PROBABILISTIC RISK EVALUATION OF CYBER-ATTACKS ON A NUCLEAR POWER PLANT SAFETY

Jong Woo Park and Seung Jun Lee
Department of Nuclear Engineering
Ulsan National Institute of Science and Technology (UNIST)
50 UNIST-gil, Ulju-gun, Ulsan, 44919, Republic of Korea
jongwoo822@unist.ac.kr; sjlee420@unist.ac.kr

ABSTRACT

By adopting digital technology in nuclear power plants (NPPs), the cyber-attack has been introduced as one of the newest dangerous threats. It is required to develop efficient defense strategies against possible cyber-attacks on NPPs. However, it is difficult to expect when, where, and how the cyber-attacks will occur because they are intended attacks. Therefore, the risk of cyber-attacks on NPP is getting more and more important. Probabilistic risk assessment (PRA), which is one of the most widely used methods in NPPs to assess the risk, is introduced. However, it is difficult to evaluate the risk of NPPs by cyber-attacks from only the existing PRA method. In this research, a new PSA method, which is based on existing PRA, is used for evaluating of cyber-attacks on NPP. In addition, the case studies were performed for the feasibility study using nuclear power plant PRA model. Through this study, the risk assessment of the cyber-attacks on the NPP is tried by using a previous PRA model. Also, several scenarios are assumed to proceed the risk assessment. The risk-assessment should be performed in future work containing more specific cyber-attack scenarios, and if possible, the risk assessment model should be developed. Finally, the risk of cyber-attacks on NPPs could be evaluated quantitatively. Also, the defense strategies could be deducted by a risk assessment based on the risk assessment results which is based on the proposed method.

Key Words: Cyber security, Cyber-attacks, Nuclear power plant, Probabilistic Risk Assessment, Digital I&C

1 INTRODUCTION

For several years, the analog instrumentation and control (I&C) systems have been replaced by digital systems for adopting NPPs. While reducing the risk of NPPs adopting analog systems due to digital systems which have many advantages such as high-speed calculation and fault tolerance technique, the risk of the plant was focused on digital systems. Moreover, cyber-attacks which are a new type of threats on NPPs adopting digital systems has emerged so cyber security has become more important. “Stuxnet” which is the cyber-attack accident on the Iran nuclear facility is a practical example [1]. It shows the possibility to break part of the NPPs physically with malicious software through the intra-network. Cyber security has been introduced as one of the significant issues for NPP safety. However, the research about the cyber-attacks on NPPs is not mature yet. Moreover, there is not much information such as risk assessment method for cyber-attacks and risk evaluation of the cyber security for NPPs. In this work, a newly proposed method for risk evaluation of cyber-attacks was introduced for identifying significant cyber-attack scenarios on NPPs. Case studies and scenario-based analysis were performed with proposed risk evaluation method and the existing PRA model.

2 NUCLEAR POWER PLANT CYBER SECURITY

In general, a new cyber security system is programmed to detect already known types of cyber-attacks. When a new cyber-attack is observed, the cyber security system is updated using the information of the

new cyber-attack. That means another new cyber-attack is hard to be detected. However, it is not allowable in a safety critical infrastructure such as nuclear facilities. If only known attacks are defendable, then the defense strategy using cyber security system is not useful for safety critical systems in case of new cyber-attack with malicious software. Therefore, the risk evaluation of cyber-attacks on NPP should be treated with a different point of view. The different defense strategies based on the different risk evaluation should be deducted. For deducting defense strategies, the new risk evaluation is required to identify possible paths of cyber-attacks and vulnerabilities against cyber-attacks. In this study, the risk evaluation method for cyber-attacks on NPPs was introduced.

3 RISK EVALUATION METHOD

One of the most useful methods to evaluate the risk of an NPP is PRA. This work utilizes a PRA model to evaluate the risk of a cyber-attack and to identify a high-risk scenario. Typically, the risk is represented by the product of frequency and consequence [2]. However, the risk induced by a cyber-attack is represented by the product of a cyber-attack probability, the conditional probability of an event caused by a cyber-attack and the consequence of the event. In a cyber-attack risk assessment, the probability of a cyber-attack was not measurable. Therefore the probability of a cyber-attack was assumed to be one. Following the existing PRA in where consequence analysis already exists, the frequency information is now needed for a risk evaluation in a cyber-attack condition. Level 1 PRA evaluate the core damage frequency (CDF). It is one of the measures to estimate an NPP safety. The risk of cyber-attack is measured by the changes of the CDF in this work.

