Failure Rate λ (/year)	MRT (h)	Repair Rate μ (/year)
CB	100	0.01	8	1095
MU	50	0.02	8	1095
PB	100	0.01	8	1095
Line Protection Panel	50	0.02	8	1095

In this study, only two states, UP and DOWN, are considered for each protection system element (except the process bus) listed in Table 28. The state transition diagram is shown in Figure 15. The failure and repair rates are denoted by λ and μ , respectively.

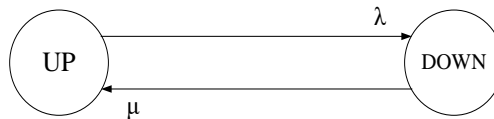


Figure 15. State transition diagram of individual element.

The exponential distribution is assumed for state residence times of each element, the probabilities of UP and DOWN can be calculated using equations (4.5) and (4.6), respectively.

$$p_{UP} = \frac{\mu}{\lambda + \mu} \quad (4.5)$$

$$p_{DOWN} = \frac{\lambda}{\lambda + \mu} \quad (4.6)$$

For the Process Bus (PB), an additional state representing DELAY is included as shown in Figure 16. The probability of delay given that the PB is not in the DOWN state is denoted by $p_d (=0.003)$. The illustration of this reliability model as well as the discussion regarding delay issues in substation communication networks have been presented in Section 2 of this part of the report.

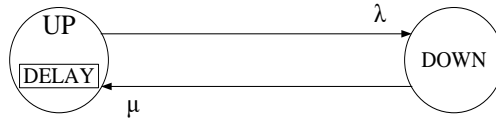


Figure 16. State transition diagram of the process bus.

Therefore, for the Process Bus, the UP, DELAY and DOWN probabilities can be calculated using equations (4.7)-(4.9).

$$p_{UP.PB} = \frac{\mu}{\lambda + \mu} (1 - p_d) \quad (4.7)$$

$$p_{DELAY.PB} = \frac{\mu}{\lambda + \mu} p_d \quad (4.8)$$

$$p_{DOWN.PB} = \frac{\lambda}{\lambda + \mu} \quad (4.9)$$

In reality, the process bus is a network consisting of basic elements that are connected with each other in various topologies, and thus more sophisticated technical details are involved [34], [38]-[41]. The consideration of these technical details in composite system reliability evaluation is beyond the scope of this part of the report.

The assumptions regarding other protection elements (e.g., CTs, PTs, and cable links) and protection issues such as backup tripping follow those stated in Section 2 of this part of the report. The CTs, PTs, and cable links are assumed not to fail. In addition, based on the features of this particular test system, several more assumptions are made:

- 1) The failure of an MU that is connected to a PT will result in the failure of acquired voltage information and thus will disable the primary protection of this line. As a result, multiple breakers associated with the primary protection will fail to trip and backup protections will be triggered. For example, in the bus 3 protection system (shown in Figure 12), if a primary fault happens at Line 6 but MU 3-6 fails, then the Line 6 Protection Panel will fail to issue trip signals to both breakers 3-1 and 3-4. As a result, backup protection zones will be triggered and breakers 3-2 and 3-5 will trip to isolate Line 6.
- 2) Since this study focuses on transmission system reliability evaluation, the details of a load branch can be extended in the distribution system. Therefore, primary faults that occur at load branches are not considered. However, the isolation of a load branch resulting from undesired trips due to primary faults that occur at adjacent transmission lines will be considered.

4.4 Reliability Analysis

The overall analysis mainly consists of two stages: the reliability analysis of protection systems at the substation level and the reliability evaluation from the system-wide perspective. The CPIM, which bridges the two stages, is a critical idea of this methodology. It decouples the analysis at the substation level from the evaluation of the composite system and makes the overall analysis more tractable.

4.4.1 Substation Level Reliability Analysis

The substation level reliability analysis follows the procedures described in Section 2 of this part of the report with the objective of obtaining CPIMs.

This section improves the CPIM that was described in Section 2 by eliminating the off-diagonal zeroes to make it more compact. In this section, each row in a CPIM represents a physical component (transmission line). Each column provides the probability of a

consequent event given that a primary fault occurred on this physical component. Therefore, the probabilities in each row sum up to 1. If the protection system is perfectly reliable, then the first column would have probabilities 1 and other columns zero.

In addition, another matrix, Consequent Event Matrix (CEM), is developed in accordance with a CPIM. A CEM provides detailed information about consequent events in which some lines go out of service while some are not affected. In a CEM, each event is coded as a 12-digit binary number, of which the left 9 digits correspond to the 9 transmission lines and the last 3 digits correspond to load branches 3, 4, and 5, respectively. A “1” digit indicates the corresponding component is going out of service whereas a “0” means this component is not affected. For example, an entry “100001100110” denotes a consequent event in which line 1, line 6, line 7, load branch 3, and load branch 4 are going out of service. A complete row of a CEM summarizes all possible consequent events when a primary fault occurs at this transmission line.

To illustrate how the malfunctions of cyber elements affect transmission line tripping behaviors, the detailed analysis for the consequent events resulting from cyber element failures at substation (bus) 3 following a primary fault occurs at line 1 is shown below as an example. The analysis for the primary faults at other lines can be performed similarly. In the analysis, the failure modes of individual cyber elements are assumed independent since they are located in different units in a substation. Therefore, the probability of a consequent event can be obtained by multiplying the probabilities of individual element states in this event.

Suppose a primary fault occurs at line 1, all possible consequent events can be categorized as follows:

- 1) All protection elements operate as intended.

If all protection elements operate as intended, then only line 1 will be isolated. The action of line 1 tripping associated with these elements at substation 3: MU 3-9, CB 3-3, MU 3-3, CB 3-5, MU 3-5, Process Bus, and Line 1 Protection Panel. Multiply the UP probabilities of all these elements, the corresponding probability of this consequent event can be obtained, which is 0.996899850569.

- 2) The Process Bus (PB) fails.

If the PB fails, then the entire substation will be affected by this fault. All lines connected to this substation will be isolated by tripping the breakers at remote substations. The corresponding probability of this consequent event can be calculated using Equation (4.9). This is an extreme case; therefore, the probability is very low. However, once this event happens, the impact is tremendous.

- 3) One or both of MU 3-3, CB 3-3 fail(s), while all other associated elements operate as intended.

In this case, CB 3-3 fails to trip while CB 3-5 trips as intended. The fault will be cleared by opening CB 3-2 and CB 3-5. As a result, Lines 1 and 4 will be isolated.

- 4) One or both of MU 3-5, CB 3-5 fail(s), while all other associated elements operate as intended.

In this case, CB 3-5 fails to trip while CB 3-3 trips as intended. The fault will be cleared by opening CB 3-3 and CB 3-4. As a result, Line 1 and load branch 3 will be isolated.

- 5) The Process Bus (PB) doesn't fail, but both CB 3-3 and CB 3-5 fail to trip due to various combinations of element states, such as Line 1 Protection Panel fails or the PB is in a DELAY state.

In this case, the fault will be cleared by opening CB 3-2 and CB 3-4. As a result, Line 1, Line 4, and load branch 3 will be isolated.

The results of all the 5 cases above are summarized in the first row of Table 29 and Table 30. It should be noted that these consequent events are the results from cyber element failures. If all the associated cyber elements are perfectly reliable, then the first case would have probability one while all other cases zero.

Following similar procedures performed above, the complete Cyber-Physical Interface Matrices (CPIMs) and Consequent Event Matrices (CEMs) for buses 3, 4, and 5 are obtained and are shown from Table 29 to Table 34.

Table 29 The cyber-physical interface matrix for bus 3

Fault Location	Probabilities				
Line 1	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 4	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 5	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 6	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2

Table 30 The consequent event matrix for bus 3

Fault Location	Events				
Line 1	100000000000	100111000000	100100000000	100000000100	100100000100
Line 4	000100000000	100111000000	000110000000	100100000000	100110000000
Line 5	000010000000	100111000000	000011000000	000110000000	000111000000
Line 6	000001000000	100111000000	000001000100	000011000000	000011000100

Table 31 The cyber-physical interface matrix for bus 4

Fault Location	Probabilities				
Line 2	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 4	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 7	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 8	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2

Table 32 The consequent event matrix for bus 4

Fault Location	Events				
Line 2	010000000000	010100110000	010000000010	010000010000	010000010010
Line 4	000100000000	010100110000	000100010000	000100100000	000100110000
Line 7	000000100000	010100110000	000100100000	000000100010	000100100010
Line 8	000000010000	010100110000	010000010000	000100010000	010100010000

Table 33 The cyber-physical interface matrix for bus 5

Fault Location	Probabilities				
Line 5	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 8	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2
Line 9	0.99689985056 9	0.00000913233 7	0.00002731249 1	0.00002731249 1	0.00303639211 2

Table 34 The consequent event matrix for bus 5

Fault Location	Events				
Line 5	000010000000	000010011000	000010001000	000010000001	000010001001
Line 8	000000010000	000010011000	000000010001	000000011000	000000011001
Line 9	000000001000	000010011000	000000011000	000010001000	000010011000

4.4.2 System-wide Reliability Evaluation

The next event sequential Monte Carlo simulation [37] forms the main framework for the reliability evaluation at this stage. The detailed steps, including illustrations on how to utilize the results of a CPIM in the composite system reliability evaluation, have been illustrated in Section 2 and are summarized as follows:

- 1) Initialize.
- 2) Determine a primary event: Find the minimum time to the next event, update the corresponding element's state, and update the total time.
- 3) Determine consequent events: If the state change in step 2) indicates a primary fault occurring at a transmission line, then use CPIMs and CEMs to determine the consequent events and update elements' states accordingly. If a CPIM row corresponding to this transmission line has n consequent events, the probabilities of these events (p_1, p_2, \dots, p_n) sum up to 1. Draw a random number ranging from 0 to 1. The value of this random number determines which consequent event is going to happen. It should be noted that a transmission line connects two substations. Therefore, two random numbers

should be drawn independently to determine the consequent event at each substation.

- 4) Effects of switching and repair: For the elements whose states have been changed in step 2) or in step 3), draw new random numbers to determine the time of their next transitions. Appropriate transition rates should be used according to situations.
- 5) Evaluate system state: Perform the network power flow analysis to assess system operation states. Update reliability indices.
- 6) Repeat steps 2) to 5) until convergence is achieved.

In step 5), the following DC power flow-based linear programming model [49]-[51] is used with the objective of minimizing total load curtailment.

$$\text{Objective: } y = \text{Min} \sum_{i=1}^{N_b} C_i$$

subject to:

$$\hat{B}\theta + G + C = L \quad (4.10)$$

$$G \leq G^{max}$$

$$C \leq L$$

$$DA\theta \leq F^{max}$$

$$-DA\theta \leq F^{max}$$

$$G, C \geq 0$$

$$\theta_1 = 0$$

$$\theta_{2...N_b} \text{ unrestricted}$$

where,

N_b	Number of buses
C	$N_b \times 1$ vector of bus load curtailments
C_i	Load curtailment at bus i

\hat{B}	$N_b \times N_b$ augmented node susceptance matrix
G	$N_b \times 1$ vector of bus actual generating power
G^{max}	$N_b \times 1$ vector of bus maximum generating availability
L	$N_b \times 1$ vector of bus loads
D	$N_t \times N_t$ diagonal matrix of transmission line susceptances, with N_t the number of transmission lines
A	$N_t \times N_b$ line-bus incidence matrix
θ	$N_b \times 1$ vector of bus voltage angles
F^{max}	$N_t \times 1$ vector of transmission line power flow capacities

In equation (4.10), the variables are vectors θ , G , and C . Thus, the total number of variables is $3N_b$. This problem can be solved by using the *linprog* function provided in MATLAB software.

The convergence is measured by the coefficient of variation of a chosen index, as defined in [47]. A simulation with 200 simulated years is performed and the coefficient of variation for the system EENS drops below 5%.

4.4.3 Results and Discussions

The simulation results of LOLP, LOLE, EENS, and EFLC for each bus and for the overall system are tabulated in Table 35. The simulated transmission line failure rates due to primary faults and protection system malfunctions are tabulated in Table 36. In Table 36, for line 3, the simulated line failure rate due to protection system malfunctions equals 0. This is because line 3 links bus 1 to bus 2 and protection malfunctions are not considered for either of the two buses.

To make a comparison, the situation in which protection systems are assumed perfectly reliable is also simulated with results tabulated in Table 37. The comparison is also displayed in Figure 17.

Table 35 Reliability indices for buses

	LOLP	LOLE (hours/year)	EENS (MWh/year)	EFLC (/year)
Bus 1	0	0	0	0
Bus 2	0.00015926	1.395	2.655	0.260
Bus 3	0.00017063	1.495	8.597	0.300
Bus 4	0.00019288	1.690	10.095	0.315
Bus 5	0.00016786	1.470	3.729	0.275
Bus 6	0.00124176	10.878	116.104	1.305
Overall System	0.00128584	11.264	141.180	1.395

Table 36 Simulated transmission line failure rates

Line No.	Failure rate resulting from primary faults (/year)	Failure rate resulting from protection malfunctions (/year)
1	1.455	0.010
2	4.850	0.005
3	3.870	0
4	1.030	0.075
5	0.925	0.010
6	1.570	0.010
7	5.100	0.005
8	1.080	0.010
9	1.030	0.010

Table 37 EENS comparison

	EENS (MWh/year)		Δ
	If protection systems are perfectly reliable	Considering protection malfunctions	
Bus 1	0	0	N/A
Bus 2	1.862	2.655	42.59%
Bus 3	2.828	8.597	204.00%
Bus 4	1.950	10.095	417.69%
Bus 5	2.145	3.729	73.85%
Bus 6	103.947	116.104	11.70%
Overall System	112.732	141.180	25.24%

For each row, Δ is defined as the percentage increment of the EENS from not considering to considering protection malfunctions.

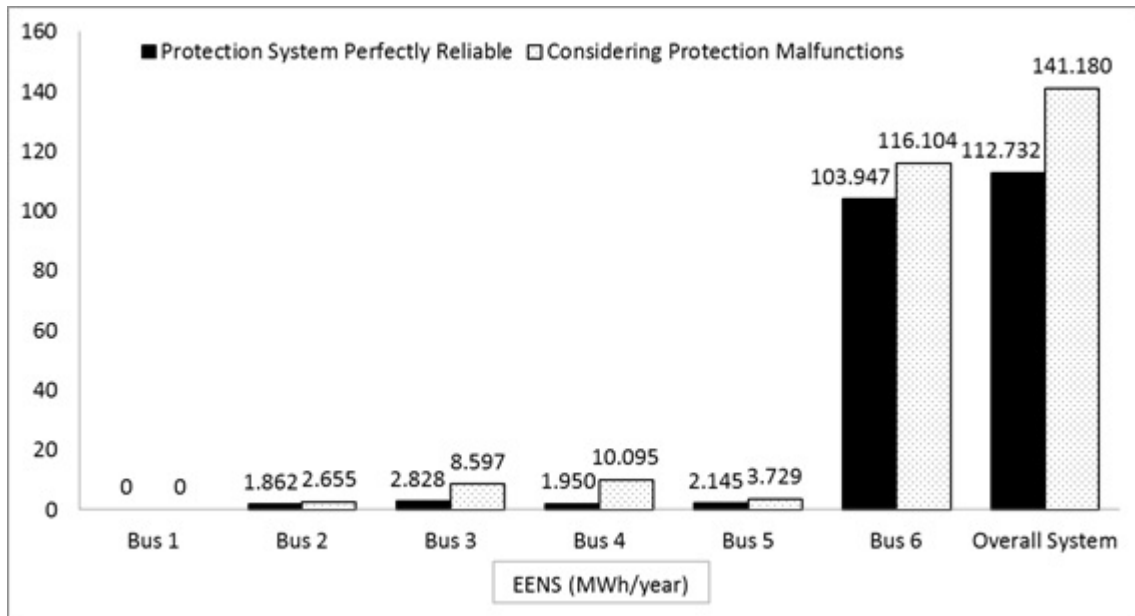


Figure 17. EENS comparison at each bus.

The results in Tables 36 and 37 show that protection system malfunctions have significant impact on energy unavailability even though they do not have much impact on individual line failure rates. Compared with not considering protection malfunctions, the percentage increment of the EENS for individual buses can be quite significant.

The effects of protection system malfunctions on EENS are noticeable for buses 3, 4, and 5, with increments of 204.00%, 417.69%, and 73.85%, respectively. These three buses are also the ones in which we have modeled and considered protection system malfunctions. This further points to the impact of protection malfunctions on energy unavailability.

4.4.4 The Effects of Switching Time

A value of 4.0 hours is assumed as the switching time for all transmission lines, and this value has been used in the analysis in Sections 4.4.1 and 4.4.2.

In engineering practice, a switching process may be accelerated with the aid of smart grid technologies, or may be prolonged due to other factors. The quantitative relationship between switching time and system EENS are studied and the results are shown in Table 38. In each case, same value of switching time is assumed for all transmission lines, and the system EENS is compared with the case in which the switching time is 4.0 hours. This relationship is also displayed in Figure 18.

Table 38 Effect of switching time on system EENS

Switching Time (hours)	System EENS (MWh/year)	Percentage increment/decrement compared with the value of 141.180 MWh/year in Table 37
0.2	115.089	-18.48%
0.5	120.941	-14.34%
1	126.945	-10.08%
2	132.675	-6.02%
4	141.180	0
10	157.615	+11.64%
24	178.986	+26.78%
48	190.628	+35.02%

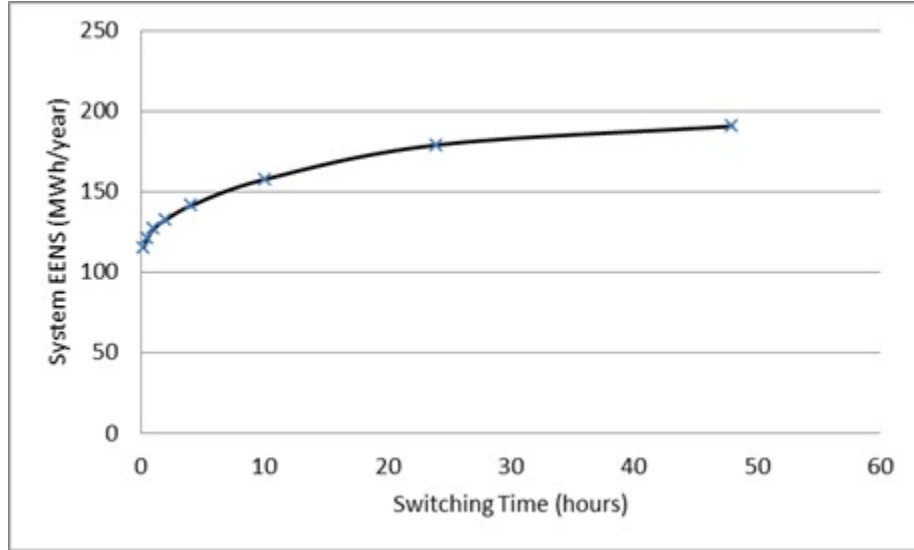


Figure 18. Relationship between switching time and system EENS.

The information in Table 38 and Figure 18 indicate a close relationship between the line switching time, and the system EENS. The value of system EENS increases considerably with prolonged switching time. This also signifies the importance of using advanced technologies with which the process of fault location and cyber failure identification would be accelerated so that healthy lines can be switched back to service more promptly.

4.5 The Scalability of the Overall Methodology

As shown in Section 4, the overall methodology consists of two stages:

- 1) Reliability analysis at the substation level (i.e., the work performed in Section 4.4.1).
- 2) System-wide reliability evaluation (i.e., the work performed in Section 4.4.2).

In the first stage, the detailed analysis depends on the actual protection architecture of a substation. The analysis may seem to be tedious for a substation with complex architecture. However, the analysis in this stage can be performed locally at each substation and the computations can be performed offline. The increased workload for more complex substations does not change the framework of the overall methodology.

It should be noted that although in this stage, the analysis is needed for each individual substation, with more experience in the analysis at substation level, it may be possible to generate classifications into types of substations and thus expedite the process.

Once the CPIMs and CEMs are established in the first stage, they can be permanently stored and can be directly plugged into the reliability evaluation in the second stage. The Monte-Carlo simulation performed in the second stage is generic and applicable for large power systems.

The CPIM decouples the first stage of analysis from the second stage and makes the overall analysis more tractable.

4.6 Considerations in Software Implementation for Large Power Systems

The proposed methodology establishes a framework for power system reliability evaluation considering cyber-malfunctions in substations. Some implementation considerations are important to its application for large power systems. This section discusses two major considerations, the CPU time for Monte Carlo simulation and the storage of matrices.

4.6.1 CPU Time for Monte Carlo Simulation

The convergence in a Monte Carlo simulation is measured by the coefficient of variation of a chosen index. In this section, simulation is performed for 200 years and the coefficient of variation for the system EENS drops below 5%. The simulation is performed in MATLAB running on a computer with a 3.10 GHz processor and the running time for a simulation is approximately 8 minutes. It should be noted that this software implementation of the simulation is only research grade to illustrate the concept and is therefore not the most efficient as far as the running time is concerned. The running time is largely consumed by the *linprog* function in MATLAB for DC power flow based linear programming to evaluate system operation states. In the development of a commercial grade program, the running time can be drastically reduced by several means as described below.

- 1) The linear programming incorporating DC power flow can be performed less frequently with the use of heuristic algorithms for screening, thus reducing the CPU time.
- 2) Simulation can be custom coded in more efficient programming languages. Custom programs are generally more efficient than generic ones coded in MATLAB.
- 3) Much more efficient methods such as interior point methods can be used for linear programming.

- 4) Monte Carlo simulation is readily amenable to parallel and distributed processing environments [52], [53] to reduce the CPU time.

It is important to mention that Monte Carlo simulation has already been successfully used for large composite power systems, but without considering the cyber malfunctions. The major contribution of this research is to develop a methodology to include cyber-induced dependent failures in such Monte Carlo programs. The cyber-induced dependent failures can be included in Monte Carlo simulation by using CPIMs. This does not significantly alter the number of times linear programming is called for and thus does not alter the CPU time much.

4.6.2 Storage of Matrices

For a given power system, let m be the total number of rows in all CPIMs (or CEMs), n be the number of columns in a CPIM (or a CEM).

The value of m depends on the number of transmission lines. The number of transmission lines in an actual power system is typically 1.2 to 2 times of the number of buses. Each transmission line contributes a row in two CPIMs (or CEMs) corresponding to both of the two buses it connects to. Consider a power system with 1000 buses (substations), which is a typical size of an actual transmission grid, it is reasonable to estimate the number of transmission lines as 2000 and thus the value of m is estimated to be 4000.

The value of n is determined by the row with maximum number of consequent events, which depends on the transmission line having maximum number of adjacent lines. Assuming a transmission line has maximum 10 adjacent lines, thus the maximum possible value of n is 2^{10} , which is 1024. Of course, it is possible that a few transmission lines may have more than 10 adjacent lines. For such lines, only the 1024 most likely consequent events are considered since the remaining consequent events have negligible probabilities. Thus, it is reasonable to estimate n as 1024.

Each entry in a CPIM can be stored as a 64-bit double-precision floating-point number. Therefore, the total storage space needed for all CPIMs is $8 \cdot m \cdot n$ Bytes, which equals 31.25 MB (32,768,000 Bytes).

Each entry in a CEM is a binary number corresponding to a consequent event. For a system with 2000 transmission lines, 2000 bits are needed to represent such an event.

Thus, each entry uses 250 Bytes (2000 bits) and the total storage space needed for all CEMs is $250 \cdot m \cdot n$ Bytes, which equals 976.5625 MB (1,024,000,000 Bytes).

Therefore, for a power system with 1000 buses, the total space needed to store all CPIMs and CEMs is estimated to be 1007.8125 MB, which is approximately 0.9842 GB. It is feasible to claim such space on hard drive or Random Access Memory (RAM).

4.7 Summary

In this section, a systematic reliability evaluation methodology is enhanced and implemented in a composite power system consisting of current-carrying components and protection systems with modern architecture. The quantitative relationship between switching time and system EENS is also studied. The results clearly indicate the impact of protection failures on system-wide reliability indices and also signify the importance of accelerating line switching process.

The methodology implemented in this section is scalable and provides an option for the reliability evaluation of large cyber-physical power systems. For such systems, in spite of more efforts needed in detailed analysis, the main procedures remain similar to those performed in this section.

It should be noted that the methodology performed in this section is based on a sequential Monte Carlo simulation. It would be significantly beneficial for its application in large systems if the efficiency is further improved with the use of non-sequential techniques since non-sequential techniques generally require less computational and storage resources compared to sequential ones. Pertinent studies will be performed in the following section.

5. Non-sequential Monte Carlo simulation for Power System Reliability Analysis Considering Dependent Failures*

5.1 Introduction

Cyber-induced dependent failures affect power system reliability and thus are important to be considered in composite system reliability evaluation. A scalable methodology is proposed in Section 2 [42] with the use of Cyber-Physical Interface Matrix (CPIM) that decouples the analysis of the cyber part from the physical part and provides the means of performing the overall analysis in a tractable fashion. In Section 4 [54], this methodology is further enhanced and implemented on an extended Roy Billinton Test System (RBTS) with the illustration of its applicability for large power systems.

The techniques used in Monte Carlo simulations can be basically classified into two categories known as sequential and non-sequential techniques [55]-[57]. The methodology presented in Section 2 and Section 4 is based on a sequential Monte Carlo technique and establishes the main framework for reliability evaluation of cyber-physical power systems. Non-sequential methods are typically easier to implement and require much less computational and storage resources as compared to sequential methods [3], [23], [24]. Therefore, it would be significantly beneficial for the application of the proposed methodology in large systems if the efficiency is further improved with the use of non-sequential techniques. However, the failure and repair processes in cyber-induced events are inherently sequential involving dependent failures, making it very difficult to utilize a non-sequential sampling throughout the complete process in the same manner as in independent components by sampling the component states individually.

It has been recognized that the basic idea in sampling is to sample states from a state space proportional to their probabilities [23]. For a large system it is not possible to first find the probabilities of all the states in the state space. Therefore, an approach is proposed in this section with the basic idea of developing a representative state space from which states can be sampled. This is achieved here by the use of sequential Monte Carlo simulation as the computational effort needed for only generating a chronological state sequence is negligible compared to the effort of evaluating the states using DC power flow based linear programming. A similar approach, called *pseudo-sequential simulation*, has been

* Part of this section is reprinted from copyrighted material with permission from IEEE. © IEEE. Reprinted, with permission, from Hangtian Lei and Chanan Singh, “Non-Sequential Monte Carlo Simulation for Cyber-Induced Dependent Failures in Composite Power System Reliability Evaluation,” *IEEE Transactions on Power Systems*, (accepted for publication).

proposed in [24] but for completely different purposes. In [24], the approach is proposed with the purpose of computing duration-related reliability indices. In this section, the purpose is to include the characteristics of cyber-induced dependent failures. Also a method is used in this section such that the state space generated by sequential simulation does not need to be stored. Furthermore, it should be noted that the sequential simulation may not be the only option for state space generation thus there may be alternative solutions. The major difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in the presence of dependent failures are also thoroughly explored in this section with the intention of enlightening future studies on this subject as composite power system reliability evaluation in the presence of dependent failures has not received sufficient attention.

The remainder of this section is organized as follows: Section 5.2 describes the overall problem. Section 5.3 illustrates the major difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in presence of dependent failures. Section 5.4 introduces the proposed method with the use of sequential technique for state space generation. The implementation of the proposed method is demonstrated in Section 5.5. Section 5.6 is the summary of this section.

5.2 Origination of Dependent Failures

Traditional power system reliability evaluation focuses on the physical part only and assumes perfect reliability for the cyber part, which neglects cyber-induced failures and results in too optimistic evaluation. For realistic evaluation, it is necessary to consider cyber-induced failures as well as their impact on composite systems.

Due to the complexity of cyber-physical interdependencies, it is hard and impractical to explicitly model and analyze the entire system, cyber and physical, in one step. A methodology has been proposed in Section 2 with the use of Cyber-Physical Interface Matrix (CPIM) to decouple the analysis of the cyber part from the physical part so that the overall evaluation is performed in a tractable fashion.

The format of a Cyber-Physical Interface Matrix is shown in Table 39. It can be obtained by examining the failure modes of cyber components and their impact on the physical system.

