PTL Volume 11-1 ISSN 2288-8195

2024 (사)ICT플랫폼학회 하계학술대회 자료집

2024년 7월 5일 금요일 13시
가천대학교 반도체대학 117호

주최/주관 : (사)아이씨티플랫폼학회 (www.ictps.or.kr)

후 원 : SK브로드밴드㈜, 쌍용정보통신㈜, 클라우데라코리아, ㈜누리아이티, 대신정보통신㈜, 세림TSG㈜, ㈜시야인사이트, ㈜온더라이브, ㈜올포랜드, ㈜인라스, 지니언스㈜, 지엔소프트㈜, ㈜진인프라, ㈜한국IT컨설팅

ICTPS
ICT Platform Society

2024
(사)ICT플랫폼학회 하계학술대회 논문집
사단법인 아이씨티플랫폼학회
서울 서초구 서초중앙로 63 (서초동) 15
ictps.staff@gmail.com

2024.07.05

PTL Volume 11-1 , ISSN 2288-8195

세부 프로그램

2024년 7월 5일(금) 13:00		
13:00~ 13:30	등록 [기념품 제공]	
시간	논문발표: 반도체대학 117호 - 좌장 : 구본근(한국교통대), 박용범(단국대)	논문발표: 반도체대학 301호 - 좌장 : 김시호(연세대)
13:30~ 15:00	◆ **멜 스펙트로그램을 사용해서 딥 페이크 오디오 탐지** 에르난데스 산티아고 루이스 에두아르도, 박용범 (단국대) ◆ **선박 대상 사이버 복원력 평가 요소 도출** 고아름, 이주현, 서정택 (가천대) ◆ **마이크로파 에너지 투과에 의한 아까시 종자의 발아율 향상에 관한 연구** 박동희 (한국교통대) ◆ **스마트 컨트랙트 취약점 정보 공유 프레임워크 제안** 최동빈, 박용범 (단국대) ◆ **산업제어시스템 환경 대상 사이버사고 대응 체계 지침 비교 분석** 최희원, 서정택 (가천대) ◆ **페르소나 기법을 통한 Intent Translation 방법론** 장재원, 이소연, 김대영 (순천향대) ◆ **ICS 네트워크 패킷의 불확실성을 학습하기 위한 특징 수준 융합 기반의 멀티모달 학습 분석** 이주현, 전승호, 서정택 (가천대) ◆ **비밀번호 강도 측정 기법 비교 분석** 박원상, 서승희, 이창훈 (서울과기대) ◆ **HMAC 기반 메시지인증을 위한 키관리 기법** 강윤희 (백석대), 권태언 ((주)하스퍼)	◆ **영-한 다국어 단어 임베딩을 통한 효율적인 문서 검색 시스템** 강어진, 유준 (가천대) ◆ **휴머노이드 로봇 제어를 위한 ROS 노드 설계** 이동완, 안해은, 지혜원, 손애은, 구본근 (한국교통대) ◆ **ix2pix-Swin: CGAN을 이용한 RGB-to-NIR 변환** 박인철, 진영완, 김시호 (연세대) ◆ **생성형 AI 기반 플랫폼 서비스에서의 UX 어포던스 및 지속사용의도간 관계연구** 박민혁, 구자준 (성균관대) ◆ **Windows VBS 악성코드 공격 동향 및 대응 방안 연구** 전규현, 이주현, 서정택 (가천대) ◆ **생성형 대형 언어 모델(LLM) 활용 영상의 날씨 조건 자동 인식 및 분류 방법** 주형진, 송한빈, 김시호 (연세대) ◆ **PNC 모델을 활용한 하천수위 예측 딥러닝 네트워크 개발** 이은서, 박귀만, 배영철 (전남대) ◆ **Jester를 활용한 핸드제스처 감지 방법 비교 연구** 이창용, 이종윤, 권소영, 이용환 (금오공과대) ◆ **IoT 환경 암호화 트래픽 대상 사이버공격 탐지 시스템 제안** 지일환, 이주현, 서정택 (가천대)
기술세미나: 반도체대학 117호 사회: 김백기 (강릉원주대)		
15:00~ 17:30	기술세미나 진행	
17:30~ 18:00	우수논문상 시상 및 경품 추첨	김현 (부천대 교수)
18:20~	저녁만찬 및 정보교류 장소 : 가천한마당 (경기 성남시 수정구 성남대로 1334 경원프라자, 031-721-8291)	

2024
(사)ICT플랫폼학회
하계학술대회

- Session 1 -

- Session 2 -

Session 1

논문발표

[반도체대학 117호]

- 좌장 -

구본근(한국교통대), 박용범(단국대)

멜 스펙트로그램을 사용해서 딥 페이크 오디오 탐지

*에르난데스 산티아고 루이스 에두아르도, **박용범

Deep Fake Audio Detection with Mel Spectrogram

*Hernandez Santiago Luis Eduardo, **Young B. Park

Abstract

In recent years, the rapid development of deep learning techniques has made easier the creation of deep fake audio. This paper proposes a method to identify deepfake audio, using long exposition methods applied to Mel spectrograms. By leveraging the temporal and spectral features captured in Mel spectrograms. Our approach demonstrates high accuracy in detecting deep fake audio, preventing the problems caused when people misuse deep fake audio.

Key words

Deep Fake, Mel-spectrogram, VLM, Long Exposure, Short-Time Fourier Transform

Ⅰ. Introduction

The rise of generative technology has transformed the way audio is produced. Deep fake audio involves using deep learning algorithms to generate human-like speech that can convincingly mimic a person's voice. Deep fake is currently originating significant concerns regarding spreading of misinformation, identity theft, and privacy breaches.[1]. Like In 2019, criminals used deep fake audio to mimic the voice of a CEO, convincing a subordinate to transfer $243,000 to a fraudulent account. Therefore, as these audio techniques become more sophisticated, the need for reliable detection mechanisms has become paramount.

Traditional audio detection methods often struggle with capturing nuances of deep fake audio, needing more advanced approaches [2].

This paper explores the use of long exposition preprocessing applied to a Mel spectrogram audio image to address this

* 단국대학교, 석사과정 luislalo3100@gmail.com
** 단국대학교, 교수 ybpark@dankook.ac.kr 교신저자

challenge. Chapter 2 will explain the dataset used for training the model. Chapter 3 will lead you through the preprocessing techniques we applied to the audio samples. Chapter 4 will explain the “TinyVGG” architecture, a small implementation of a VGG model. Chapter 5 will depict the training results and Chapter 6 will conclude and explain further research.

Ⅱ. Data Composition

2.1 In the Wild Dataset

The dataset used, known as “In the Wild” dataset, comprises 40 hours distributed over thirty thousand audio files of 16kHz sample rate from 58 notable celebrities and politicians.

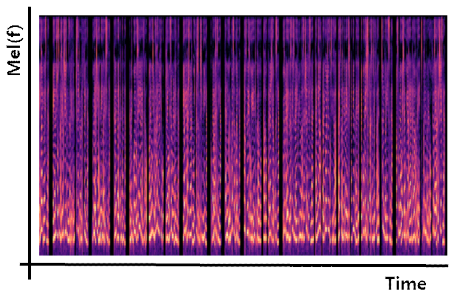
These files are divided into two categories: "Bona-fide," which contains authentic audio recordings, and "Spoofed-audio” containing deepfake generated audio files. Each speaker is represented by approximately 23 minutes of bona-fide audio and 18 minutes of spoofed audio.

Ⅲ. Audio Preprocessing

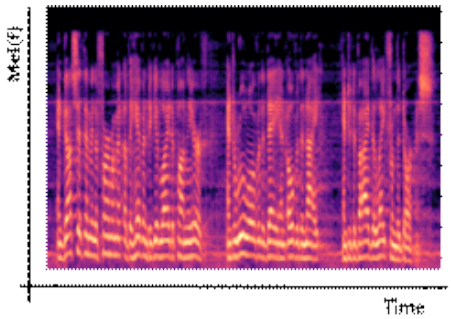
3.1 Applying Mel Spectrogram

Mel spectrograms transform audio files from the time domain to the frequency domain using the Short-Time Fourier Transform (STFT). This process captures spectral information essential for identifying anomalies introduced by deepfake audio. The x-axis refers to time, and the y-axis represents the frequency on the Mel scale, which approximates human auditory perception [3]. This allows us to see patterns in the spectrogram that reveal anomalies introduced by deepfake audio.

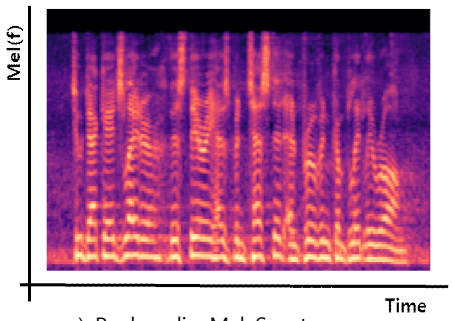
We extracted Mel Spectrogram features from TTS (Text to Speech) fig.1(a), Deepfake fig.1(b) and Real audio samples fig.1(c). We can notice a constantly generated horizontal line in both TTS and Deep fake audio samples.



a) TTS Mel Spectrogram



b) Deepfake Mel Spectrogram



c) Real audio Mel Spectrogram

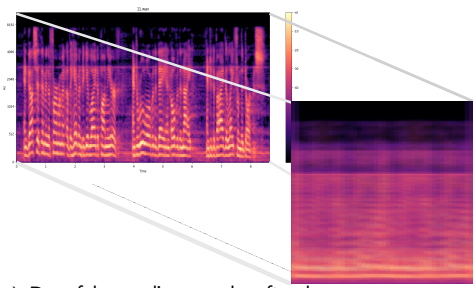
Fig.1 Mel Spectrogram Preprocessing

3.2 Long Exposure Technique

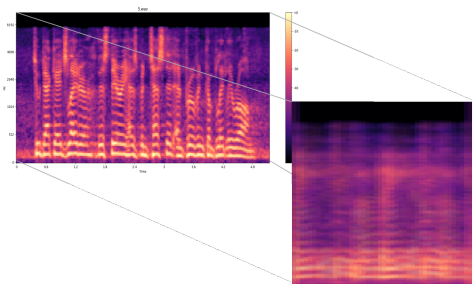
Preprocessing images with long exposure effects involves applying techniques that enhance motion within a scene. Long exposure technique captures the movement of elements within the scene. Constant elements appear sharp and clear while moving elements appear blurred. [4]

3.3 Windowing Preprocessing

To apply the long exposure effect to our audio images, we applied a 1-second window screening followed by a .5 millisecond overlap between windows. We then combined these segmented spectrograms to achieve a long exposure effect. [5] The results of this process are depicted in Fig. 2.



a) Deepfake audio sample after long exposure



b) Real audio sample after long exposure

Fig.2 Audio sample after long exposure

For machine learning purposes, audio files were transformed into 64x64 pixel images, representing spectrogram data, to facilitate pattern recognition in audio characteristics.

IV. Model Architecture

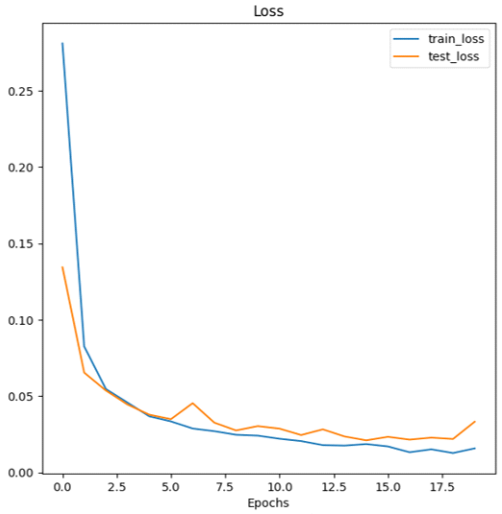
The TinyVGG model is a simplified version of the VGG neural network architecture, designed for image classification tasks. The input requires 3-Channel images for feature extracting with the convolutional operations and a classifier in charge of the feature mapping and flattening into a fully connected layer for classification. TinyVGG ensures simplicity and is less resource intensive than traditional VGG models, while maintaining robustness and accuracy [6].

V. Results

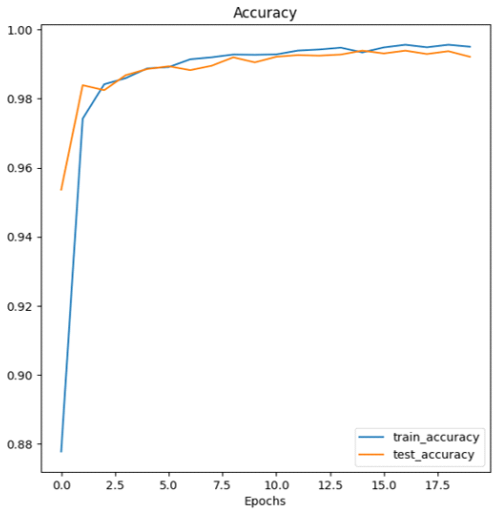
The model was trained for 20 epochs with a batch size of 32 and the Adam optimizer. The Cross Entropy Loss function was used. Results were summarized in Table 1; loss function Fig.3 a) and accuracy graph Fig.3 b) are depicted in Fig 3.

Table 1. Training Results

Train Loss	0.0158
Train Accuracy	0.9950
Test Loss	0.0333
Test Accuracy	0.9830



a) Loss Function



b) Accuracy

Fig 3. Loss Function & Accuracy Graph

VI. Conclusion and Future Research

This study presents an effective method for detecting deep fake audio using long exposition techniques applied to Mel spectrograms. By capturing detailed temporal and spectral features, our approach accurately distinguishes genuine audio from synthetic samples.

Due to insufficient data for TTS audio file we expect to collect enough data to add TTS deepfake detection

Acknowledgement

“This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (RS-2023-00259099) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)”

참 고 문 헌

[1] Wijethunga, R. L. M. A. P. C., Matheesha, D. M. K., Al Noman, A., De Silva, K. H. V. T. A., Tissera, M., & Rupasinghe, L. (2020, December). Deepfake audio detection: a deep learning bas ed solution for group conversations. In 2020 2nd International conference on advancement s in computing (ICAC) (Vol. 1, pp. 192-197). IEEE.