For instance, assume that the reactor protection system (RPS), which is in the NPP adopting digital system, is not working because of a cyber-attack with malicious software. Then, the cyber-attack causes automatic trip signal generation failure. Due to the operator manual backup and diverse protection system (DPS), this cyber-attack cannot cause core damage in NPP directly. However, the cyber-attack increase the CDF due to failure of automatic reactor trip. In this case of a cyber-attack, the failure probability of RPS trip signal generation should be one. Then the results of cyber-attack are represented by the CDF changes which are evaluated in level 1 PRA. Therefore, the risk of cyber-attacks on NPPs can be evaluated by the CDF changes using level 1 PRA method. Also, the significant cyber-attack scenarios are identified by analyzing minimal cut sets (MCSs) for the CDF evaluation.

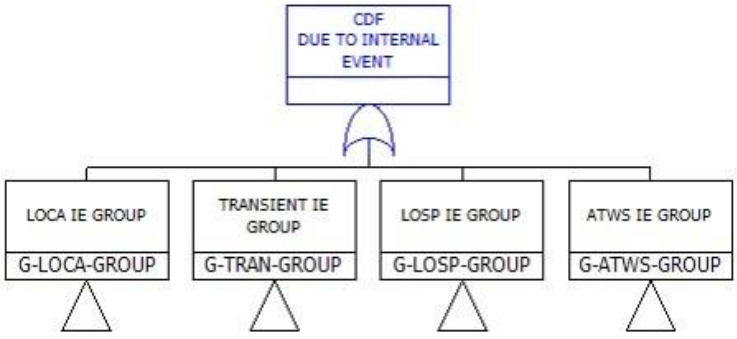

Figure 1. A part of PRA model

Figure 1 shows a part of level 1 PRA model. For level 1 PRA, event tree (ET) and fault tree (FT) models are used. ET is represented by the scenario of events. FT is represented by the system reliability analysis. ET/FT modeling is one of the popular methods to develop the CDF estimation models.

To reflect the cyber-attacks on a PRA model, there are several types of cyber-attacks, which are categorized as the following:

- Type 1: Attacks causes initiating events such as loss of coolant accident (LOCA) and station black out (SBO).
- Type 2: Attacks on digital system to make that system unavailable or to cause abnormal behavior (direct attacks)
- Type 3: Attacks on control logics for not digitalized component such as pumps and valves (indirect attacks)
- Type 4: Attacks on information systems to block the information or to switch it with wrong information

The cyber-attacks on NPP can be modeled as a basic event in the PRA model. All necessary data for risk evaluation such as component failure probability, human error probability, initiating event frequencies and accident scenarios are in the PRA model. In this work, the NPP is assumed that digital I&C, adopted in RPS, engineered safety features actuation system (ESFAS), DPS and plant control system (PCS).

3.1 Basic Events Categorization

To analyze a cyber-attack, MCSs were categorized in level 1 PRA results. A large number of MCSs came out from the PRA model. Therefore, basic events categorization is necessary to distinguish whether basic events were related to a cyber-attack or not.

Table I. The Part of MCSs in the CDF Table of NPP Model

VALUE	F-V	EVENT#1	#2	#3	#4	#5	#6
2.76E-08	0.004533	%ILSSB	AFMPS02BB	AFTPS02AB	FLAG-ID-NR-AC1HR	SDOPHEARLY	
2.70E-08	0.004436	%ISL	HCCQWHPP	MXOPHDPLI			
2.64E-08	0.004335	%ILOOP-SBO	EGDGW01ABET	MSOPHEVADV-2	NR-AC7HR		
2.51E-08	0.004122	%ILOFW	AFMPW01A2B	AFTPW01B2A	FLAG-ID-NR-AC1HR	SDOPHEARLY	
1.60E-09	0.000262	%IGTRN	DPTCAMG2	FLAG-ID-IATWS	MTC	RPOPVTRIP	RPPMWBP
1.60E-09	0.000262	%IGTRN	DPTCAMG1	FLAG-ID-IATWS	MTC	RPOPVTRIP	RPPMWBP
1.35E-09	0.000222	%ILODC	DPSKAPLC2	FLAG-ID-NR-AC1HR	FSQPFCLP2A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC1	FLAG-ID-NR-AC1HR	FSQPFCLV1A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC1	FLAG-ID-NR-AC1HR	FSQPFCLP2A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC1	FLAG-ID-NR-AC1HR	FSQPFCLP1A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC2	FLAG-ID-NR-AC1HR	FSQPFCLP1A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC2	FLAG-ID-NR-AC1HR	FSQPFCLV2A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC1	FLAG-ID-NR-AC1HR	FSQPFCLV2A	SDOPHEARLY	
1.35E-09	0.000222	%ILODC	DPSKAPLC2	FLAG-ID-NR-AC1HR	FSQPFCLV1A	SDOPHEARLY	
1.11E-09	0.000182	%IGTRN	DPTCAMG1	FLAG-ID-IATWS	MTC	RPOMWCP	RPOPVTRIP
1.11E-09	0.000182	%IGTRN	DPTCAMG2	FLAG-ID-IATWS	MTC	RPOMWCP	RPOPVTRIP
3.66E-10	0.00006	%ISL	CXPMAHSMV0675A	FLAG-ID-NR-AC8HR	HSMVO0676B		
3.66E-10	0.00006	%ISL	CXPMAHSMV0676B	FLAG-ID-NR-AC8HR	HSMVO0675A		