Table 39 The format of a cyber-physical interface matrix

	Event 1	Event 2	Event n
Component 1	p_{11}	p_{12}	p_{1n}
Component 2	P_{21}	P_{22}	P_{2n}
.....
Component m	P_{m1}	p_{m2}	P_{mn}

The number of rows in a CPIM, denoted by m , corresponds to the number of transmission lines. The number of columns, denoted by n , corresponds to the number of consequent events. Consider a transmission line l ($0 < l \leq m$) with k adjacent lines. When a fault occurs at line l , $0-k$ of its unfaulted adjacent lines can be isolated due to cyber malfunctions with 2^k maximum possible cases. The complete row l in a CPIM summarizes the probabilities of all possible consequent events (cases) given that a fault occurs at transmission line l . These probabilities are obtained by analyzing cyber element failure cases and their impact on transmission line tripping behaviors. It should be noted that these probabilities are a result of cyber failures. If the cyber part is perfectly reliable, there is only one consequent event with probability 1 in each row corresponding to the isolation of the faulty transmission line only.

In applications, another matrix called Consequent Event Matrix (CEM) is used as an auxiliary matrix for CPIM to identify specific components involved in a consequent event. The detailed illustrations of the CPIM and CEM have been presented in Section 4.

Once the CPIM is obtained, its results can be directly utilized in the composite system reliability evaluation without the necessity of considering cyber element details. The steps of a sequential Monte Carlo simulation can be outlined as follows:

- 1) Generate a chronological sequence of system state by drawing random numbers for each physical component and applying the load profile;
- 2) Evaluate each system state and update reliability indices;
- 3) If a state transition associates with a fault occurs on a transmission line, such as line i , then the i^{th} row of the CPIM is used to determine the consequent event. The probabilities of n possible events ($p_{i1}, p_{i2}, \dots, p_{in}$) in the i^{th} row sum up to 1. Draw a random number r ($0 < r < 1$). The value of r determines which consequent event is going to happen. Update all affected component states in the determined consequent event;

- 4) For the components whose states have been changed in step 2) or step 3), draw new random numbers to determine the time of their next transitions to update the chronological system state sequence;
- 5) Repeat steps 2) to 4) until the convergence criterion is achieved.

The steps shown above are completely based on sequential techniques. Non-sequential techniques typically require much less CPU time and memory as compared to sequential techniques, and thus are preferable in applications for large systems. In order to seek possible non-sequential techniques that could be applied in the situation described above, it is important to first discuss the difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in presence of cyber-induced dependent failures or other types of dependent failures, as presented in the following section.

5.3 Problem of Applying Non-sequential Sampling for Dependent Failures

It is important to first examine the basic idea of sampling a state in non-sequential simulation in order to appreciate the problem of applying it to systems with dependent failures. Let us say that the state space S is defined by all states $x \in S$ where the state x is represented by the states of components:

$$x = (x_1, x_2, \dots, x_n)$$

where x_i represents the state of component i .

Now if the probabilities of all the states x of the system (i.e., the joint probability distribution of components) are known, then we can sample states by drawing a random number between 0 and 1 and formulating a process as follows. The interval between 0 and 1 can be divided into segments equal to the number of system states, the length of a particular segment being equal to the probability of the state that the segment represents. Then one has to determine which segment the random number drawn falls in and that determines the state that is sampled. In practice it is more convenient to construct the probability distribution function of x and determine the state. For example let us assume that the system has five states with their probabilities as shown in Figure 19 (a). These can be organized as a distribution function shown in Figure 19 (b). Now a random number between 0 and 1 can be drawn and put on the probability axis and the random variable axis to give the system state sampled. One should remember that the basic idea is that the states should be sampled proportional to their probabilities.

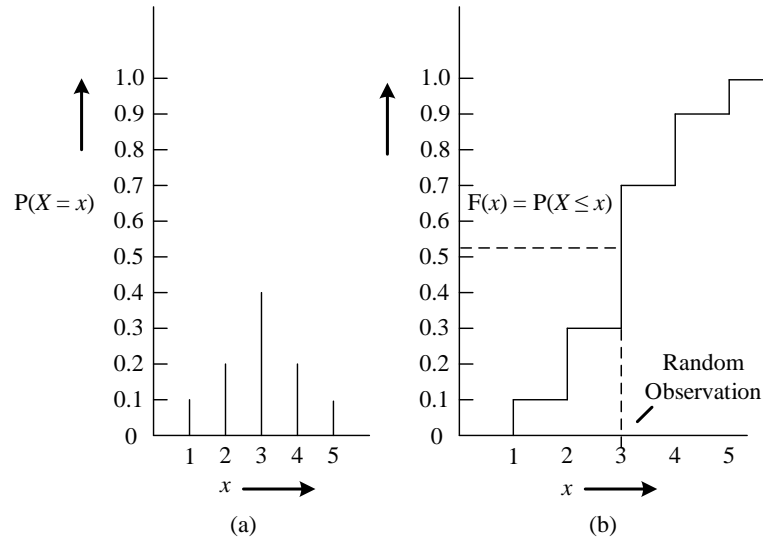


Figure 19. An example of sampling.

The difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in presence of dependent failures are discussed in this section.

A small power system with three components is shown in Figure 20. Each component has two states, UP and DOWN. The scenarios in which the components are completely independent, partially independent, and fully dependent are illustrated respectively as follows.

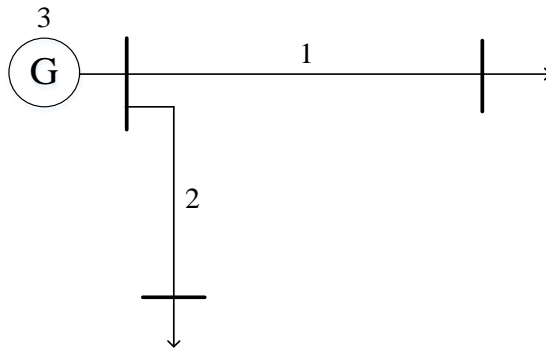


Figure 20. A three-component system.

5.3.1 Completely Independent Scenario

In this scenario, the failure modes of all three components shown in Figure 20 are independent from each other. The corresponding state space transition diagram for the system is shown in Figure 21.

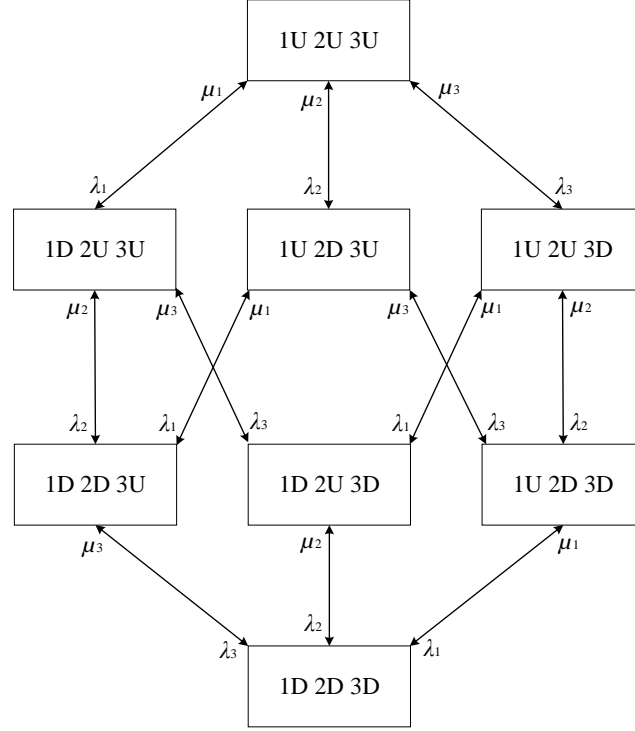


Figure 21. The system state space diagram for completely independent scenario.

The letters “U” and “D” represent Up and Down states, respectively. λ_i ($1 \leq i \leq 3$) represents the failure rate of component i . μ_i ($1 \leq i \leq 3$) represents the repair rate of component i .

Because of the independence of individual component failure modes, the system state space can be decoupled into component states, as shown in Figure 22. When performing a non-sequential sampling, the states of individual components can be sampled independently and their combination yields a system state. For example, if the sampled states of individual components are 1U, 2D, and 3D, then the system state is (1U, 2D, 3D). It is actually this ability to sample a system state by a combination of component states that makes the non-sequential method powerful as the probabilities of all the system states

do not need to be known before sampling but only the probabilities of component states need to be known. The components can be two-state or multi-state but so long as they are independent, their states can be sampled at the component level.

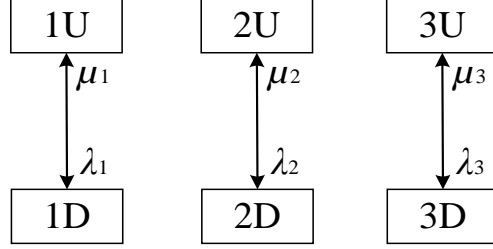


Figure 22. The state transition diagrams for individual components.

The UP and DOWN probabilities of individual components are readily computable. With the assumption of exponential reliability distribution, the UP and DOWN probabilities can be computed using equations (5.1) and (5.2), respectively.

$$p_{Up-i} = \frac{\mu_i}{\lambda_i + \mu_i} \quad (5.1)$$

$$p_{Down-i} = \frac{\lambda_i}{\lambda_i + \mu_i} \quad (5.2)$$

5.3.2 Partially Independent Scenario

In this scenario, components 1 and 2 have a common failure mode with failure rate λ_{c12} , as shown in Figure 23.

Component 3 is still independent from components 1 and 2. The corresponding state space transition diagram for the system is shown in Figure 24.

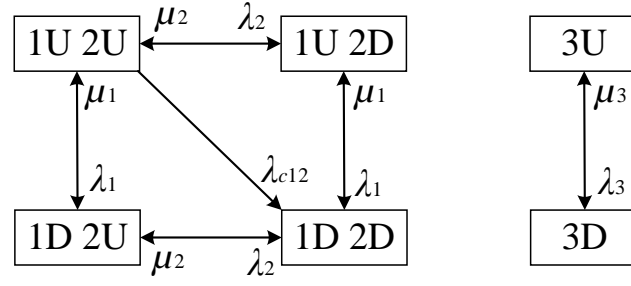


Figure 23. The state transition diagrams for partially independent scenario.

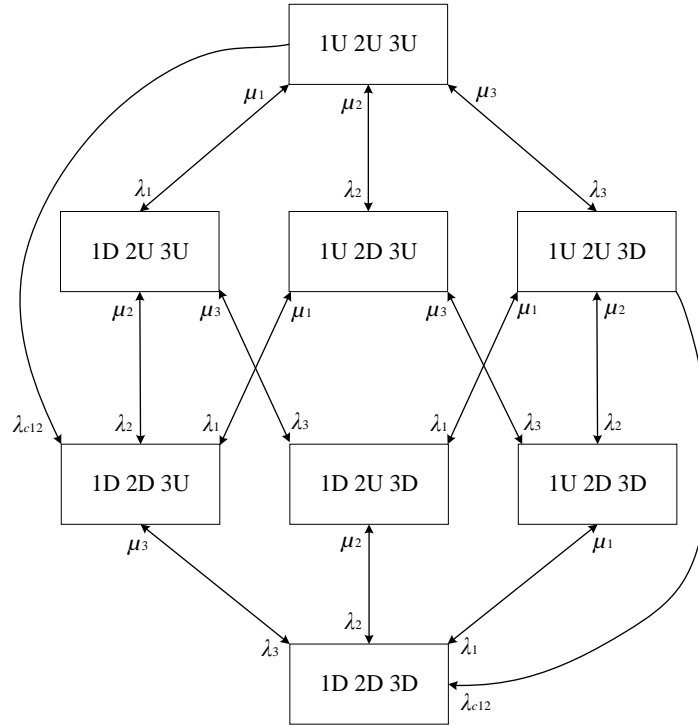


Figure 24. The system state space diagram for partially independent scenario.

In this scenario, the system state space can still be decoupled if components 1 and 2 are treated as a joint unit with four states. However, the probabilities of the four states need to be computed before performing a non-sequential sampling to generate the system state space. The four states, denoted by p_{1U2U} , p_{1U2D} , p_{1D2U} , p_{1D2D} , can be obtained by solving

equation (5.3). This requires more effort in analytical analysis compared to the previous scenario.

$$\mathbf{BP} = \mathbf{C} \quad (5.3)$$

where,

$$\mathbf{P} = [p_{1U2U} \quad p_{1U2D} \quad p_{1D2U} \quad p_{1D2D}]^T$$

$$\mathbf{B} = \begin{bmatrix} -(\lambda_2 + \lambda_1 + \lambda_{c12}) & \mu_2 & \mu_1 & 0 \\ 1 & 1 & 1 & 1 \\ \lambda_1 & 0 & -(\mu_1 + \lambda_2) & \mu_2 \\ \lambda_{c12} & \lambda_1 & \lambda_2 & -(\mu_1 + \mu_2) \end{bmatrix}$$

$$\mathbf{C} = [0 \quad 1 \quad 0 \quad 0]^T$$

Once the four state probabilities are obtained, components 1 and 2 can be treated as a joint unit. The UP and DOWN probabilities of component 3 are readily computable from equations (5.1) and (5.2). A non-sequential sampling can be performed accordingly.

5.3.3 Fully Dependent Scenario

In this scenario, components 1 and 2 have a common failure mode with transition rate λ_{c12} and components 2 and 3 have a common failure mode with transition rate λ_{c23} . All three components also have a common failure mode with transition rate λ_{c123} . The corresponding state space transition diagram for the system is shown in Figure 25.

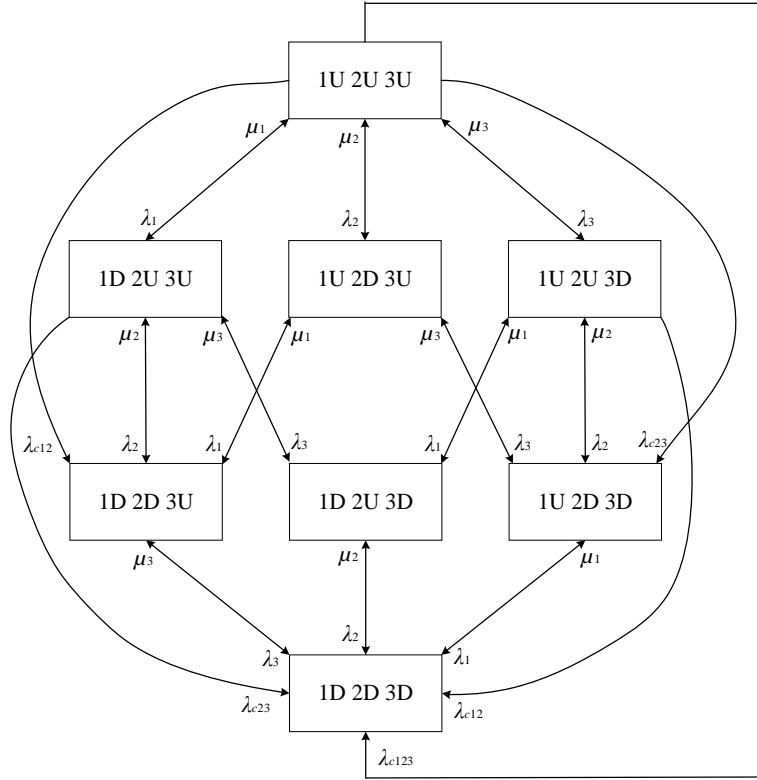


Figure 25. The system state space diagram for fully dependent scenario.

In this scenario, it requires that the probabilities of all the system states be known before sampling can be performed. Due to the dependencies existing among all components, it is very difficult to decouple the system state space into mutually independent disjoint subsets with state probabilities readily available so that a non-sequential sampling can be conveniently performed. It is possible to analytically compute the system state probabilities for a small system with only a few components. However, for a large system consisting of highly dependent components, it is impractical to analytically compute the probabilities of all the system states before sampling is performed. This is what makes the application of non-sequential methods to systems with dependent failures challenging.

5.4 Proposed Method

The state space for an actual power system may consist of numerous states, as shown in Figure 26. Due to the effects of cyber-induced failures, various transitions exist among these system states. Examples of transitions can be a primary fault, a repair process, a cyber-induced expanded outage, or a switching process.

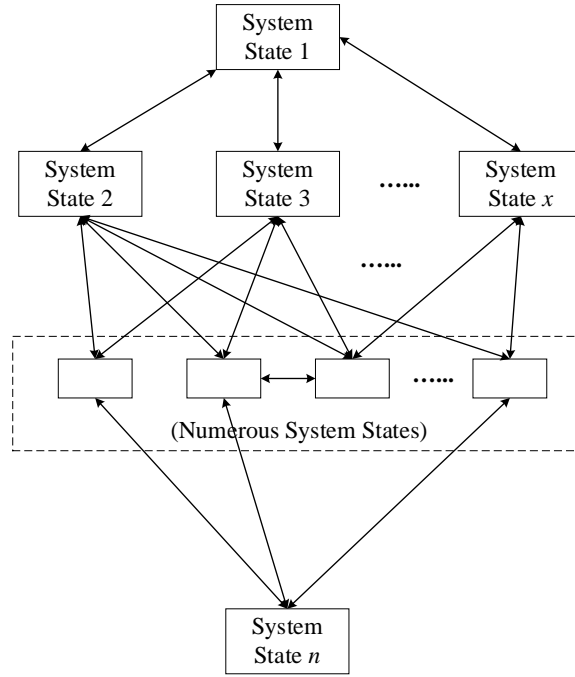


Figure 26. The system state space diagram for an actual power system.

One approach of generating the system state space is to use a completely analytical method to compute the probabilities of all system states. However, the size of a large system as well as the complexity of various transitions may prohibit such an analytical analysis.

Another approach is to decouple the entire system into mutually independent subsets so that system states can be sampled as combinations of subset states, as illustrated in the previous section. Each subset should consist of only a small number of components so that the state space for this subset can be readily developed with state probabilities computed. However, the high dependencies among system states in the presence of cyber-induced failures make such a decoupling very difficult.

It is, therefore, necessary to find some “smart” means of generating a representative system state space that preserves the causal and chronological features without insurmountable analytical computations. It is expected that this state space will not represent the entire state space but will be its representative. With this objective, a novel method is proposed in this section.

The method is proposed based on the fact that in a sequential Monte Carlo simulation, the computational effort for only generating a chronological state sequence is small compared

to the effort of evaluating the states using DC power flow based linear programming. Actually the method outlined below is valid irrespective of this observation but the computational advantage over pure sequential simulation is obtained only if the state evaluation process is computationally time consuming.

- 1) Perform a sequential Monte Carlo simulation for N years to generate a chronological state sequence of system states. One hour is considered as the unit of time advancement so that a system state, which is a combination of all component states and load states, remains unchanged within each hour. Therefore, there are $8760N$ system states when the simulation is finished. These system states, each with probability $1/(8760N)$, establish a representative state space. They are generated sequentially therefore they preserve the chronological, causal, and dependent characteristics. For more precision, smaller time advancement like 0.25 hour can be selected.
- 2) Randomly generate a list of $8760N_e$ ($1 \leq N_e < N$) integer numbers with each random number having a value between 1 to $8760N$. Pick up the $8760N_e$ states from the state space generated in the previous step and evaluate them. It should be noted that since time unit of each state is the same, all of the states are equally probable.
- 3) Calculate the reliability indices.

To implement the procedure outlined above, a very large number of system states need to be stored first. This can make the storage and retrieval cumbersome and time consuming. To overcome this difficulty of storing the system states generated by the sequential Monte Carlo simulation, the list of $8760N_e$ random numbers can be generated first. The sequential Monte Carlo simulation is then performed and as it proceeds, it only evaluates the states existing in the list. Other states are discarded and therefore the states do not need to be stored.

An intuitive expression of the proposed method is shown in Figure 27. A mark “**X**” indicates one of the $8760N_e$ hours generated in the list by random sampling. Then the sequential process is started and as the sampled hour (**X**) is encountered, the state is evaluated. Therefore, only the marked states are evaluated as the sequential simulation proceeds. All unmarked states are discarded.

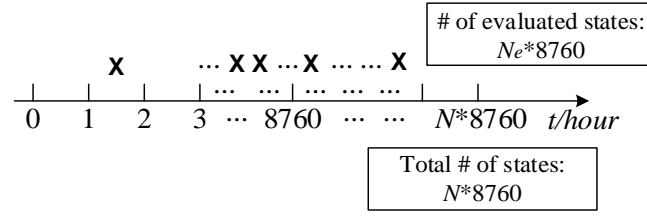


Figure 27. Expression of the proposed method.

5.5 Case Studies

The proposed method is implemented in comparison with a completely sequential methodology, as described in Section 4. The test system is established based on the Roy Billinton Test System (RBTS) [46] with extensions of cyber part at buses 3, 4, and 5. The placements and connections of cyber elements such as Current/Potential Transformers, Merging Units, and Protection Intelligent Electronic Devices, are defined in the extensions. The reliability data of the extended system are presented in Section 4. The hourly load profile is created based on the data in Tables 1, 2, and 3 from [48].

Currently, the obstacle to performing the case studies in a larger system (such as an extended IEEE Reliability Test System [48]) is the unavailability of such a system with detailed configuration descriptions on the cyber part. The size of the RBTS permits us to perform such extensions, as described in Section 4, while the effort needed to extend the IEEE RTS is very significant and far beyond the scope of this research. The applicability of the completely sequential methodology for large power systems is illustrated in Section 4. Once a large test system with detailed descriptions on the cyber part is available, the method proposed in this section can be implemented as well.

5.5.1 Index Definitions

In the completely sequential methodology, the reliability indices Loss of Load Probability (LOLP), Loss of Load Expectation (LOLE), and Expected Energy Not Supplied (EENS) are defined in Section 4. The coefficient of variation (β) is defined as:

$$\beta = \frac{\sqrt{V(F)/N_y}}{E(F)} \quad (5.4)$$

where $V(F)$ is the variance of the test function, N_y is the number of simulated years, and $E(F)$ is the expected estimate of the test function.

For the method proposed in this section, reliability indices are defined as follows.

$$LOLP = \frac{1}{N_k} \sum_{i=1}^{N_k} H_i \quad (5.5)$$

where N_k is the total number of samples randomly selected from the state space. N_k equals $8760N_e$ with N_e defined in the previous section. H_i equals 1 if load curtailment occurs in the i^{th} sample, otherwise it equals 0.

$$LOLE = LOLP \cdot 8760 \quad (5.6)$$

with unit of hours/year.

$$EENS = \frac{8760}{N_k} \sum_{i=1}^{N_k} C_i \cdot (1 \text{ hour}) \quad (5.7)$$

with unit of MWh/year, where N_k is the total number of samples randomly selected from the state space. C_i is the load curtailment in the i^{th} sample, with unit MW.

For the method proposed in this section, the coefficient of variation (β') is defined as:

$$\beta' = \frac{\sqrt{V(F)/N_k}}{E(F)} \quad (5.8)$$

where N_k is the total number of samples randomly selected from the state space. $V(F)$ is the variance of the test function. $E(F)$ is the expected estimate of the test function.

5.5.2 Results and Discussions

The completely sequential methodology is simulated for 300 years. For the method proposed in this section, $8760N$ samples are generated as the representative state space, of which $8760N_e$ samples are randomly selected and evaluated. $N = 300$ and $N_e = 50$ are used.

The simulations are performed in MATLAB on a computer with a 3.20 GHz processor. The results are shown in Table 40.

Table 40 Estimated reliability indices

Index	Completely Sequential	Proposed Method with $N = 300$, $N_e = 50$
LOLP	0.00131989	0.00134932
LOLE (hours/year)	11.562	11.820
EENS (MWh/year)	144.901	140.440
CPU Time (seconds)	501	212
β/β' of EENS	0.0340	0.0429

The comparison between the proposed method and the completely sequential methodology clearly indicates an efficiency improvement in evaluation. With an acceptable coefficient of variation ($<5\%$), the CPU time of the proposed method is reduced to less than 50% of the time consumed by the completely sequential methodology. It should be noted that the relative CPU times are important but the absolute values of CPU times can be significantly improved in both cases by using more efficient linear programming modules.

5.6 Summary

In this section, a non-sequential approach is proposed for systems involving dependent failures. A representative state space is generated from which sampling can be performed. The sequential technique is used to generate this representative state space that preserves the chronological and causal characteristics of cyber-induced incidents. The non-sequential technique is used with the major objective of accelerating evaluation process. The results clearly show an efficiency improvement by the proposed method compared to the completely sequential methodology, with an acceptable precision. The efficiency improvement will be even more significant for applications in large cyber-physical power systems.

It should be noted that the sequential simulation may not be the only option to generate the representative state space, which is basically a joint probability distribution of component probability distributions. The major difficulties of applying conventional non-sequential sampling methods to generating appropriate state space in presence of dependent failures are thoroughly discussed in this section with the intention of enlightening future studies on this subject.

6. Conclusions

This research extends the scope of bulk power system reliability modeling and analysis with the consideration of cyber elements. The major contributions, research conclusions, and outlook are summarized as follows.

In Section 2, a novel methodology for composite cyber-physical power system reliability analysis is proposed. The concept of Cyber-Physical Interface Matrix (CPIM) is introduced. A typical substation protection system with modern features is designed and analyzed as an example to illustrate the construction and utility of CPIM. The CPIM is the critical idea in the proposed methodology. It decouples the analysis of the cyber part from the physical part and provides the means of performing the overall analysis in a tractable fashion.

In Section 3, the consideration of cyber-link failures is included in substation modeling. The results clearly indicate a degradation of system reliability due to cyber-link failures.

In Section 4, the CPIM is realized into a concrete application from an abstract concept. The overall methodology for composite system reliability evaluation with the use of CPIM is enhanced and implemented on an extended Roy Billinton Test System (RBTS). The results clearly indicate the impact of cyber-induced failures on system-wide reliability indices and also signify the importance of accelerating switching process. With its scalability illustrated, the overall methodology provides a scalable option for reliability evaluation of large cyber-physical power systems.

In Section 5, the difficulties of using non-sequential techniques when there are dependent failures are thoroughly explored. An approach is proposed to overcome the difficulties by generating a representative state space from which states can be sampled. Furthermore, a method is introduced such that the state space generated by the sequential process does not need to be stored. The results clearly show an efficiency improvement of the proposed approach compared to the completely sequential methodology, with an acceptable precision. The efficiency improvement will be even more beneficial for applications in large cyber-physical power systems.

The overall methodology proposed in this research establishes the main framework for reliability evaluation of large cyber-physical power systems. More technical details, such as the internal structure of the Process Bus, can be considered in cyber part modeling without changing the framework of the proposed methodology. Consideration of these details will provide more precise probabilities in the CPIM and thereby yield more realistic reliability indices.

Furthermore, it has been recognized in this research that the unavailability of a large reliability test system with detailed description on the cyber part is an obstacle for testing developed methodologies. It is worthwhile to develop such a system as this is of crucial significance for future cyber-physical reliability studies. The development of such a test system requires substantial efforts as well as sound judgment from industry.

References

- [1] R. Billinton, *Power System Reliability Evaluation*. New York, USA: Gordon and Breach, 1970.
- [2] R. Billinton and R. N. Allan, *Reliability Assessment of Large Electric Power Systems*. Boston, USA: Kluwer, 1988.
- [3] R. Billinton and W. Li, *Reliability Assessment of Electric Power Systems using Monte Carlo Methods*. New York, USA: Plenum Press, 1994.
- [4] B. Falahati, Y. Fu, and L. Wu, "Reliability assessment of smart grid considering direct cyber-power interdependencies," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1515-1524, Sep. 2012.
- [5] B. Falahati and Y. Fu, "Reliability assessment of smart grids considering indirect cyber-power interdependencies," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1677-1685, Jul. 2014.
- [6] C. Singh and A. Sprintson, "Reliability assurance of cyber-physical power systems," in *Proc. IEEE Power and Energy Society General Meeting*, pp. 1-6, Jul. 2010.
- [7] A. G. Phadke and J. S. Thorp, "Expose hidden failures to prevent cascading outages," *IEEE Computer Applications in Power*, vol. 9, no. 3, pp. 20-23, Jul. 1996.
- [8] X. Yu and C. Singh, "A practical approach for integrated power system vulnerability analysis with protection failures," *IEEE Trans. Power Systems*, vol. 19, no. 4, pp. 1811-1820, Nov. 2004.
- [9] D. C. Elizondo, J. De La Ree, A. G. Phadke, and S. Horowitz, "Hidden failures in protection systems and their impact on wide-area disturbances," in *Proc. IEEE Power Engineering Society Winter Meeting*, vol. 2, pp. 710-714, Jan./Feb. 2001.
- [10] F. Yang, A. P. S. Meliopoulos, G. J. Cokkinides, and Q. B. Dam, "Effects of protection system hidden failures on bulk power system reliability," in *Proc. 2006 IEEE 38th North American Power Symp.*, pp. 517-523, Sep. 2006.
- [11] S. H. Horowitz and A. G. Phadke, *Power System Relaying*, 3rd ed. Chichester, West Sussex, UK: John Wiley & Sons Ltd, 2008.
- [12] J. L. Blackburn and T. J. Domin, *Protective Relaying: Principles and Applications*, 3rd ed. Boca Raton, FL, USA: CRC, 2007.
- [13] P. M. Anderson, *Power System Protection – Part VI: Reliability of Protective Systems*. New York, USA: McGraw-Hill/IEEE Press, 1999.