[2] Yi, J., Fu, R., Tao, J., Nie, S., Ma, H., Wang, C., ... & Li, H. (2022, May). Add 2022: the first audio deep synthesis detection challenge. In ICASSP 2022-2022 IEEE International Conferen ce on Acoustics, Speech and Signal Processing (ICASSP) (pp. 9216-9220). IEEE.

[3] Hwang, Y., Cho, H., Yang, H., Won, D. O., Oh, I., & Lee, S. W. (2020). Mel-spectrogram augme ntation for sequence to sequence voice convers ion. arXiv preprint arXiv:2001.01401.

[4] Mugnier, L. M., Fusco, T., & Conan, J. M. (2004). MISTRAL: a myopic edge-preserving image rest oration method, with application to astronomic al adaptive-optics-corrected long-exposure i mages. JOSA A, 21(10), 1841-1854.

[5] Tan, K. C., Lim, H., & Tan, B. T. G. (1991). Windowi

ng techniques for image restoration. CVGIP: Graphical Models and Image Processing, 53(5), 491-500.

[6] Truong, T. D., Nguyen, V. T., & Tran, M. T. (2018, January). Lightweight Deep Convolution al Network for Tiny Object Recognition. In ICPR AM (pp. 675-682).

선박 대상 사이버 복원력 평가 요소 도출

*고아름, **이주현, ***서정택

Development of Cyber Resilience Evaluation Criteria for Ship

*Areum Ko, **JuHyeon Lee and ***Jung Taek Seo

요 약

해양선박 환경에서 사이버 위협의 증가로 인해 사이버 복원력의 중요성이 더욱 커지고 있다. 이에 따라 사이버 복원력을 확보하고 평가하기 위한 지침 연구의 필요성이 대두되고 있다. 그러나 해양선박 환경에 특화된 사이버 복원력 평가 체계에 관한 연구는 아직 활발히 이루어지고 있지 않다. 이에 본 논문에서는 선박 대상 사이버 복원력 요구사항을 제시하는 IACS(International Association of Classification Societies) UR(Unified Requirements) E26(Cyber resilience of ships)을 통하여 선박 특화 사이버 복원력 평가 요소를 도출하며, 중요 핵심 인프라를 보호하고 복원력을 제고하기 위한 주요 방법론을 제시하는 NIST(National Institute of Standards and Technology) CSF(Cyber Security Framework)를 활용하여 사이버 복원력의 일반적인 측면까지 고려하는 사이버 복원력 평가 요소를 도출하였다. 본 논문에서 도출한 평가 요소를 활용하여 평가 체계를 구축하고 사이버 복원력 평가가 이루어질 경우, 선박 대상 사이버 복원력 향상에 도움이 될 것으로 기대된다.

Key words

Maritime Cybersecurity, Cyber Resilience, Evaluation Criteria

Ⅰ. 서 론

사이버 복원력(Cyber Resiliency)이란 사이버 자산을 사용하거나 이를 활용하는 시스템에 대한 공격을 예측하고, 견디며, 이로부터 회복하고 적응하는 능력을 말한다[1]. 정보통신기술(Information & Communications Technology, ICT)의 발전으로 사이버 위협의 빈도와 사이버 피해가 증가함에 따라 다양한 산업 분야에서는 사이버 복원력의 중요성을 강조하고 있다[2]. 특히 해양선박 환경의 사이버 위협이 증가하면서 해양선박 환경에서의 사이버 복원력의 중요성은 더욱

* 가천대학교 정보보호학과 석사과정 (ko6103@gachon.ac.kr)
** 가천대학교 정보보호학과 박사과정 (202240226@gachon.ac.kr)
*** 가천대학교 컴퓨터공학부 스마트보안전공 교수 (seojt@gachon.ac.kr)

커지고 있다[3]. 국제해사기구(International Maritime Organization, IMO)는 이러한 필요성을 인식하고 결의안 MSC.428(98)을 통해 사이버 위협에 대한 운영상 복원력의 중요성을 강조하였으며[4], 이를 기반으로 국제선급협회(International Association of Classification Societies, IACS)에서는 선박의 사이버 복원력을 위한 최소한의 요구사항인 UR(Unified Requirements) E26(Cyber resilience of ships)을 발행하였다[5].

이렇듯 사이버 복원력의 중요성이 커짐에 따라 이에 대한 평가 체계의 개발 필요성 또한 대두되고 있다[2]. 사이버 복원력 평가는 사이버보안 상태를 객관적으로 파악하고, 잠재적인 취약점을 사전에 식별하는 데 도움을 주며, 평가를 통한 개선점을 도출 및 반영하여 사이버 복원력을 강화할 수 있다. 이에 최재혁 외 2명[6]은 국방정보시스템의 미비한 분야를 식별 및 보완하기 위해 사이버 복원력을 평가할 수 있는 성숙도 모델을 제안하였다. 또한 Haque, M. A. 외 3명[7]은 견고성, 자원성, 중복성 및 신속성의 4가지 지표로 구성된 R4 재해 복원력 프레임워크[8]를 활용하여 ICS 대상 사이버 복원력 평가모델을 제안하였다. 이렇듯 사이버 복원력 평가를 위한 연구가 이루어지고 있으나 해양선박 환경 특화 사이버 복원력 평가에 관한 연구는 부족한 실정이다.

이러한 배경을 바탕으로 본 논문에서는 선박 대상 사이버 복원력 평가 요소를 개발한다. 해양선박 환경에 특화된 평가 요소를 개발하기 위해 IACS UR E26을 활용하였으며, 포괄적인 사이버 복원력 평가를 위해 NIST(National Institute of Standards and Technology) CSF(Cyber Security Framework)를 적용하였다. 이를 통해 UR E26의 기능 요소를 보완하고 더욱 체계적인 사이버 복원력 평가를 위한 평가 요소를 도출하였다.

본 논문의 구성은 다음과 같다. 2장에서 IACS UR E26과 NIST CSF에 대해서 분석하고, 3장에서 이를 바탕으로 사이버 복원력 평가 요소를 개발한다. 그리고 4장에서 결론 및 향후 연구방향을 제시한다.

Ⅱ. 선행연구

2.1 IACS UR E26 분석

IACS UR E26은 선박 사이버 복원력을 위한 최소 요구사항을 제시하며, 선박의 설계, 건조, 시운전, 운항 등 선박의 운용 주기 전반에 걸쳐 운영기술 및 정보기술 장비를 선박 네트워크에 안전하게 통합하는 것을 목표로 하고 있다[9].

[표 1] 선박 사이버 복원력 기능 요소

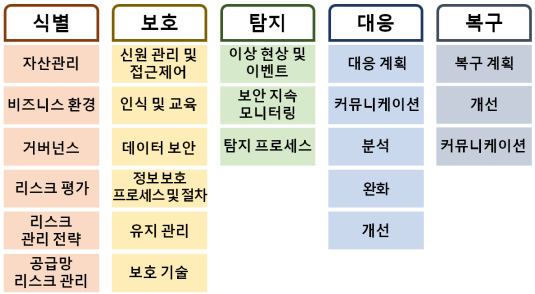
기능 요소	내용
식별	• 선내 시스템, 사람, 자산, 데이터 및 기능(Capabilities)에 대한 사이버보안 위험을 관리하기 위한 조직적 이해를 개발
보호	• 사이버 사고로부터 선박을 보호하고 해운 운영의 연속성을 최대화하기 위한 적절한 보호 장치를 개발하고 이행
탐지	• 선내 사이버 사고의 발생을 탐지하고 식별하기 위한 적절한 조치를 개발하고 이행
대응	• 선내에서 탐지된 사이버 사고에 대한 조치를 취하기 위한 적절한 조치 및 활동을 개발하고 이행
복구	• 사이버 사고로 인해 손상된 선박 운항에 필요한 모든 기능 또는 서비스를 복구하기 위한 적절한 조치 및 활동을 개발하고 이행

UR E26은 [표 1]과 같이 식별, 보호, 탐지, 대응, 복구의 5가지 기능 요소를 정의하며, 각 기능 요소별 요구사항을 제시한다. 총 17가지의 요구사항이 있으며, 각 기능 요소별 요구사항의 목표는 다음과 같다.

- **식별.** 선내 컴퓨터 기반 시스템(Computer Based System, CBS)의 상호의존성 및 관련 정보 흐름, 그리고 선박의 관리, 운영, 거버넌스, 역할 및 책임과 관련된 주요 자원을 식별하는 것을 목표로 한다.
- **보호.** 잠재적 사고의 영향을 제한하거나 억제하는 능력을 지원하는 적절한 보호 장치의 개발 및 구현을 목표로 한다.
- **탐지.** 선내 CBS 및 네트워크에서 이상 활동을 인식하고 사이버 사고를 식별하는 능력을 지원하는 적절한 수단의 개발 및 구현을 목표로 한다.
- **대응.** 선내 CBS 및 네트워크에서 발생 가능한 사이버 사고의 영향을 최소화할 수 있는 능력을 지원하는 적절한 수단의 개발 및 구현을 목표로 한다.
- **복구.** 사이버 사고로 영향을 받은 선내 CBS 및 네트워크를 복구하는 기능을 지원하는 적절한 수단의 개발 및 구현을 목표로 한다.

2.2 NIST CSF 분석

NIST CSF는 사이버보안 리스크를 관리하고 줄이기 위한 프레임워크로 중요 핵심 인프라를 보호하고 복원력을 제고하기 위한 주요 방법론을 제시한다[10, 11]. 또한 사이버 리스크 관리의 우선순위 결정과 유연하고 효율적인 접근 방식을 통해 조직이 사이버보안 프로그램을 시작하거나 개선하는 데 도움을 준다. NIST CSF는 [그림 1]과 같이 5개의 주요 단계 및 23개의 구성요소를 포함한다.



[그림 1] NIST CSF의 단계 및 구성요소

- **식별.** 중요한 시스템, 자산, 데이터 및 관련 사이버보안 위험을 파악하여 조직의 상황을 이해하는 것을 목표로 한다. 해당 단계는 조직이 비즈니스 맥락을 이해하고, 핵심 기능을 지원하는 자원과 관련된 사이버보안 위험을 평가하여 리스크 관리 전략과 비즈니스 필요에 따라 우선순위를 정하는 데 필수적이다.
- **보호.** 파악된 자산과 데이터를 보호하기 위한 방어적 조치를 개발하고 구현하는 것을 목표로 한다. 해당 단계는 중요한 서비스의 제공을 보장하기 위한 적절한 방어 조치를 통해 잠재적인 사이버보안 사건의 영향을 제한하거나 억제하는 능력을 지원한다.
- **탐지.** 보안 이벤트의 발생을 신속하게 식별하기 위한 활동을 실행하는 것을 목표로 한다. 해당 단계는 보안 이벤트를 신속히 발견하고 분석할 수 있도록 하여 효과적인 대응 및 복구 활동을 지원한다.
- **대응.** 탐지된 사이버보안 이벤트에 효과적으로 대응하기 위한 계획과 절차를 수립하고 실행하는 것을 목표로 한다. 해당 단계는 잠재적인 사이버보안 사건의 영향을 최소화하기 위한 활동을 포함한다.
- **복구.** 사이버보안 사건 이후 정상 운영 상태로 복구하기 위한 활동을 수행하는 것을 목표로 한다. 해당 단계는 사이버보안 사건의 이해관계자와 소통을 하며, 신속한 복원을 통해 사이버보안 사건의 영향을 줄이고, 학습을 통해 향후

활동을 개선한다.

UR E26과 NIST CSF는 식별, 보호, 탐지, 대응, 복구로 이루어져 있지만, 세부 구성요소가 다르고 NIST CSF의 구성요소가 UR E26에 적용되어 선박 대상 사이버 복원력을 강화할 수 있다. 예를 들어, NIST CSF은 선박 사이버 복원력을 평가하는 데 있어 중요한 리스크 기반 접근 방식을 제공한다. NIST CSF의 리스크 관리 전략을 UR E26에 통합하면, 선박의 사이버보안 위험을 체계적으로 관리하고 평가할 수 있다. 이에 따라 UR E26에 NIST CSF의 구성요소를 도입하여 선박 대상 사이버 복원력 평가 요소를 도출하고자 한다.

Ⅲ. 선박 대상 사이버 복원력 평가 요소

본 장에서는 UR E26의 요구사항과 각 기능 요소별로 적용가능한 NIST CSF의 구성 요소를 식별하여 선박 대상 사이버 복원력 평가 요소를 정의하고 그 도출 근거를 설명한다. [표 2]에 제시된 바와 같이, UR E26의 17가지 요구사항과 NIST CSF의 11가지 구성 요소를 도입하여 총 28가지 사이버 복원력 평가 요소를 도출하였다.

3.1 식별

식별 단계는 조직이 관리해야 할 자산, 데이터 및 기능에 대한 사이버보안 위험을 파악하는 데 중점을 둔다. 조직이 잠재적 위협을 체계적으로 식별하고 우선순위를 설정하여 효과적으로 대응하고 예방하기 위해서는 리스크 평가가 필수적이다[12]. 따라서 NIST CSF의 '리스크 평가'와 '리스크 관리 전략'을 도입하였다. 또한 선박 ICT 공급망 사이버공격 위협이 증가함에 따라 이를 대응하기 위한 '공급망 리스크 관리'를 도입하여 총 4가지 평가 요소를 도출하였다[13]. NIST CSF의 '자산관리'는 UR E26의 ‘선박 자산 목록 작성 및 유지’ 평가 요소와 동일하여 제외하였으며, '비즈니스 환경'과 '거버넌스'는 내부 정책 설정에 중점을 두고 있어 제외하였다.