Table I shows the part of MCSs in the CDF Table of NPP model. For example, the basic events were categorized as the following:

- The Red Events: The failure of digitalized components in RPS by a cyber-attack
- The Purple Events: The failure of digitalized components in ESFAS by a cyber-attack
- The Green Events: The failure of digitalized components in DPS by a cyber-attack
- The Blue Events: The failure of operation of operators due to a cyber-attack on the information system
- The Yellow Events: The failure of PCS card for control of analog part by a cyber-attack

Basic events which possibly induced by a cyber-attack were categorized [3]. Using this categorization, basic events induced by cyber-attacks can be grouped in risk evaluation.

3.2 Importance Analysis

Risk achievement worth (RAW) is one of the most important measures in PRA [4]. It is for observing a total system failure probability when the failure probability is assumed to be one. In a cyber-attack case, the probability cannot be predictable. Therefore, the RAW analysis is the most reasonable method to assess the risk of a cyber-attack on NPP.

Table II. The part of RAW chart

ID	Event	Probability	RAW	# of MCS
1	%IISLOCA	1.77E-09	156,119.30	1
2	%IRVR	2.66E-07	156,119.20	1
3	AFCVW	2.34E-06	6,422.39	17,104
	...			
1109	FSOPVAFAS2	3.68E-03	1.001	901
1110	NLBSYLC03N	3.35E-06	1	5

Table II shows a part of the RAW chart. The number of basic events in RAW was related to cut off value. Through the RAW analysis, the vulnerability caused by cyber-attacks on NPP could be analyzed.

3.3 Case Study

The basic events in the RAW analysis were categorized following basic events categorization method. In this work, the case studies for RPS, ESFAS, and information system were conducted.

- Case 1: RPS failure by a cyber-attack
- Case 2: ESFAS failure by a cyber-attack
- Case 3: Information system for operators failure by a cyber-attack

Table III. CDF changes of each case by a cyber-attack

	Case 1	Case 2	Case 3
CDF changes	450 times increases	200 times increases	35000 times increases

Table III shows the CDF changes of each case study by a cyber-attack. In case 1, it is assumed that the cyber-attacks affect all basic events which have probability for cyber-attacks. As a result, the total CDF was increased about 450 times. In the same way, case 2 and 3 increases the CDF about 200 and 35,000 times each. It is not possible to the whole system by a cyber-attack. This work was just examples for influence analysis. The risk of cyber-attacks on NPP can be assessed in terms of the CDF with realistic scenarios.

4 SCENARIO-BASED ANALYSIS

Scenario analysis was based on basic events categorization and RAW analysis. Through the case study, the risk of cyber-attacks on NPP can be assessed in terms of the CDF. For assessing the risk of feasible cyber-attacks, scenarios selection and basic events grouping are necessary.

4.1 Failure Modes

In this work, failure modes of RPS and ESFAS by cyber-attacks were analyzed. They were based on basic events categorization. Table IV and V show failure modes of RPS and ESFAS by cyber-attacks. RPS actuates only one signal, it is a reactor trip signal. Therefore, RPS has only one trip failure mode. However, in ESFAS case, there were several failure modes with different signals. Safety injection signal (SIAS), aux feed water actuation signal (AFAS), recirculating actuation signal (RAS) and containment spray actuation signal (CSAS) had a similar trend in failure mode. Also, the error of commission (EOC) was considered in this study. CIAS and MSIS are related to isolation. Therefore, it can be considered initiating events. In case of CIAS and MSIS failure, it can be considered both core damage and radioactive material release. Following the failure modes, scenarios can be developed.