- [14] T. Skeie, S. Johannessen, and C. Brunner, "Ethernet in substation automation," *IEEE Control Syst. Mag.*, vol. 22, no. 3, pp. 43-51, Jun. 2002.
- [15] B. Kasztenny, J. Whatley, E. A. Udren, J. Burger, D. Finney, and M. Adamiak, "Unanswered questions about IEC 61850 – What needs to happen to realize the vision?," presented at the 32nd Annual Western Protective Relay Conf., Spokane, WA, USA, Oct. 2005.
- [16] T. S. Sidhu and P. K. Gangadharan, "Control and automation of power system substation using IEC 61850 communication," in *Proc. 2005 IEEE Conference on Control Applications*, pp. 1331-1336, Aug. 2005.
- [17] J. D. McDonald, *Electric Power Substations Engineering*, 3rd ed. Boca Raton, FL, USA: CRC, 2012.
- [18] F. Yang, "A comprehensive approach for bulk power system reliability assessment," Ph.D. dissertation, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, 2007.
- [19] R. Billinton and J. Tatla, "Composite generation and transmission system adequacy evaluation including protection system failure modes," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-102, no. 6, pp. 1823-1830, Jun. 1983.
- [20] M. C. Bozchalui, M. Sanaye-Pasand, and M. Fotuhi-Firuzabad, "Composite system reliability evaluation incorporating protection system failures," in *Proc. 2005 IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 486-489, May 2005.
- [21] K. Jiang and C. Singh, "New models and concepts for power system reliability evaluation including protection system failures," *IEEE Trans. Power Systems*, vol. 26, no. 4, pp. 1845-1855, Nov. 2011.
- [22] K. Jiang, "The impact of protection system failures on power system reliability evaluation," Ph.D. dissertation, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA, 2012.
- [23] C. Singh and R. Billinton, *System Reliability Modeling and Evaluation*. London, UK: Hutchinson, 1977.
- [24] J. C. O. Mello, M. V. F. Pereira, and A. M. Leite da Silva, "Evaluation of reliability worth in composite systems based on pseudo-sequential Monte Carlo simulation," *IEEE Trans. Power Systems*, vol. 9, no. 3, pp. 1318-1326, Aug. 1994.

- [25] I. Ali and M. S. Thomas, "Substation communication networks architecture," in *Proc. Joint International Conf. on Power System Technology (POWERCON) and IEEE Power India Conf.*, pp. 1-8, Oct. 2008.
- [26] P. Zhang, L. Portillo, and M. Kezunovic, "Reliability and component importance analysis of all-digital protection systems," in *Proc. IEEE Power Eng. Soc. Power Systems Conf. and Exposition*, pp. 1380-1387, Oct. 2006.
- [27] T. S. Sidhu, M. G. Kanabar, and P. P. Parikh, "Implementation issues with IEC 61850 based substation automation systems," in *Proc. 15th National Power Systems Conf.*, Mumbai, India, pp. 473-478, Dec. 2008.
- [28] M. G. Kanabar and T. S. Sidhu, "Reliability and availability analysis of IEC 61850 based substation communication architectures," in *Proc. IEEE Power & Energy Society General Meeting*, Calgary, AB, Canada, pp. 1-8, 2009.
- [29] Y. Zhang, A. Sprintson, and C. Singh, "An integrative approach to reliability analysis of an IEC 61850 digital substation," in *Proc. IEEE Power & Energy Society General Meeting*, San Diego, CA, USA, pp. 1-8, 2012.
- [30] A. F. Sleva, *Protective Relay Principles*. Boca Raton, FL, USA: CRC, 2009, pp. 112-113.
- [31] *IEC Standard for Communication network and systems in substations, IEC 61850*, 2003.
- [32] K. Jiang and C. Singh, "Reliability modeling of all-digital protection systems including impact of repair," *IEEE Trans. Power Delivery*, vol. 25, no. 2, pp. 579-587, Apr. 2010.
- [33] K. P. Brand, V. Lohmann, and W. Wimmer, *Substation Automation Handbook*. Bremgarten, Switzerland: Utility Automation Consulting Lohmann, 2003.
- [34] V. Skendzic, I. Ender, and G. Zweigle, "IEC 61850-9-2 process bus and its impact on power system protection and control reliability," Schweitzer Engineering Laboratories, Inc., Pullman, WA, USA, Tech. Rep., 2007.
- [35] T. M. Lindquist, L. Bertling, and R. Eriksson, "Circuit breaker failure data and reliability modeling," *IET generation, transmission & distribution*, vol. 2, no. 6, pp. 813-820, Nov. 2008.
- [36] P. M. Kanabar, M. G. Kanabar, W. El-Khattam, T. S. Sidhu, and A. Shami, "Evaluation of communication technologies for IEC 61850 based distribution automation system with distributed energy resources," in *Proc. IEEE Power & Energy Society General Meeting*, Calgary, AB, Canada, pp. 1-8, 2009.

- [37] C. Singh and J. Mitra, "Monte Carlo simulation for reliability analysis of emergency and standby power systems," in *Proc. 30th IEEE Industry Applications Society Conf.*, Orlando, FL, USA, pp. 2290-2295, Oct. 1995.
- [38] M. G. Kanabar and T. S. Sidhu, "Performance of IEC 61850-9-2 process bus and corrective measure for digital relaying," *IEEE Trans. Power Delivery*, vol. 26, no. 2, pp. 725-735, Apr. 2011.
- [39] P. Ferrari, A. Flammini, S. Rinaldi, and G. Prytz, "Mixing real time Ethernet traffic on the IEC 61850 process bus," in *Proc. 9th IEEE International Workshop on Factory Communication Systems (WFCS)*, Lemgo, Germany, pp. 153-156, May 2012.
- [40] L. Andersson, K.-P. Brand, C. Brunner, and W. Wimmer, "Reliability investigations for SA communication architectures based on IEC 61850," in *Proc. 2005 IEEE Russia Power Tech*, St. Petersburg, Russia, pp. 1-7, Jun. 2005.
- [41] A. Apostolov and B. Vandiver, "IEC 61850 process bus—principles, applications and benefits," in *Proc. 63rd Annual Conf. Protective Relay Engineers*, pp. 1-6, 2010.
- [42] H. Lei, C. Singh, and A. Sprintson, "Reliability modeling and analysis of IEC 61850 based substation protection systems," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2194-2202, Sep. 2014.
- [43] H. Hajian-Hoseinabadi, "Impacts of automated control systems on substation reliability," *IEEE Trans. Power Delivery*, vol. 26, no. 3, pp. 1681-1691, Jul. 2011.
- [44] H. Hajian-Hoseinabadi, M. Hasanianfar, and M. E. H. Golshan, "Quantitative reliability assessment of various automated industrial substations and their impacts on distribution reliability," *IEEE Trans. Power Delivery*, vol. 27, no. 3, pp. 1223-1233, Jul. 2012.
- [45] J. J. Grainger and W. D. Stevenson, *Power System Analysis*. New York, USA: McGraw-Hill, 1994, pp. 337-338.
- [46] R. Billinton, S. Kumar, N. Chowdhury, K. Chu, K. Debnath, L. Goel, E. Khan, P. Kos, G. Nourbakhsh, and J. Oteng-Adjei, "A reliability test system for educational purposes-basic data," *IEEE Trans. Power Systems*, vol. 4, no. 3, pp. 1238-1244, Aug. 1989.
- [47] A. Sankarakrishnan and R. Billinton, "Sequential Monte Carlo simulation for composite power system reliability analysis with time varying loads," *IEEE Trans. Power Systems*, vol. 10, no. 3, pp. 1540-1545, Aug. 1995.

- [48] IEEE Committee Report, "IEEE reliability test system," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 2047-2054, Nov./Dec. 1979.
- [49] C. Singh and J. Mitra, "Composite system reliability evaluation using state space pruning," *IEEE Trans. Power Systems*, vol. 12, no. 1, pp. 471-479, Feb. 1997.
- [50] EPRI, "Composite system reliability evaluation methods," Final Report on Research Project 2473-10, EPRI EL-5178, Jun. 1987.
- [51] J. Mitra and C. Singh, "Incorporating the DC load flow model in the decomposition-simulation method of multi-area reliability evaluation," *IEEE Trans. Power Systems*, vol. 11, no. 3, pp. 1245-1254, Aug. 1996.
- [52] N. Gubbala and C. Singh, "Models and considerations for parallel implementation of Monte Carlo simulation methods for power system reliability evaluation," *IEEE Trans. Power Systems*, vol. 10, no. 2, pp. 779-787, May 1995.
- [53] C. L. T. Borges, D. M. Falcão, J. C. O. Mello, and A. C. G. Melo, "Composite reliability evaluation by sequential Monte Carlo simulation on parallel and distributed processing environments," *IEEE Trans. Power Systems*, vol. 16, no. 2, pp. 203-209, May 2001.
- [54] H. Lei and C. Singh, "Power system reliability evaluation considering cyber-malfunions in substations," *Electric Power Systems Research*, vol. 129, pp. 160-169, Dec. 2015.
- [55] R. Billinton and W. Li, "A system state transition sampling method for composite system reliability evaluation," *IEEE Trans. Power Systems*, vol. 8, no. 3, pp. 761-770, Aug. 1993.
- [56] A. M. Rei and M. T. Schilling, "Reliability assessment of the Brazilian power system using enumeration and Monte Carlo," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1480-1487, Aug. 2008.
- [57] Z. Shu and P. Jirutitijaroen, "Non-sequential simulation methods for reliability analysis of power systems with photovoltaic generation," in *Proc. 2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pp. 703-709, Jun. 2010.

Part II

Reliability Assessment and Modeling of Cyber-Enabled Power Distribution Systems

Dr. Visvakumar Aravinthan

Mohammad Heidari Kapourchali, Graduate Student

Mojtaba Sepehry, Graduate Student

Wichita State University

For information about this project, contact:

Visvakumar Aravinthan, Assistant Professor
Electrical Engineering and Computer Science
Wichita State University
1845 Fairmount
Wichita, Kansas 67260-0083
Email: visvakumar.aravinthan@wichita.edu
Phone: (316) 978-6324

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: <http://www.pserc.org>.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
551 E. Tyler Mall
Engineering Research Center #527
Tempe, Arizona 85287-5706
Phone: (480) 965-1643
Fax: (480) 965-0745

Notice Concerning Copyright Material

PSERC members are given permission to copy without fee all or part of this publication for internal use if appropriate attribution is given to this document as the source material. This report is available for downloading from the PSERC website.

© 2016 Wichita State University.

All rights reserved.

Table of Contents

1.	Introduction	1
1.1.	Distribution Network Reliability Assessment.....	1
1.2.	Smart Grid Improves Reliability or Puts the Grid at Higher Risk?	1
1.3.	On-Demand Failure of Cyber Components.....	2
1.4.	Cyber Security Threats	2
2.	An Algorithm for Reliability Evaluation of Radial Networks	3
2.1.	Model	3
2.2.	Load Type Detection	7
2.2.1.	Network without alternative supply path.....	7
2.2.2.	Network with alternative supply path	8
2.3.	Interruption time detection of load points.....	10
2.3.1.	Network without alternative supply path.....	10
2.3.2.	Network with alternative supply path	10
2.4.	Failure rate (λ) and unavailability (U) calculation	10
2.5.	Case Study.....	11
3.	Cyber-Enabled Power Distribution System.....	16
3.1.	Problem Statement.....	16
3.1.1.	Single-path end-to-end communication framework	16
3.1.2.	Multi-path end-to-end communication framework	17
3.2.	Cyber-Enabled Device Availability	18
3.3.	Proposed Reliability Modeling	20
3.4.	Reliability computation of cyber-enabled feeders	21
3.5.	Analyzing impact of manually operated protective devices.....	23
3.6.	Impact of Cyber Attacks	25
3.7.	Common Cause Failures.....	28
3.8.	Optimal Placement of Devices	30
3.8.1.	Objective:	30
3.8.2.	Model Validation and Analysis.....	31
3.8.2.1.	RBTS - Bus 2.....	31
3.8.2.2.	Typical 27-node distribution feeder.....	35
4.	Conclusion and Future Work	38

List of Figures

Fig. 1.	A sample 7-bus radial distribution network	4
Fig. 2.	Equivalent of a line in series with a distribution Transformer	6
Fig. 3.	Protection Zones.....	8
Fig. 4.	Detached Protection Zones.....	8
Fig. 5.	Modified RBTS- Bus 2 test system.....	12
Fig. 6.	Flowchart of the proposed algorithm	15
Fig. 7.	Smart fault detector	16
Fig. 8.	Layout of RCSs on distribution feeder.....	17
Fig. 9.	Multi-hop data transmission.....	18
Fig. 10.	State space diagram for physical component and communication link	18
Fig. 11.	Illustrative layout of distribution feeder with flow of data	19
Fig. 12.	Scenario 1: Successful fault detection and isolation	20
Fig. 13.	Scenario 2: Unsuccessful fault detection and isolation.....	22
Fig. 14.	Scenario 3: Unsuccessful switch operation.....	22
Fig. 15.	Impact of branch B ₉ failure on load point 2	22
Fig. 16.	Impact of branch B ₂ failure on load point 8	23
Fig. 17.	Impact of branch B ₆ failure on load point 4	24
Fig. 18.	Sample network containing manual and remote-controlled switches	25
Fig. 19.	Abstract view of a Power system SCADA.....	26
Fig. 20.	SCADA system from reliability point of view of a power distribution system.....	26
Fig. 21.	The component of Fig. 4a with non-exponentially distributed up state.....	27
Fig. 22.	Different outage durations experienced by load points.....	28
Fig. 23.	Common cause component groups	29
Fig. 24.	Integer string representation for optimal planning problem.....	32
Fig. 25.	RBTS bus 2 test feeder.....	32
Fig. 26.	Load point outage time.....	34
Fig. 27.	Typical 27-node practical distribution feeder	36
Fig. 28.	Total cost incurred for different successful data transmission probabilities and switches and FDs cost.....	37
Fig. 29.	Total cost incurred for different successful data transmission probabilities and cyber attack rates.....	37
Fig. 30.	Impact on (a) customers' interruption cost and (b) total cost.....	38

List of Tables

Table I. Load point reliability indices for modified RBTS-bus 2.....	13
Table II. Three boundary cases for first case study	33
Table III. Two sets of successful data transmission probabilities	33
Table IV. Optimal decisions for experiment 1.....	34
Table V. Results for experiment 2	34
Table VI. Boundary values for case 2.....	36
Table VII. Optimal solutions for case 2 with three link probabilities	36

Overview:

Wichita State University contribution to this project pursued the following steps of research:

1. The team developed a new algorithm for reliability evaluation of radial distribution networks. New analytical model based on physical characteristics of distribution system was developed to pave the way for new reliability model developments in the presence of cyber enabled devices. *(Results were published in North American Power Symposium and received first prize in the student paper contest).*
2. The team investigated the ways emerging cyber enabled devices and logics put power system at risk. These vulnerabilities were categorized into unavailability of data and cyber security threats. This investigation was required before embarking on new model developments.
3. The team also investigated the necessity of updating the traditional reliability models to incorporate cyber enabled logic. The developed model led to more accurate and realistic reliability indices at distribution feeder level.
4. Common mode failures were introduced as a potential vulnerability in cyber enabled power distribution network as multiple devices can fail due to a common cause.
5. The proposed probabilistic model was incorporated into a traditional power distribution network planning problem to illustrate the effectiveness of the developed model. *(Outcome of research in steps 2-3-4-5 was published in Transaction on Smart Grid)*

1. Introduction

1.1. Distribution Network Reliability Assessment

Distribution system is the final link in the delivery of electric power to consumers which plays a pivotal role in the overall system reliability. About 80% of power interruptions result from power distribution system failure [1]. Therefore, the study of distribution system reliability has been of keen interest [1-4]. There are two major approaches to assess the reliability of system: Analytical and simulation. Although the simulation methods have been pervasively used, their effectiveness highly relies on the number of simulations and is subjected to time constraints. On the other hand, analytical methods determine the exact solution to the problem, but they are concerned with the complications of the system.

There are different analytical methods for reliability evaluation of radial distribution network (RDN). Conventional FMEA method [2] considers all possible failure modes and evaluates their effects on the load points. Basic load point indices are acquired by the summation of all effects on a load point. Zone-Branch method [3] is the FMEA method which is presented in matrix form. To reduce the calculation burden, the equivalent approach was introduced in [1]. In this approach an equivalent element is used instead of a portion of distribution network with two equivalent parameters named equivalent failure rate and equivalent mean time to repair. Consequently, a large network can be reduced to a smaller one which can be used to evaluate reliability indices. The aim of the aforementioned analytical methods is to ease the involving calculation procedure. However, the realization of such methods can be a time-consuming task. One algorithm which is in use for reliability calculation of RDNs is presented in [4]. This algorithm utilizes breadth-first or depth-first search algorithm. The two search algorithms are the basis for the rest of reliability calculations: upstream restoration, downstream restoration and incorporation of unsuccessful operation of protection devices. As stated in [4] a breadth-first search becomes computationally intensive for large systems and a depth-first search is computationally efficient but is memory intensive.

1.2. Smart Grid Improves Reliability or Puts the Grid at Higher Risk?

The smart grid calls for ubiquitous deployment of advanced monitoring and automated control equipment known as cyber-enabled devices. It places a higher requirement on the dependable two-way flow of information to report events and issue commands. Cyber-enabled technology deployment has gained momentum, but its pervasive use at the distribution level is hindered by several factors. First, lack of a transparent measure for the expected rate of return per unit of money invested [5]. Second, vulnerability of the communication infrastructure due to cyber-attacks and its impact on the power grid [6]. Third, unavailability of cyber-enabled technologies upon demand [7]. Addressing the impact of these factors paves the way for mobilizing investment in order to bring intelligence to the existing power grid. The culmination of retrofitting the existing grid and adopting smart technologies is to realize a self-healing power grid. Such grid promptly responds to and recovers from contingencies and thereby improves electric grid efficiency, reliability, and safety [8].

Currently, distribution networks are limited by monitoring and communication technologies. Therefore, the operation of a power distribution grid requires human intervention for both normal and abnormal operations. Reducing human intervention and the duration of customer interruption is of a great interest in power distribution system management. It has always been an ever-present

planning and management goal in the integration of monitoring and protective devices into the existing grid.

Manual switch placement has been considered in [9-12], in order to improve reliability by limiting interrupted loads. To mitigate the fault isolation time, an optimal FD placement is proposed in [13, 14]. FDs visually or remotely aid the field crew in finding the fault location. Installing remote-controlled switch (RCS) was a further step to expedite restoration of service to as many customers as possible [15]. This trend along with technological improvement in communication equipment has reached the point whereby the grid acts in a self-healing manner. In a self-healing smart grid, faults should be automatically detected and isolated, and service should be restored to non-faulty regions with little or no human intervention. In a smart grid setup, the communication network serves as a backbone structure, providing interoperability of smart equipment and the control center. This brings new challenges for maintaining a secure connection in order to transmit data. Since a cyber-infrastructure is not failure free, its integration with the existing grid increases the failure modes of power delivery. New failure modes need to be included into the reliability calculation so that more precise planning and operating decisions can be made.

1.3. On-Demand Failure of Cyber Components

Several studies have been published to incorporate the impact of unavailability of the communication infrastructure into reliability of the power grid. Falahati and Fu [16] categorized the cyber-power interdependencies into two types: direct and indirect. They quantitatively evaluated failure modes of a smart micro-grid [17] and high-voltage substation [18] to model the immediate and potential impacts of a cyber-network on power equipment. Lei et al. [19] decoupled the reliability analysis of an IEC 61850-based substation protection system into separate entities—namely a cyber-network and power systems—by defining a cyber-physical interface matrix. The matrix makes it possible to incorporate failure modes of the substation cyber-equipment into a composite power system model. In [17-19], reliability of geographically confined cyber-enabled power systems, such as a substation and a micro-grid, were investigated.

Impact analysis of communication network unavailability was extended to feeder FDISR in [20-22]. Authors in [22] integrated the information and communication technology representation of a distribution feeder including RCSs into Monte Carlo Simulation (MCS) to numerically assess their interdependencies. Reference [22] is one of the initial attempts to quantify the reliability of power distribution feeder considering on-demand communication failure. All the above works only focus on on-demand failure of cyber components and its impact on power system reliability. However, this is not the only concern that cyber-enabled grid gives rise to.

1.4. Cyber Security Threats

Result of a survey on cyber security of electric utilities [23] shows many of them are target of constant and frequent cyber-attacks. Cyber-attacks against components of supervisory control and data acquisition (SCADA) system can lead to malfunction of intelligent electronic devices and degrade the system reliability. These attacks are performed by gaining higher privilege level on cyber components and sending false control commands or status data. Cyber-attacks are generally represented by the mean time expended by an attacker in an effort to elevate the privilege level and compromise the system. This time is called mean time to compromise (MTTC).

McQueen et al. [24] proposed a model to estimate MTTC of cyber components of a SCADA system. In this work, actions of an attacker trying to exploit known vulnerabilities of a component

are categorized in three cases. These cases depend on whether the attacker is aware of a vulnerability in the component and its exploit. Nzoukou et al. [25] used Bayesian network-based attack graph to represent the potential attack sequences to attain the root privilege of a cyber-network. In each step of a sequence, the vulnerability of a component is exploited to elevate the privilege level. Both known and unknown (zero-day) vulnerabilities are included in MTTC evaluation. Sommestad et al. [26] presented an attack graph based on Bayesian network to represent the potential sequences for man in the middle attack against SCADA communication links. The attacker can eavesdrop on the intercepted messages, modify them or inject new messages in the network. This work also incorporates possible countermeasures into the graph. Stamp et al. [27] applied MCS to assess the impact of cyber-attack scenarios on reliability degradation of power system. The first scenario is the attack against generation unit and transmission line protection. In the second scenario, the SCADA system is targeted. If the attacker gains higher privilege level, trip signal can be sent to multiple breakers. Time to compromise is assumed as exponentially distributed random variable. Mean time to repair (MTTR) is defined as the required time for cyber forensics and physical restoration of the unit. Zhang et al. introduced different scenarios to attack SCADA [23] and wind farm energy management systems [28]. These scenarios include attacks against SCADA communication links and local area network (LAN) of control center, corporate and substation. MTTC and MTTR of each scenario are used in MCS to evaluate loss of load probability and expected energy not supplied of a benchmark power system. In [29] a modified semi Markov process was utilized to evaluate MTTC of attack scenarios. In this work a game theoretic approach is used for optimal allocation of offensive and defensive resources to different targets. In [30] the model in [29] was extended to evaluate MTTC of normal and penetration attacks against intrusion tolerant SCADA systems. References [29] and [30] also make use of MCS for reliability assessment of power system. Reviewing recent works shows the necessity of an all-inclusive and scalable reliability model for power distribution network. Recognizing what has been missing in the literature, this project aims to figure this out by analytically assessing system reliability indices incorporating both on-demand communication failures and cyber-attacks.

2. An Algorithm for Reliability Evaluation of Radial Networks

2.1. Model

This section proposes a new algorithm for reliability evaluation of complex radial distribution networks. The proposed algorithm is based on failure mode and effect analysis (FMEA) method. The paper talks about the definition of network topology using a matrix which is modified and new matrices are derived. These matrices are used for evaluation of load point reliability indices. The rows and columns of these matrices correspond to branches and buses of the network respectively. Each element of these matrices attest a specific relation between a faulted branch and the network bus. The above mentioned salient features make the whole algorithm straight forward to implement in a real distribution system. Thus, the results obtained from this method are verified with the Bus 2 of RBTS.

A network structure representation of a distribution feeder is illustrated in this section, using a simple 7-bus radial distribution network as shown in Fig. 1.

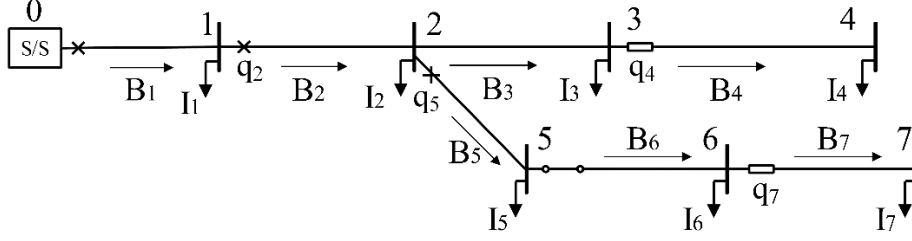


Fig. 1. A sample 7-bus radial distribution network

The relationship between the bus current injection and branch currents is given by [31]:

$$\begin{aligned}
 I_1 &= B_1 - B_2 \\
 I_2 &= B_2 - B_3 - B_5 \\
 I_3 &= B_3 - B_4 \\
 I_4 &= B_4 \\
 I_5 &= B_5 - B_6 \\
 I_6 &= B_6 - B_7 \\
 I_7 &= B_7
 \end{aligned}$$

Where I_i is the equivalent current injection for bus i and B_i is the current flowing into the branch i . The set of equations is represented by branch current to bus injection (*BCBI*) matrix as follows:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \\ I_7 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{bmatrix}$$

By inverting the *BCBI* matrix we get the branch current to bus injection (*BIBC*) matrix.

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \\ I_7 \end{bmatrix}$$

In order to build *BIBC* matrix, the network data are stored in network data (*ND*) matrix.

$$ND = \begin{bmatrix} 0 & 1 & 2 & 3 & 2 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{bmatrix}$$

Each column of ND matrix represents a branch, where first row gives the sending bus number and the second row gives the receiving bus number. Using the ND matrix and the cue information below, the $BCBI$ matrix can be constructed.

$$BCBI(i, j) = \begin{cases} -1 & \text{if } i \text{ is the sending bus of branch } j \\ & \text{except the substation bus} \\ 1 & \text{if } i \text{ is the receiving bus of branch } j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The calculation procedure of load point failure rate and unavailability is initiated through updating the ND matrix by appending two rows to it: row one gives the information about the type of protection device and row two specifies the probability of unsuccessful operation of the devices respectively. The following terms are used along with their designated numerical notations: (a) 0: without any protection or switching device, (b) 1: circuit breaker (c) 2: disconnected switch and (d) 3: fuse. Using the above notation the revised ND matrix for the sample 7-bus system is given below.

$$ND = \begin{bmatrix} 0 & 1 & 2 & 3 & 2 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 0 & 2 & 1 & 3 & 2 \\ 0 & q_2 & 1 & q_4 & q_5 & 1 & q_7 \end{bmatrix} \quad (2)$$

If the distribution transformer is considered in the primary side of distribution feeder, equivalent failure rate and repair time are calculated as:

$$\lambda_{eq} = \lambda_{line} + \lambda_{Tran} \quad (3)$$

$$U_{eq} = r_{eq} \lambda_{line} + r_{Tran} \lambda_{Tran} \quad (4)$$

$$r_{eq} = U_{eq} / \lambda_{eq} \quad (5)$$

Using the equivalent failure and repair rate the distribution system can be simplified as shown in figure 2.

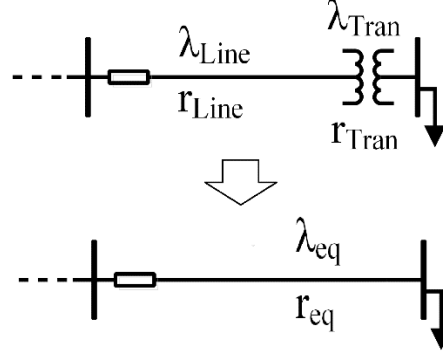


Fig. 2. Equivalent of a line in series with a distribution Transformer

BIBC matrix for the network shown in figure 1 has been formed in section II. Each row and column in *BIBC* matrix represents the network branches and buses, respectively. Ones in each row represent the buses being fed by the corresponding branch and zeroes for the buses which are not being fed by the corresponding branch: buses in the upstream network. In other words, the downstream impact of a fault on each branch can be shown by *BIBC* matrix.