도출된 평가 요소는 UR E26에 따라 선박 자산 목록의 최신화 여부를 평가하고, NIST CSF에 따라 리스크 관리 전략이 명확히 정의

[표 2] UR E26 요구사항 및 NIST CSF 구성 요소 기반 선박 사이버 복원력 평가 요소

단계	평가 요소	
	UR E26	NIST CSF
식별	• 선박 자산 목록 작성 및 유지	• 리스크 평가 • 리스크 관리 전략 • 공급망 리스크 관리
보호	• 보안 구역 및 네트워크 분할 • 네트워크 보호 안전장치 • 안티바이러스 및 멀웨어 보호 • 접근 통제 • 무선 통신 • 원격 접근 통제 및 비신뢰 네트워크에서 통신 • 모바일 및 휴대용 장치의 사용	• 데이터 보안 • 정보보호 프로세스 및 절차 • 유지 보수
탐지	• 네트워크 운영 감시 • CBS 및 네트워크 검증 및 진단 기능	• -
대응	• 사고 대응 계획 • 로컬, 독립 및/또는 수동 운전 • 네트워크 격리 • 최소 위협 상태로의 대비책	• 분석 • 커뮤니케이션 • 개선
복구	• 복구 계획 • 백업 및 복구 기능 • 제어된 셧다운, 재설정, 롤백 및 재시작	• 개선 • 커뮤니케이션

되어 주기적으로 검토 및 업데이트되는지 평가한다. 또한, 이를 기반으로 선박의 자산과 공급망에 대한 주기적인 리스크 평가 여부를 평가한다.

3.2 보호

보호 단계는 조직이 파악한 자산과 데이터를 보호하기 위한 방어적 조치를 개발하고 구현하는 데 중점을 둔다. 이를 위해 구체적인 개발 및 구현 프로세스와 절차를 수립하는 것이 중요하며, 지속적인 유지 보수를 통해 가용성의 침해를 방지하는 것도 필수적이다. 따라서 NIST CSF의 '데이터 보안', '정보보호 프로세스 및 절차', '유지 보수'를 도입하여 총 10가지 평가 요소를 도출하였다. NIST CSF의 '신원 관리 및 접근제어' 및 '보호 기술'은 UR E26의 '접근 통제' 및 ' 네트워크 보호 안전장치' 평가 요소의 내용과 동일하여 제외하였으며, '인식 및 교육'은 보호 단계의 구체적인 기술적 조치와 직접적인 관련이 없으므로 제외하였다.

도출된 평가 요소는 UR E26에 따라 CBS가 보안 구역으로 그룹화되고 해당 보안 구역의 격리 여부를 평가한다. 또한, 방화벽과 같은 안전장치를 통한 바이러스, 웜 등 악성코드로부터의 보호 여부를 평가한다. 보안 구역에 대한 적절한 접근 권한 설정 여부를 평가하며, 무선 통신 네트워크의 보안 설계, 구현 및 유지 관리 여부, 모바일 및 휴대용 장치의 사용 권한에 따른 사용 제한 여부를 평가한다. NIST CSF의 평가 요소는 조직의 정보보호 조치와 데이터 보안 및 유지 보수 절차의 실행 여부를 평가한다.

3.3 탐지

탐지 단계는 조직의 정보 시스템과 자산에서 발생하는 비정상 활동을 신속히 탐지하는 데 중점을 둔다. NIST CSF의 '이상 현상 및 이벤트', '보안 지속 모니터링', '탐지 프로세스'는 UR E26의 '네트워크 운영 감시', 'CBS 및 네트워크 검증 및 진단 기능'에 포함되므로, 탐지 단계에서는 NIST CSF의 추가적인 평가 요소를 도입하지 않고 UR E26의 평가 요소 2가지를 도출하였다.

도출된 평가 요소는 네트워크에 대한 지속적인 실시간 모니터링 작동 여부 및 이벤트 경보를 알림 시스템의 작동 여부를 평가한다. 또한, 모든 보안 기능의 성능과 가용성 및 무결성을 평가한다.

3.4 대응

대응 단계는 탐지된 사이버보안 사고에 대해 신속하고 효과적으로 대응하기 위한 계획과 절차를 수립하고 실행하는 데 중점을 둔다. 효과적인 대응을 위해 사이버보안 사건 분석 및 이해관계자들 간의 효율적인 의사소통이 필요하다[12][14]. 이에 따라 NIST CSF의 '분석', '커뮤니케이션'을 도입하였으며, 교훈을 학습하고 개선함으로써 사이버보안 사고의 재발을 방지하도록 '개선'을 도입하여 총 7개의 평가 요소를 도출하였다. NIST CSF의 '대응 계획'은 UR E26의 '사고 대응 계획' 평가 요소와 동일하여 제외하였으며, '완화'는 UR E26의 평가 요소에 이미 포함된 개념을 포괄하는 상위 개념이므로 제외하였다.

도출된 평가 요소는 UR E26에 따라 사고 대응 계획의 수립 여부와 그 실행 가능성을 평가한다. 또한 선박의 주요 제어 시스템이 고장 날 경우를 대비한 로컬, 독립 및 수동 운전 시스템의 준비 여부를 평가하며, 사이버보안 사고 발생 시 보안 구역과의 네트워크 기반 통신 종료 가능 여부와 시스템의 최소 기능 유지 및 안전한 상태로의 전환 가능 여부를 평가한다. NIST CSF 평가 요소는 대응 과정에서 효과적인 의사소통과 분석이 이루어지는지 평가하며, 해당 과정에서 학습과 개선 절차가 적절히 이루어지는지 평가한다.

3.5 복구

복구 단계는 사이버보안 사건 이후 정상 운영 상태로 복구하기 위한 계획 및 절차를 수립하고 실행하는 데 중점을 둔다. 사이버보안 사건은 동일한 이유로 재발하지 않도록 교훈을 학습하고 개선해야 한다[14]. 이에 NIST CSF의 '개선'을 도입하였으며, 이해관계자 간의 효율적인 의사소통을 통한 효과적인 복구 활동을 위해 '커뮤니케이션'을 도입하여 총 5개의 평가 요소를 도출하였다. NIST CSF의 '복구 계획'은 UR E26의 '복구 계획' 평가 요소와 동일하여 제외하였다.

도출된 평가 요소는 UR E26에 따라 사이버보안 사고 이후 시스템을 복구하기 위한 구체적인 복구 계획 수립 여부를 평가하며 이를 기반으로 복구 절차가 효과적으로 운영되는지 평가한다. 또한, CBS 및 네트워크 백업의 정기적인 수행 여부를 평가하며, 사이버보안 사고 발생시 시스템 자체적으로 네트워크를 종료하고 재설정 및 롤백을 수행하여 재시작 되는 절차의 체계적인 수립 여부를 평가한다. NIST CSF의 평가 요소는 복구 과정을 통한 학습과 개선 프로세스의 수행 여부 및 복구를 수행하는 동안 이해관계자들의 효율적인 의사소통이 이루어지는지 평가한다.

IV. 결론 및 향후 연구방향

본 논문에서는 해양선박 환경에서 사이버 복원력의 중요성이 커짐에 따라, 이에 대한 평가 체계의 필요성이 증가하고 있음을 인식하고, 선박 대상 사이버 복원력 평가 요소를 개발하였다. 해양선박 환경에 특화된 평가 요소를 개발하기 위해 IACS UR E26의 요구사항을 기반으로 평가 요소를 도출하였으며, IACS UR E26의 기능 요소를 보완하기 위해 NIST CSF의 구성요소를 적용하여 평가 요소를 개발하였다. 향후 연구로는 본 논문에서 개발한 평가 요소를 활용하여 선박 대상 사이버 복원력 평가 체계 구축 연구를 하고자 한다.

Acknowledgment

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2024-00400955, 스마트선박 국제규정 대응을 위한 핵심 보안 기술 개발, 50%)과 원자력안전위원회의 재원으로 한국 원자력안전재단의 지원을 받아 수행한 원자력안전 연구사업의 연구결과입니다. (RS-2021-KN051410, 50%)

참 고 문 헌

[1] NIST, Cyber Resiliency, https://csrc.nist.gov/glossary/term/cyber_resiliency, [last access 2024/06/24]

[2] KISA, 침해사고 등 사이버 공격 대응 관점에서의 사이버 복원력(CyberResilience) 정책 이슈 분석 및 시사점, 2024.

[3] 보안뉴스, ICT 기반 차세대 선박의 사이버 위협 증가.'사이버 복원력' 중요성 부각, https://m.boannews.com/html/detail.html?idx=126909, [last access 2024/06/24]

[4] International Maritime Organization, "Resolution MSC.428(98) - Maritime Cyber Risk Management in Safety Management Systems", https://www.im o.org/en/OurWork/Security/Pages/Cyber-security. aspx

[5] IACS, IACS adopts new requirements on cyber safety, https://iacs.org.uk/news/iacs-adopts-new-requirements-on-cyber-safety, [last access 2024/06/24]

[6] 최재혁, et al. "국방정보시스템 사이버복원력 수준 평가를 위한 성숙도모델에 관한 연구." 정보보호학회논문지 29.5 2019.

[7] Haque, Md Ariful, et al. "Cyber resilience frame

work for industrial control systems: concepts, metrics, and insights." 2018 IEEE international conference on intelligence and security inform atics (ISI). IEEE, 2018.
[8] K. Tierney and M. Bruneau, "Conceptualizing and measuring resilience: A key to disaster loss reduction", TR news, no. 250, 2007.
[9] IACS, UR E26 Cyber resilience of ships, 2023.
[10] Cybersecurity, Critical Infrastructure. "Frame work for improving critical infrastructure cybe rsecurity." URL: https://nvlpubs. nist. gov/nist pubs/CSWP/NIST. CSWP 4162018, 2018., [last access 2024/06/24]
[11] 황재호, et al. "사이버 복원력 도입을 위한 평가모델 연구." 한국정보처리학회 학술대회논문집 30.2, 2023.
[12] 한국선급, 해상 사이버보안 시스템 지침, 2021
[13] 데일리시큐, 선박 ICT 공급망 사이버 공격 위협 증가, https://www.dailysecu.com/news/articl eView.html?idxno=142881, [last access 2024/ 06/24]
[14] IOSCO, Guidance on Cyber Resilience for Fina ncial Market Infrastructures, 2016, [last access 2024/06/24]

마이크로파 에너지 투과에 의한 아까시 종자의 발아율 향상에 관한 연구

*박동희

A Study on the Improvement of Germination Rate of Acacia(Robinia pseudoacacia) Seeds by Microwave Energy Penetration

*Dong-Hee Park

요 약

본 논문에서는 아까시 종자의 발아율 향상을 위하여 종자의 외부적 자극으로 2.45 GHz 마이크로파에너지를 이용하였다. 마이크로파 에너지는 습식방법으로 200 W와 400 W의 전력을 30s에서 180s까지 30s 단위로 투과시켜 발아율을 비교하였다. 실험결과 가장 높은 발아율은 200 W에 180s에서 53 %의 결과가 나타났다. 또한 참고문헌[5]에서 제시된 결과의 합당함을 입증하였다. 결과적으로 본 논문에서는 마이크로파 에너지의 크기보다는 투과시간을 조절하는 것이 발아율 향상에 유리하다는 것을 확인하였다.

Key words

Seed Germination, Robonica Pseudoacacia Seed, Microwave Energy

I. 서 론

기후변화의 영향이 전 지구적으로 확대됨에 따라 대응전략도 강화되고 있다. 2050 탄소중립은 2015년 파리기후협약의 목표에 부합하도록 2050년까지 온실가스 순 배출량을 제로로 만드는 것이다. 탄소중립의 목표는 재생에너지 사용을 확대함으로 이산화탄소 배출량을 축소하고, 또 식물의 광합성을 통해 흡수량을 증대하는 전략이다[1,2]. 기후변화대응에서 자연적 해결방법으로 산림면적확대의 중요성은 크게 강조되고 있다. 2020년을 기준으로 전 세계 산림 면적은 총 40.6억 헥타르(ha)로서, 전체 육지 면적의 약 31%를 차지하고 있다[3].

우리나라는 제6차 산림기본계획(2018 ~ 2037)을 수립하여 기후변화에 대응하기 위한 산림분야 주요정책을 추진 중이다. 우리나라의 여러 활엽수종 중 탄소 저장량을 비교한

* 한국교통대학교, 교수 (dhpark@ut.ac.kr)

결과는 아까시나무에서 가장 높게 나타났다 [4].

아까시나무의 묘목을 대량으로 생산하기 위한 씨앗 발아율 실험은 참고문헌 [5]에서 진행되었다. 이 논문에서 발아율 향상을 위한 외부 투과에너지는 전자기장에너지다. 전자기장에너지는 세 종류로 정자기장, 정전기장, 그리고 전자기장인 마이크로파장이다. 이 논문에서 정자기장의 크기는 6 mT, 정전기장의 크기는 190 mV/cm에서 680 mV/cm 범위 값, 그리고 2.45 GHz의 400 W 마이크로파에너지의 크기는 습식방법으로 0,8 W이다. 이 논문의 결과는 마이크로파 에너지 400 W를 180초간 투과시킨 종자에서 높은 발아율을 보였다[5].

본 논문에서는 참고문헌 [5]의 확장으로 마이크로파에너지의 크기와 투과시간을 세분화하여 아까시 종자의 발아율 특성을 고찰하였다.

Ⅱ. 재료 및 방법

2.1 재료

본 논문에서 사용된 아까시 종자는 2023년에 중국에서 수확한 종자로 11월에 구입한 것을 사용하였다. 이들 종자를 물에 60분간 침적시킨 후 10분간 3%의 과산화수소에서 소독하였다. 매회 실험에 사용된 종자 수는 100립씩 선택하였다.

2.2 실험방법

전자기장에너지로 마이크로파주파수는 2.45 GHz이고 에너지의 크기는 200 W와 400 W를 선택하였다. 사용된 용기는 깊이 8.8 cm를 갖는 500 mL 유리그릇이다. 에너지는 30초 단위로 180초까지 투과시켰으며, 이 때 씨앗에 가해진 실질적 에너지는 대략 200 W에서 0.4 W 그리고 400 W에서 0.8 W이며, 물의 온도는 대략 최소 25℃에서 최대 41℃사이가 된다.

Ⅲ. 결과 및 고찰

참고문헌 [5]에서 아까시 종자의 최종적인 발아율(GR)은 그림 1과 같다.

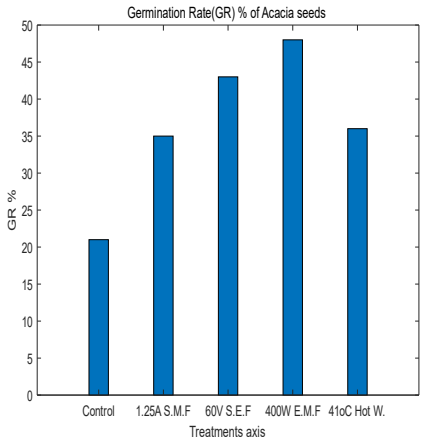


Figure 1. Comparison of germination rates of acacia seeds[5].