Table IV. Failure modes of RPS by cyber-attacks

System	Signal of Function	State	Failure Mode 1 (Direct CA)	Failure Mode 2 (Indirect CA)	Result
RPS	Trip signal	Abnormal	OK	OK	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	ATWS

Table V. Failure modes of ESFAS by cyber-attacks

System	Signal of Function	State	Failure Mode 1 (Direct CA)	Failure Mode 2 (Indirect CA)	Result
ESFAS	SIAS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD
	AFAS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD
	RAS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD
	CSAS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD
	CIAS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD +release
		Normal	Released Actuation Signal by CA	Operation backup failed by CA	IE
	MSIS	Abnormal	OK	OK	OK
		Abnormal	OK	EOC induced by CA	CD
		Abnormal	Digital modules failed by CA	Operation backup	OK
		Abnormal	Digital modules failed by CA	Operation backup failed by CA	CD +release
		Normal	Released Actuation Signal by CA	Operation backup failed by CA	IE

4.2 Risk of Cyber-attack Analysis Based on Scenario

Scenario based analysis is important because there are numerous unpredictable ways of Cyber-attacks that can happen. Therefore, the risk of cyber-attacks should be assessed with a feasible scenario. From the failure mode analysis, scenarios can be created by the combination of failure modes caused by direct and indirect cyber-attacks. Moreover, the specific case that cyber-attacks such as malware can be activated when the initiating event occurs and it should be considered.

To show feasibility, scenario-based risk analysis was performed. The target system was only RPS which has only one failure mode. The scenarios of cyber-attacks on RPS was two caused by both direct and indirect cyber-attacks. Totally, the four scenarios were constructed as following:

- Scenario 1: RPS output modules failure occurred by a cyber-attack
- Scenario 2: Reactor trip involved RPS is failed by a cyber-attack
- Scenario 3: RPS fail to trip with small loss of coolant accident (SLOCA) which is an initial event occurred by cyber-attacks
- Scenario 4: Operator cannot manually backup due to indirect cyber-attacks when scenario 3 occurred

In scenario 3 and 4 cases, cyber-attacks were considered with initiating event SLOCA. Therefore, initiating event were used as conditional event. Therefore risk metric should be changed to core damage probability (CDP) instead of CDF.

Table VI. CDF changes or CDP table by scenario

	Scenario 1	**Scenario 2**
CDF changes	CDF increases 35 times	CDF increases 450 times
	Scenario 3	**Scenario 4**
CDP	4.044%	4.769%

Table VI shows the CDF changes or CDP by scenario 1 to 4. In case of scenario 1, the cyber-attack on RPS output modules assumed as CCF of RPS output module failure. In case of scenario 2, assumed that all possible cyber-attacks on RPS occurs, the CDF increased 35times and 450 times each. However, in case of scenario 3 and 4, initiating event SLOCA which was induced by cyber-attacks was considered. As previously stated, in case of scenario 3 and 4, they should be obtained as a result of CDP. The initiating event is adopted as conditional probability. As a result, CDP of scenario 3 and 4 were 4.044% and 4.769% each. In the same way, the risk of cyber-attacks on NPP will be assessed with the various scenario in future work.

5 CONCLUSIONS

The cyber-attacks on the NPP were already a realizable threat. Through this work, the risk assessment of the cyber-attacks on the NPP was tried by using an existing PRA model. Also, several scenarios are assumed to proceed the risk assessment. The risk-assessment should be performed in future work containing more specific and feasible cyber-attack scenarios. In addition, if possible, the risk assessment model should be developed. Finally, the defense strategies could be deducted by a risk assessment based on a scenario analysis.

6 REFERENCES

1. J. Park, Y. Suh, and C. Park, “Implementation of cyber security for safety systems of nuclear facilities,” *Prog. Nucl. Energy*, vol. 88, pp. 88–94, 2016.
2. Y. Cherdantseva *et al.*, “A review of cyber security risk assessment methods for SCADA systems,” *Comput. Secur.*, vol. 56, pp. 1–27, 2015.
3. J. Song, J. Lee, C. Lee, K. Kwon, and D. Lee, “a Cyber Security Risk Assessment for the Design of I & C Systems in Nuclear Power Plants,” vol. 44, no. 8, pp. 919–928, 2012.
4. M. Van Der Borst and H. Schoonakker, “An overview of PSA importance measures,” *Reliab. Eng. Syst. Saf.*, vol. 72, no. 3, pp. 241–245, 2001.