Next, we evaluate the upstream effect of fault on each branch in two steps. First, consider a case without the impact of protection devices in the upstream network. For example, in figure 1, *ND* matrix shows that branch 7 is protected through a fuse with the probability of unsuccessful operation q_7 , i.e. all the buses that are upstream to branch 7 are affected by the probability of q_7 . Hence, all the zeroes in row 7 are replaced by q_7 . As another example branch 3 has no protection device and branch 6 has a disconnect switch. Now, assuming all the branches upstream to branch 3 and 6 have no protection, resulting in a probability of 1. In other words, all the 0s in rows 3 and 6 are replaced by 1. *BIBC* matrix then called failure mode and effect coefficient matrix (*FMECM*) is modified now as follows:

$$FMECM_{s1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_4 & q_4 & q_4 & 1 & q_4 & q_4 & q_4 \\ q_5 & q_5 & q_5 & q_5 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_7 & q_7 & q_7 & q_7 & q_7 & q_7 & 1 \end{bmatrix}$$

In a second step, the impact of upstream protection devices of a branch on load point failure rate is calculated and stored in $FMECM_{s2}$. Each element e_{ij} of $FMECM_{s2}$ equals to the multiplication of all the elements in the j th column of $FMECM_{s1}$ which belong to the path from i th branch to the source. As an example consider $FMECM_{s2}(7,4)$:

$$\begin{aligned}
FMECM_{s2}(7,4) &= \prod_{k=1,2,5,6,7} FMECM_{s1}(k,4) \\
&= 1 \times 1 \times q_5 \times 1 \times q_7
\end{aligned}$$

Applying step 2 to the previous $FMECM_{s1}$ results in:

$$FMECM_{s2} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_2 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_2 & 1 & 1 & 1 & 1 & 1 & 1 \\ q_2 q_4 & q_4 & q_4 & 1 & q_4 & q_4 & q_4 \\ q_2 q_5 & q_5 & q_5 & q_5 & 1 & 1 & 1 \\ q_2 q_5 & q_5 & q_5 & q_5 & 1 & 1 & 1 \\ q_2 q_5 q_7 & q_5 q_7 & q_5 q_7 & q_5 q_7 & q_7 & q_7 & 1 \end{bmatrix}$$

BIBC matrix can be used to detect the upstream path of a specific branch to the source. The implemented procedure has been depicted in the flowchart of figure 6. In this flowchart *nb* and *Rb* represent the number of buses except the substation bus and receiving bus, respectively.

2.2. Load Type Detection

Consider a case where there is no alternative supply. Occurrence of fault divides the system into two areas: Faulted area and the area that can be restored by switching. In another case consider the network with alternative supply, occurrence of fault may divide it into three areas. Faulted area, area restored by main substation and the area restored by an alternative supply. Detection of the load point type is presented below.

2.2.1. Network without alternative supply path

First, it is assumed that all the protection and switching devices are automatic and 100% reliable. This assumption means that the fourth row of *ND* matrix should be set to zero for all branches with protection or switching device. Then, step 1 and step 2 of the section III are executed. The resulted matrix is called Load Type Matrix (*LTM*). The *LTM* for the network of figure 1 is written below:

$$LTM = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Rows and columns of this matrix represent branches and buses of the network, respectively. In each row, elements that are represented by one are the buses that cannot be restored during the repair of the corresponding faulted branch. The elements that are represented by zero are the buses that can be restored after switching and isolation of the faulted branch.

2.2.2. Network with alternative supply path

In this case, at first, *LTM* is formed. Then it is modified through following procedure: When fault occurs on each branch, considering the location of protection and switching devices, some load points will be in the faulted area and the rest of them can be detached from the faulted area. As an example, Fig. 3 shows how the sample 7- bus network can be classified into zones when fault occurs.

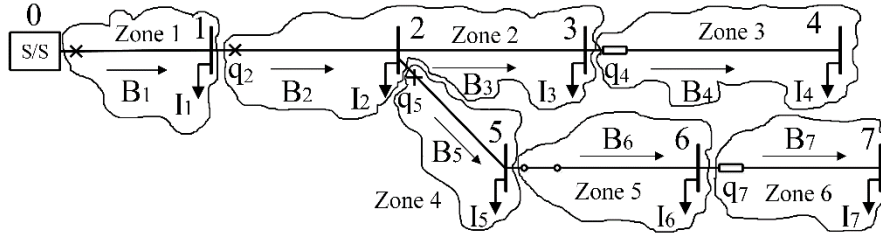


Fig. 3. Protection Zones

For the system shown in Fig. 3, Fig. 4 shows these zones separately.

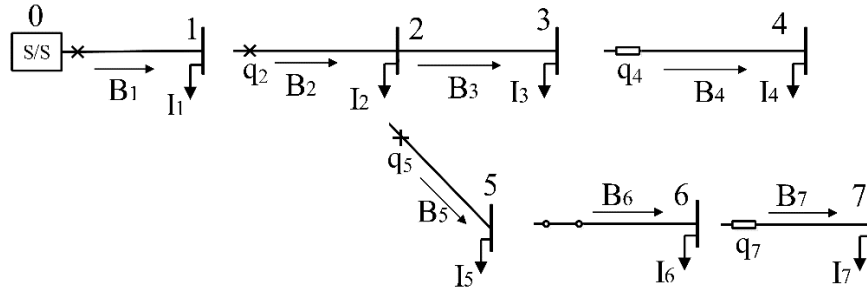


Fig. 4. Detached Protection Zones

According to this figure the relation between injected bus currents and the currents flowing through branches are as follows:

$$\begin{aligned} I_1 &= B_1 \\ I_2 &= B_2 - B_3 \\ I_3 &= B_3 \\ I_4 &= B_4 \\ I_5 &= B_5 \\ I_6 &= B_6 \\ I_7 &= B_7 \end{aligned}$$

These relations can be shown in the form of a square matrix called Zone Branch Current to Bus Injection (*ZBCBI*) and Zone Bus injection to branch current (*ZBIBC*) matrix is obtained by inverting *ZBCBI* matrix. *ZBCBI* can be formed using *ND* matrix and the following information:

$$ZBCBI(i, j) = \begin{cases} 1 & \text{if } i \text{ is the receiving bus of the branch } j \\ -1 & \left\{ \begin{array}{l} \text{if } i \text{ is sending bus of branch } j, \text{ except} \\ \text{substation bus and if no protection or} \\ \text{switching device at branch } j \end{array} \right. \\ 0 & \text{otherwise} \end{cases}$$

After formation of *ZBCBI*, it is assumed that fourth row of *ND* matrix is zero for all branches with protection or switching device. Then, step 1 and step 2 of the section III are executed. The resulting matrix is called Isolated Load Matrix (*ILM*). This matrix for the network of figure 1 is written below. One in each row represents a faulted bus and zero for the bus that can be detached from the faulted area.

$$ILM = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

Then, the modified *LTM* (*LTM_{alt.s}*) is formed using (7).

$$LTM_{alt.s} = LTM + ILM \quad (7)$$

The elements of *LTM* matrix that are zero, one and two represent upstream buses that can be detached from faulted area, downstream buses that can be restored through alternative supply and the buses that are in faulted area. After the detection of the load types, the desired downstream restoration algorithm is executed and if some load points should be shed due to the violation of voltage or current constraints or the inadequacy of the alternative feeder capacity, corresponding elements of this load points in *LTM_{alt.s}* are replaced with 2 as given below.

$$LTM_{alt.s.} = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

2.3. Interruption time detection of load points

After detection of load types, interruption time of load points is calculated using the following procedure and stored in failure mode and interruption time matrix (FMITM).

2.3.1. Network without alternative supply path

In this case, consider LTM as $FMITM_{initial}$. Then, in each row of $FMITM_{initial}$ matrix if an element is zero, it is replaced by the switching time (T_{sw}) and if the element is one, it is replaced by the repair time of corresponding branch (T_R). Final $FMITM$ is calculated using (8).

2.3.2. Network with alternative supply path

In this case, consider $LTM_{alt.s.}$ as $FMITM_{initial}$. Then, in each row of $FMITM_{initial}$ matrix if an element is zero, it is replaced by the switching time (T_{sw}) and if the element is one it is replaced by the required time for switching and closing alternative supply switch ($T_{sw} + T_{a.s.sw.}$). If the element is two, it is replaced by the repair time of the corresponding branch (T_R). Then, final $FMITM$ is calculated using (8).

$$FMRTM = FMRTM_{initial} \circ FMECM \quad (8)$$

Where, the operator ‘ \circ ’ indicates element wise product of two matrices in which corresponding elements of matrices are multiplied.

2.4. Failure rate (λ) and unavailability (U) calculation

Finally, failure rate and unavailability of the load points are calculated using (9) and (10).

$$\begin{bmatrix} \lambda_{Load-1} \\ \vdots \\ \lambda_{Load-n} \end{bmatrix} = \begin{bmatrix} FMECM^T \end{bmatrix} \begin{bmatrix} \lambda_{Branch-1} \\ \vdots \\ \lambda_{Branch-n} \end{bmatrix} \quad (9)$$

and

$$\begin{bmatrix} U_{Load-1} \\ \vdots \\ U_{Load-n} \end{bmatrix} = \begin{bmatrix} FMRTM^T \end{bmatrix} \begin{bmatrix} \lambda_{Branch-1} \\ \vdots \\ \lambda_{Branch-n} \end{bmatrix} \quad (10)$$

It is worth mentioning that the proposed algorithm is applicable in networks with any bus and branch numbering scheme. Overall flowchart of the algorithm has been depicted in Fig. 6.

2.5. Case Study

The proposed algorithm has been verified on bus 2 of RBTS [32] in which all main sections have disconnect switch and the lateral fuses were considered 100% reliable. Results for load point reliability indices match with the ones in [32].

Furthermore, to incorporate the probability of failure of protection devices into the reliability evaluation, some modifications have been made and the resulting network is shown in Fig. 5. These changes include the elimination of disconnect switches of branches 7, 18, 21, 24 and 32 and introduction of a circuit breaker on branches 7 and 18. Furthermore, all fuses and added circuit breakers are considered to have the probability of failure equal to 0.2 and 0.1, respectively.

The obtained results for load point reliability indices have been shown in table I. Since a disconnect switch of feeder 4 has been eliminated and all fuses are prone to failure, failure rate and unavailability of all load points in these feeders have been deteriorated, compared to the base case. In feeders 1 and 3 the load points upstream to the circuit breaker show an improvement in failure rate and unavailability to the corresponding values in base case. However, the load points located downstream of the circuit breakers have increased failure rate and unavailability values due to lack of disconnects and totally reliable fuses.

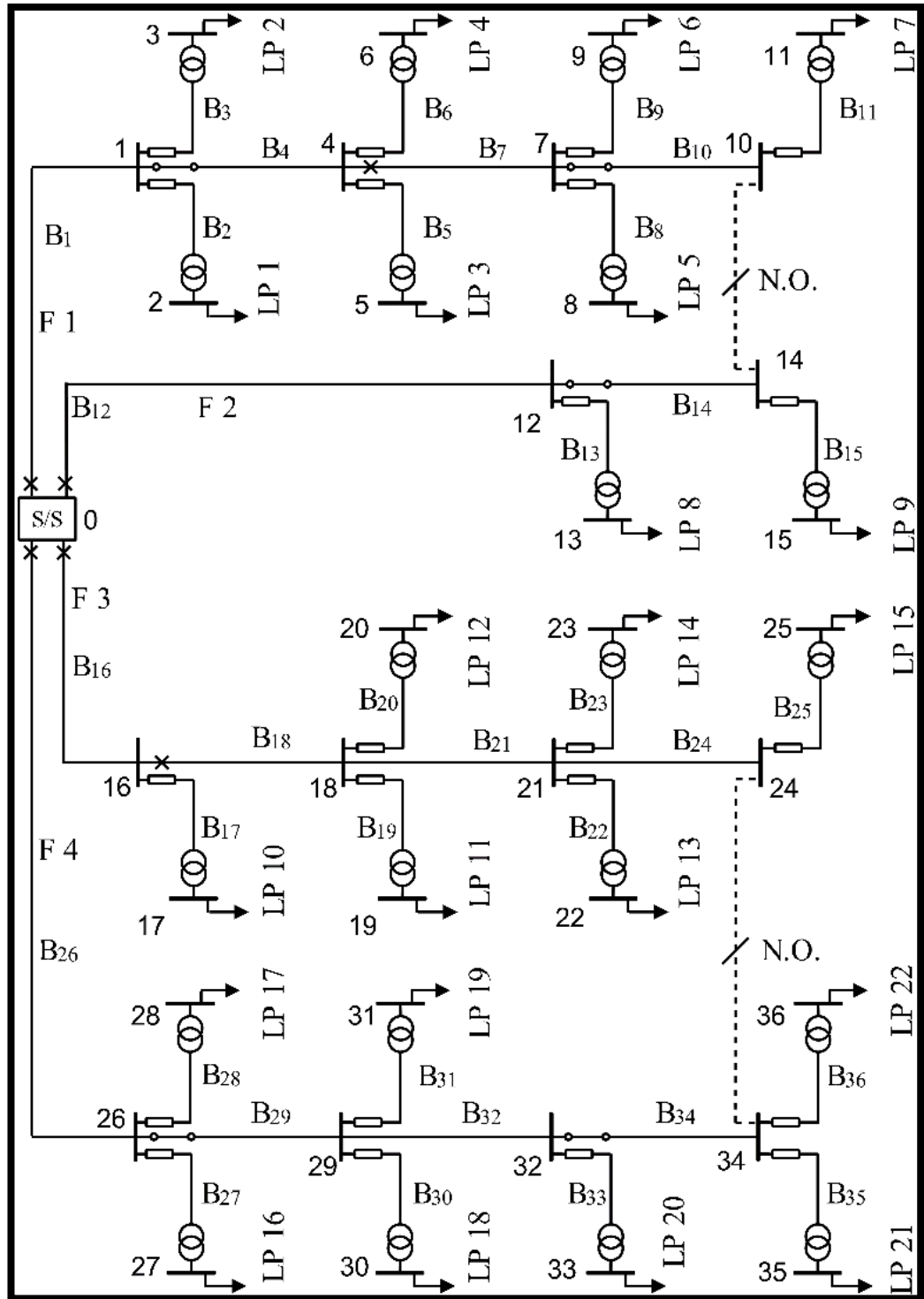
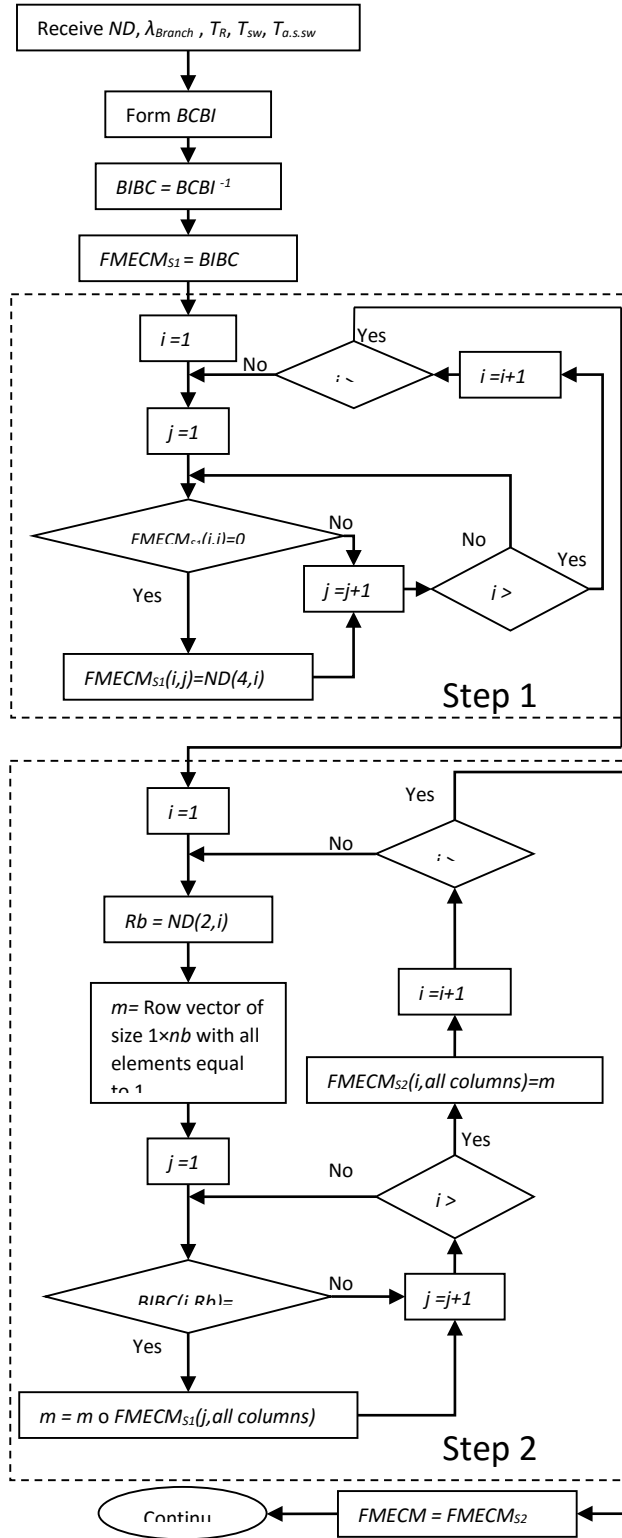


Fig. 5. Modified RBTS- Bus 2 test system

TABLE I
LOAD POINT RELIABILITY INDICES FOR MODIFIED RBTS-BUS 2

Load Point	λ	r	U
	f/yr	hr	hr/yr
Feeder 1			
1	0.1562	22.3604	3.4922
2	0.1689	21.0572	3.5569
3	0.1689	21.0572	3.7519
4	0.1562	22.3604	3.6872
5	0.2597	14.0456	4.1937
6	0.2565	14.1569	4.1775
7	0.2597	13.8954	4.1937
Feeder 2			
8	0.1408	3.8624	0.5438
9	0.1408	3.5854	0.6998
Feeder 3			
10	0.1042	33.0121	3.4402
11	0.2584	14.1132	4.1924
12	0.2616	14.0032	4.2086
13	0.2584	13.9120	4.1924
14	0.2616	13.8044	4.2086
15	0.2488	14.4078	4.1438
Feeder 4			
16	0.2597	14.0456	3.6477
17	0.2502	14.3880	3.5992
18	0.2502	14.3360	3.9892
19	0.2629	13.8875	4.0539
20	0.2629	13.8875	4.0539
21	0.2597	13.8453	4.1937
22	0.2629	13.7392	4.2099



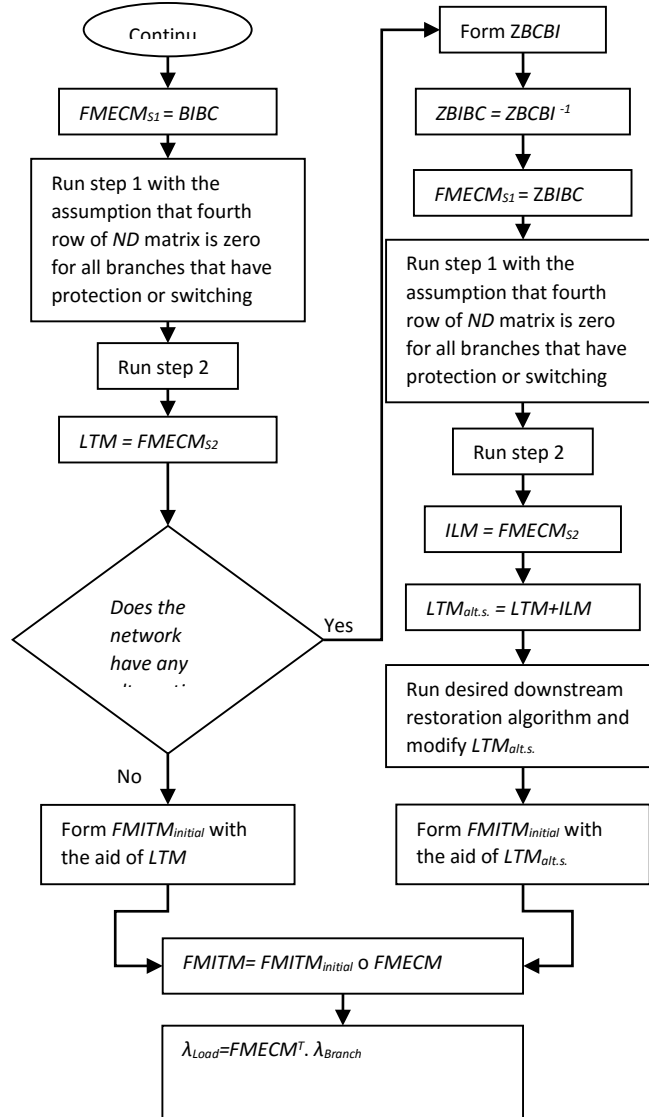


Fig. 6. Flowchart of the proposed algorithm

3. Cyber-Enabled Power Distribution System

3.1. Problem Statement

IEEE PES Distribution automation working group [36], defines distribution automation as, “a system that enables an electric utility to remotely monitor, coordinate and operate distribution components in a real-time mode from remote locations”. Continued advance in devices with information sending and receiving capabilities is the key enabler for distribution automation. In line with this trend manual switches and light emitting FDs have been substituted by cyber dependent RCSs and FDs. This has been typically associated with the deployment of SCADA technology along the distribution network [37], [38]. Wireless sensor network is also a promising technology to realize power distribution network automation [39]. In this work, the grid equipped with cyber-enabled devices is called cyber-enabled distribution network.

Fig. 7 shows a schematic diagram of a smart FD with its multi-hop communication link to the control center. Once a fault occurs in distribution feeder, it will be reported to the control center by detectors located at the source side of the fault location. Fig. 8 depicts a RCS scheme for a typical distribution feeder. After receiving the fault signal at the control center, the feeder will analyze and send the appropriate open/close signal to the local control unit (LCU) of normally closed/open switches. This will isolate the faulted area and backfeed as many interrupted loads as possible.

The above-mentioned arrangements can be integrated in a typical distribution feeder to bring about smart FDISR. Monitoring and control signals are always sent between two nodes through a communication channel. Success of consecutive post-fault actions is highly reliant on the performance of cyber-enabled devices (nodes) and underlying communication network. Since communication nodes and links are neither failure free nor fully predictable, their probability of failure is discussed in the following subsections.

3.1.1. Single-path end-to-end communication framework

The connection between two nodes at a fixed distance may fail because of the unsuccessful operation of any node or link between the two ends [40]. Communication devices are represented as nodes. Node unavailability is a function of Mean Time to Failure (MTTF) and MTTR of a device:

$$P_{failure}^{node} = 1 - \left[\frac{MTTF}{MTTF + MTTR} \right] \quad (11)$$



Fig. 7. Smart fault detector

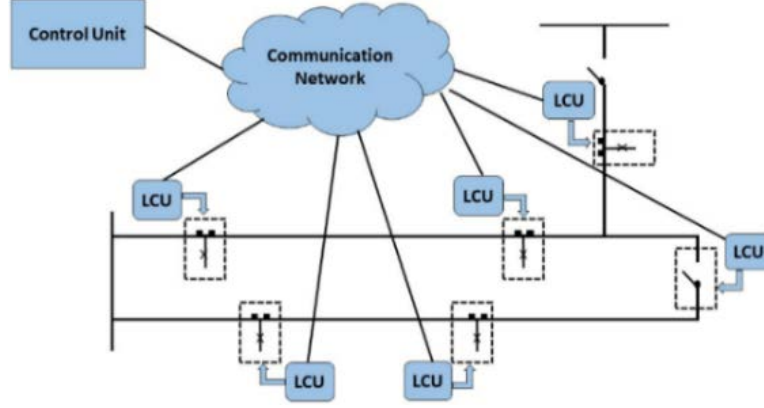


Fig. 8. Layout of RCSs on distribution feeder

In a typical distribution system, due to the length of distribution feeders, a direct communication link between the sender and the receiver may not be feasible. In order to transfer data between two points, multiple hops between sender and receiver may be required, as shown in Fig. 9. A sequence of nodes through which information is relayed must operate simultaneously in order to connect sending and receiving nodes.

In wireless systems, there is typically a target minimum received power level P_{min} below which performance becomes unacceptable. The probability of unsuccessful end-to-end transmission of a wireless link is [41], [42]:

$$P_{failure}^{link}(P_{min}, d) = P(P_r(d) < P_{min}) \quad (12)$$

From Fig. 9, the transmitted data from node S to node D should pass through route $(r_0, r_1 \dots, r_h)$, hopping h times. The route is identical to a sequence of h point-to-point links. Assuming ideal nodes, the event of successful end-to-end transmission is the event that all h transmissions are successful. The probability of successful end-to-end transmission is:

$$P_{success}^{multi.hops} = \prod_{i=1}^h (1 - P_{failure}^{link_i}) \quad (13)$$

3.1.2. Multi-path end-to-end communication framework

To enhance data transmission probability, multiple paths to re-route information is necessary. In case of more than one path between source and destination, the probability of successful operation of a communication path is characterized by success of at least one path between two specified nodes. Having discussed node and link data transmission probabilities, the availability evaluation of a cyber-enabled device is discussed in the following section.

3.2. Cyber-Enabled Device Availability

Physical devices can be in either operating or failure modes during their lifetime. Fig. 10a shows a state-space diagram of a single component. The probability of being in state “UP” or “DOWN” can be calculated using Markov rules [43]. P_{UP} and P_{DOWN} are referred to as the steady-state or limiting availability (state **A**) and unavailability of the system, respectively. In communication network, sensors are fragile and they may fail by environmental causes (e.g. lightning), hardware malfunction, battery depletion, and destruction by extreme events [39].

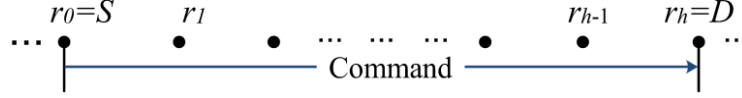


Fig. 9. Multi-hop data transmission

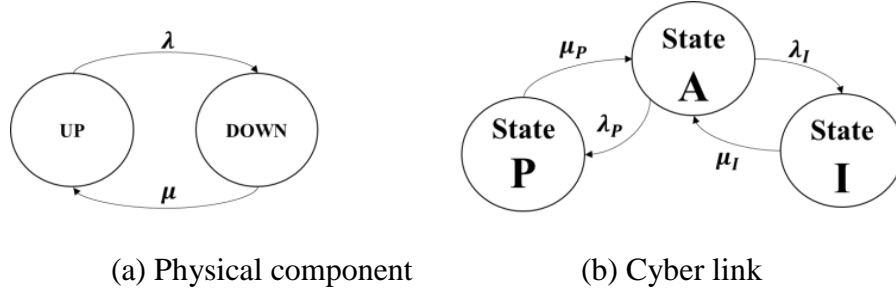


Fig. 10. State space diagram for physical component and communication link

Fig. 10b illustrates a simple state-space diagram of a communication link. Connection between the two nodes can be represented by a three-state diagram. The nature of the communication link is different than simple physical devices. It may undergo intermittent failures, which are transient and may occur as the result of events such as network congestion, and interferences (defined as state **I**). On the other hand, some failures are persistent, which may occur due to permanent blockage or environmental condition (precipitation) [39], [41]. High signal attenuation due to transmitter-receiver distance may also lead to a permanent failure. Recovery from these types of failures takes more time; hence, these failure modes are categorized into another state (state **P**). Depending on the communication protocol the intermittent state will contribute to the permanent or available states. Communication protocols with re-transmission allow control center to re-send the signal in case of an intermittent failure.

Therefore, the signal may reach destination without the need of human intervention. In this case the intermittent state contributes to the available state because load points are re-energized without undergoing manual restoration. In case of protocols without re-transmissions, an intermittent failure contributes to the permanent state since smart restoration is impossible. The cause of on-demand failure in communication network may sustain for a short period of time which also places protocols with limited number of re-transmissions into permanent states. In this case maximum number of transmission attempts is reached without successful reception, resulting in message loss

[41]. The availability and unavailability of communication data transmission from a power system point of view can be calculated as:

$$P_{success}^{link} = P_A + P_I \quad (14)$$

$$P_{failure}^{link} = P_P \quad (15)$$

where $P_{success}^{link}$ is the link availability, P_A and P_I are the probability of finding a link in available and intermittent failure states, $P_{failure}^{link}$ is the link unavailability, and P_P is the probability of permanent failure in a link. The term service availability is a long term measure in the context of timescale. It refers to the percentage of time that a communication link or network is up and running [41]. Fig. 11 shows a comprehensive view of the communication of smart devices with control center.