그림 1. 아까시 종자의 발아율 비교[5].

그림 1의 결과로부터 아까시 종자에 대한 대조군1의 최종 발아율은 21%인 것에 비교하여 정자기장 에너지를 투과한 종자의 발아율 35%, 정전기장 에너지를 투과한 종자의 발아율 43%, 마이크로파 에너지를 투과한 종자의 발아율은 48%, 그리고 열처리법에 의한 종자의 발아율 36%를 나타낸다. 마이크로파 에너지를 투과한 종자의 발아율이 가장 높게 나타났다. 이는 대조군1과 비교하여 27% 그리고 대조군2와 비교하여 12% 향상된 결과이다.

본 논문의 실험에 의한 발아율(GR) 결과는 그림 2에 제시하였다.

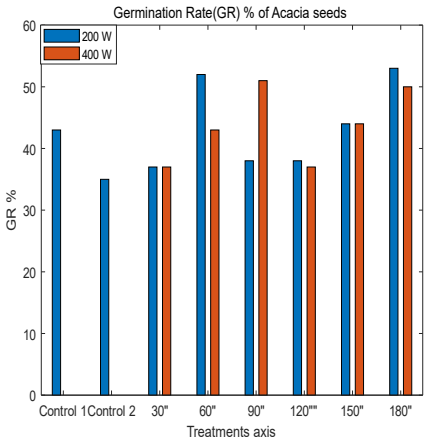


Figure 2. Comparison of germination rates of acacia seeds according to the magnitude and penetration time of mircowave energy.

그림 2. 마이크로파에너지의 크기 및 투과시간에 의한 아까시 종자의 발아율 비교.

그림 1과 2의 비교에서 대조군의 결과는 2차 실험에서 22 %가 높게 나타났다. 또 그림 1에서 마이크로파 에너지 400 W를 3분간 투과시킨 발아율 결과는 48 %인 반면에 그림 2에서 동일한 조건에서 발아율은 50 %이다. 그림 2에서 최대 발아율은 53 %로 200 W에 180s에서 나타났다. 다음으로는 60s에 52 %, 또 400 W 90s에 51 %, 180s에 50 %를 나타낸다.

이 결과로부터 마이크로파 에너지의 크기보다는 투과시간을 조절함으로 종자의 발아율을 향상시킬 수 있다.

IV. 결론

지구의 온도를 높이는 주된 원인인 이산화탄소를 저렴한 비용으로 포집하는 방법 중 하나는 자연적 방법으로 산림을 조성하고 가꾸는 일이다. 주요 산림 및 조림 수종에 있어 아까시나무는 탄소 흡수량이 우수하고 또 속성수로써 탄소 저장량 역시 매우 높다. 본 논문은 우수한 조림수종인 아까시 종자의 발아율을 향상시키기 위하여 외부적 자극으로 200 W 및 400 W의 크기를 갖는 마이크로파에너지를 30s에서 180s까지 투과시켜 발아율을 실험하였다. 가장 높은 발아율은 200 W 180s에서 53 %로 나타났다. 결과적으로 발아율을 향상시키기 위해서는 투과 에너지의 크기보다는 투과시간을 조절하는 것이 효과적이다.

본 논문의 후속 연구는 이 연구방법을 다른 수종으로 확대하여 종자의 발아율과 새싹의 성장에 대한 비교를 연구할 계획이다.

V. 감사의 글

2024년 한국교통대학교 산학협력단 지원을 받아 수행하였음.

참 고 문 헌

[1] 안현진, 이상민, 정호근, 김동욱, 탄소중립 실현을 위한 기후 스마트 산림경영 연구, 연구보고 R955, 한국농촌경제연구원, 2022.

[2] Park, Go Eun, Jung, Jong bin, Choi, Won Il, Kim, Eun-sook, and Yang, Hee Moon, "Considerations for enhancing the effectiveness of climate change adaptation information in the forest sector," Journal of Climate Change Research, Vol. 14, No. 6-2, pp. 957~964, 2023.

[3] Sun-jeoung Lee, Jong-su Yim, Jin-take Kang, Rae-hyun Kim, Yow-han Son, Gawn-su Park, Yeong-mo Son, "Application and Development of Carbon Emissions Factors for Deciduous Species in Republic of Korea", J. of Climate Change Research, Vol. 8, No. 4, pp. 393-399, Dec. 2017.

[4] Yeong-mo Son, So-won Kim, Sun-jeoung Lee and Jeong-soo Kim, "Estimation of Stand Yield and Carbon Stock for Robinia pseudoacacia Stands in Korea", J. of Korean Forest Society, Vol. 103, No. 2, pp. 264-269, Jun. 2014.

[5] Park, Donghee, Kwak, Yoonsik, Ko, Kyunbyoung, Kim, Hagwone, Mun, Cheol, Park, Manbok, Song, Seokil, Song, Changick, Lim, Sungmuk,

Jung, Hogi "A Study on the Efficient Germination of Acacia(Robinia pseudoacacia) Seeds using Electromagnetic Fields Energy", Journal of Platform Technology Vol. 12, No. 1, Feb. 2024.

스마트 컨트랙트 취약점 정보 공유 프레임워크 제안

*최동빈, **박용범

Smart Contract Vulnerability Information Sharing Framework Proposal

*Dong Bin Choi, **Young B. Park

요 약

스마트 컨트랙트가 사용되는 분산원장기술(DLT) 시스템에서 스마트 컨트랙트의 취약점에 대한 정보를 논문이나 보고서 등, 시스템 외적인 매체를 통해서 공유되고 있다. 시스템 외에서 정보가 공유되고 있기에 같은 시스템에 참여하는 참여자들이라고 할지라도 스마트 컨트랙트 취약점 정보를 각자 다르게 가지고 있을 가능성이 높다. 따라서 본 논문은 DLT 시스템 참여자들이 모두 동일한 스마트 컨트랙트 취약점 정보를 공유할 수 있는 프레임워크를 제안한다. 또한 프레임워크를 이용하여 스마트 컨트랙트 취약점 정보를 공유하기 위한 절차 및 해당 정보를 가지고 취약점 탐지기를 구성하는 등 활용할 수 있는 시나리오를 제시한다.

Key words

Smart Contract, Vulnerability, DLT, Classification, Information sharing framework

I. 서 론

분산원장기술(DLT) 시스템은 블록체인과 같이 참여자가 합의 알고리즘에 따라 합의된 내용은 원장에 기록하여 공유하는 시스템을 의미하며, 스마트 컨트랙트는 DLT 시스템에서 사용되는 프로그램으로, 일반적으로 원장에 기록되며, 특정 조건을 만족 시 해당 프로그램이 동작하는 특성을 가지고 있다.

원장에 기록된다는 특성으로 인해서 스마트 컨트랙트가 기록되면 이후 해당 스마트 컨트랙트는 수정이 어려우며, 특정 조건 만족 시 바로 실행된다는 특성으로 인해 만약 스마트 컨트랙트가 특정 취약점이 있을 경우 그 취약점을 지속해서 공격하여 피해의 규모가 다른 소프트웨어에 비해서 크게 발생한다는 특징을 지닌다[1][2].

스마트 컨트랙트가 많이 사용되는 이더리움의 경우 몇몇 연구 들을 통해서 많은 기록된 스마트 컨트랙트 중 많은 수가 취약점이 있음을 발견하였다[3][4][5][6][7]

* 단국대학교 컴퓨터학과, 박사수료 (dbchoi85@gmail.com)

** 단국대학교 소프트웨어학과, 교수, 교신저자 (ybpark@dankook.ac.kr)

[8][9][10][11].

다만 해당 연구들은 각 연구마다 다른 데이터 셋, 취약점 정보를 사용했기에, 한 연구에서 탐지된 취약점이 다른 연구에서는 제외되거나, 데이터 셋이 다르기에 정확도 면에서도 차이가 나는 문제가 발생한다.

해당 문제를 해결하기 위해서 본 논문은 동일한 DLT 시스템을 사용하는 참여자들이 스마트 컨트랙트 취약점 정보를 모두 동일하게 공유할 수 있는 프레임워크를 제안하여, 이를 활용하여 스마트 컨트랙트 취약점을 탐지할 수 있는 분류기 생성하는 시나리오를 제시한다.

Ⅱ. 기존연구

2.1 취약점 탐지기 별 성능

Thomas et al[1]은 8개의 취약점을 정의하고 해당 취약점이 존재하는 스마트 컨트랙트를 수집 후 9개의 스마트 컨트랙트 취약점 탐지기의 성능을 비교하였다.

그 결과 탐지기 별로 성능이 천차만별이었으며, 특정 탐지기의 경우 준비된 모든 취약점을 탐지하지 못하는 결과를 보여주었다.

2.2 데이터 셋 차이

Thomas et al[1]은 성능 비교와 마찬가지로 각 탐지기를 학습 혹은 제작하는데 사용한 데이터 셋을 비교하였으며, 그 결과 모든 취약점을 가지고 학습 혹은 제작된 탐지기는 없었으며, 특정 탐지기의 경우 매우 빈약한 데이터 셋이 사용되었음을 밝히고 있다.

Ⅲ. 제 안

3.1 개요

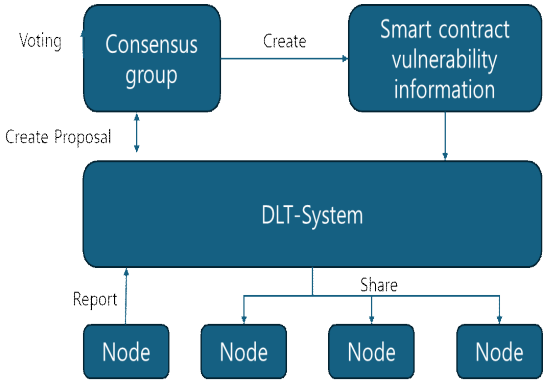


그림 1 프레임워크 개요도

그림 1은 본 논문이 제안하고자 하는 프레임워크의 개요도를 나타내며 구성요소는 아래와 같다.

a) 합의체(Consensus group) : 블록을 생성하기 위한 합의체가 아닌 스마트 컨트랙트 취약점 정보에 대한 합의를 위한 구성체

b) 참여자(Node) : 해당 분산원장기술 시스템의 참여자로, 스마트 컨트랙트 취약점 정보를 공유 받거나, 신규 스마트 컨트랙트 취약점을 보고하는 주체

c) 스마트 컨트랙트 취약점 정보(Smart contract vulnerability information) : 공유 되는 스마트 컨트랙트 취약점 정보의 총 합

3.2 공유 절차

프레임워크를 통해서 스마트 컨트랙트 취약점이 공유되는 절차는 다음과 같다.

1. 노드가 원장으로 취약점을 보고
2. 합의체는 원장에서 보고된 취약점 정보를 수집
3. 합의체는 수집된 취약점 정보에 관해서 투표를 통해 합의를 진행
4. 합의체는 합의된 취약점 정보를 원장에 게시
5. 노드는 게시된 취약점 정보를 확인

3.3 활용 예시

주소	소스 코드	컴파일된 코드	취약점 일련번호	취약점 명칭	취약점 설명
0x006699d34AA3013605d468d2755A2Fe59A16B12B	pragma solidity 0.5.4; interface IERC20 { function balanceOf(address account) external ...	0x6080604052348015610010576000080fd5b506004361061020257600035 7c0100000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000900...	4	오버플로우/언더플로우	산술 연산이 데이터 유형의 최대 또는 최소 크기에 도달할 때 발생

그림 2 스마트 컨트랙트 취약점 정보 예시

그림 2는 프레임워크를 통해서 공유되는 취약점 정보를 나타내며 해당 정보를 이용하여 그림 3과 같은 절차를 통해서 각 노드는 취약점 탐지기를 제작하여 활용할 수 있다.

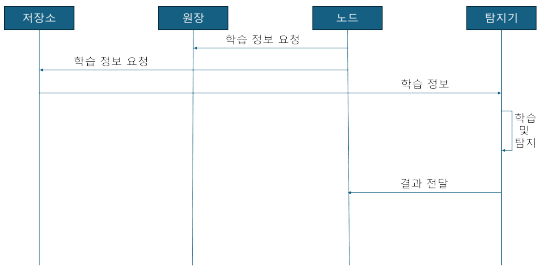


그림 3 정보 활용을 통한 탐지기 생성

감사의 글

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터 사업의 연구결과로 수행되었음 “(IITP-2024-RS-2023-00259099)

참 고 문 헌

[1] Thomas Durieux, João F. Ferreira, Rui Abreu, Pedro Cruz, Empirical Review of Automated Analysis Tools on 47,587 Ethereum Smart Contracts, Proceedings of the 2nd International Conference on Software Engineering

[2] Josselin Feist and Gustavo Grieco and Alex Groce, Slither: A Static Analysis Framework for Smart Contracts, 2019 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB)

[3] Christof Ferreira Torres, Mathis Steichen, and Radu State. 2019. The Art of The Scam: Demystifying Honeypots in Ethereum Smart Contracts. In 28th USENIX Security Symposium (USENIX Security 19). USENIX Association, Santa Clara, CA, 1591–1607. https://www.usenix.org/conference/usenixsecurity19/presentation/ferreira

[4] Ivica Nikolić, Aashish Kolluri, Ilya Sergey, Prateek Saxena, and Aquinas Hobor. 2018. Finding the greedy, prodigal, and suicidal contracts at scale. In Proceedings of the 34th Annual Computer Security Applications Conference. ACM, New York, NY, USA, 653–663.

[5] Mark Mossberg, Felipe Manzano, Eric Hennenfent, Alex Groce, Gustavo Grieco, Josselin Feist, Trent Brunson, and Artem Dinaburg. 2019. Manticore: A User Friendly Symbolic Execution Framework for Binaries and Smart Contracts. arXiv:1907.03890

[6] Bernhard Mueller. 2018. Smashing ethereum smart contracts for fun and real profit. In 9th Annual HITB Security Conference (HITBSecConf). HITB, Amsterdam, Netherlands, 54.