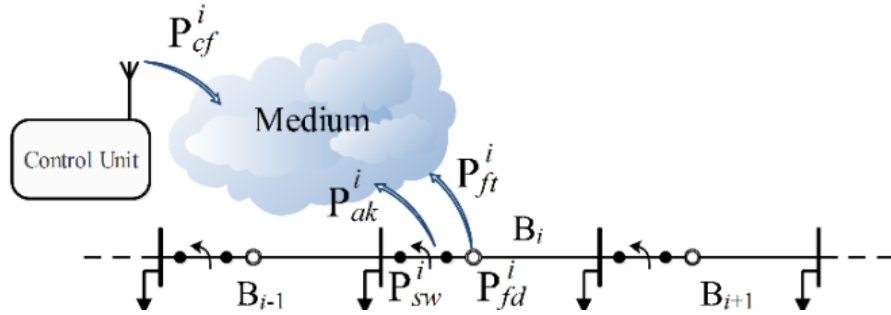


Fig. 11. Illustrative layout of distribution feeder with flow of data

When the sending node is a FD and the receiving node is a control center, then the probability that a smart FD on the i^{th} branch fails to operate is given by:

$$P_{f_i} = P_{fd}^i + (1 - P_{fd}^i)P_{ft}^i \quad (16)$$

The term $(1 - P_{fd}^i)P_{ft}^i$ means that the FD works; however, its transmitter and communication medium fail to transmit notification to the control center. The probability that the smart FD of the i^{th} branch operates correctly $(1 - P_{f_i})$ is calculated as:

$$(1 - P_{f_i}) = (1 - P_{fd}^i)(1 - P_{ft}^i) \quad (17)$$

The probability that the RCS of the i^{th} branch fails to operate (P_{s_i}) is given as:

$$P_{s_i} = \underbrace{P_{cf}^i}_{18.1} + \underbrace{\left(1 - P_{cf}^i\right) P_{sw}^i}_{18.2} + \underbrace{\left(1 - P_{cf}^i\right) \left(1 - P_{sw}^i\right) P_{ak}^i}_{18.3} \quad (18)$$

where the term (18.1) is for the probability of unsuccessful reception of signal at switch i . (18.2) means the switch receives the signal; however, it fails to change its status. The term (18.3) stands for the scenario whereby the signal is received and the switch operates; however, the backward signal conveying acknowledgment has not reached the control center. Therefore, the control center is uncertain of the operation. The probability that the RCS of the i^{th} branch works successfully is

$$(1 - P_{s_i}) = \left(1 - P_{cf}^i\right) \left(1 - P_{sw}^i\right) \left(1 - P_{ak}^i\right) \quad (19)$$

If FDs and switches operate in time, then FDISR is automatic. Under this scenario, the time required to perform this post-fault operation is assumed to be r_{sg} . According to the re-transmission procedure defined in different protocols, RCS switching time may vary from fraction of a second to a couple of minutes. Depending on the type of protocol, different average RCS switching time are possible.

3.3. Proposed Reliability Modeling

In this section, possible fault detection and isolation scenarios are investigated, and a reliability evaluation model is proposed. When a fault occurs on a feeder section, the substation circuit breaker trips. In the case of successful operation, all fault-indicating sensors located between the fault and the substation will report the fault.

The farthest FD that senses and reports the fault indicates the faulted section. Accordingly, the control center will send the trip signal to nearby RCSs. After isolating the fault, upstream loads are fed by reclosing the substation breaker, and downstream loads can be restored through an alternative supply path. For example, consider scenario 1 shown in Fig. 12. If the fault occurs in the i^{th} branch, f_i will report the fault, and the control center will send the opening signal to S_i and S_{i+1} . If all operations are coordinated successfully, then the isolated zone is as indicated in Fig. 12.

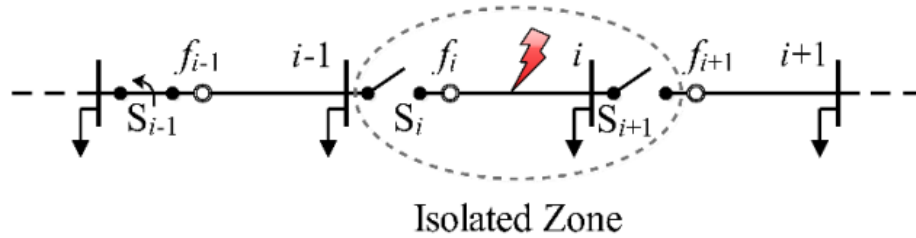


Fig. 12. Scenario 1: Successful fault detection and isolation

In scenario 2, shown in Fig. 13, the FD f_i fails to operate. Thus, the next immediate upstream FD (f_{i-1}) will report the occurrence of a fault. In this situation, it is assumed that the fault has occurred in $(i-1)^{\text{th}}$ section, which would result in a wrong fault location. The control center will then send the trip signal to S_{i-1} and S_i . If both of these switches operate successfully, then the wrong section will be isolated.

As the third scenario, if there exists an alternative supply path, it will be closed to serve the downstream loads. However, in the case shown in Fig. 13 (scenario 2), the faulty section will be energized through the alternative supply path, thus triggering its breaker to trip. Since the FD sensor has bi-directional fault detection capability, the control center will send the trip signal to S_{i+1} . With successful operation of this switch, the isolated zone is as shown in Fig. 14.

Generally, if a switch fails to operate, then the control center will send the trip signal to the next closest switch and the isolated zone is expanded.

In the aforementioned scenarios, all branches are assumed to be equipped with a FD and a RCS. If there is one or more RCS(s) between the fault location and a load point and if none of them operate successfully, then the interruption time is equal to the manual switching time (r_{sw}). If there is one or more RCS(s) between the fault location and a load point and if at least one of them operates successfully, then the interruption time is equal to the required time for remote switching (r_{sg}).

3.4. Reliability computation of cyber-enabled feeders

Considering the above-mentioned scenarios, the method of calculating reliability parameters of load points is investigated in this subsection. Three general cases that require distinct reliability computation are the following: (i) impact of fault on downstream loads, (ii) impact of fault on upstream loads, and (iii) a combination of both (i) and (ii).

Case i: Impact of fault on downstream loads

Assume a fault occurs on branch B_9 , as shown in Fig.15, and its impact on load point 2 is to be calculated.

In this case, the contributing devices to restoration time are f_3, f_8, f_9, S_3, S_8 , and S_9 .

- The failure rate of load point 2 due to the fault on B_9 is:

$$\lambda_{2-9} = \lambda_9$$

- The interruption duration of load point 2 due to the fault on branch B_9 is:

$$r_{2-9} = \Lambda_1 r_{sw} + (1 - \Lambda_1) r_{sg}$$

Fig. 13. Scenario 2: Unsuccessful fault detection and isolation

Fig. 14. Scenario 3: Unsuccessful switch operation

Fig. 15. Impact of branch B₉ failure on load point 2

where

$$\Lambda_1 = \left(1 - P_{f_9}\right) P_{S_9} P_{S_8} P_{S_3} + P_{f_9} \left(1 - P_{f_8}\right) P_{S_8} P_{S_3} \\ + P_{f_9} P_{f_8} \left(1 - P_{f_3}\right) P_{S_3} + P_{f_9} P_{f_8} P_{f_3}$$

Case *ii*: Impact of fault on upstream loads

The impact of a fault on branch B₂ on load point 8 is used as an illustrative example in this case, as shown in Fig. 16. In this case, the contributing devices to restoration time are $f_2, f_3, S2, S3$.

- The failure rate of load point 8 due to failure in branch B₂ is:

$$\lambda_{8-2} = \lambda_2$$

- The interruption duration of load point 8 due to failure in branch B₂ is

$$r_{8-2} = \Lambda_2 r_1 + (1 - \Lambda_2) r_2$$

where

$$\begin{aligned} \Lambda_2 = & \left(1 - P_{f_2}\right) P_{S_3} + P_{f_2} P_{S_2} P_{S_3} + P_{f_2} \left(1 - P_{S_2}\right) P_{f_3} + \\ & P_{f_2} \left(1 - P_{S_2}\right) \left(1 - P_{f_3}\right) P_{S_3} \\ r_1 = & (1 - P_{as}) r_{sw} + P_{as} r_{l2} , \quad r_2 = (1 - P_{as}) r_{sg} + P_{as} r_{l2} \end{aligned}$$

where P_{as} stands for alternative supply path failure probability.

Case iii: Combining both (i) and (ii)

All other possible cases are analyzed using a combination of the upstream and downstream effects. For example, assume that the fault occurs on branch B₆, and the effect on load point 4 is analyzed. In this case, the contributing devices to restoration time are $f_2, f_3, f_4, f_6, S_2, S_3, S_4$, and S_6 . To analyze the impact of this contingency, upstream and downstream effects are calculated through paths I and II, as shown in Fig. 17. If the FD at branch B₆ successfully conveys the location of the fault, then its impact on load point 4 is calculated through path I and the analysis has been given in case *i*. On the other hand, in the case of unsuccessful operation of detector at branch B₆, impact of the fault on load point 4 is analyzed through path II, and the analysis is given in case *ii*.

3.5. Analyzing impact of manually operated protective devices

Impact of adding fuse on laterals:

Successful or unsuccessful operation of a fuse affects the load point failure rate. In the case of unsuccessful operation of the fuse, the type of equipment on upstream branches influences the load point restoration time. For example, the failure rate and repair time of the load point at node 3 due to a fault on branch B₁₀ shown in Fig. 11 are given as follows:

The failure rate of load point 3 due to a fault on B₁₀ is

$$\lambda_{3-10} = P_f \cdot \lambda_{10}$$

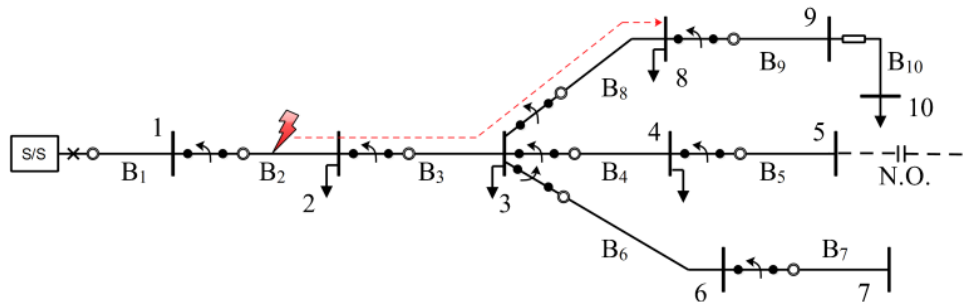


Fig. 16. Impact of branch B₂ failure on load point 8

P_f is the probability of unsuccessful operation of the fuse.

- The interruption duration of load point 3 due to a fault on branch B₁₀ is

$$r_{3-10} = \Lambda_3 r_{sw} + (1 - \Lambda_3) r_{sg}$$

where

$$\Lambda_3 = \left(1 - P_{f_9}\right) P_{S_9} P_{S_8} + P_{f_9} \left(1 - P_{f_8}\right) P_{S_8} + P_{f_9} P_{f_8}$$

In reality, *not all branches may be equipped with FDs and automated switches*. Typically, the number of FDs and automated switches are limited. This might require some modifications to the general equations:

No fault detector on branch:

If the i^{th} branch does not have a FD, then $P_{f_i} = 1$.

No remote-controlled switch:

If the i^{th} branch does not have a RCS, but has a manual switch, then $P_{S_i} = 1$.

If there is at least a manual switch between the fault location and a load point, then the interruption time is equal to the switching time (r_{sw}).

No switch:

If there is no switch (either manual or RCS) between the fault location and a load point, then the interruption time is equal to the repair time of the faulted line (r_{li}).

The following example illustrates how the general equations are modified to incorporate sections without FDs and switches. Fig. 18 shows a simple radial feeder that does not have all of its sections equipped with smart devices.

The fault is assumed on section 5 and its effect on load 1 is of interest in this example. The probability of unsuccessful smart operation and restoration time are given below:

$$\begin{aligned} \Lambda &= P_{f_5} P_{f_4} P_{f_3} P_{f_2} + \left(1 - P_{f_5}\right) P_{S_5} P_{S_4} P_{S_3} P_{S_2} + P_{f_5} \left(1 - P_{f_4}\right) P_{S_4} P_{S_3} P_{S_2} \\ &+ P_{f_5} P_{f_4} \left(1 - P_{f_3}\right) P_{S_3} P_{S_2} + P_{f_5} P_{f_4} P_{f_3} \left(1 - P_{f_2}\right) P_{S_2} = 1 \\ r_{1-5} &= \Lambda \times r_{sw} + (1 - \Lambda) \times r_{sg} = 1 \times r_{sw} + 0 \times r_{sg} = r_{sw} \end{aligned}$$

From Fig. 18, it is expected that load point 1 is exposed to the switching time, in case of this contingency.

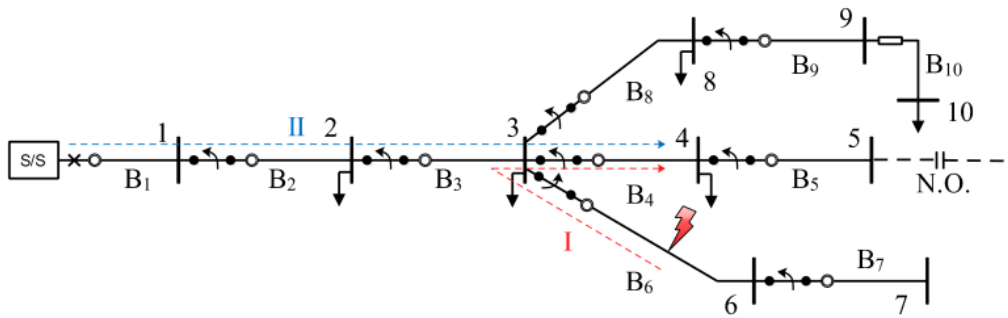


Fig. 17. Impact of branch B₆ failure on load point 4

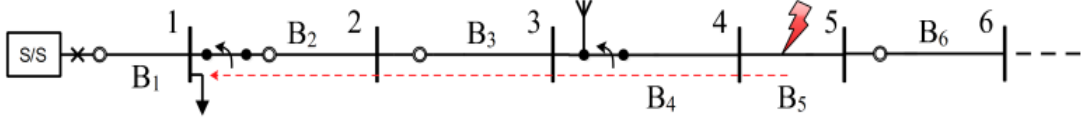


Fig. 18. Sample network containing manual and remote-controlled switches

3.6. Impact of Cyber Attacks

Fig. 19 shows an abstract view of a power system SCADA. As discussed in section I, vulnerabilities of control center LAN, corporate LAN, substation LAN or SCADA communication links can be exploited by an attacker to gain higher privilege level and compromise the system. Each attack scenario is represented by an MTTC and an MTTR [23, 28-30]. Having gained higher privilege level, an attacker can create a fake outage scenario and send false trip signals to RCSs. In order for a fake outage to seem authentic, an attacker should imitate a real outage scenario of a power distribution system. As discussed above, in a real outage scenario, circuit breaker of the substation trips upon occurrence of a failure. Then, the farthest FD that senses and reports the fault shows the closest region to the fault location. Trip signals are issued to nearby RCSs to isolate the faulted region.

Similarly, an attacker can create a fake outage scenario. To avoid any unnecessary effort, the attacker only needs to send false trip signal to substation breaker and manipulate the status data of some FDs. By doing so, the restoration process is automatically initiated. The last manipulated FD along the feeder shows the closest region to the fake fault location. Subsequently, trip signals are sent to nearby RCSs and crew members are dispatched. The curtailed load points in this region experience outage time equal to manually inspection of the region plus required time for cyber forensics (r_c).

Such an attack against power distribution system can be launched through control center LAN, corporate LAN, substation LAN or eavesdropping on the SCADA communication links and then injecting false trip and status signals. Therefore, from distribution system reliability perspective, these cyber-attack paths are viewed as series connected components as shown in Fig. 20. Down and up states of the components are construed as being under cyber-attack or not. Successful cyber-attack against each of the components can lead to interruption of supply to some load points or the whole power distribution feeder.

The question here is, whether the analytical approach used for reliability assessment of series systems (as in section V) is applicable for the system of Fig. 20. The underlying assumption in extracting the analytical approach for a series system is constant failure and repair rates of the components. In other words, time to failure and time to repair of components are considered to be exponentially distributed.

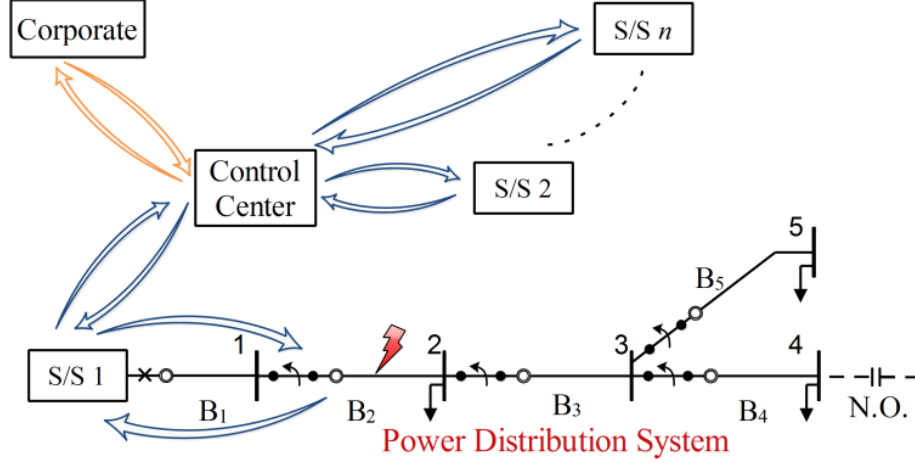


Fig. 19. Abstract view of a Power system SCADA

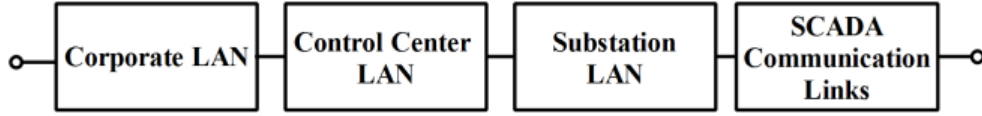


Fig. 20. SCADA system from reliability point of view of a power distribution system

In case of a cyber-network, a non-exponential distribution for time to compromise and time to repair are expected. This results in a non-Markovian system. However, in [41] it is proved that the same analytical approach is applicable as long as the long term average values (not the time dependent values) are being evaluated for systems of statistically independent components.

The first step to prove this is considering the fact that combination of two or more exponentially distributed states results in a non-exponentially distributed state. The shape of the resulting distribution is defined based on the number of combining states, the way they are combined and their parameters. Therefore, a non-exponentially distributed state can be represented as the combination of some exponentially distributed states. This process is called the method of stages [44].

Fig. 10a shows the state space diagram of a repairable component with constant failure and repair rates. The limiting state probabilities and frequencies of this component are

$$P_{up} = P_0 = \frac{\mu}{\lambda + \mu} \quad , \quad P_{down} = P_1 = \frac{\lambda}{\lambda + \mu} \quad , \quad f_{up} = f_{down} = \frac{\lambda\mu}{\lambda + \mu}$$

In Fig. 21 the up state of the component is considered as a non-exponentially distributed state represented by two series-connected exponentially distributed states. The departure rates are chosen as 2λ to give the same MTTF as in Fig. 10a. This can be easily shown considering state 2 as the absorbing state.

The stochastic transitional probability matrix of this system is

$$\begin{bmatrix} 1-2\lambda & 2\lambda & 0 \\ 0 & 1-2\lambda & 2\lambda \\ \mu & 0 & 1-\mu \end{bmatrix}$$

Using this matrix the following limiting state probabilities are obtained:

$$P_0 = \frac{\mu}{2(\lambda + \mu)} \quad , \quad P_1 = \frac{\mu}{2(\lambda + \mu)} \quad , \quad P_2 = \frac{\lambda}{\lambda + \mu}$$

The probability of down state is equal to P_2 which is the same as in Fig. 10a. The probability and frequency of the up state are

$$P_{up} = P_1 + P_2 = \frac{\mu}{\lambda + \mu}$$

$$f_{up} = f_{down} = P_0 \cdot 2\lambda + P_1 \cdot 2\lambda - P_0 \cdot 2\lambda = \frac{\lambda\mu}{\lambda + \mu}$$

Which are equal to the probability and frequency of encountering up state in Fig. 10a.

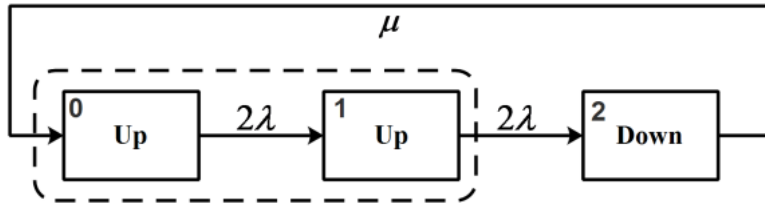


Fig. 21. The component of Fig. 4a with non-exponentially distributed up state

This proof can be extended to more complex systems. It shows in a system with non-exponential distributions, limiting state probabilities and frequencies are equal to those obtained under the assumption of exponential distribution. Therefore, to obtain long term average values, the same analytical approach for series system is applicable. So, each cyber network of Fig. 20 is represented by constant cyber-attack and repair rate which give the same MTTC and MTTR of the network as follows.

$$\lambda = \frac{1}{MTTC} \quad , \quad \mu = \frac{1}{MTTR} \quad (20)$$

In each cyber-attack of this kind, the status of an arbitrary set of FDs can be manipulated. As an example, in Fig. 22, if the status of fault detectors f_1 and f_4 are manipulated, trip signals are sent to RCSs S2 and S6. The dispatched crew members can then limit the isolated zone using the manual switch S3. Therefore, load point1, load point 2, load points 3-5 and load point 6 will experience

outage duration equal to r_{sg} , r_{sw} , r_c and $\{(1 - P_{as})r_{sg} + P_{as}r_c\}$, respectively. Without the alternative supply path, load point 6 will also experience outage duration equal to r_c . The other scenarios can be manipulating the status of fault detector f_1 alone and fault detectors f_1, f_4 and f_6 , simultaneously. The probability of occurring each scenario is considered as $1/w$, where w represents the total number of fault detectors. The same probability for different scenarios shows the indifference of an attacker to load points. However, the attacker can bring more financial harm by attacking load points with higher interruption cost (like commercial or industrial load points) more frequently. Then, higher probability of occurrence can be assigned to the scenarios including such load points.

3.7. Common Cause Failures

Multiple communication devices may fail due to a shared cause, simultaneously. Taking these failures into account in addition to independent failures will result in a more precise reliability estimation. Reference [45] presents guidelines on modeling Common Cause Failures (CCFs) in probabilistic risk assessment.

Two main factors are required for a CCF to occur: a root cause and a coupling factor that makes multiple components susceptible to the same cause. Examples of a coupling factor can be the same location for components, same environment, function or procedures. At power distribution network, environment is a coupling factor for line and installed devices. Environment can be the source of different root causes for a CCF. Cyber-attack in the form of denial of service can be another root cause whereby the time-critical data sent by a FD or the trip signals sent to a RCS can be blocked or corrupted. This type of attack can also lead to failure of multiple communication devices at the same time. It should be noted that the main target of the denial of service attack may be other infrastructures that coincidentally lead to the malfunction of fault detectors and RCSs. CCFs have become an important topic for complex cyber-physical systems design [46]. An identified group of components vulnerable to CCFs is called a common cause component group (CCCG).

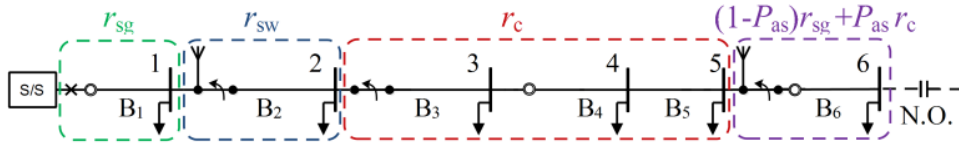


Fig. 22. Different outage durations experienced by load points

The first step in incorporation of CCFs into reliability analysis is to identify the basic failure events of a CCCG. For instance, the basic events for a CCCG of three components A, B, and C are C_A , C_B , C_C , C_{AB} , C_{BC} , C_{AC} , and C_{ABC} . The first, fourth, and last events represent independent failure of component A, CCF of A-B, and CCF of A-B-C, respectively. Any of events C_A , C_{AB} , C_{AC} and C_{ABC} causes A to fail. The independent failure of a component itself is sum of the true independent failure and the failure of only that component due to a common cause. The basic events that lead to failure of a component are considered to be mutually exclusive which implies zero probability for the occurrence of events such as $C_A C_{AB}$ or $C_{AB} C_{AC}$. The reason is the indiscernibility of events $C_A C_{AB}$ from C_{AB} and $C_{AB} C_{AC}$ from C_{ABC} .

Next step is to quantify basic events probabilities. Commonly, the probabilities of similar events with similar components are assumed to be the same (symmetry assumption). For a CCCG of size m , probability of simultaneous failure of k components is represented by $P_k^{(m)}$. A practical model to quantify $P_k^{(m)}$ is known as alpha factor model [45]. Two parameters are used in this model: P_t and α_k . P_t denotes the total failure probability of each component. α_k represents the probability of failure of k components given a CCF occurs. Using these two parameters, $P_k^{(m)}$ is:

$$P_k^{(m)} = \frac{k}{\binom{m-1}{k-1}} \frac{\alpha_k}{\sum_{k=1}^m k \alpha_k} P_t \quad (21)$$

In order to incorporate CCFs into reliability of power distribution system, two facts should be taken into account. First, simultaneous failure of higher number of components has lower chance to occur. Second, when a failure occurs along the feeder, first, the closest cyber-enabled devices are used to isolate the fault. If they do not operate properly, either due to independent failure or a CCF, the fault propagates and the next immediate cyber-enabled devices are tried. Therefore, the load points close to the fault location have higher chance of being affected. This chance for the load points far from the fault location reduces significantly. Taking these facts into account justifies limiting the CCCGs to the communication devices close to the fault location. The consequence is significant reduction in the number of basic events at the cost of negligible underestimation of unavailability of cyber devices. In the following, for the simplicity of representation the CCCGs is limited to the fault location itself and the immediate adjacent cyber-enabled devices as shown in Fig.23.

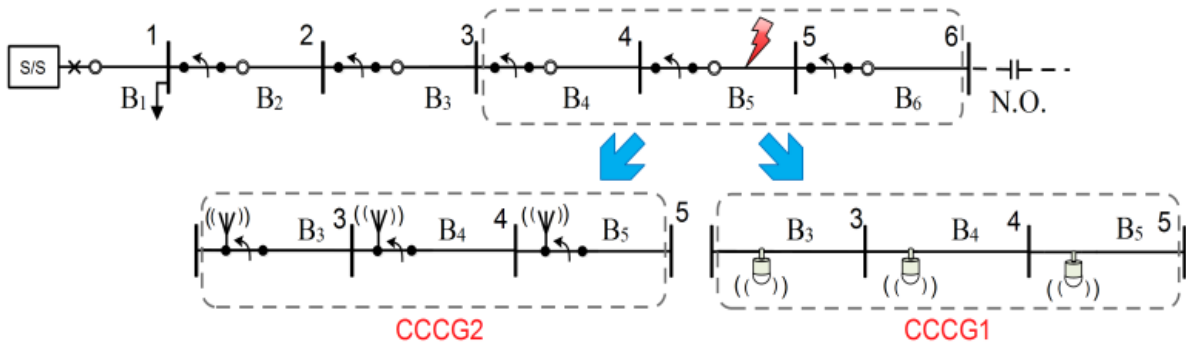


Fig. 23. Common cause component groups

In this figure, two CCCGs are discernible: CCCG1 including FDs and CCCG2 including RCSs. In CCCG1, FDs are identical. The only source of dissimilarity among these components is different probability of link failures which only affects the true independent failure of FDs. Taking these facts into account, the symmetry assumption can be applied to failure probability of the components in this CCCG. This is the case for components of CCCG2. Cyber-enabled devices outside the CCCGs fail independently.

After the probability of basic events is quantified, the following steps are carried out. In section V, analytical expressions were developed to calculate reliability parameters of load points considering only independent failures. To incorporate CCFs into the expressions, the following four steps need to be taken successively:

- Each failure probability is replaced by all its relevant basic events (the complement of event C is represented by C')
- Boolean algebra properties are used to combine these basic events
- Indiscernible events are eliminated
- Finally, each basic event is replaced by its probability of occurrence.

As an example, in Fig. 23, the term $P_{f_5}(1-P_{f_4})P_{s_4}P_{s_3}$ is used in the expression for calculating the impact of failure on load point 2. Applying the first step, the following expression is formed.

$$\left(C_{f_5} + C_{f_{4,5}} + C_{f_{5,6}} + C_{f_{4,5,6}} \right) \left(C_{f_4} + C_{f_{4,5}} + C_{f_{4,6}} + C_{f_{4,5,6}} \right)' \left(C_{s_4} + C_{s_{4,5}} + C_{s_{4,6}} + C_{s_{4,5,6}} \right) C_{s_3}$$

If the second to fourth steps are carried out the following expression is obtained.

$$\left\{ P_{f_5} \left(1 - P_{f_4} - P_{f_{4,6}} \right) + P_{f_{5,6}} \left(1 - P_{f_4} \right) \right\} \left(P_{s_4} + P_{s_{4,5}} + P_{s_{4,6}} + P_{s_{4,5,6}} \right) P_{s_3}$$

where,

$$\begin{aligned} P_{f_4} &= P_{f_5} = P_{f_6} = P_1^{(3)} \\ P_{f_{4,5}} &= P_{f_{4,6}} = P_{f_{5,6}} = P_2^{(3)} \\ P_{f_{4,5,6}} &= P_3^{(3)} \end{aligned}$$

The same equalities are hold for switches.