[7] Christof Ferreira Torres, Julian Schütte, et al. 2018. Osiris: Hunting for integer bugs in ethereum smart contracts. In Proceedings of the 34th Annual Computer Security Applications Conference. ACM, New York, NY, USA, 664–676.

[8] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. 2016. Making smart contracts smarter. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security. ACM, New York, NY, USA, 254–269.

[9] Petar Tsankov, Andrei Dan, Dana Drachsler-Cohen, Arthur Gervais, Florian Buenzli, and Martin Vechev. 2018. Securify: Practical security analysis of smart contracts. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. ACM, New York, NY, USA, 67–82.

[10] Josselin Feist, Gustavo Greico, and Alex Groce. 2019. Slither: A Static Analysis Framework for Smart Contracts. In Proceedings of the 2Nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB '19). IEEE Press, Piscataway, NJ, USA, 8–15. https://doi.org/10.1109/WETSEB.2019.00008

[11] Sergei Tikhomirov, Ekaterina Voskresenskaya, Ivan Ivanitskiy, Ramil Takhaviev, Evgeny Marchenko, and Yaroslav Alexandrov. 2018. Smartcheck: Static analysis of ethereum smart contracts. In 2018 IEEE/ACM 1st International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB). IEEE, Gothenburg, Sweden, Sweden, 9–16.

산업제어시스템 환경 대상 사이버사고 대응 체계 지침 비교 분석

*최희원, **서정택

A Comparative Analysis of Guidelines for Cyber Incident Response Systems in Industrial Control System Facilities

HeeWon Choi, Jung Taek Seo***

요 약

과거의 산업제어시스템(Industrial Control Systems, ICS) 환경은 폐쇄된 네트워크망 내에서 독점적인 제어 프로토콜을 사용했으나, 현대에는 IT 시스템과의 통합으로 외부 네트워크와의 접점이 증가하면서 사이버공격의 위험이 커지고 있다. 이에 따라 ICS 환경에 대한 사이버사고 발생 빈도는 매년 증가하고 있으다. 세계 각국은 이러한 사이버공격에 대응하기 위해 지침 및 가이드라인을 제공하고 있으며, 미국 NCCIC(National Cybersecurity and Communication Integration Center)는 NIST(National Institute of Standards and Technology)는 컴퓨터보안 영역을 포함하는 사이버사고 대응에 대한 지침으로 현재 발간되고 있는 ICS 환경 대상 사이버사고 대응 및 훈련에 관한 지침의 밑바탕이 되었다. 또한, DHS(Department of Homeland Security)에서 발간한 지침은 ICS 환경을 대상으로 하는 사이버사고 대응 체계에 대한 지침을 제시한다. 두 지침은 사이버사고 대응 체계의 모델을 구성하고, 각 단계별 요구사항을 권장하고 있으나, 두 모델의 구성과 목적성에 따른 차이점이 존재한다. 본 논문은 사이버사고 대응 체계에 대한 두 가지 지침에 대해 비교 분석하였다.

Key words

Industrial Control System, Computer security incident, Incident response, Incident handling

Ⅰ. 서 론

IT 시스템과 통합되면서 외부 네트워크와의 접점이 증가하게 됨에 따라 산업제어시스템(Industrial Control Systems, ICS) 환경에 대한 사이버공격과 사이버사고 발생 빈도가 매년 증가하고 있으며, 기술적인 측면에서도 지속적으로 발전하고 있다[1]. 이러한 사이버공격에 대응하기 위해 세계 각국의 기관들은 사이버사고 대응 체계 관련 지침 및 가이드라인을 제공하여, 조직이 사이버사고가 발생했을 때 적절한 대응을 신속하고 체계

* 가천대학교 정보보호학과, 석사과정 (chw1226@gachon.ac.kr)

** 가천대학교 컴퓨터공학부, 교수, 교신저자 (seojt@gachon.ac.kr)

적으로 수행할 수 있도록 훈련을 권장한다.

미국의 사이버보안 및 통신 통합 센터(National Cybersecurity and Communications Integration Center, NCCIC)는 NIST(National Institute of Standards and Technology)와 DHS(Department of Homeland Security)의 사이버사고 대응 체계를 대표적인 ICS 환경 대상 지침으로 지정하였다[2]. NIST에 발간한 NIST 800-61은 컴퓨터보안 전반을 포함하는 사이버사고 대응 체계에 대한 지침으로, 이후 발간된 ICS 대상 지침의 기초가 된다. DHS 지침은 ICS를 대상으로 하는 대표적인 사이버사고 대응 체계이다. ICS 환경을 대상으로 하는 사이버사고 대응 훈련은 NIST 800-61과 DHS 지침과 같은 사이버사고 대응 체계 지침을 참고하여 잘 준비되고 훈련의 목적과 적합하게 구성된 체계를 바탕으로 이루어져야 한다. 본 논문은 두 지침을 비교 분석하여 목적성, 사이버사고 대응 체계 절차 및 사이버사고 대응 팀 구성 및 역할 정의에 대한 차이점을 도출하였다.

본 논문의 구성은 다음과 같다. 2장에서 사이버사고 대응 체계에 대한 개요를 기술하며, 3장에서 NIST 800-61과 DHS 지침을 비교분석하여 도출된 비교항목을 설명한다. 그리고 4장에서 결론으로 마무리한다.

Ⅱ. 사이버사고 대응 체계 개요

사이버사고는 컴퓨터 시스템 및 네트워크에 실제 또는 잠재적인 영향을 미치는 모든 사고와 명시적 및 묵시적 보안정책을 위반하는 행위를 의미한다[3].

이에 대응하기 위해 ICS 환경에서의 사이버사고 대응 체계는 단일 조치가 아닌 컴퓨터보안 사고의 탐지와 사고의 완화 및 복구까지 포함하는 프로세스로 구성된다. 미국 NRC(Nuclear Regulatory Commission)[4] 및 영국 ONR(Office for Nuclear Regulation)[5] 등의 규제기관에 따르면, 사이버사고 대응 체계는 준비 및 예방, 식별, 완화(억제/근절/복구), 사후 활동의 4가지 단계를 포괄한다. [표 1]는 각 단계에 대해 세부 내용을 나타낸다.

[표 1] 사이버사고 대응 체계 단계별 세부 설명

단계	세부 내용
준비 및 예방	• 사이버사건 대응에 대한 모든 이해관계자의 역할 및 책임 정의, 정책 수립, 절차 구현, 자산 식별 • 시스템을 사이버공격으로부터 보호하기 위한 표준, 가이드, 소프트웨어 도구 등의 자원 확보
식별	• 네트워크, 시스템 또는 애플리케이션에서 비정상적인 활동 또는 사고의 징후를 탐지하고 식별하는 과정 • 신속하게 문제를 발견하여 대응할 수 있도록 하는 것을 목표
완화 (억제/근절/복구)	• 식별된 사고를 조사하고, 사고로 인한 원인을 제거하며 영향을 받은 시스템을 복구하여 정상 상태로 돌아가기 위한 조치를 실행하는 과정
사후 활동	• 사고 대응 과정을 검토하고, 사고의 원인 및 대응의 효과성을 분석하여 미래의 대응을 개선하는 단계 • 사고 대응 절차, 정책, 도구를 개선하여 잠재적인 보안 취약점을 해결하는 것을 목표

Ⅲ. NIST와 DHS 사이버사고 대응 체계 지침 비교 분석

본 장에서는 사이버사고 대응 체계에 대한 NIST 및 DHS 사이버사고 대응 체계 지침에 대해 목적성, 사이버사고 체계 절차, 사이버사고 대응 팀 구성 및 역할 항목에 대해 비교 분석한다. NIST에서 발간한 NIST 800-61은 컴퓨터보안 영역에 대한 사고와 관련된 데이터를 분석하고, 각 사고에 대한 적절한 대응을 결정할 수 있는 사고 대응 사이클(IRC, Incident Response Cycle)

모델을 제공하는 지침이다[6]. DHS에서 발행한 Developing an Industrial Control Systems Cybersecurity Incident Response Capability 지침은 ICS 환경을 대상으로 하는 잠재적인 공격에 대해 대응하기 위한 대응 절차의 핵심 요소를 제공한다[7].

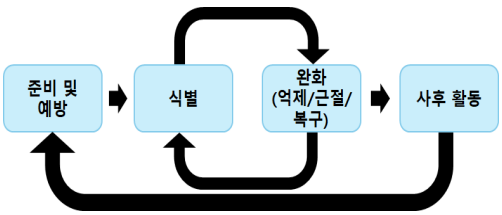
3.1 사이버사고 대응 체계 목적성

NIST는 기업, 정부 기관, 표준 기관 등을 대상으로 사이버보안 지침 및 교육을 지원하는 기관으로써[8], IT 환경 전반에 걸쳐 발생할 수 있는 다양한 사이버사고에 대한 포괄적인 대응 체계를 제공한다. 모든 종류의 조직이 활용할 수 있는 일반적인 가이드라인을 제공하며, 광범위한 사이버사고를 다루기 위한 목적으로 설계되었다.

반면에, DHS는 자연 및 인위적 재난에 대한 긴급 대응, 반테러 업무 및 사이버 보안과 관련된 지침을 다루는 기관으로써[9], ICS 환경에 특화된 사이버사고 대응 체계를 제공한다. ICS 환경의 특수한 요구사항과 운영 환경을 고려하여 시스템의 안정성과 가용성을 보장하는 데 중점을 두고 있다.

3.2 사이버사고 대응 체계 절차

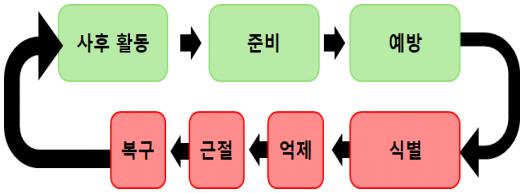
NIST 800-61에서 제공하는 IRC 모델은 [그림 1]과 같이 순환적인 구조를 포함하는 순서로 나타난다. 해당 모델은 2장에서 언급한 4가지 단계를 포괄한다.



[그림 1] NIST IRC 모델

해당 모델은 식별 단계와 억제/근절/복구 단계를 추가적인 사고 징후가 나타나지 않을 때까지 반복한다. 이는 추가적인 사고 징후가 나타날 때까지 이 두 단계를 반복함으로써 문제를 완전히 해결하고자 한다. 또한, 반복 주기를 통해 조직이 다양한 사이버사고 시나리오에 유연하고 대응할 수 있도록 한다. 이 과정은 사이버사고의 근본적인 원인을 제거하고, 시스템의 무결성을 확보하는 데 중점을 둔다.

반면에, DHS 지침에서 제공하는 사이버사고 대응 체계는 [그림 2]와 같이 직선적인 접근 방식으로 구성된다.



[그림 2] DHS 사이버사고 대응 체계

DHS에서 제시하는 사이버사고 대응 체계는 각 단계를 명확히 구분하고, 한 단계가 끝나면 다음 단계로 넘어가는 직선적인 접근 방식을 채택한다. 이는 ICS 환경이 IT 환경과 달리 시스템의 안정성과 가용성을 최우선으로 고려해야 하기 때문이다. 따라서 반복적인 사이클보다는 명확한 단계별 접근을 통해 신속하고 명확하게 사고를 해결하는 데 중점을 둔다. 특히 ICS 환경에서 시스템 가동 중단은 물리적 안전에 큰 영향을 미칠 수 있어, 문제를 신속히 탐지하고 해결한 후 복구하는 것이 중요하다.

3.3 사이버사고 대응 팀 구성 및 역할 정의

NIST는 다양한 IT 전문가로 구성된 사이버사고 대응 팀을 권장한다. 이는 네트워크 전문가, 시스템 전문가, 보안 분석가 등이 포함되며, 각자의 전문성을

바탕으로 사고 대응의 각 단계를 협력하여 수행한다. 이는 다양한 사이버사고를 다루기 위한 폭넓은 기술적 역량과 경험이 요구된다. 반면에, DHS는 ICS 전문가와 엔지니어가 포함된 팀 구성을 권장한다. 이는 ICS 환경의 특수한 요구사항과 운영 환경을 깊이 이해하는 전문가들이 필요하기 때문이다. ICS 시스템의 안정성과 물리적 장비의 운영을 보장하기 위해, 해당 시스템에 대한 깊은 이해와 경험을 가지는 전문가들이 중요하다.

[표 2]는 앞서 분석한 세 가지 비교 항목에 대한 NIST와 DHS의 두 지침의 내용을 요약한 표이다.

IV. 결 론

본 논문에서는 NIST 800-61과 DHS의 “Developing an Industrial Control Systems Cybersecurity Incident Response Capability” 지침을 비교 분석하여 세 가지 주요 차이점을 도출하였다. 첫 번째 주요 차이점은 사이버사고 대응 체계의 목적성이다. NIST 800-61은 IT 환경 전반의 광범위한 사이버사고에 대응하도록 설계된 반면, DHS 지침은 ICS 환경의 특수성을 반영하여 시스템의 안정성과 가용성을 보장하는 데 중점을 둔다. 또한, NIST의 사이버사고 대응 체계는 반복적인 사이클을 통해 지속적인 위협 모니터링과 대응을 강조하여 설계되었지만, DHS는 각 단계를 명확히 구분하여 직선적인 접근 방식을 채택함으로써 신속하고 명확한 사고 해결을 목표로 한다. 마지막으로 큰 차이점이 나타난 항목은 사이버사고 대응 팀 구성 및 역할이다. NIST는 다양한 IT 전문가로 구성된 팀을 권장하는 반면에, DHS는 ICS 전문가와 엔지니어가 포함된 팀 구성을 권장하여 ICS 환경의 특수한 요구사항을 충족시킨다.

ICS 환경이 IT 시스템과 통합되면서 외부 네트워크와의 접점이 증가하게 되었으며, 이에 따라 이메일 서버 및 웹 서버와 같은 IT 시스템을 통한 사이버공격이 주로 발생하고 있다[10][11]. 따라서, IT 환경 전반의 광범위한 사이버사고에 대응하도록 설계된 NIST 800-61과 ICS 환경의 특성을 반영하여 사이버사고에 대응하도록 설계된 DHS 지침을 각 환경의 특성에 맞게 적절히 반영하여 사이버사고 대응 체계를 구축해야 한다.