3.8. Optimal Placement of Devices

3.8.1. Objective:

Improving system reliability as a whole will eventually lead to customer satisfaction. However, it should be economically justified in order to convince decision-makers and stakeholders to invest more resources and adopt more intelligent devices. In this work, the optimal number and location of FDs along with manual and automatic switches is determined to reduce the EIC to customers.

This cost is computed as

$$EIC = \sum_{y=1}^{ny} \left(\frac{1}{(1+\rho)^y} \sum_{n=1}^{NC} \sum_{j=1}^{LP} \sum_{k=1}^{LT} \left(L_{jky} \cdot \lambda(n) \cdot C_{r_{nj}}^k \right) \right) \quad (22)$$

EIC presents an estimation of customers' interruption cost. Eq. 22 shows that EIC is a function of type of interrupted load, interruption duration, and average load power. Using this index instead of common reliability indices, it is possible to better estimate customers' true interruption cost. A survey was conducted in [47] to estimate interruption losses to customers from different sectors which is used in this work. The objective function for this planning problem is:

$$\min C_F + C_S + EIC \quad (23)$$

This problem is solved using genetic algorithm (GA). GA by integer representation of individual solutions is a suitable optimization tool to solve this problem [48].

Each chromosome gives a possible solution in which genes represent candidate locations for the placement problem. As shown in Fig. 24, the first string represents the location of FDs, which can be either 0 (no FD placed) or 1 (FD is placed). Each bit in the string corresponds to the potential candidate location. The second string represents the switch placement. Each bit could contain either 0 (no switch placed), 1 (manual switch is placed), or 2 (RCS is placed).

3.8.2. Model Validation and Analysis

This work utilizes RBTS - bus2 [32] and a typical 27-node distribution feeder [35] to evaluate the proposed model. These two testbeds have been selected to better demonstrate the effectiveness of the proposed model. In order to implement the analytical reliability model and make it applicable to different scenarios, the methodology was realized on a Pentium-IV personal computer using MATLAB software. GA is used to find the optimal planning decisions. For both placement problems, the elitism rate is assumed to be 0.1. Crossover and mutation rates are considered to be 0.95 and 0.01, respectively. CCCGs are limited to the fault location itself and the immediate adjacent cyber devices. 90% of RCS or FD total probability is considered to be independent failure, 5% to be CCF including two components and 5% to be CCF of more than 2 components.

3.8.2.1. RBTS - Bus 2

Bus 2 of the RBTS is a well-known distribution testbed with four feeders and 22 load points. Two normally open switches connect buses 10 to 14 and 24 to 34, as indicated in Fig. 25. Candidate locations for switches and FDs have been highlighted in red in Fig. 25. The substation is assumed to have a RCS and a FD for each of four feeders. For this planning problem, the project horizon year is equal to five years with five years of load growth. The discount rate is assumed to be 7%. The manual and smart grid switching times are 60 minutes and 1 minute, respectively. The repair time for any faulted section is assumed to be five hours. All the other required information has been given in [32].

Three boundary values are investigated to illustrate the effectiveness of deploying manual and remote-controlled switches as well as FDs on this testbed. The cost of each FD is assumed to be \$540. Manual and remote-controlled switches cost \$580 and \$5,200, respectively.

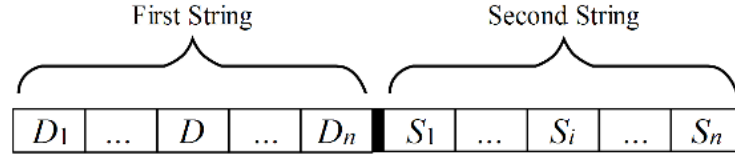


Fig. 24. Integer string representation for optimal planning problem

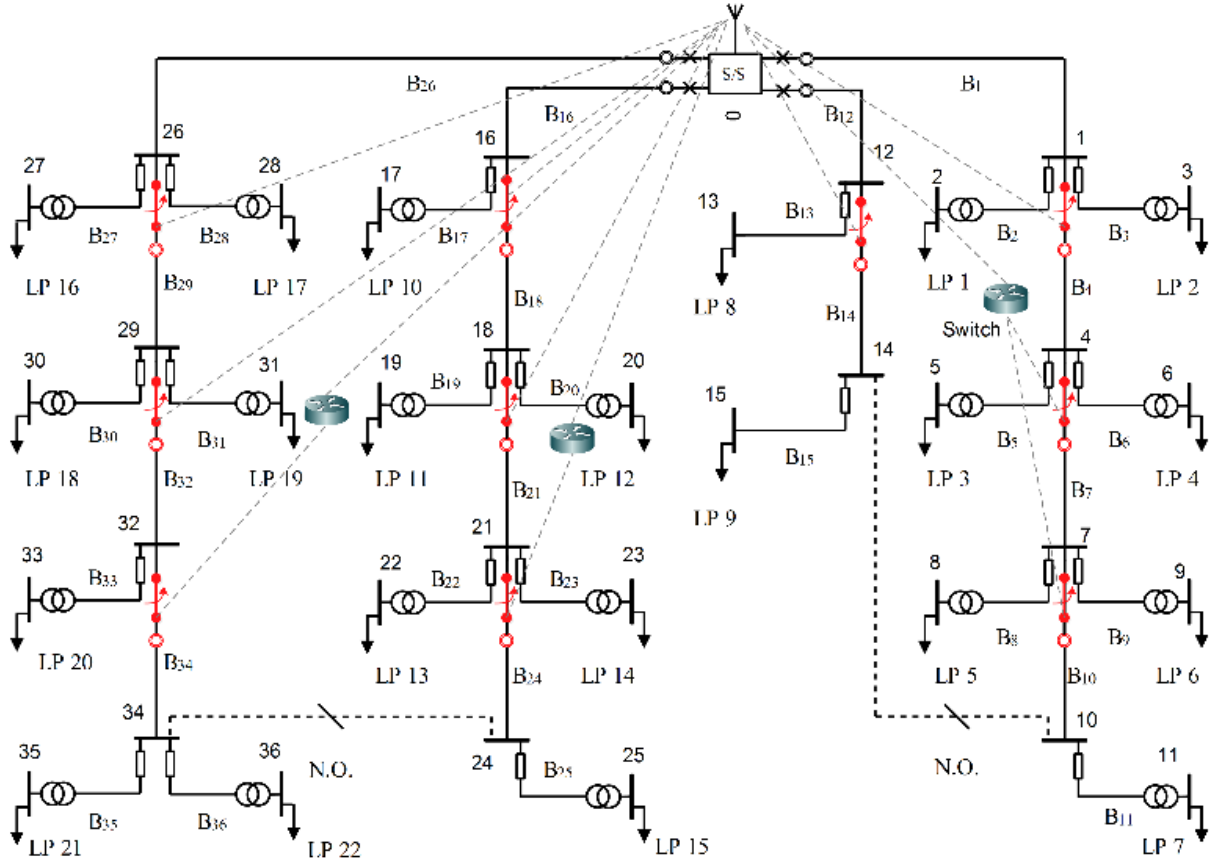


Fig. 25. RBTS bus 2 test feeder

From Table II, it can be seen that with only four breakers and their FDs, EIC is equal to 0.3 million dollars. Total investment cost is equal to zero in this scenario, since no switch or FD has been selected.

In the second case, candidate locations are filled by manual switches that show a significant improvement in EIC. If all locations are filled by FDs and RCSs, then the best values are obtained for EIC, meaning that the most expensive investment is made. Table II provides a comparison of the three above-mentioned scenarios.

In order to show the impact of communication infrastructure failure on planning decisions, two

sets of data transmission probabilities are assigned as shown tabulated in Table III. Using these two probability sets, two experiments are performed on RBTS-bus 2.

The failure probability for fault-detector and switching-device hardware is assumed to be 0.01. With the two sets of probabilities, the planning problem is solved using GA to find optimal solutions. Optimal decisions for the first experiment are shown in Table IV. For the first and second sets of probabilities, three and four RCSs and FDs were selected, respectively. Higher probabilities in the second set economically justify this additional switch and FD. EIC for the second group is also lower than the first group, which implies lower customer outage time.

Fig. 26 shows a comparison of load point outage time for the two sets of probabilities given in Table III. As expected, in the second feeder, there is no change in load point outage time, since no RCSs and FDs have been chosen. For the first and third feeders there is no change in numbers and locations of FDs and switches. However, due to an improvement in probabilities, load-point outage time has been reduced. For the fourth feeder, without adding the switch, the load point outage time should have been reduced. However, a more significant improvement is seen in outage time after adding this switch.

In order to show the impact of changing cost of smart devices on the planning problem, the second experiment was conducted, and results are tabulated in Table V.

All scenarios are performed for two sets of successful data transmission probabilities. Table V shows that increasing the price can reduce the number of switches selected.

TABLE II. THREE BOUNDARY CASES FOR FIRST CASE STUDY

Type of devices	EIC $\times(10^5)$ (\$)	Investment cost (\$)
No switch& no FD	3.098	0
Manual switches	2.22	5,800
RCS & FDs	1.753	57,400

TABLE III. TWO SETS OF SUCCESSFUL DATA TRANSMISSION PROBABILITIES

Candidate location	Set #1	Set #2
B₄	0.967	0.999905
B₇	0.936	0.999989
B₁₀	0.941	0.999989
B₁₄	0.967	0.999905
B₁₈	0.956	0.999902
B₂₁	0.939	0.999989
B₂₄	0.911	0.999989
B₂₉	0.954	0.999902
B₃₂	0.937	0.999989
B₃₄	0.903	0.999989

TABLE IV. OPTIMAL DECISIONS FOR EXPERIMENT 1.

	Set #1	Set #2
Location of RCSs	7-21-29	7-21-29-34
Location of FDs	7-21-29	7-21-29-34
EC	\$195,510	\$186,857
Investment cost	\$21,280	\$26,440

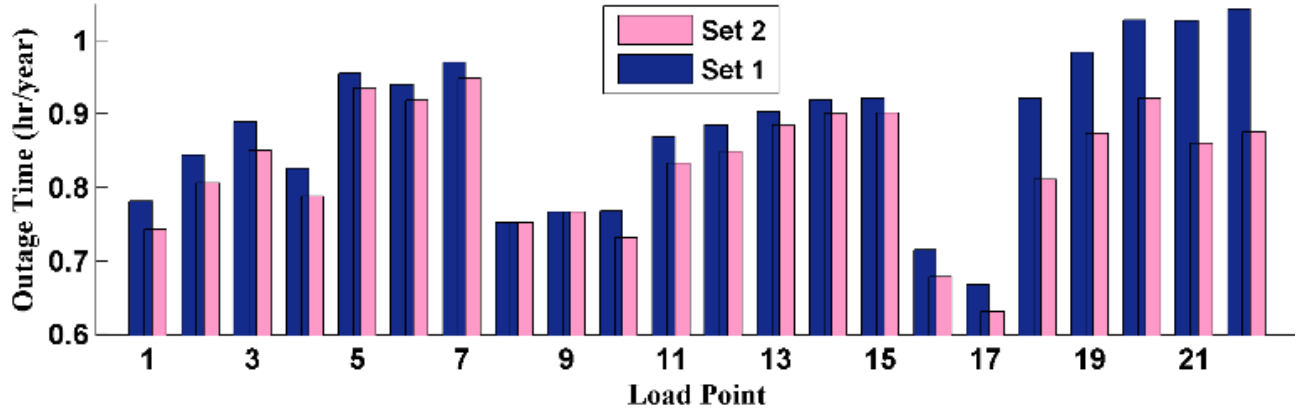


Fig. 26. Load point outage time

TABLE V. RESULTS FOR EXPERIMENT 2

FD and RCS Costs	Set	EIC (K\$)	FDs Location	RCSs Location
FD=\$740 RCS=\$4,200	1	192	B ₇ , B ₂₉	B ₇ , B ₂₁ , B ₂₉
	2	183	B ₇ , B ₂₁ , B ₂₉	B ₇ , B ₂₁ , B ₂₉ , B ₃₄
FD=\$1,040 RCS=\$6,200	1	202	B ₇ , B ₂₉	B ₇ , B ₂₉
	2	192	B ₂₁ , B ₂₉	B ₇ , B ₂₁ , B ₂₉
FD=\$1,540 RCS=\$10,200	1	222	-	-
	2	210	B ₃₂	B ₃₂

RBTS includes four feeders spreading out from the substation. It also has a small number of candidate locations for smart devices compared to practical distribution feeders. Two experiments performed in the first case study show the difference made by communication infrastructure failure

probability. However, in order to better illustrate the importance of grid modernization and including communication unavailability into the decision-making process, it is necessary to validate the proposed model using a more practical distribution network in the second case study.

3.8.2.2. Typical 27-node distribution feeder

Fig. 27 shows a typical 27-node distribution feeder. In this figure, candidate locations for FDs and manual/remote-controlled switches have been highlighted in red. FDs are assumed to cost \$10,000. Each remote-controlled and manual switch cost \$25,000 and \$5,000, respectively. The role of cyber-attacks is also investigated in this case. The same MTTR is used for different cyber-attack scenarios. Therefore, an equivalent cyber-attack rate is considered for all the attack paths. In our work the attacker is assumed to be indifferent to load points. The planning horizon is five years. GA is applied to this case with the same setting as given in subsection A. Boundary condition values are provided in Table VI. Changing successful data transmission probabilities from 0.9 to 0.9999 shows a significant improvement in system EIC.

Table VI shows optimal solutions after solving the placement problem for three different probabilities. For the sake of comparison, total cost incurred by different switching and FDs placement are obtained and shown in Fig. 28. As this figure shows, total cost reduces with either the increase of probability of successful data transmission or decrease of cost of switches and FDs. The same comparison has been done in Fig. 29 considering different successful data transmission probabilities and different cyber-attack rates. Cyber-attack rate is function of attacker's skill level. As one last experiment three factors were given to Design-Expert® Software Version 9 [49] to further evaluate the impact of input data on the objective function. Design of experiment is a statistical tool to analyze joint effects of input parameters on the output [49]. Effects of cyber-attack rate (A), successful data transmission probability (B) and equipment cost (C) on customers' EIC and total cost are shown in Fig. 24a and Fig. 24b, respectively. Each factor is represented by two levels. Maximum and minimum values of the three factors given in Fig. 28 and Fig. 29 are considered as up and down levels. The standardized effect is the difference of average output for up and down levels. The points on the far right of the straight line are identified as significant effects. The vertical axis represents the Half-Normal probability value for each effect.

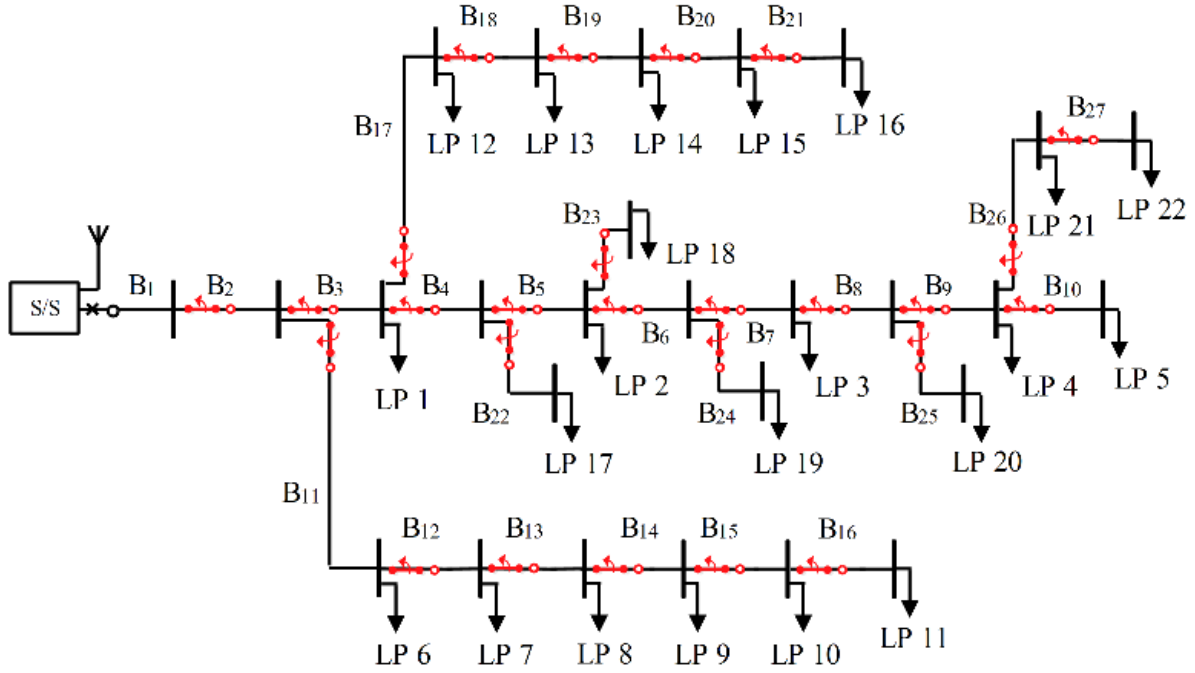


Fig. 27. Typical 27-node practical distribution feeder

TABLE VI. BOUNDARY VALUES FOR CASE 2

Type of devices	Successful data trans. Probability	EIC (K\$)	Investment cost (K\$)
No switch & no FD	-	2,229	35
Manual switches		1,498	165
RCSs & FDs	.9999	692	945
	.999	693	
	.99	710	
	.9	872	

TABLE VII. OPTIMAL SOLUTIONS FOR CASE 2 WITH THREE LINK PROBABILITIES

Link success probability	0.9999	0.99	0.9
EIC (K\$)	794	837	1,061
Invest. cost(K\$)	355	365	365
RCS locations	B ₁ , B ₈ , B ₁₂ , B ₁₅ , B ₁₇ , B ₂₂ B ₂₃ , B ₂₄ , B ₂₆	B ₁ , B ₇ , B ₁₂ , B ₁₅ , B ₁₇ , B ₂₂ , B ₂₃ , B ₂₄ , B ₂₅	B ₁ , B ₆ , B ₁₂ , B ₁₅ , B ₁₇ , B ₁₈ , B ₂₂ , B ₂₃ , B ₂₆
Manual switch locations	B ₅ , B ₆ , B ₁₀ , B ₁₁ , B ₁₆ , B ₁₉ , B ₂₀ , B ₂₀	B ₄ , B ₅ , B ₈ , B ₁₁ , B ₁₆ , B ₁₈ , B ₁₉ , B ₂₁ , B ₂₅ , B ₂₆	B ₄ , B ₈ , B ₁₀ , B ₁₁ , B ₁₄ , B ₁₆ , B ₂₀ , B ₂₁ , B ₂₄ , B ₂₅
FD locations	B ₁ , B ₈ , B ₁₂ , B ₁₅ , B ₁₇ , B ₂₂ B ₂₃ , B ₂₄ , B ₂₆	B ₁ , B ₇ , B ₁₂ , B ₁₅ , B ₁₇ , B ₂₂ , B ₂₃ , B ₂₄ , B ₂₅	B ₁ , B ₆ , B ₁₂ , B ₁₅ , B ₁₇ , B ₁₈ , B ₂₂ , B ₂₃ , B ₂₆

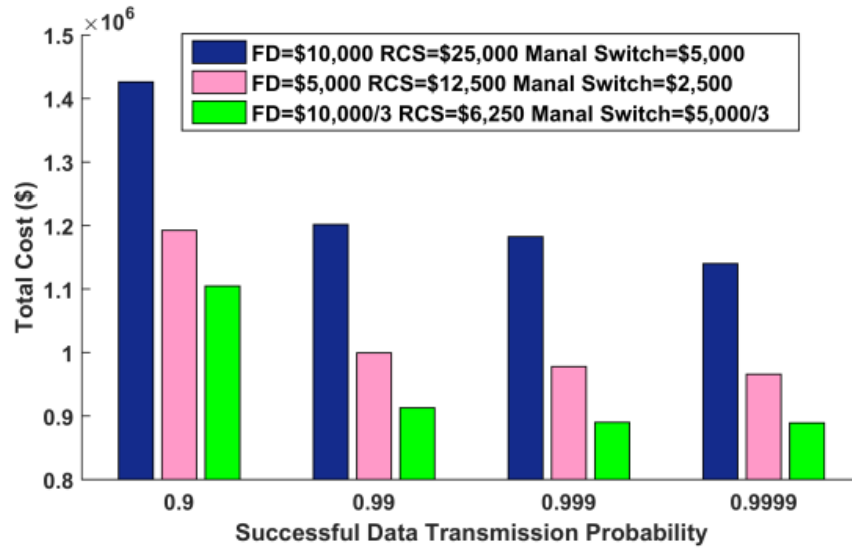


Fig. 28. Total cost incurred for different successful data transmission probabilities and switches and FDs cost

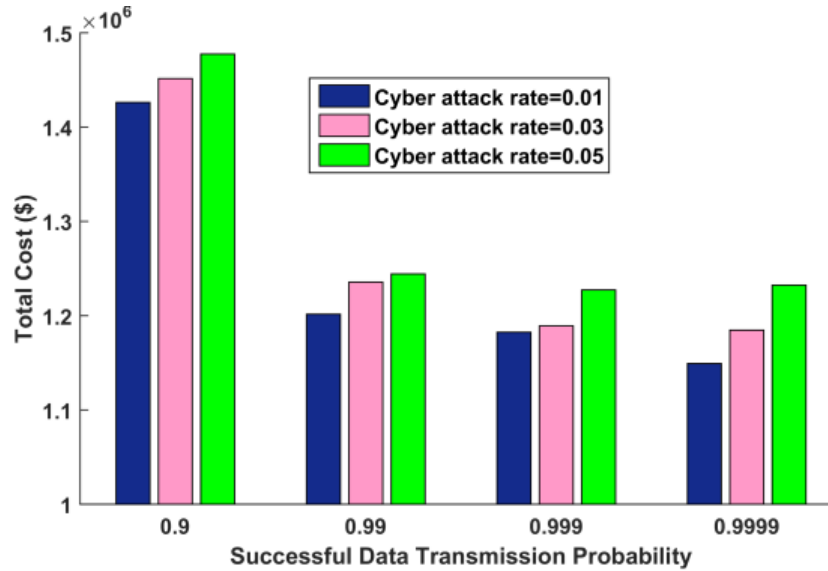


Fig. 29. Total cost incurred for different successful data transmission probabilities and cyber attack rates

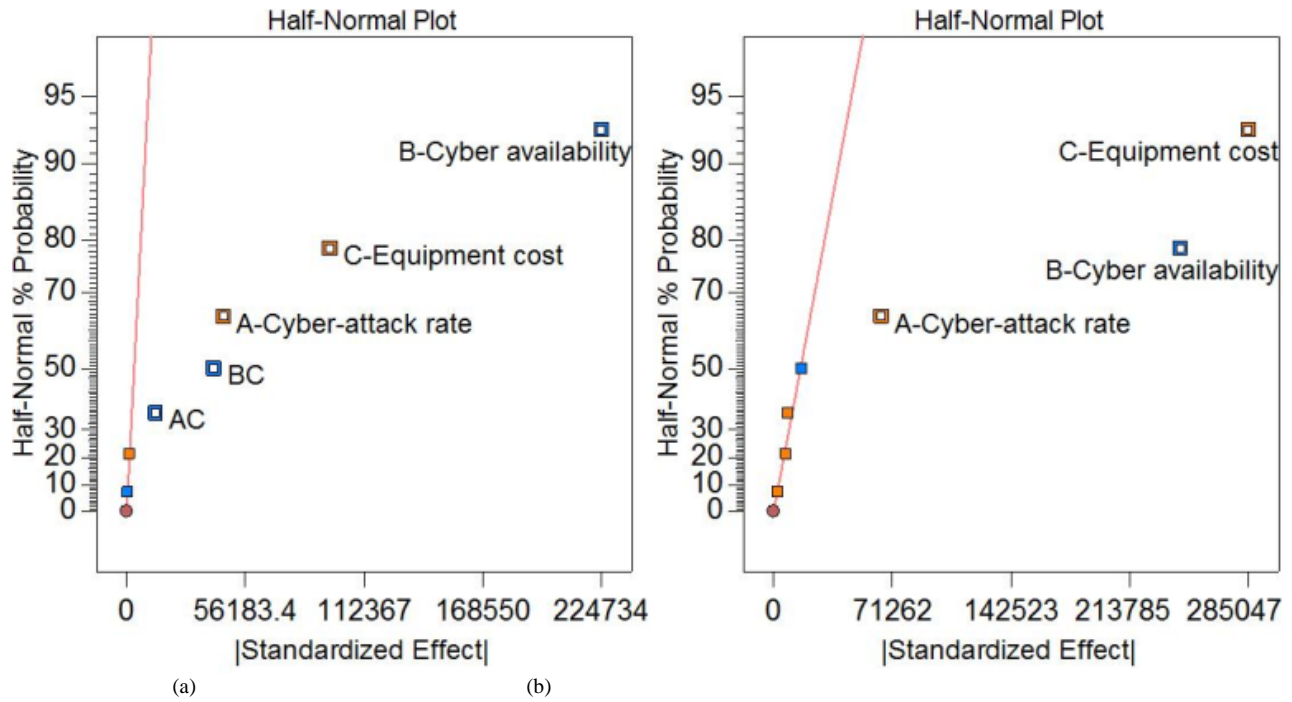


Fig. 30. Impact on (a) customers' interruption cost and (b) total cost

4. Conclusion and Future Work

In this work, an analytical model was developed to measure distribution network reliability indices. The model was extended to incorporate cyber enabled logic into power system. A new analytical model was proposed to analyze the impact of cyber-enabled FDISR on the reliability of power distribution system. The analytical model was employed to calculate the reliability of the system incorporating on-demand network failure and cyber-attack in a switch and FD placement problem. The optimal location and number of switches and FDs were determined to minimize the sum of the total cost of customer service interruption and investment cost. Results of case studies show the importance of communication vulnerabilities in cyber-enabled power distribution system design. In this research fault detection, isolation and service restoration were selected to show the impact of cyber-vulnerability on a typical power system planning problem. This example has both sensing and actuating characteristics so that cyber-vulnerability could be properly analyzed. Cyber vulnerability would affect data transfer between the two ends which has not been fully investigated in this work. The future work would be proposing an advanced state estimation tool to fix data unavailability and detect abnormality in measured data. This is the target of our research now as we are trying to develop a highly accurate prediction software.

References:

- [1] R. Billinton and P. Wang, "Network-equivalent Approach to Distribution System Reliability Evaluation," *Proc. Inst. Elect. Eng. Gen. Transm. Distrib.*, vol. 145, no. 2, pp. 149–153, 1998.
- [2] BILLINTON, R., and ALLAN, R.N., "Reliability Evaluation of Power Systems", (Plenum Press, New York, 1984)
- [3] Don O. Koval, "Zone-Branch Reliability Methodology for Analyzing Industrial Power Systems", *IEEE Transactions on Industry Applications*, Vol. 36, No. 5, pp. 1212-1218, 2000
- [4] Richard E. Brown, "Electric Power Distribution Reliability- Second Edition", CRC Press, Taylor and Francis Group, 2009
- [5] K. Mets, J. A. Ojea, and C. Develder, "Combining power and communication network simulation for cost-effective smart grid analysis," *IEEE Communications Surveys & Tutorials*, vol.16, pp. 1771-1796, 2014.
- [6] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Communications Magazine*, vol. 50, no. 8, pp. 38-45, 2012.
- [7] G. Celli, E. Ghiani, F. Pilo, and G. G. Soma, "Reliability assessment in smart distribution networks," *Electric Power Systems Research*, vol. 104, pp. 164-175, 2013.
- [8] M. Amin, "A smart self-healing grid: In Pursuit of a more reliable and resilient system [In My View]," *IEEE Power and Energy Magazine*, vol. 12, no. 1, pp. 112-110, 2014.
- [9] J. R. Bezerra, G. C. Barroso, R. P. S. Leao, and R. F. Sampaio, "Multiobjective optimization algorithm for switch placement in radial power distribution networks," *IEEE Trans. Power Delivery*, vol. 30, no. 2, pp. 545-552, 2015.
- [10] H. Falaghi, M.-R. Haghifam, and C. Singh, "Ant colony optimization-based method for placement of sectionalizing switches in distribution networks using a fuzzy multiobjective approach," *IEEE Trans. Power Delivery*, vol. 24, no. 1, pp. 268-276, 2009.
- [11] A. Moradi and M. Fotuhi-Firuzabad, "Optimal switch placement in distribution systems using trinary particle swarm optimization algorithm," *IEEE Trans. Power Delivery*, vol. 23, no. 1, pp. 271-279, 2008.
- [12] W. Tippachon and D. Rerkpreedapong, "Multiobjective optimal placement of switches and protective devices in electric power distribution systems using ant colony optimization," *Electric Power Systems Research*, vol. 79, pp. 1171-1178, 2009.
- [13] D.-P. Cong, B. Raison, J.-P. Rognon, S. Bonnoit, and B. Manjal, "Optimization of fault indicators placement with dispersed generation insertion," in *Proc. IEEE Power Engineering Society General Meeting*, pp. 355-362, 2005.
- [14] A. Shahsavari, S. M. Mazhari, A. Fereidunian, and H. Lesani, "Fault indicator deployment in distribution systems considering available control and protection devices: a multi-objective formulation approach," *IEEE Trans. Power Systems*, vol. 29, no. 5, pp. 2359-2369, 2014.
- [15] C.-S. Chen, C.-H. Lin, H.-J. Chuang, C.-S. Li, M.-Y. Huang, and C.-W. Huang, "Optimal placement of line switches for distribution automation systems using immune algorithm," *IEEE Trans. Power Systems*, vol. 21, pp., no. 3, 1209-1217, 2006.
- [16] B. Falahati and Y. Fu, "A study on interdependencies of cyber-power networks in smart grid applications," in *Proc. IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1-8, 2012.
- [17] B. Falahati, Y. Fu, and L. Wu, "Reliability assessment of smart grid considering direct cyber-power interdependencies," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1515-1524, 2012.
- [18] B. Falahati and Y. Fu, "Reliability assessment of smart grids considering indirect cyber-power interdependencies," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1677-1685, 2014.
- [19] H. Lei, C. Singh, and A. Sprintson, "Reliability Modeling and Analysis of IEC 61850 Based Substation Protection Systems," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2194-2202, 2014.
- [20] W. Ling, D. Liu, Y. Lu, P. Du, and F. Pan, "IEC 61850 Model Expansion Toward Distributed Fault Localization, Isolation, and Supply Restoration," *IEEE Trans. Power Delivery*, vol. 29, pp. 977-984, 2014.
- [21] N. Kashyap, C.-W. Yang, S. Sierla, and P. G. Flikkema, "Automated Fault Location and Isolation in Distribution Grids With Distributed Control and Unreliable Communication," *IEEE Trans. Industrial Electronics*, vol. 62, no. 4, pp. 2612-2619, 2015.
- [22] T. Chaudonneret, H. Decroix, and J. McDonald, "Representation of the influence of telecommunications on electrical distribution network reliability," in *Proc. IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pp. 258-263, 2012.
- [23] Y. Zhang, L. Wang, Y. Xiang, and C. W. Ten, "Power system reliability evaluation with SCADA cybersecurity considerations," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1707-1721, 2015.
- [24] M. A. McQueen, W. F. Boyer, M. A. Flynn, and G. A. Beitel, "Time-to-compromise model for cyber risk reduction estimation," in *Proc. 1st Workshop Qual. Prot.*, Milan, Italy, pp. 49–64, 2005.
- [25] W. Nzoukou, L. Wang, S. Jajodia, and A. Singhal, "A unified framework for measuring a network's mean time-to-compromise," in *Proc. 32nd Int. Symp. Reliable Distrib. Syst.*, pp. 215–224, 2013.
- [26] T. Sommestad, M. Ekstedt, and L. Nordström, "Modeling security of power communication systems using defense graphs and influence diagrams," *IEEE Trans. Power Del.*, vol. 24, no. 4, pp. 1801–1808, 2009.
- [27] J. Stamp, A. McIntyre, and B. Ricardson, "Reliability impacts from cyber attack on electric power systems," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, pp. 1–8, 2009.
- [28] Y. Zhang, Y. Xiang, and L. Wang, "Power system reliability assessment incorporating cyber attacks against wind farm energy management systems," *IEEE Trans. Smart Grid*, Early Access, pp. 1–15, 2016.
- [29] Y. Zhang, L. Wang, Y. Xiang, and C. W. Ten, "Inclusion of SCADA cyber vulnerability in power system reliability assessment considering optimal resources allocation," *IEEE Trans. Power Systems*, Early Access, pp. 1-16, 2015.

- [30] Y. Zhang, L. Wang, and Y. Xiang, "Power system reliability analysis with intrusion tolerance in SCADA systems," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 669-683, 2016.
- [31] Abdellatif Hamuda, Khaled Zehar, "Improved Algorithm for Radial Distribution Networks Load Flow Solution", *Electrical power and energy systems*. 33(2011) 508-514.
- [32] R. N. Allan, R. Billinton, I. Sjarief, L. Goel, K.S. So, "A Reliability Test System for Educational Purposes- Basic Distribution System Data and Results", *IEEE Transaction on Power Systems*, Vol. 6, No. 2, pp. 813-820, 1991.
- [33] V. Aravinthan, B. Karimi, V. Namboodiri, and W. Jewell, "Wireless communication for smart grid applications at distribution level—Feasibility and requirements," in *Proc. 2011 IEEE Power and Energy Society General Meeting*, pp. 1-8, 2011.
- [34] R. N. Allan, R. Billinton, I. Sjarief, L. Goel, and K. So, "A reliability test system for educational purposes-basic distribution system data and results," *IEEE Trans. Power Systems*, vol. 6, no. 2, pp. 813-820, 1991.
- [35] M. Moeini-Aghtaie, P. Dehghanian, and S. H. Hosseini, "Optimal distributed generation placement in a restructured environment via a multi-objective optimization approach," in *Proc. 2011 16th Conference on Electrical Power Distribution Networks (EPDC)*, pp. 1-6, 2011.
- [36] Kyriakides, E., and Polycarpou, M., *Intelligent monitoring, control, and security of critical infrastructure systems*, Springer, 2015.
- [37] A. Abdrabou, "A wireless communication architecture for smart grid distribution networks," *IEEE Systems journal*, vol. 10, no. 1, pp. 251-261, 2016.
- [38] F.Salvadori, C.S. Gehrke, A. de Oliveira, M. de Campos, and P. S. Sausen, "Smart grid infrastructure using a hybrid network architecture," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1630-1639, 2013.
- [39] L. Paradis, and Q. Han, "Smart grid infrastructure using a hybrid network architecture," *journal of network and system management*, vol. 15, pp. 171-190, 2007.
- [40] A. E. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Reliability and route diversity in wireless networks," *IEEE Trans. Wireless Communications*, vol. 7, no. 12, pp. 4772-4776, 2008.
- [41] E. Hossain, Z. Han, and H. Vincent poor, "Smart grid communications and networking," Cambridge, 2012.
- [42] A. Goldsmith, "Wireless communications," Stanford, 2004.
- [43] R. Billinton and R. N. Allan, "Reliability evaluation of engineering systems," Springer, 1992.
- [44] C. Singh, and R. Billinton, *System reliability modelling and evaluation*, Hutchinson, 1997.
- [45] A. Mosleh, D. M. Rasmuson, and F.M. Marshall, "Guidelines on modeling common-cause failures in probabilistic risk assessment," INEEL/EXT-97-01327, 1998.
- [46] S. Sierla, B. M. O'Hallorn, T. Karhela, N. Papakonstantinou, and I. Y. Tumer, "Common cause failure analysis of cyber-physical systems situated in constructed environments," *Journal of Research in Engineering Design*, Springer, pp. 375-394, 2013.
- [47] R. Billinton and P. Wang, "Distribution system reliability cost/worth analysis using analytical and sequential simulation techniques," *IEEE Trans. Power Systems*, vol. 13, no. 4, pp. 1245-1250, 1998.
- [48] Y. Cao and Q. Wu, "Teaching genetic algorithm using MATLAB," *International Journal of Electrical Engineering Education*, vol. 36, pp. 139-153, 1999.
- [49] D. Montgomery, "Design and analysis of experiments," Wiley, 2013.

Part III

An Efficient Algorithm for Approximating Failure Frequency with Provable Guarantees Using Near- Minimum Cutsets

Anoosheh Heidarzadeh,

Alex Sprintson, and

Chanan Singh

Texas A&M University

For information about this project, contact

Chanan Singh, Regents Professor and Irma Runyon, Chair Professor
Department of Electrical and Computer Engineering
Texas A&M University
3128 TAMU
College Station, Texas 77843-3128
Phone: (979) 845-7589
Email: singh@ece.tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: <http://www.pserc.org>.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
551 E. Tyler Mall
Engineering Research Center #527
Tempe, Arizona 85287-5706
Phone: (480) 965-1643
Fax: (480) 965-0745

Notice Concerning Copyright Material

PSERC members are given permission to copy without fee all or part of this publication for internal use if appropriate attribution is given to this document as the source material. This report is available for downloading from the PSERC website.

© 2016 Texas A&M University. All rights reserved.

Table of Contents

1	Notation and Nomenclature.....	i
2	Introduction.....	1
3	System Model and Problem Setup.....	2
4	Preliminaries.....	3
4.1	Cutsets and Near-Minimum Cutsets.....	3
4.2	Inclusion-Exclusion Formulas.....	4
4.3	Bounding Technique.....	5
5	Proposed Approximation Algorithm.....	6
5.1	Motivation.....	6
5.2	Main Idea.....	7
5.3	Approximation of P	8
5.4	Approximation of P_f	12
5.5	Approximation of F_f	13
5.6	Computational Complexity.....	14
6	Conclusion.....	14

List of Figures

1	An example of a system with 16 terminals and 24 components.	2
---	--	---

List of Tables

1	Minimal Cutsets of Size 2 and 3 in the System of Fig. 1	4
2	Probabilities and Frequencies of Failure of the System in Fig. 1	6

1 Notation and Nomenclature

S	a composite system
m	number of components in S
n	number of terminals in S
$[i]$	set of integers $\{1, \dots, i\}$ (for any integer i)
λ_i	failure rate of component i
μ_i	repair rate of component i
p_i	unavailability probability of component i , $p_i = \lambda_i / (\lambda_i + \mu_i)$
w_i	weight of component i , $w_i = -\ln p_i$
\mathbb{C}	set of all minimal cutsets in S
N	number of minimal cutsets in S , $N = \mathbb{C} $
C_i	i th minimal cutset in S
$s(C_i)$	number of components in C_i
i_j	index of the j th component in C_i ($j \in [s(C_i)]$)
\mathcal{C}_i	index set of components in C_i , $\mathcal{C}_i = \{i_j\}_{j=1}^{s(C_i)}$
$w(C_i)$	weight of C_i , $w(C_i) = \sum_{j \in \mathcal{C}_i} w_j$
w^*	minimum cutset weight, $w^* = \min_{i \in [N]} w(C_i)$
p^*	probability of all components of a minimum-weight cutset being unavailable, $p^* = \exp(-w^*)$
$p(\mathcal{C})$	probability of unavailability of all components in a collection \mathcal{C} of cutsets
ε	maximum approximation error factor, an arbitrary $\varepsilon > 0$ (independent of m and n)
δ	maximum approximation error probability, an arbitrary $0 < \delta < 1$ (independent of m and n)
ξ	an arbitrary $\xi > 0$ (not necessarily independent of m and n)
P_f	probability of system failure in steady-state
P_f^+, P_f^-	first-order upper- and lower-bound on P_f
d_P	maximum number of decimal places up to which P_f^+ and P_f^- match
\hat{P}_f	value of P_f^+ or P_f^- truncated to d_P decimal places
\tilde{P}_f	an (ε, δ) -approximation of P_f
F_f	frequency of system failure in steady-state
F_f^+, F_f^-	first-order upper- and lower-bound on F_f

d_F	maximum number of decimal places up to which F_f^+ and F_f^- match
\hat{F}_f	value of F_f^+ or F_f^- truncated to d_F decimal places
\tilde{F}_f	an (ε, δ) -approximation of F_f
P	probability of all components of some cutset being unavailable and unexposed
\tilde{P}	an (ε, δ) -approximation of P
p_{\min}	minimum unavailability probability, $p_{\min} = \min_{i \in [m]} p_i$
w_{\max}	maximum component weight, $w_{\max} = -\ln p_{\min}$
$\mathbb{C}^{(\alpha)}$	set of all α -min cutsets for a constant $\alpha \geq 1$
M	number of all α -min cutsets, $M = \mathbb{C}^{(\alpha)} $
s_{α}^*	minimum α -min cutset size, $s_{\alpha}^* = \min_{C \in \mathbb{C}^{(\alpha)}} s(C)$
$P^{(\alpha)}$	probability of all components of some α -min cutset being unavailable and unexposed
$\tilde{P}^{(\alpha)}$	a $(\xi/2, \delta/2)$ -approximation of $P^{(\alpha)}$
$P_f^{(\alpha)}$	probability of all components of some α -min cutset being unavailable
$\tilde{P}_f^{(\alpha)}$	a $(\xi/2, \delta/2)$ -approximation of $P_f^{(\alpha)}$
α -min cutset	any minimal cutset of weight no larger than α times the minimum cutset weight
(ε, δ) -approx.	any multiplicative approximation with error factor at most ε and error probability at most δ
$\text{trunc}(x, d)$	value of x truncated to d decimal places (for any real number $x \geq 0$ and integer $d \geq 1$), $\text{trunc}(x, d) = \lfloor 10^d \cdot x \rfloor / 10^d$
$\mathbb{E}(\cdot)$	expected value of a random variable
$\text{var}(\cdot)$	variance of a random variable
$\text{median}(\cdot)$	median of an array
$f(n), g(n)$	arbitrary functions of n
$f(n) = O(g(n))$	for sufficiently large n , there exists $k > 0$ (independent of n) such that $ f(n) \leq k \cdot g(n) $
$f(n) = \Omega(g(n))$	for sufficiently large n , there exists $k > 0$ (independent of n) such that $f(n) \geq k \cdot g(n)$
$f(n) = \tilde{O}(g(n))$	$f(n) = O(g(n) \log^k g(n))$, for some $k > 0$ (independent of n)

2 Introduction

Consider a composite system whose terminals are connected through components that are subject to statistically independent stationary failure and repair processes over time. At any given time, each component is either operational or not, and the system fails if the surviving system of operational components does not connect all terminals. Probability and frequency of failure are two important measures of reliability of such systems [1–3]. These quantities are very useful to derive other reliability measures such as mean down-time and mean cycle-time of the system.

Numerous algorithms were previously designed for computing the failure probability [2, 4–16] and the failure frequency [2, 6, 10, 17–20], in transient or steady-state regime. The computations however become intractable in large-scale systems because the computational complexity grows very quickly with the size of the system. Specifically, the exact computation of these quantities was shown to be NP-hard [21, 22]. To overcome this challenge, various methods were developed to approximate the probability and frequency of failure. The existing techniques for approximating the failure probability are based on the Monte Carlo and rare-events simulations [23–27] or enumerating only a subset of the system states (instead of the set of all system states) and computing the sum of their probability of occurrence [28–31]. Similar methods were used to approximate the failure frequency in [32] and [33]. The only existing technique which can provably approximate the failure probability within an arbitrary multiplicative error and runs in polynomial time was proposed in [34]. To the best of our knowledge, however, no such computationally efficient algorithm with provable guarantees was previously proposed for approximating the failure frequency.

The algorithm of [34] estimates the probability of failure of a system in the steady state where the failure and repair processes of each component are stationary. This is equivalent to say that in the steady state, each component fails or it survives at any time instant, independently from other elements, with some constant probability (independent of time). The probability of failure of the system is then equal to the probability that some cutset in the system (i.e., some collection of components whose failure results in a loss of connectivity of some terminals from the rest) fails.

The main ideas behind the algorithm of [34] can be summarized as follows: (i) the number of weak cutsets in a system (i.e., those cutsets with higher probability of failure) is polynomial in the number of terminals and the number of components, and the enumeration of all such cutsets can be done in polynomial time; (ii) the probability of the union of failures of weak cutsets provides a multiplicative approximation of the failure probability of the system; and (iii) this probability can be written as the truth probability of a disjunctive normal form (DNF) formula, and a multiplicative approximation of this probability can be computed in polynomial time.

The frequency of failure, however, cannot be directly linked to the truth probability of a DNF formula. This implies the need for a novel approximation technique. In this work, we provide a polynomial-time algorithm using near-minimum cutsets to approximate the failure frequency with high probability within an arbitrary multiplicative error. Moreover, our numerical results show that

of the system in the steady-state regime, defined as follows. At any given time, the (sub-) system of S including all n terminals, restricted only to the set of available components (i.e., the original system excluding unavailable components), is referred to as the *surviving system*. At any given time, the system is said to be *unavailable* if the surviving system fails to connect all n terminals. The (steady-state) *probability of failure* is the probability that the system is unavailable, and the (steady-state) *frequency of failure* is the expected number of times per unit time (i.e., the expected rate) that the system becomes unavailable [2]. For arbitrary $\varepsilon > 0$ and $0 < \delta < 1$ (independent of m and n), the problem is to compute (ε, δ) -approximations of P_f and F_f , denoted by \tilde{P}_f and \tilde{F}_f , respectively, defined as

$$\Pr \{ |\tilde{P}_f - P_f| \geq \varepsilon P_f \} \leq \delta,$$

and

$$\Pr \{ |\tilde{F}_f - F_f| \geq \varepsilon F_f \} \leq \delta.$$

Without loss of generality, we assume that the repair rates $\{\mu_i\}$ are all equal to μ . No assumption is however made on the failure rates $\{\lambda_i\}$. We notice that our results are generalizable to the cases with unequal repair rates due to the following lemma.

Lemma 1 *Any system with unequal repair rates can be transformed to an equivalent system (with the same probability of failure and frequency of failure) with equal repair rates by replacing each component i with $n_i = \mu_i/\mu$ independent components in parallel (for sufficiently small $\mu > 0$), each with repair rate μ and failure rate $\mu p/(1 - p)$, where $p = p_i^{1/n_i}$.*

Proof: The proof follows easily from the following facts: (i) the unavailability probability of a component with failure and repair rates λ and μ is equal to $\lambda/(\lambda + \mu)$; (ii) the unavailability probability of a set of independent components in parallel is equal to the product of their unavailability probabilities, and (iii) the equivalent repair rate of a set of independent components in parallel is equal to the summation of their repair rates. \square

4 Preliminaries

4.1 Cutsets and Near-Minimum Cutsets

Any collection of components which, if all unavailable, results in the unavailability of the system is referred to as a *cutset*. A cutset is *minimal* if it does not contain any other cutsets. Hereafter, we often use the term “cutset” as a shorthand for “minimal cutset.” Let $\mathbb{C} = \{C_1, \dots, C_N\}$ be the set of all (minimal) cutsets in S . Let $\mathcal{C}_i = \{i_1, \dots, i_{s(C_i)}\}$ denote the (index) set of components in C_i , where $s(C_i)$, the *size* of C_i , is the number of components in C_i (i.e., $s(C_i) = |\mathcal{C}_i|$). For example, in

Table 1: Minimal Cutsets of Size 2 and 3 in the System of Fig. 1

cutsets of Size 2	cutsets of Size 3	
{1, 4}	{1, 2, 5}	{11, 15, 18}
{3, 7}	{1, 8, 11}	{11, 15, 22}
{18, 22}	{2, 3, 6}	{14, 17, 21}
{21, 24}	{2, 4, 5}	{14, 17, 24}
	{2, 6, 7}	{18, 19, 23}
	{3, 10, 14}	{19, 22, 23}
	{4, 8, 11}	{20, 21, 23}
	{7, 10, 14}	{20, 23, 24}

the system of Fig 1, there exist 4, 16, 30, 64, 120, 112, 137, 96, 48 cutsets of size 2, 3, ..., 10, respectively. For example, the cutsets of size 2 and 3 are enumerated in Table 1.

We define the *weight of cutset* C_i , denoted by $w(C_i)$, as the sum of the weights of all components in C_i , i.e., $w(C_i) = \sum_{j \in C_i} w_j$. Let $w^* = \min_{C \in \mathbb{C}} w(C)$. For any constant $\alpha \geq 1$, let $\mathbb{C}^{(\alpha)}$ be the set of all (minimal) cutsets in S of weight less than or equal to αw^* , i.e., $\mathbb{C}^{(\alpha)} = \{C \in \mathbb{C} : w(C) \leq \alpha w^*\}$. Let $s_\alpha^* = \min_{C \in \mathbb{C}^{(\alpha)}} s(C)$. It is easy to see that

$$s_\alpha^* \geq \frac{w^*}{w_{\max}} \alpha \geq \frac{w^*}{w_{\max}},$$

where $w_{\max} = -\ln p_{\min}$ and $p_{\min} = \min_{i \in [m]} p_i$. We refer to the cutsets in $\mathbb{C}^{(\alpha)}$ as α -min cutsets. For example, Table 1 enumerates the 1.5-min cutsets of the system in Fig. 1. For simplicity, we refer to 1-min cutsets as *min-cutsets*. By definition, for every min-cutset C , $w(C) = w^*$.

4.2 Inclusion-Exclusion Formulas

For any arbitrary $\mathcal{J} \subseteq [N]$, the (joint) probability of failure of cutsets $\{C_i\}_{i \in \mathcal{J}}$, denoted by $p(\cap_{i \in \mathcal{J}} C_i)$, is equal to the probability that all components in C_i , for all $i \in \mathcal{J}$, are unavailable, i.e., $p(\cap_{i \in \mathcal{J}} C_i) = \prod_{j \in \cup_{i \in \mathcal{J}} C_i} p_j$. Then, by the cutset approach [2], the probability of failure (P_f) and the frequency of failure (F_f), based on the inclusion-exclusion principle, can be written as:

$$\begin{aligned}
P_f &= p(C_1) + \cdots + p(C_N) \\
&\quad - p(C_1 \cap C_2) - \cdots - p(C_{N-1} \cap C_N) \\
&\quad \cdots \\
&\quad (-1)^{2N-1} p(C_1 \cap \cdots \cap C_N),
\end{aligned} \tag{1}$$

and

$$\begin{aligned}
F_f = & p(C_1)|\mathcal{C}_1|\mu + \dots + p(C_N)|\mathcal{C}_N|\mu \\
& - p(C_1 \cap C_2)|\mathcal{C}_1 \cup \mathcal{C}_2|\mu - \dots \\
& - p(C_{N-1} \cap C_N)|\mathcal{C}_{N-1} \cup \mathcal{C}_N|\mu \\
& \dots \\
& (-1)^{2N-1} p(C_1 \cap \dots \cap C_N)|\mathcal{C}_1 \cup \dots \cup \mathcal{C}_N|\mu.
\end{aligned} \tag{2}$$

4.3 Bounding Technique

The number of minimal cutsets (N) is generally exponential in the number of terminals (n), and hence the number of terms in equations (1) and (2) is doubly-exponential in n . Thus, there is no polynomial-time algorithm in order to compute P_f and F_f from (1) and (2), directly. The bounding technique, described bellow, is one of the most common approaches to provide upper and lower bounds on each of these quantities via truncating the corresponding inclusion-exclusion formula.

Let

$$P_f^+ = \sum_{i \in [N]} p(C_i)$$

and

$$P_f^- = \sum_{i \in [N]} p(C_i) - \sum_{1 \leq i < j \leq N} p(C_i \cap C_j).$$

Similarly, let

$$F_f^+ = \sum_{i \in [N]} p(C_i)|\mathcal{C}_i|\mu$$

and

$$F_f^- = \sum_{i \in [N]} p(C_i)|\mathcal{C}_i|\mu - \sum_{1 \leq i < j \leq N} p(C_i \cap C_j)|\mathcal{C}_i \cup \mathcal{C}_j|\mu.$$

Then, P_f^+ and P_f^- (or F_f^+ and F_f^-) are the *first-order upper-bound* and *lower-bound* on P_f (or F_f), respectively. Note that none of these bounds is computable in polynomial time (because N is exponential in n). Let d_P (or d_F) be the maximum number of decimal places up to which P_f^+ and P_f^- (or F_f^+ and F_f^-) match. Let $\hat{P}_f = \text{trunc}(P_f^+, d_P) = \text{trunc}(P_f^-, d_P)$ and $\hat{F}_f = \text{trunc}(F_f^+, d_F) = \text{trunc}(F_f^-, d_F)$, where $\text{trunc}(x, d) = \lfloor 10^d \cdot x \rfloor / 10^d$, for any real number $x \geq 0$ and integer $d \geq 1$. Then, \hat{P}_f and \hat{F}_f are the best estimators of P_f and F_f based on the (first-order) upper and lower bounds.

For example, \hat{P}_f and \hat{F}_f for the system in Fig. 1 for various system parameters are listed in Table 2. (The exact values of P_f and F_f , however, are not computable efficiently and hence not

Table 2: Probabilities and Frequencies of Failure of the System in Fig. 1

$\varepsilon = 10^{-6}$ and $\delta = 10^{-2}$						
Failure rate (λ)	Repair rate (μ)	Component unavailability (p)	Failure Probability (P_f)		Failure Frequency (F_f)	
			\hat{P}_f	\tilde{P}_f	\hat{F}_f	\tilde{F}_f
0.5	10^3	0.49975×10^{-3}	1.00099×10^{-6}	1×10^{-6}	0.200×10^{-2}	0.2003×10^{-2}
1	10^3	0.99900×10^{-3}	0.40079×10^{-5}	4×10^{-6}	0.803×10^{-2}	0.8031×10^{-2}
1.5	10^3	1.49775×10^{-3}	0.09026×10^{-4}	9×10^{-6}	1.810×10^{-2}	1.8107×10^{-2}
2	10^3	1.99600×10^{-3}	0.16063×10^{-4}	16×10^{-6}	3.225×10^{-2}	3.2254×10^{-2}
\tilde{P}_f and \tilde{F}_f are truncated at 6 decimal places						

presented.) For these results, all components are assumed to be identical with failure rate λ and repair rate μ (and hence the unavailability probability $p = \lambda / (\lambda + \mu)$), and four cases with $\lambda = 0.5, 1, 1.5$, and 2 , and $\mu = 10^3$ are considered. Note that $\lambda \ll \mu$ in the cases of our interest because in such cases the failure of the system is a rare event and approximating P_f and F_f is computationally more expensive.

5 Proposed Approximation Algorithm

5.1 Motivation

Interestingly, it was previously shown in [35] that the number of *min-cutsets* is only polynomial, and even further, there are only a polynomial number of *near-minimum cutsets* whose weight is not larger than any given constant factor of the weight of the min-cutsets. Such cutsets can all be enumerated in polynomial time (see Lemma 2). However, even for a polynomial number of cutsets, there are an exponential number of terms in (1) and (2), and consequently, using (1) and (2) it is only possible to provide a series of upper and lower bounds on P_f and F_f via restricting the computations up to some odd-order terms or even-order terms, respectively [32]. This, however, fails to give (ε, δ) -approximations of P_f and F_f , for arbitrary ε and δ . Note that using the crude Monte Carlo technique, one requires exponentially many runs of simulation to obtain (ε, δ) -approximations of P_f and F_f [34].

In [34], Karger proposed a polynomial-time algorithm to give an (ε, δ) -approximation of P_f using near-minimum cutsets. This algorithm, discussed in detail shortly in Section 5.4 as part of the proposed algorithm in the present work, exploits an unbiased estimator, referred to as KLM,

due to Karp, Luby and Madras [36]. The KLM estimator is useful for approximating (with high probability) the truth probability of any disjunctive normal form (DNF) formula (for definition, see Section 5.3), within an arbitrary multiplicative error. As can be seen in (1), P_f is the probability of the union of a set of events, and consequently, it can be thought of as the probability of satisfying some DNF formula with random Boolean variables. However, F_f , as one can see in (2), is not the probability of the union of any set of events. In the sequel, we propose a polynomial-time algorithm to give an (ε, δ) -approximation of F_f .

5.2 Main Idea

It is easy to see that F_f cannot be directly written as the probability of the union of a set of events. To overcome this challenge, we define a set of events Ω such that F_f can be written as a linear combination of P_f and the probability P of union of the events in Ω . Then, in order to approximate F_f , we need to approximate P_f and P . Each of these quantities, P_f and P , is the probability of union of a set of events, and thus can be approximated by the KLM estimator. We formalize this idea in the following.