향후 연구로는 각 ICS 환경의 특성에 따라 여러 지침의 요구사항을 설계할 수 있는 방법론을 제시하고자 한다.

Acknowledgment

이 논문은 2024년도 정부(산업통상 자원부)의 재원으로 한국에너지 기술평가원의 지원을 받아 수행된 연구 (20224B10100020,

[표 2] NIST와 DHS의 사이버사고 대응 체계 지침 요구사항 비교 분석표

비교 항목	NIST 사이버사고 대응 체계 지침	DHS 사이버사고 대응 체계 지침
사이버사고 대응 체계의 목적성	• IT 환경 전반에 걸쳐 발생할 수 있는 다양한 사이버사고에 대한 포괄적인 대응 체계 제공	• ICS 환경의 특수한 요구사항 및 운영 환경을 고려한 사이버사고 대응 체계 제공
사이버사고 대응 체계 절차	• 조직 내 시스템의 무결성을 고려한 식별 단계와 완화 단계의 반복 주기 형성	• 조직 내 시스템의 안정성을 고려한 각 명확한 단계 구분 • 조직 내 시스템의 가용성을 고려한 직선 형태의 사이버사고 대응 체계 제시
사이버사고 대응 팀 구성 및 역할	• 다양한 사이버사고를 다루기 위한 폭넓은 기술적 역량을 가진 다양한 IT 전문가로 구성된 사이버사고 대응 팀 권장	• ICS 환경의 특수한 요구사항과 운영 환경을 깊이 이해하는 전문가들로 구성된 사이버사고 대응 팀 권장

원전 사이버위협 대처 시스템 설계 및 시험검증 기술개발 (APR1400국산화 MMIS), 50%)이며, 원자력안전위원회의 재원으로 한국원자력 안전재단의 지원을 받아 수행한 원자력안전연구사업의 연구결과 (RS-2021-KN051410, 50%)임.

참 고 문 헌

[1] 송재구 외 10인, "국가원자력시설 사이버공격 대응체계 구축 및 운영", 한국원자력연구원, 2021.
[2] NCCIC, "Preparing for NCCIC ICS Cyber Incident Analysis", https://www.cisa.gov/sites/default/files/FactSheets/NCCIC%20ICS_FactSheet_Cyber_Incident_Analysis_S508C.pdf, Accessed on June, 2024.
[3] NIST, https://csrc.nist.gov/glossary/term/cyber_incident, Accessed on June, 2024.
[4] NRC, https://www.nrc.gov/reading-rm/doc-collections/cfr/part073/full-text.html, Accessed on June, 2024.
[5] ONR, "Security Assessment Principles". https://www.onr.org.uk/publications/regulatory-guidance/regulatory-assessment-and-permissioning/security-assessment-principles-syaps/security-assessment-principles-syaps/, Accessed on June, 2024.
[6] NIST, "Computer Security Incident Handling Guide", https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf, Accessed on June, 2024.
[7] U.S. Department of Homeland Security, "Developing an Industrial Control Systems Cybersecurity Incident Response Capability", https://www.cisa.gov/sites/default/files/2023-01/final-RP_ics_cybersecurity_incident_response_100609.pdf, Accessed on June, 2024.
[8] NIST, https://www.nist.gov/standardsgov/what-we-do, Accessed on June, 2024.
[9] U.S. Department of Homeland Security, https://www.usa.gov/agencies/u-s-department-of-homeland-security, Accessed on June, 2024.
[10] Kaspersky, "BlackEnergy APT Attacks in Ukraine", https://www.kaspersky.com/resource-center/threats/blackenergy, Accessed on June, 2024.
[11] Industrialcyber, "Oldsmar water treatment plant incident allegedly caused by human error, not remote access cybersecurity breach", https://industrialcyber.co/utilities-energy-power-water-waste/oldsmar-water-treatment-plant-incident-allegedly-caused-by-human-error-not-remote-access-cybersecurity-breach, Accessed on June, 2024.

페르소나 기법을 통한 Intent Translation 방법론

*장재원, **이소연, ***김대영

Intent Translation Methodology with Personas Techniques

*Jae Won Jang, **So-Yeon Lee and ***Dae-Young Kim

요 약

기존의 네트워크 관리 방식은 복잡한 설정과 많은 인력 투입으로 인해 오류가 자주 발생하고, 동적인 네트워크 요구사항 변화에 신속하게 대응하기 어려운 한계가 있다. 이를 해결하기 위해 IBN이 도입되었으며, 이는 네트워크 관리자가 원하는 바를 고수준의 의도로 표현하면 네트워크 설정이 자동으로 이루어지는 기술이다. 본 논문에서는 스마트 팩토리 환경에서 페르소나를 도출하고, 각 페르소나의 요구사항을 Intent Translation을 통해 저수준의 시스템 명령어로 변환하는 과정을 제안한다. 이러한 접근법은 사용자 의도를 명확히 파악하는 과정에서 페르소나 기법 적용의 효과성을 보여준다.

Key words

Intent-based Networking, Intent Translation, Persona, Scenario, Smart Factory

Ⅰ. 서 론

최근 네트워크 기술의 빠른 발전과 함께, 복잡한 네트워크 환경에서의 관리와 운영의 중요성이 더욱 부각되고 있다. 기존의 네트워크 관리 방식은 서비스 품질을 보장하기 위해 복잡한 설정과 많은 인력의 투입이 요구되며, 이는 오류를 유발하고 효율적인 네트워크 운영을 저해한다[1]. 이러한 문제를 해결하기 위해 등장한 Intent-based Networking (IBN)은 사용자의 의도를 네트워크 구성과 운영에 반영하여 자동화된 네트워크 관리를 가능하게 한다. IBN은 네트워크 관리자가 원하는 바를 고수준의 의도로 표현하면, 이를 기반으로 네트워크 설정이 자동으로 이루어지게 하는 기술이다[2]. IBN의 핵심 요소 중 하나인 Intent Translation은 사용자의 고급 네트워크 요구사항을 구체적이고 실행 가능한 네트워크 설정으로 변환하는 과정을 의미한다[3]. 네트워크 환경이 복잡해짐에 따라, 네트워크

* 순천향대학교 소프트웨어융합학과 석사과정 (fndl48@sch.ac.kr)

** 순천향대학교 소프트웨어융합학과 박사과정 (lsy8647@sch.ac.kr)

*** 순천향대학교 컴퓨터소프트웨어공학과 교수, 교신저자 (dyoung.kim@sch.ac.kr)

관리자들이 일일이 수동 설정을 수행하는 것은 매우 어려운 일이다. 이에 따라 요구사항을 세밀하게 설정하는 것이 필요하며, 본 논문에서는 페르소나 기법을 통한 Intent Translation 방법론을 제안한다. 페르소나 기법은 다양한 사용자 유형과 그들의 요구사항을 기반으로 한 가상의 시나리오를 만드는 기법으로 시스템 디자인 과정에서 흔히 사용되는 도구이다[4]. 따라서 사용자 의도를 명확하게 이해하고 이를 네트워크 설정으로 변환하기 위한 Intent Translation 과정에 페르소나 기법을 적용하는 것은 효과적이다.

Ⅱ. 제안하는 방법론

본 연구에서는 다양한 IoT 기기들이 상호 연결되어 실시간 데이터 교환이 이루어지고, 높은 자동화와 생산성 향상을 목표로 하는 스마트 팩토리 환경에서 Intent Translation을 위해 페르소나와 이에 따른 시나리오를 도출하는 과정을 담고 있다. 스마트 팩토리 네트워크 환경엔 센서, 로봇, 기계 및 각종 생산 설비들이 다양하게 연결되어 있어 효율적이고 안정적인 네트워크 관리가 필수적이다. 이러한 환경에서는 네트워크 관리자들이 수동적으로 모든 장치의 설정을 수행하는 데 한계가 있기에, IBN과 같은 자동화된 네트워크 관리 접근법이 필요하다.

2.1 네트워크 기능 유형화

가상의 스마트 팩토리 생산라인의 목표는 제품의 불량을 자동으로 검출하는 것이다. 생산라인에는 제품 불량 검출을 위한 카메라 센서와 생산라인 정상 작동 감지를 위한 진동 센서, 온도 센서, 광 센서가 부착되어 있다. 이러한 환경에 요구되는 네트워크 기능은 표 1과 같다.

표 1. 스마트 팩토리 기능 유형화

네트워크 유형화	세부 유형
네트워크 개수	총 네트워크 개수: 3 네트워크 1, 네트워크 2, 네트워크 3
네트워크 환경	이더넷, WiFi
토폴로지	총 토폴로지 개수: 3 tree, bus, ring
대역폭	초기 대역폭 값 100Mbps, 50Mbps, 10Mbps

스마트 팩토리의 네트워크 기능이 표 1과 같은 형태로 구성됨에 따라, 이를 효율적으로 관리하고 운영하기 위해서는 각각의 역할에 맞춘 접근이 필요하다. 이러한 요구를 충족시키기 위해, 업무에 따라 도출된 페르소나는 표 2와 같다.

표 2. 업무에 따른 페르소나

페르소나	특성
생산관리자	공장의 생산 과정 전체 총괄 역할 자동화된 생산라인의 지속적인 모니터링
IoT 기술자	여러 센서와 IoT 장치 설치 및 유지보수 역할 센서 연동을 통해 IoT 장비 구동
데이터 분석가	생산 데이터 수집 및 분석 역할 데이터 관리 및 생산 프로세스 패턴 분석

2.2 페르소나 시나리오

도출된 페르소나에 따른 구체적인 시나리오를 통해 이들의 요구사항을 보다 명확히 이해할 수 있다. 표 3은 스마트 팩토리에서 각 페르소나가 직면할 수 있는 구체적인 시나리오를 설명한다.

표 3. 페르소나 시나리오

생산 관리자
"생산관리 시스템은 실시간으로 데이터를 모니터링해야 하며, 최소 500개의 센서 데이터를 1초마다 수집할 수 있어야 한다. 이를 위해 네트워크의 최소 대역폭은 200Mbps여야 한다."
"생산 과정의 모든 데이터를 클라우드 서버로 전송하여 중앙 집중식으로 관리해야 한다. 이를 위해 네트워크는 최소 1Gbps의 대역폭을 제공해야 한다."
"두 개의 온도 센서(T1, T2)와 압력 센서(P1) 는 각각 다음과 같은 고정 IP 주소로 설정되어 네트워크에 연결된다. T1 - 192.168.1.10, T2 – 192.168.1.11, P1 – 192.168.1.12.
IoT 기술자
"스마트 팩토리의 IoT 장치는 수천 개의 센서를 동시에 관리해야 한다. 모든 장치의 원활한 작동을 위해 네트워크는 최소 1Gbps의 대역폭을 지원해야 하며, 안정적인 연결을 위해 이더넷 환경이 필요하다."
"IoT 장치들은 사이버 보안을 고려하여 안전하게 네트워크에 접근해야 한다. 이를 위해 네트워크에 엄격한 보안 프로토콜이 필요하다."
"각 센서는 동적으로 IP 주소를 할당받으며, 네트워크 관리자는 DHCP 서버를 통해 센서의 IP 주소를 관리한다. 온도 센서(T1, T2)와 압력 센서(P1)는 동적 IP 주소 할당을 받고, 진동 센서(V1)와 광센서(L1)는 정적 IP 주소를 할당받는다."
데이터 분석가
"데이터 분석 시스템은 매일 1TB 이상의 데이터를 처리해야 하며, 이를 위해 데이터 전송 속도가 중요하다. 따라서 네트워크 대역폭은 최소 500Mbps여야 하며, 데이터 손실 없이 안정적인 전송을 위해 Ring 토폴로지가 필요하다."
"버스 토폴로지를 사용하는 네트워크 2는 WiFi 기반으로 구성되어 사무실 내에서 데스크톱 컴퓨터와 모바일 장치 간의 데이터 공유를 지원한다. 사무실 내의 오피스 어플리케이션 및 클라우드 서비스 접근을 위해 최소 50Mbps의 대역폭이 필요하다.""
"네트워크 3은 이더넷 기반의 링 토폴로지를 사용하여 제품 불량 카메라 센서와 연결된다. 설치된 카메라는 실시간으로 영상 데이터를 전송하며, AI 모델에 입력되어 불량을 판단한다. 각 카메라가 안정적으로 연결되어야 하므로 최소 10Mbps의 대역폭이 필요하다."

2.3 Intent Translation 결과

표 3과 같이 도출된 시나리오는 IBN에 입력되는 고수준의 사용자 의도로, 이를 저수준의 시스템 언어로 번역하는 Intent Translation 과정이 필요하다. 따라서, 고수준의 사용자 의도를 저수준의 시스템 언어로 변환하기 위해 의도 속 핵심 키워드를 Entities라 칭하고 변환 결과는 표 4와 같다.

표 4. Intent Translation 결과

순번	Intent	Entities
1	"생산관리 시스템은 실시간으로 데이터를 모니터링해야 하며, 최소 500개의 센서 데이터를 1초마다 수집할 수 있어야 한다. 이를 위해 네트워크의 최소 대역폭은 200Mbps여야 한다."	• 실시간 • 1초 • 200Mbps
2	"생산 과정의 모든 데이터를 클라우드 서버로 전송하여 중앙 집중식으로 관리해야 합니다. 이를 위해 네트워크는 최소 1Gbps의 대역폭을 제공해야 한다."	• 클라우드 서버 • 중앙 집중식 • 1Gbps
3	"두 개의 온도 센서(T1, T2)와 압력 센서(P1) 는 각각 다음과 같은 고정 IP 주소로 설정되어 네트워크에_ 연결된다. T1 – 192.168.1.10 T2 – 192.168.1.11 P1 – 192.168.1.12	• 온도 센서 • 압력 센서 • IP 주소 • 192.168.1.10 • 192.168.1.11 • 192.168.1.12 • 고정 IP

표 4는 고수준의 의도를 Entities로 변환한 예시다. 이를 통해 각 의도는 구체적인 네트워크 엔티티로 매핑되며, 이러한 엔티티는 시스템 언어로 변환되어 실제 네트워크 설정에 적용된다. 다음으로, 도출된 Entities에 따라 시스템 언어로 매칭되는 예시를 표 5에서 확인할 수 있다. 표 5는 표 4의 3번 Intent에서 나온 Entities에 따른 예시를 보여준다.