Let

$$\begin{aligned} P = & p(C_1) \left(1 - \frac{|\mathcal{C}_1|}{m}\right) + \dots + p(C_N) \left(1 - \frac{|\mathcal{C}_N|}{m}\right) \\ & - p(C_1 \cap C_2) \left(1 - \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{m}\right) - \dots \\ & - p(C_{N-1} \cap C_N) \left(1 - \frac{|\mathcal{C}_{N-1} \cup \mathcal{C}_N|}{m}\right) \\ & \dots \\ & (-1)^{2N-1} p(C_1 \cap \dots \cap C_N) \left(1 - \frac{|\mathcal{C}_1 \cup \dots \cup \mathcal{C}_N|}{m}\right). \end{aligned}$$

It is easy to see that

$$F_f = (P_f - P)m\mu,$$

where F_f and P_f are given by (1) and (2), respectively. We now define the set of events Ω such that P , defined as above, is the probability of union of the events in Ω .

In the steady state, at any given time, each component i is unavailable or available, with probability p_i or $1 - p_i$, respectively, independent of time [2]. Thus, the random process under consideration (Section 3) is equivalent to the following *one-shot* random process over S . Each component i , statistically independently from other components, is set to be unavailable with probability p_i , or it is set to be available with probability $1 - p_i$. We refer to this process as *sampling*.

We further introduce an auxiliary *one-shot* random process over S as follows. Each component is assumed to have two states: *exposed* and *unexposed*. One component is chosen uniformly at

random and is set to be exposed, and the rest of the components are set to be unexposed. We refer to this process as *exposure*. The sampling and exposure processes are assumed to be statistically independent.

The intuition behind the sampling and exposure processes is as follows. Each term in P is the joint probability that all components in a collection of cutsets are unavailable and unexposed. Moreover, P is expressed by an inclusion-exclusion formula. Thus, it should not be hard to see that P is equal to the probability that all components of some cutset of S (under the sampling and exposure processes) are unavailable and unexposed (due to the statistical independence of the underlying processes).

We, first, propose an algorithm for approximating P (Section 5.3), and next, we slightly modify the proposed algorithm in order to approximate P_f (Section 5.4). Finally, we use the approximations of P_f and P , and give an approximation of F_f (Section 5.5).

5.3 Approximation of P

For every min-cutset C , $p(C) = \exp(-w^*)$. Let $p^* = \exp(-w^*)$. It is obvious that the probability that all components of some cutset are unavailable and unexposed, P , is bounded from below by the probability that all components of a given min-cutset are unavailable and unexposed, $p^*(1 - s_1^*/m)$, i.e.,

$$P \geq p^* \left(1 - \frac{s_1^*}{m}\right).$$

If $p^* > n^{-4}$, then $P = \Omega(n^{-6})$ since $m = O(n^2)$ and $s_1^* = O(m)$. Thus, in this case, we resort to the Monte Carlo simulation. Specifically, we simulate the system S under the sampling and exposure processes (simultaneously) $O(n^6 \log(2/\delta)/\xi^2)$ times, for any given $\xi > 0$ and $0 < \delta < 1$, and subsequently, compute $\tilde{P}^{(\cdot)}$, the fraction of times all components of some cutset are unavailable and unexposed. (We use the superscript “ \cdot ” to highlight the fact that the approximation based on the Monte Carlo simulation is not restricted to any specific subset of cutsets.) Applying the Chernoff bound, we get

$$\Pr\left\{|\tilde{P}^{(\cdot)} - P| \geq \xi P\right\} \leq \frac{\delta}{2},$$

and so $\tilde{P}^{(\cdot)}$ gives a $(\xi, \delta/2)$ -approximation of P .

We notice that in each simulation run, one needs to check whether there exists a cutset whose components are all unavailable and unexposed. Naively checking for such an event, however, cannot be performed in polynomial time if the exposed component is also unavailable. (This comes from the fact that there might exist exponentially many cutsets including the exposed component.) We can alternatively consider the exposed component to be available in each simulation run, regardless of its true state (available or unavailable), and check the connectivity of the surviving system. This approach resolves the issue of high computational complexity of the native approach because

the connectivity can be tested in polynomial time, e.g., using the breadth first search or the depth first search in $O(m+n)$ time. Thus, $\tilde{P}^{(\cdot)}$ is computable in polynomial time so long as $\xi = \Omega(n^{-k})$, for some constant $k > 0$ (independent of n), and δ is a constant. (As will be shown in Section 5.5, $\xi = \Omega(n^{-2})$ for the purpose of this work.)

Now, assume $p^* \leq n^{-4}$. Let $\gamma = w^*/\ln n - 2$. Since $p^* = n^{-2-\gamma} \leq n^{-4}$, obviously $\gamma \geq 2$. In this case, P might be very small and so it cannot be approximated in polynomial time using the Monte Carlo simulation. The procedure of approximating P is as follows.

Fix a particular constant $\alpha \geq 1$. Let $\{C_1, \dots, C_M\}$ be the set of all α -min cutsets in S , i.e., $\mathbb{C}^{(\alpha)} = \{C_1, \dots, C_M\}$. (We will shortly specify a proper choice of α .) The recursive contraction algorithm for unweighted graphs due to Karger and Stein [35] can be generalized to the case of weighted graphs to show the following result by using the same proof technique as in [34].

Lemma 2 *The number of α -min cutsets (M) is bounded from above by $n^{2\alpha}$, and all such cutsets can be enumerated with high probability in $\tilde{O}(n^{2\alpha})$ time.*

Proof: [Sketch] The proof of the first part of the result consists of two steps. The first step is to use the contraction algorithm to find a min-cutset, and consequently, compute w^* . This algorithm runs in $O(n^2)$ time, and finds a min-cutset with probability $\Omega(n^{-2})$. Thus $O(n^2)$ runs of the contraction algorithm are sufficient to find a min-cutset with high probability. By a clever recursive implementation of the contraction algorithm, as shown in [35, Lemma 4.1] and [35, Lemma 4.3], a min-cutset can be found in $O(n^2 \log^2 n)$ time (instead of $O(n^4)$ time of the obvious implementation) with high probability. Furthermore, it is not hard to see that the number of min-cutsets is at most n^2 . The second step is to generalize this result by using a similar technique (via slightly modifying the contraction algorithm to output α -min cutsets, instead of min-cutsets), and show that the number of α -min cutsets is at most $n^{2\alpha}$ [35, Theorem 8.4]. The proof of the second part of the result follows from a coupon-collector paradigm: if there are k bins, and potentially an infinite number of balls are thrown independently and uniformly one at a time, $O(k \log k)$ balls suffice with high probability to have every bin contain at least one ball. Thinking of α -min cutsets as bins and the number of runs of the contraction algorithm as balls, then it should be obvious that one can enumerate all α -min cutsets with high probability in $O(n^{2\alpha} \log^2 n)$ time (by running the recursive contraction algorithm $O(n^{2\alpha} \log n)$ times) [35, Theorem 8.5]. \square

By the result of Lemma 2 together with the application of union bound, it can be shown that the probability that all components of some cutset of weight αw^* are unavailable and unexposed is bounded from above by $n^{-\alpha\gamma}(1 - s_\alpha^*/m)$, and subsequently, $n^{-\alpha\gamma}(1 - w^*/(w_{\max}m))$. This result is generalizable for all cutsets of weight greater than αw^* as follows.

Lemma 3 *The probability that all components of some cutset of weight greater than αw^* are unavailable and unexposed is bounded from above by*

$$n^{-\alpha\gamma} \left(1 - \frac{w^*}{w_{\max}m}\right) \left(1 + \frac{2}{\gamma}\right).$$

Proof: The proof follows the same line as [34, Theorem 2.9], and hence omitted. \square

By the result of Lemma 3, for any arbitrary $\xi > 0$, it is easy to show that

$$n^{-\alpha\gamma} \left(1 - \frac{w^*}{w_{\max}m}\right) \left(1 + \frac{2}{\gamma}\right) \leq \frac{\xi}{2}P$$

so long as

$$\alpha \geq \frac{1}{\gamma} \left(4 - \frac{\ln \left(\frac{\xi}{2} \left(\frac{\gamma}{\gamma+2} \right) \left(\frac{w_{\max}m}{w_{\max}m-w^*} \right) \right)}{\ln n} \right).$$

(We will specify our choice of ξ , depending on m and n , in Section 5.5.) Since $\gamma \geq 2$, it is not hard to see that

$$\frac{1}{\gamma} \left(4 - \frac{\ln \left(\frac{\xi}{2} \left(\frac{\gamma}{\gamma+2} \right) \left(\frac{w_{\max}m}{w_{\max}m-w^*} \right) \right)}{\ln n} \right) \leq 2 - \frac{\ln \frac{\xi}{4}}{2 \ln n}.$$

Taking

$$\alpha = \frac{1}{\gamma} \left(4 - \frac{\ln \left(\frac{\xi}{2} \left(\frac{\gamma}{\gamma+2} \right) \left(\frac{w_{\max}m}{w_{\max}m-w^*} \right) \right)}{\ln n} \right), \quad (3)$$

it follows that α is bounded from above by some constant (independent of m and n), particularly for ξ polynomially small in n . (Such a choice of ξ is of particular interest in this work as will be shown in Section 5.5.)

Let $P^{(\alpha)}$ be the probability that all components of some α -min cutset are unavailable and unexposed. Then,

$$\left(1 - \frac{\xi}{2}\right)P \leq P^{(\alpha)} \leq P. \quad (4)$$

Thus, $P^{(\alpha)}$ is a $(\xi/2, 0)$ -approximation of P . Now, the main idea is to enumerate all the α -min cutsets in S , and compute $P^{(\alpha)}$. Such cutsets can be all enumerated in polynomial time (Lemma 2). However, one cannot compute $P^{(\alpha)}$ in polynomial time using the inclusion-exclusion formula due to the exponential number of terms. Instead, we give a $(\xi/2, \delta/2)$ -approximation of $P^{(\alpha)}$, denoted by $\tilde{P}^{(\alpha)}$, i.e.,

$$\Pr \left\{ |\tilde{P}^{(\alpha)} - P^{(\alpha)}| \geq \frac{\xi}{2}P^{(\alpha)} \right\} \leq \frac{\delta}{2}, \quad (5)$$

for any given $0 < \delta < 1$. As a consequence, $\tilde{P}^{(\alpha)}$ gives a $(\xi, \delta/2)$ -approximation of P .

Before moving forward with the approximation algorithm, let us recall some definitions from Boolean algebra. Let $\Phi = Z_1 \vee Z_2 \vee \dots \vee Z_M$ be a formula on M Boolean variables $\{Z_i\}_{i=1}^M$, where the *clause* Z_i is a conjunction of some *literals* $\{z_{ij}\}$, i.e., $Z_i = \wedge_j z_{ij}$. (The symbols \vee and \wedge are

logical conjunction (AND) and logical disjunction (OR), respectively.) Each literal z_{ij} is either a Boolean variable or the negation of a Boolean variable, and it takes two values: “true” and “false.” The formula Φ of such form is said to have *disjunctive normal form* (DNF).

The approximation algorithm proceeds as follows. Let $\Phi = \vee_{i \in [M]} Z_i$, where Z_i is the conjunction of two *sub-clauses* X_i and Y_i (i.e., $Z_i = X_i \wedge Y_i$). Let $X_i = x_{i_1} \wedge x_{i_2} \wedge \dots \wedge x_{i_{|\mathcal{C}_i|}}$ and $Y_i = y_{i_1} \wedge y_{i_2} \wedge \dots \wedge y_{i_{|\mathcal{C}_i|}}$, where $i_1, \dots, i_{|\mathcal{C}_i|}$ are the labels of the components in the cutset C_i , and the literals x and y are Boolean (random) variables defined as follows. For every (random) *assignment* of the literals x_j and y_j , for every $j \in [m]$, we assume: x_j is true with probability p_j , and it is false with probability $1 - p_j$; and y_j is true for one and only one j chosen uniformly at random from the set $[m]$, and it is false for every other j . Thus, it follows that (i) a random assignment $x = \{x_1, \dots, x_m\}$ satisfies X_i (i.e., X_i is true) with probability $P_X(i) = p(C_i) = \prod_{j \in \mathcal{C}_i} p_j$, and (ii) a random assignment $y = \{y_1, \dots, y_m\}$ satisfies Y_i (i.e., Y_i is true) with probability $P_Y(i) = (1 - |\mathcal{C}_i|/m)$. By (i) and (ii), a random assignment $z = (x, y)$ satisfies Z_i (i.e., Z_i is true) with probability

$$P_Z(i) = p(C_i) \left(1 - \frac{|\mathcal{C}_i|}{m}\right),$$

since X_i and Y_i are statistically independent. The formula Φ is true so long as some clause Z_i is true, and thus by the inclusion-exclusion principle, it immediately follows that the truth probability of Φ is equal to $P^{(\alpha)}$. Thus, it suffices to approximate the truth probability of Φ .

By the definition, it is obvious that Φ , constructed as above, is a DNF formula, and as a consequence, the KLM estimator, proposed in [36], is applicable to approximate the truth probability of Φ . The KLM estimator with inputs $(\Phi, P_Z; \xi, \delta)$ proceeds in steps as follows:

0. Initialization: $t = 1$ and $l = 1$.
1. Choose a random clause Z_i , with probability of selecting Z_i being equal to $P_Z(i)/Q_Z$, where $Q_Z = \sum_i P_Z(i)$.
2. Choose a random assignment z satisfying clause Z_i .
3. Compute $P_t = Q_Z/N(z)$, where $N(z)$ is the number of clauses Z that z satisfies;
4. $t \leftarrow t + 1$
5. Repeat steps 1-4 $T = 16M/\xi^2$ times, and output the mean $\tilde{P}_l = (\sum_{t=1}^T P_t)/T$.
6. $l \leftarrow l + 1$
7. Repeat steps 1-6 $T^* = \log(2/\delta)$ times, and output the median of the means $\{\tilde{P}_l\}_{l=1}^{T^*}$.

Let $\tilde{P}^{(\alpha)} = \text{median}(\{\tilde{P}_l\}_l)$. Then, the following result holds.

Lemma 4 $\tilde{P}^{(\alpha)}$ is a $(\xi/2, \delta/2)$ -approximation of $P^{(\alpha)}$.

Proof: By a simple chain-rule analysis as in [36], it follows that $\mathbb{E}(P_t) = P^{(\alpha)}$ and $\text{var}(P_t) \leq M(P^{(\alpha)})^2$. Applying the Chebychev's inequality, we get

$$\Pr \left\{ |\tilde{P}_l - P^{(\alpha)}| \geq \frac{\xi}{2} P^{(\alpha)} \right\} \leq \frac{1}{4},$$

and consequently,

$$\Pr \left\{ |\tilde{P}^{(\alpha)} - P^{(\alpha)}| \geq \frac{\xi}{2} P^{(\alpha)} \right\} \leq \frac{\delta}{2},$$

where $\tilde{P}^{(\alpha)} = \text{median}(\{\tilde{P}_l\}_l)$. □

5.4 Approximation of P_f

If $p^* > n^{-4}$, the Monte Carlo simulation can be used to compute $\tilde{P}_f^{(\cdot)}$, a $(\xi, \delta/2)$ -approximation of P_f (similarly as before for computing $\tilde{P}^{(\cdot)}$ in Section 5.3).

Now, assume $p^* \leq n^{-4}$. Consider the system S under the sampling process, yet not the exposure process. Let $P_f^{(\alpha)}$ be the probability that all components of some α -min cutset are unavailable. Similarly as before, it can be shown that

$$\left(1 - \frac{\xi}{2}\right) P_f \leq P_f^{(\alpha)} \leq P_f, \quad (6)$$

for some $\xi > 0$ (the proper choice of ξ is given shortly), and thus $P_f^{(\alpha)}$ is a $(\xi/2, 0)$ -approximation of P_f . Similar to the probability $P^{(\alpha)}$, the exact computation of the probability $P_f^{(\alpha)}$ cannot be performed in polynomial time, and instead we aim at approximating it.

The method of approximating $P_f^{(\alpha)}$ is similar to that of $P^{(\alpha)}$, except that the formula Φ needs to be modified slightly. This technique was previously used in [34]. We define the DNF formula $\Phi_f = Z_1 \vee Z_2 \vee \dots \vee Z_M$, where $Z_i = X_i$, and X_i is defined as before (Section 5.3). The formula Φ_f is true so long as X_i is true for some i . The probability that Φ_f is true is equal to $P_f^{(\alpha)}$. Let $\tilde{P}_f^{(\alpha)}$ be the output of the KLM estimator with inputs $(\Phi_f, P_Z; \xi, \delta)$. Then, the following result is immediate.

Lemma 5 $\tilde{P}_f^{(\alpha)}$ is a $(\xi/2, \delta/2)$ -approximation of $P_f^{(\alpha)}$.

Proof: The proof follows the same line as in the proof of Lemma 4, and hence omitted. □

5.5 Approximation of F_f

By Lemma 5, it immediately follows that

$$\Pr \left\{ |\tilde{P}_f^{(\alpha)} - P_f^{(\alpha)}| \geq \frac{\xi}{2} P_f^{(\alpha)} \right\} \leq \frac{\delta}{2}. \quad (7)$$

For any given $\varepsilon > 0$, choose

$$\xi = \frac{\varepsilon}{1 + 2 \left(\frac{w_{\max} m}{w^*} \right)} \quad (8)$$

in the Monte Carlo simulation or the KLM estimator, and compute $\tilde{P}^{(\cdot)}$ and $\tilde{P}_f^{(\cdot)}$ or $\tilde{P}^{(\alpha)}$ and $\tilde{P}_f^{(\alpha)}$, respectively. Then, the following result holds.

Theorem 1 *If $p^* > n^{-4}$ or $p^* \leq n^{-4}$, then $\tilde{F}_f = (\tilde{P}_f^{(\cdot)} - \tilde{P}^{(\cdot)})m\mu$ or $\tilde{F}_f = (\tilde{P}_f^{(\alpha)} - \tilde{P}^{(\alpha)})m\mu$ gives an (ε, δ) -approximation of F_f , respectively.*

Proof: We only give the proof for the case of $p^* \leq n^{-4}$, and the proof of the other case follows the exact same lines (and hence omitted).

By combining (4) and (5), we get

$$\Pr \left\{ |\tilde{P}^{(\alpha)} - P| \geq \xi P \right\} \leq \frac{\delta}{2}, \quad (9)$$

and combining (6) and (7), we get

$$\Pr \left\{ |\tilde{P}_f^{(\alpha)} - P_f| \geq \xi P_f \right\} \leq \frac{\delta}{2}. \quad (10)$$

By putting together (9) and (10), one can see

$$\Pr \left\{ \left| (\tilde{P}_f^{(\alpha)} - \tilde{P}^{(\alpha)}) - (P_f - P) \right| \geq \varepsilon (P_f - P) \right\} \leq \delta,$$

so long as

$$\varepsilon \geq \xi \left(1 + 2 \left(\frac{P}{P_f - P} \right) \right).$$

Since

$$\frac{P_f}{P} \geq 1 + \frac{s_1^*}{m} \geq 1 + \frac{w^*}{w_{\max} m},$$

it follows that

$$\xi \left(1 + 2 \left(\frac{P}{P_f - P} \right) \right) \leq \xi \left(1 + 2 \left(\frac{w_{\max} m}{w^*} \right) \right).$$

Taking

$$\xi = \frac{\varepsilon}{1 + 2 \left(\frac{w_{\max} m}{w^*} \right)},$$

$\tilde{P}_f^{(\alpha)} - \tilde{P}^{(\alpha)}$ is an (ε, δ) -approximation of $P_f - P$, and consequently, $\tilde{F}_f = (\tilde{P}_f^{(\alpha)} - \tilde{P}^{(\alpha)})m\mu$ is an (ε, δ) -approximation of F_f , for any given $\varepsilon > 0$ and $0 < \delta < 1$. \square

Table 2 shows (ε, δ) -approximations of P_f and F_f (being truncated at 6 decimal places) for the system in Fig. 1 with the approximation parameters $\varepsilon = 10^{-6}$ and $\delta = 10^{-2}$. A simple comparison of \hat{P}_f and \tilde{P}_f or \hat{F}_f and \tilde{F}_f shows that not only the proposed approximation technique is computationally more efficient than the bounding technique of Section 4.3 (polynomial-time vs. exponential-time), but it can also perform superior in terms of precision. Moreover, these advantages become more profound in larger systems. This however comes at a price: \tilde{P}_f and \tilde{F}_f are not correct always, but with high probability. This is in contrast to \hat{P}_f and \hat{F}_f that are always correct (up to d_P and d_F decimal places).

5.6 Computational Complexity

From (8), it follows that $\xi = \Omega(n^{-2})$. By using (3), one can see that it suffices to enumerate α -min cutsets for some $\alpha \leq 3$, and the number of such cutsets is $M = O(n^6)$. (For computing \tilde{P}_f , it was previously shown in [34] that it suffices to enumerate only α -min cutsets for some $\alpha \leq 2$, and there are $O(n^4)$ such cutsets.) By Lemma 2, one can enumerate all the M α -min cutsets with high probability in $\tilde{O}(n^6)$ time. Since $M \leq n^{2\alpha} \leq n^6$, for computing each of the estimates $\tilde{P}^{(\alpha)}$ and $\tilde{P}_f^{(\alpha)}$, the KLM estimator can be run in $\tilde{O}(M \cdot m / \xi^2) = \tilde{O}(n^{12})$ time [36]. The estimates $\tilde{P}^{(\cdot)}$ and $\tilde{P}_f^{(\cdot)}$ can be computed in $O(n^6 \cdot (m + n) / \xi^2) = O(n^{12})$ time using the Monte Carlo simulation. This concludes that the proposed algorithm for computing \tilde{F}_f runs in polynomial time.

6 Conclusion

We considered the problem of estimating the failure frequency of large-scale composite systems. It was previously shown that the failure probability can be efficiently approximated with provable guarantees. However, no such result was known for the failure frequency. In this work, we proposed a polynomial-time algorithm based on near-minimum cutsets for approximating the failure frequency with high probability with an arbitrary multiplicative error factor. Moreover, comparing the proposed approximation technique with the commonly-used bounding technique, our numerical results show that the proposed technique provides a more accurate approximation with less computational complexity.

References

- [1] D. P. Gaver, F. E. Montmeat, and A. D. Patton, "Power systems reliability I-measures of reliability and methods of calculation," *IEEE Trans. Power App. Syst.*, vol. 83, no. 7, pp. 727–737, Jul. 1964.
- [2] C. Singh and R. Billinton, *System Reliability Modeling and Evaluation*. London, U.K.: Hutchinson Educational, 1977.
- [3] C. L. Hwang, F. K. Tillman, and M. H. Lee, "System-reliability evaluation techniques for complex/large systems—A review," *IEEE Trans. Reliab.*, vol. 30, no. 5, pp. 416–423, Dec. 1981.
- [4] R. Billinton and K. E. Bollinger, "Transmission system reliability evaluation using Markov processes," *IEEE Trans. Power App. Syst.*, vol. PAS-87, no. 2, pp. 538–547, Feb. 1968.
- [5] J. A. Abraham, "An improved algorithm for network reliability," *IEEE Trans. Reliab.*, vol. R-38, no. 1, pp. 58–61, 1979.
- [6] C. Singh, "Markov cut-set approach for the reliability evaluation of transmission and Distribution Systems," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 6, pp. 2719–2725, Jun. 1981.
- [7] G. S. Fishman, "A Monte Carlo sampling plan for estimating network reliability," *Operations Research*, vol. 34, no. 4, pp. 581–594, 1986.
- [8] M. O. Locks, "A minimizing algorithm for sum of disjoint products," *IEEE Trans. Reliab.*, vol. R-36, no. 4, pp. 445–453, 1987.
- [9] F. Beichelt and L. Spross, "An improved Abraham-method for generating disjoint sums," *IEEE Trans. Reliab.*, vol. R-36, no. 1, pp. 70–74, Apr. 1987.
- [10] C. Dichirico and C. Singh, "Reliability analysis of transmission lines with common mode failures when repair times are arbitrarily distributed," *IEEE Trans. Power Syst.*, vol. 3, no. 3, pp. 1012–1019, Aug. 1988.
- [11] K. D. Heidtmann, "Smaller sums of disjoint products by subproduct inversion," *IEEE Trans. Reliab.*, vol. 38, no. 3, pp. 305–311, 1989.
- [12] J. M. Wilson, "An improved minimizing algorithm for sum of disjoint products," *IEEE Trans. Reliab.*, vol. 39, no. 1, pp. 42–45, Apr. 1990.

- [13] M. O. Locks and J. M. Wilson, "Note on disjoint algorithms," *IEEE Trans. Reliab.*, vol. 41, no. 3, pp. 81–92, Mar. 1992.
- [14] J. Lin, C. Jane, and J. Yuan, "On reliability evaluation of a capacitated-flow network in terms of minimal path sets," *Networks*, vol. 25, no. 3, pp. 131–138, May 1995.
- [15] S. M. Lee and D. H. Park, "An efficient method for evaluating network reliability with variable link-capacities," *IEEE Trans. Reliab.*, vol. 50, no. 4, pp. 374–379, Dec. 2001.
- [16] A. Balan and L. Traldi, "Preprocessing minpaths for sum of disjoint products," *IEEE Trans. Reliab.*, vol. 52, no. 3, pp. 289–295, Sep. 2003.
- [17] C. Singh and R. Billinton, "A new method to determine the failure frequency of a complex system," *IEEE Trans. Reliab.*, vol. 23, no. 4, pp. 231–234, Oct. 1974.
- [18] C. Singh, "A matrix approach to calculate the failure frequency and related indices," *Microelectron. Reliab.*, vol. 19, no. 4, pp. 395–398, 1979.
- [19] —, "Effect of probability distributions on steady state frequency," *IEEE Trans. Reliab.*, vol. 29, no. 3, p. 274, 1980.
- [20] —, "Rules for calculating the time-specific frequency of system failure," *IEEE Trans. Reliab.*, vol. 30, no. 4, pp. 364–366, Oct. 1981.
- [21] J. Provan and M. Ball, "The complexity of counting cuts and of computing the probability that a graph is connected," *SIAM Journal on Computing*, vol. 12, no. 4, pp. 777–788, 1983.
- [22] M. Ball, "Computational complexity of network reliability analysis," *IEEE Trans. Reliab.*, vol. 35, no. 3, pp. 230–239, Aug. 1986.
- [23] J. C. O. Mello, M. V. F. Pereira, and A. M. Leite da Silva, "Evaluation of reliability worth in composite systems based on pseudo-sequential Monte Carlo simulation," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1318–1326, Aug. 1994.
- [24] A. M. Leite da Silva, J. G. de Carvalho Costa, L. A. da Fonseca Manso, and G. J. Anders, "Transmission capacity: Availability, maximum transfer and reliability," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 843–849, Aug. 2002.
- [25] C. Srivareeratana, A. Konak, and A. E. Smith, "Estimation of all-terminal network reliability using an artificial neural network," *Computers and Operations Research*, vol. 29, no. 7, pp. 849–868, 2002.

- [26] K.-P. Hui, N. Bean, M. Kraetzl, and D. Kroese, "The cross-entropy method for network reliability estimation," *Operations Research*, vol. 134, no. 1, pp. 101–118, 2005.
- [27] F. Altiparmak, B. Dengiz, and A. E. Smith, "A general neural network model for estimating telecommunications network reliability," *IEEE Trans. Reliab.*, vol. 58, no. 1, pp. 2–9, Mar. 2009.
- [28] P. A. Jensen and M. Bellmore, "An algorithm to determine the reliability of a complex system," *IEEE Trans. Reliab.*, vol. 18, no. 4, pp. 169–174, Nov. 1969.
- [29] A. C. Nelson, J. R. Batts, and R. L. Beadles, "A computer program for approximating system reliability," *IEEE Trans. Reliab.*, vol. 19, no. 2, pp. 61–65, May 1970.
- [30] J. D. Esary and F. Proschan, "A reliability bound for systems of maintained, independent components," *Journal of the American Statistical Association*, vol. 65, pp. 329–338, 1970.
- [31] J.-M. Won and F. Karray, "A greedy algorithm for faster feasibility evaluation of all-terminal-reliable networks," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 41, no. 6, pp. 1600–1611, 2011.
- [32] C. Singh, "On the behavior of failure frequency bounds," *IEEE Trans. Reliab.*, vol. 26, no. 1, pp. 63–66, Apr. 1977.
- [33] J. Mitra and C. Singh, "Pruning and simulation for determination of frequency and duration indices of composite power systems," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 899–905, Aug. 1999.
- [34] D. R. Karger, "A randomized fully polynomial time approximation scheme for the all terminal network reliability problem," *SIAM Review*, vol. 43, no. 3, pp. 499–522, 2001.
- [35] D. R. Karger and C. Stein, "A new approach to the minimum cut problem," *J. ACM*, vol. 43, no. 4, pp. 601–640, Jul. 1996.
- [36] R. M. Karp, M. Luby, and N. Madras, "Monte Carlo approximation algorithms for enumeration problems," *J. Algorithms*, vol. 10, no. 3, pp. 429–448, Sep. 1989.