표 5. Entities와 매핑된 시스템 언어

Entities	시스템 언어
• 192.168.1.10 • 192.168.1.11 • 192.168.1.12	ping 192.168.1.10 ping 192.168.1.11 ping 192.168.1.12
• 온도 센서 1 • 온도 센서 2 • 압력 센서	ip link set eth0 name T1 ip link set eth1 name T2 ip link set eth2 name P1
• 고정 IP	auto T1, T2, P1 iface T1, T2, P1 inet static address 192.168.1.10 address 192.168.1.11 address 192.168.1.12

표 4와 5와 같이 제시된 Intent Translation 결과와 시스템 명령어 매핑을 통해, 스마트 팩토리 환경에서의 다양한 네트워크 요구사항이 어떻게 구체적인 엔티티로 변환되고, 실제 네트워크 설정에 반영되는지 이해할 수 있다. 이와 같이 의도에서 엔티티, 그리고 엔티티에서 시스템 언어로의 변환 과정을 통해, 네트워크 관리의 자동화와 최적화가 가능하며, 이는 스마트 팩토리의 운영을 최적화하는 데 중요한 역할을 한다.

Ⅲ. 결 론

네트워크 서비스 품질 향상을 위해 복잡한 처리 로직을 수동으로 설정하는 기존 네트워크 관리 방식은 오류가 자주 발생하며, 동적인 네트워크 요구사항 변화에 신속하게 대응하기 어려운 한계가 존재한다. 이에 따라 본 연구는 기존 한계점을 해결하기 위해 등장한 IBN의 핵심 요소인 Intent Translation에 페르소나를 적용하는 이론적 방법론 제안을 목표로 한다. 제안하는 페르소나 기법을 통한 Intent Translation 방법론은 Intent 도출을 위해 다양한 사용자 유형과 요구사항을 기반으로 한 가상의 시나리오를 만드는 페르소나 기법을 적용하여, 사용자의 의도를 명확하게 이해하고 이를 구체적이고 실행 가능한 네트워크 설정으로 변환하는 데 도움이 된다. 이는 IBN의 효과적인 구현에 중요한 역할을 할 것이며, 향후 네트워크 관리의 방향성과 기술 발전에 큰 기여를 할 것으로 기대된다. 향후 연구로 제안한 접근법이 실제 네트워크 환경에 적용될 수 있도록 구현하는 방법을 개발할 것이다.

Acknowledgement

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1C1C1013133)

참 고 문 헌

[1] B. Li, X. Deng and P. Zhang, "A Policy Conflict Detection Mechanism for Intent-based Networking," 2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS), Dali, China, pp. 164-175, 2023.

[2] Y. Njah, et al. "Toward Intent-Based Network Automation for Smart Environments: A Healthcare 4.0 Use Case," in IEEE Access, vol. 11, pp. 136565-136576, 2023.

[3] A. Leivadeas and M. Falkner, "A Survey on Intent-Based Networking," in IEEE Communications Surveys & Tutorials, vol. 25, no. 1, pp. 625-655, 2023.

[4] Y. Luo and H. Zhang, "A Study of Smart Home User Personas Based on Context Theory," 2022 8th International HCI and UX Conference in Indonesia (CHIuXiD), Bali, Indonesia, pp. 88-93, 2022.

ICS 네트워크 패킷의 불확실성을 학습하기 위한 특징 수준 융합 기반의 멀티모달 학습 분석

*이주현, **전승호, ***서정택

Multimodal learning analysis based on feature-level fusion for learning uncertainty of ICS network packets

*Ju Hyeon Lee, **Seungho Jeon and ***Jung Taek Seo

요 약

Industrial Control Systems(ICS)의 다양한 사이버공격을 탐지하기 위하여 인공지능기반의 이상탐지 연구가 수행되고 있다. 대부분의 연구들은 이상탐지 성능을 높이기 위하여 대량의 ICS의 데이터를 학습하거나 인공지능 모델 구조를 변형시켰다. 하지만 운용 환경에 따라서 모델 구조는 변형이 어려운 환경이 존재하고, 데이터의 양이 한정적이다. 이러한 경우 모델의 구조를 최대한 유지하면서 성능을 개선하는 방법이 필요하다. 이에 본 논문에서는 ICS 네트워크 패킷의 불확실성을 학습하기 위한 특징 수준 융합 기반의 멀티모달 학습 방법을 분석하였다. 본 논문은 동일한 ICS 네트워크 패킷의 특징을 융합하여 학습함으로써 모델이 ICS 네트워크 패킷의 불확실성을 더 학습할 수 있게 하였다. 해당 방법으로 학습한 모델의 성능이 개선되는지를 분석하기 위하여 ICS 네트워크 패킷 데이터 셋인 ICS_PCAPS를 사용하여 실험하였다. 실험 방법은 특징 수준 융합 기반 학습을 수행한 모델과 아닌 모델의 성능을 비교하였다. 실험 결과는 모든 성능지표에서 특징 수준 융합 기반의 학습을 적용한 모델이 아닌 모델보다 성능이 높았고, F1 Score가 6.4%더 증가한 것을 확인하였다.

Key words

Industrial Control Systems, Cyber Security, Anomaly detection, Multi modal, Feature level fusion

Ⅰ. 서 론

Industrial Control Systems(ICS)는 일반적으로 발전소, 수처리시설, 석유 및 가스, 원자력발전소 등과 같은 산업에서 사용한다[1]. ICS는 4차 산업혁명으로 인하여 상용 Information & Communications Technology(ICT)가

* 가천대학교 정보보호학과 일반대학원 정보보호학과 박사과정 (202240226@gachon.ac.kr)

** 가천대학교 컴퓨터공학부 스마트보안전공 교수 (shjeon90@gachon.ac.kr)

*** 가천대학교 컴퓨터공학부 스마트보안전공 교수 (seojt@gachon.ac.kr)

적용되고 다른 네트워크와의 연결성 증가하였다[2]. ICS의 접근할 수 있는 외부망이 증가함에 따라서 ICS에 대한 사이버공격이 증가하고 있다[3].

이러한 이유로 인공지능을 이용하여 ICS 데이터를 모니터링하고 이상을 탐지하는 이상탐지 시스템 연구들이 제안되고 있다[4]. 대표적인 이상탐지 연구들은 ICS에서 사이버공격을 탐지하기 위해 ICS 네트워크 패킷을 학습하고 이상을 탐지하는 모델을 제안하였다[5]. 대부분의 연구에서는 대량의 ICS 데이터를 학습하거나 모델의 구조를 변형시켜 모델의 성능을 증가시켰다. 하지만 모델 운용 환경에 따라 모델 구조의 변형이 가능한 환경이 존재하며, 현재 설치되어 운용 중인 모델을 업그레이드하는 것은 어렵다. 또한 ICS에서 생성되는 데이터는 한정적이다.

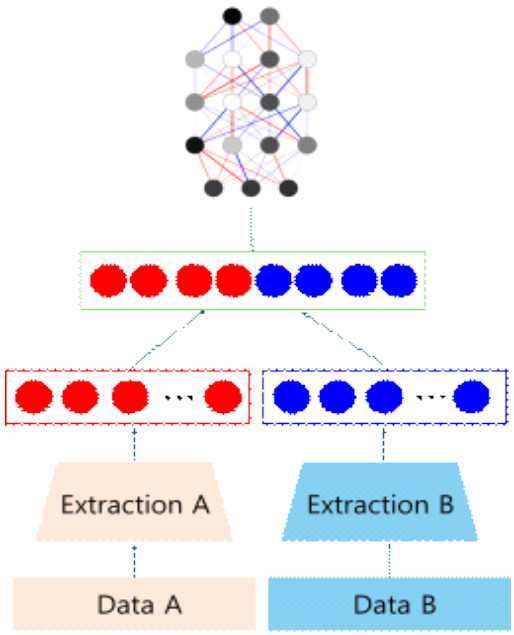
이러한 이유로 모델의 구조를 최대한 유지하면서 한정적인 ICS 데이터를 최대한 활용해야 한다. 이에 본 논문에서는 ICS 네트워크 패킷의 불확실성을 학습하기 위한 특징 수준 융합 기반의 멀티모달 학습 방식을 분석한다. 멀티모달 학습이란 여러 모달리티의 데이터를 통합하여 학습하는 방식이며, 특히 특징 수준 융합 기반 멀티모달 학습 방식은 다양한 데이터의 특징을 결합하여 모델에 학습시키는 방법이다[6]. 해당 방식을 이용하여 ICS 네트워크 패킷의 특징을 융합하고 모델이 학습함으로써 ICS의 불확실한 데이터의 구조를 모델이 이해할 수 있게 한다. 이러한 방식은 이상탐지 모델의 구조를 최대한 유지하면서 기존의 데이터로 모델의 이상탐지 성능을 개선할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 특징 수준 융합 기반의 멀티모달 학습 방식을 설명한다. 3장에서는 ICS 네트워크 패킷을 학습하는 특징 수준 융합 기반의 멀티모달 학습 방식을 논의한다. 4장에서는 실험을 통하여 이상탐지 성능을 검증한다. 마지막으로 5장에서 결론 및 향후 연구방향을 논의한다.

Ⅱ. 배경지식

Jiquan Ngiam외 5명은 다양한 형식의 특징을 학습하기 위해 멀티모달 학습 방법을 제안하였다[7]. 해당 연구에서는 오디오와 비디오와 같이 서로 다른 형식의 데이터를 학습시킴으로써 모델의 성능을 증가시켰다.

특징 수준 융합은 서로 다른 유형의 데이터의 특징을 추출하여 결합하고, 모델이 학습하는 멀티모달 학습 방법이다[8]. 특징 Level Fusion 방법의 구조는 [그림 1]과 같다[9].



[그림 1] 특징 수준 융합 방식 구조

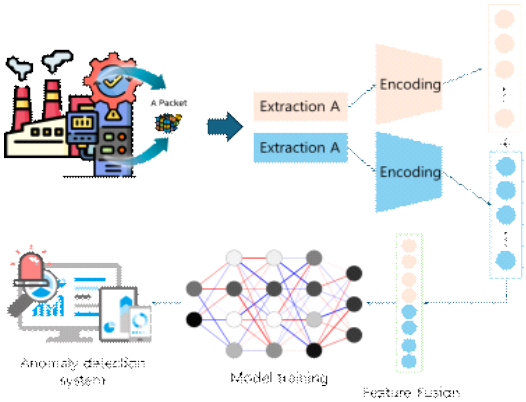
[그림 1]과 같이 특징 수준 융합 방식은 두 개의 데이터에서 특징을 추출한다. 이후 추출된 2개의 특징들을 하나의 데이터로 결합하여 모델 학습에 사용한다.

해당 방식은 서로 다른 유형의 데이터를

학습하여 다양한 특징을 학습한다. 본 논문에서는 한정적인 ICS 데이터의 문제를 고려하여 모델의 이상탐지 성능을 개선하기 위해 동일한 데이터를 특징 수준 융합 방식으로 학습한다. 해당 학습 방식은 모델이 학습할 수 있는 패킷의 특징이 많아지기 때문에 모델이 ICS 패킷의 특징을 더 이해하게 한다.

Ⅲ. ICS 패킷 대상 특징 수준 융합 기반의 멀티모달 학습 방법

본 장에서는 이상탐지 모델이 ICS 네트워크 패킷의 불확실한 구조를 고려하기 위한 ICS 네트워크 패킷 대상 특징 수준 기반의 멀티모달 학습 방법을 제안한다. 제안하는 학습 방법의 개요도는 [그림 2]와 같다.



[그림 2] ICS 네트워크 패킷 대상 특징 수준 융합 기반의 멀티모달 학습 개요도

해당 방식은 동일한 네트워크 패킷에서 추출된 특징을 결합하여 사용한다. 제안하는 ICS 네트워크 패킷 대상 특징 수준 융합 기반의 멀티모달 학습 과정은 알고리즘 1과 같다.

Algorithm 1. Feature-level fusion multimodal learning

Input: training dataset $X_{1:\beta}$, test dataset $T_{1:\gamma}$
Output: Anomaly result R
$\triangle \leftarrow$ empty set;
$\triangle_{enc1} \leftarrow$ empty set;
$\triangle_{enc2} \leftarrow$ empty set;
$\triangle_{fus} \leftarrow$ empty set;
$M \leftarrow$ model empty set;
for t=1 to β do
  $\triangle_{enc1} \leftarrow \triangle_{enc1} \cup$ encode(X_t);
  $\triangle_{enc2} \leftarrow \triangle_{enc2} \cup$ encode(X_t);
end
$\triangle_{fus} \leftarrow$ concatenate ($\triangle_{enc1}$, $\triangle_{enc2}$);
$M \leftarrow$ train M with $\triangle_{fus}$; // model training
//prediction
$\alpha \leftarrow$ threshold;
for t=1 to γ do
  $p \leftarrow p \cup M(T_t)$;
  $\triangle \leftarrow T_t - p$;
  if $\triangle > \alpha$ then
    $R \leftarrow RU\{1\}$; //anomaly
  else
    $R \leftarrow RU\{0\}$; //normal
end

ICS 네트워크 패킷 수집: ICS 환경에서 발생하는 네트워크 패킷을 수집한다. 패킷 수집은 ICS 동작에 영향을 미칠 수 있기 때문에 부담이 가장 적은 스위치 미러링 방식을 채택한다.

특징 추출: 수집한 네트워크 패킷을 파싱하고, 특징을 추출하여 개수가 β개인 학습 데이터 셋 $X_{1:\beta}$을 생성한다.

특징 인코딩: 학습 데이터 셋 $X_{1:\beta}$을 인코딩을 통하여 차원이 축소된 데이터 셋 $\triangle_{enc1}$, $\triangle_{enc2}$를 생성한다. 인코딩을 통한 차원 축소는 불필요한 특징은 제거하고 중요한 특징들로 구성되게 한다[10]. 특징을 결합하기 때문에 데이터 크기가 증가하여 모델에 부담되고 성능저하의 문제로 이어질

수 있다. 이에 인코딩을 통해 중요한 특징을 추출하고, 차원을 축소한다.

특징 결합: 인코딩된 특징을 결합하여 $\triangle_{fus}$를 생성한다.

모델 학습: $\triangle_{fus}$를 기반으로 멀티모달 모델 M을 학습한다. 이후 학습된 모델은 개수가 γ개인 테스트 데이터 셋 $T_{1:\gamma}$을 입력으로 받아 p를 예측한다. t시점의 테스트 데이터 셋 T_t와 p의 차이 계산하여 $\triangle$에 저장한다. 만약 $\triangle$가 임계치 α보다 크다면 1(이상)을 R에 저장하고, 작다면 0(정상)을 R에 저장한다. R은 모델이 정상과 이상을 예측한 결과이다.

이상탐지 시스템: 학습된 모델을 이상탐지 시스템에 탑재하여 이상을 탐지한다. 이상탐지 구간은 ICS 네트워크 패킷을 수집한 구간이다.

제안하는 방식은 모델이 동일한 ICS 네트워크 패킷의 특징을 결합하여 학습함으로써 패킷 특징의 구조를 반복적으로 학습한다. 모델은 반복된 학습과 다양화된 특징으로 인하여 ICS 네트워크 패킷의 구조를 이해하고, 탐지 성능이 향상된다.

IV. 실험 및 검증

4.1 실험 환경

본 실험은 Intel(R) Core(TM) i5-13600KF 3.50 GHz, 64GB RAM, Windows 10, NVIDIA GeForce RTX 4070 Ti인 PC에서 수행되었다.

4.2 데이터 셋

실험에 사용한 데이터 셋은 ICS_PCAPS 데이터 세트를 사용하였다[11]. ICS_PCAPS는 MODBUS/TCP를 기반으로 통신하는 PLC, Human Machine Interface (HMI), Remote Terminel Unit (RTU), Vacuum Fluorescent Display (VFD), 3-phase motor로 구성된 테스트베드에서 정상 및 사이버공격 패킷을 수집하였다. ICS_PCAPS의 사이버공격 패킷은 서비스거부공격을 수행하고 발생한 패킷이다. 네트워크 패킷은 CICFlowMeter [12]를 사용하여 파싱 및 특징 추출을 수행하고 CSV파일로 변환하였다. 학습 데이터는 정상 데이터 9,570개를 사용하였고, 테스트 데이터는 정상 데이터 1,576개와 공격 데이터 516개를 사용하였다.

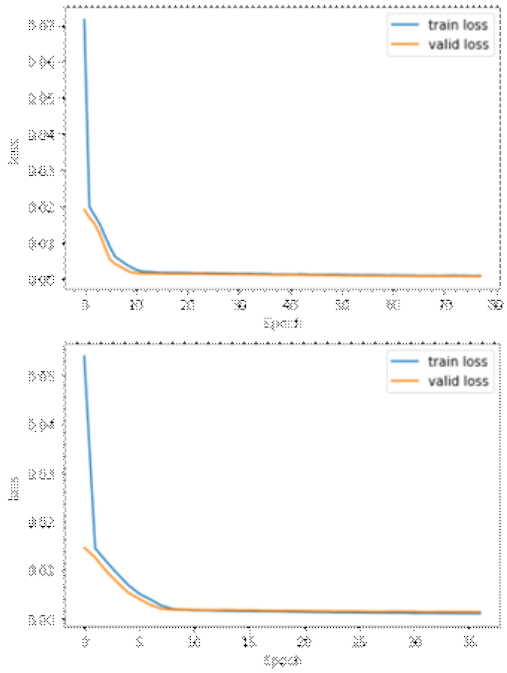
4.3 실험

제안하는 방법은 동일한 ICS 네트워크 패킷을 특징 수준 융합을 기반으로 학습할 경우, 성능에 유의미한 변화가 있는지 실험한다. 실험은 2개의 동일한 정상 데이터 셋을 인코딩하고, 융합하여 학습한다.

먼저 인코더를 학습하기 위하여 오토인코더 모델을 구현하고 학습하였다. 인코더를 통해 82개의 특징에서 40개의 특징으로 차원을 축소하였다. 해당 과정은 동일한 데이터를 융합하는 과정에서 의미있는 특징을 추출함과 동시에 데이터의 크기가 갑자기 증가하여 모델에 부담 및 성능의 저하를 예방하고자 인코딩을 수행한다.

이후 멀티모달 학습 방식을 적용하는 모델과 그렇지 않은 Long Short-Term Memory(LSTM) 모델 2개를 개발하여 성능을 비교하였다.

일반적인 학습 방식의 모델은 인코딩된 하나의 정상 데이터만 학습하였다. 멀티모달 학습 방식을 적용하는 모델은 인코딩된 두 개의 정상 데이터를 결합하여 학습하였다. 두 개의 LSTM 모델이 학습하는 과정에서 발생한 손실은 [그림 3]과 같다.



[그림 3] 일반적인 학습 방식의 모델 손실(위)과 특징 수준 융합 기반 학습 방식의 모델 손실(아래)

2개의 모델은 학습 시 발생하는 손실이 더 이상 줄어들지 않을 때까지 학습을 수행하였다.

모델의 성능지표로는 Accuracy, Precision, Recall, F1 score, Balance accuracy, ROC-AUC를 사용하였다[13]. 성능지표는 true positives (TP), true negatives (TN), false positives (FP), false negatives (FN)를 사용하여 계산된다. 각 성능지표 수식은 아래와 같다.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\mathrm{Precision} = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

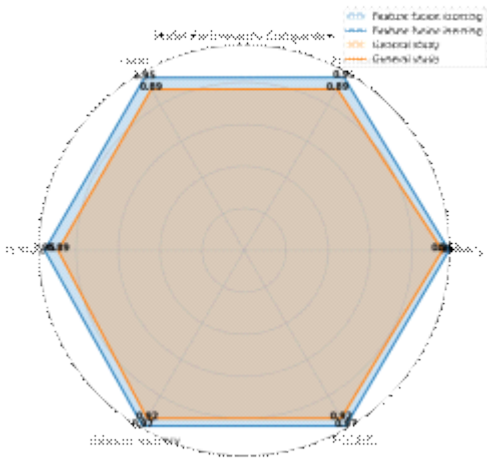
$$F1\,Score = \frac{2\times \mathrm{Precision}\times Recall}{\mathrm{Precision}+Recall}$$

$$Balance\,accuracy = \frac{1}{2}\left(\frac{TP}{TP+FN}+\frac{TN}{TN+FP}\right)$$

해당 성능지표를 이용하여 특징 수준 융합 기반 학습 방식과 그렇지 않은 학습 방식을 채택한 모델의 성능은 아래와 같다.

[표 1] 학습 방식에 따른 모델별 성능 비교 표

	특징 수준 융합 기반 학습	일반적인 학습
Accuracy	0.9756	0.9440
Precision	0.9514	0.8873
Recall	0.9496	0.8856
F1 score	0.9505	0.8865
Balance accuracy	0.9668	0.9244
ROC-AUC	0.9668	0.9244



[그림 4] 학습 방식에 따른 모델별 성능 그래프

모든 지표에서 특징 수준 융합 기반의 학습을 한 모델이 그렇지 않은 모델보다 더 좋은 성능을 보였다. [그림 4]에서 볼 수 있듯이 특징 수준 융합 기반 학습이 일반적인 학습 방법보다 모든 면에서 더 높은 성능을 보인다. 특히 F1 Score에서는 특징 수준 융합 기반 학습 모델이 일반적인 학습을 한 모델보다 6.4% 성능이 향상된 것을 확인할 수 있다.

이처럼 한가지의 데이터에 대해서 단일 데이터를 학습하는 것보다, 같은 데이터를 특징 수준 융합 기반 학습을 사용한다면 모델의 성능이 향상된다. 이러한 방식은 모델의 구조를 크게 변형하지 않고, 학습 방법만 달리하여 성능 개선이 가능하다. 또한 ICS와 같이 생성되는 데이터가 한정적인

환경에서 활용할 수 있다.

V. 결론 및 향후 연구방향

ICS의 사이버공격을 탐지하기 위해 다양한 이상탐지 방안이 제안되고 있다. 기존의 이상탐지 방안은 대량의 ICS 데이터를 학습시키거나, 모델의 네트워크 구조를 변형하여 성능을 증가시켰다, 하지만 모델 운용 환경에 따라서 모델 구조 변경이 가능한 환경이 존재하고 모델에 대한 업그레이드가 어렵다. 또한 ICS환경에서 발생하는 데이터의 양이 한정적인 문제가 있다. 이에 본 논문에서는 모델의 구조를 최대한 유지하면서 ICS 네트워크 패킷의 불확실성을 학습하기 위한 특징 수준 융합 기반의 멀티모달 학습 방식을 분석하였다. 하나의 데이터만을 학습하는 모델과 같은 데이터를 특징 수준 융합을 통해 결합하고 학습하는 모델의 성능을 비교하였다. ICS 네트워크 패킷을 사용하여 실험한 결과 특징 수준 융합 기반의 멀티모달 학습 방식을 채택한 모델이 그렇지 않은 모델보다 모든 성능지표에서 유의미한 성능 개선을 보였다, 향후 연구로는 ICS에서 발생하는 다양한 데이터에 대하여 제안하는 방식을 검증할 예정이다.

Acknowledgment

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국에너지기술평가원의 지원을 받아 수행된 연구(20224B10100020, 원전 사이버위협 대처 시스템 설계 및 시험검증 기술개발 (APR1400국산화 MMIS)과 한국서부발전(주)의 과제 "AI(인공지능)를 이용한 신재생에너지 제어시스템 사이버공격 탐지기술 개발"의 지원을 받아 수행된 연구임 (2022-위탁-11, 50%)

참 고 문 헌

[1] Drias, Zakarya, Ahmed Serhrouchni, and Olivier Vogel. "Analysis of cyber security for industrial control systems." 2015 international conferenc e on cyber security of smart cities, industrial control system and communications (ssic). IEEE, 2015.

[2] Q1 2024 – a brief overview of the main incidents in industrial cybersecurity. https://ics-cert.kaspersky.com/publications/reports/2024/06/03/q1-2024-a-brief-overview-of-the-main-incidents-in-industrial-cybersecurity/.

[3] 김우년, 박응기, and 김신규. "4 차 산업혁명 시대의 산업 제어시스템 보안성 평가 방안 연구." 한국통신학회논문지 44.5 (2019): 943-956.

[4] Jeffrey, Nicholas, Qing Tan, and José R. Villar. "A review of anomaly detection strategies to detect threats to cyber-physical systems." Electronics 12.15 (2023): 3283.

[5] Ali, Wasim A, et al. "A review of current machine learning approaches for anomaly detection in network traffic." Journal of Telecommunications and the Digital Economy 8.4 (2020): 64-95.

[6] Yoon, JunHo, GyuHo Choi, and Chang Choi. "Multimedia analysis of robustly optimized multimodal transformer based on vision and language co-learning." Information Fusion 100 (2023): 101922.

[7] Ngiam, Jiquan, et al. "Multimodal deep learning." Proceedings of the 28th international conference on machine learning (ICML-11). 2011.

[8] Sharma, Priyanka, and Manavjeet Kaur. "Multimodal classification using 특징 level fusion and SVM." International Journal of Computer Applications 76.4 (2013): 26-32.

[9] Ehatisham-Ul-Haq, Muhammad, et al. "Robust human activity recognition using multimodal feature-level fusion." IEEE Access 7 (2019): 60736-60751.

[10] Vahdat, Arash, and Jan Kautz. "NVAE: A deep

hierarchical variational autoencoder." Advances in neural information processing systems 33 (2020): 19667-19679.
[11] Frazão, I.; Abreu, P. H.; Cruz, T.; Araújo, H.; Simões, P. In Denial of Service Attacks: Detecting the Frailties of Machine Learning Algorithms in the Classification Process, Cham, 2019; Springer International Publishing: Cham, 2019; pp 230-235.
[12] Habibi Lashkari, A.; Draper-Gil, G.; Mamun, M. S. I.; Ghorbani, A. A. In Characterization of Tor Traffic using Time based features, International Conference on Information Systems Security and Privacy, 2017; 2017.
[13] Lee, Ju Hyeon, Jiho Shin, and Jung Taek Seo. "Solar Power Plant Network Packet-Based Anomaly Detection System for Cybersecurity." Computers, Materials & Continua 77.1 (2023).

비밀번호 강도 측정 기법 비교 분석

*박원상, **서승희, ***이창훈

Comparative Analysis of Password Strength Measurement Techniques

*Wonsang Park, **Seunghee Seo and ***Changhoon Lee

요 약

본 논문은 비밀번호 강도 평가 기법의 비교 분석을 통해 기법들의 장단점을 파악하고, 상황에 맞는 기법을 사용할 수 있도록 각각의 특성을 명확히 제시하는 것을 목표로 한다. 비밀번호 강도 평가는 정보 보안의 핵심 요소로, 약한 비밀번호는 시스템의 취약점을 초래할 수 있다. 본 연구에서는 비밀번호를 평가하는 여러 방법 중 단일 단어, 데이터 세트의 특성을 이용한 강도 평가들의 한계를 분석하고 전통적인 사전 매칭 기법 및 확률 모델 기반의 평가 방법인 Monte Carlo, Confident Monte Carlo 방법들의 동작 과정과 장단점을 분석하고 특성을 명확히 제시한다.

Key words

Password, Password Strength Measurement, Monte Carlo, Confident Monte Carlo

Ⅰ. 서 론

비밀번호는 디지털 보안의 첫 번째 방어선으로, 사용자의 개인 정보와 데이터를 보호하는 데 중요한 역할을 한다. 그러나 비밀번호의 강도가 낮은 경우, 해커는 이를 손쉽게 추측하여 시스템에 침입할 수 있다. 또한 사이버 수사 과정에서도 최우선으로 고려되는 대상이 용의자가 설정한 기기의 잠금을 해제하는 것이며, 이에 사용되는 첫 번째 수단이 비밀번호다. 이러한 이유로 비밀번호의 강도와 보안성을 평가하는 것이 중요한 이슈로 떠올랐고, 이에 관한 다양한 방법론이 개발되고 있다.

본 연구는 비밀번호 강도 평가 방법을 분류하고 이들의 장단점을 비교 및 평가하여 보안 정책 개선을 위한 가이드를 제공하고자 한다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 단일 비밀번호 및 비밀번호 세트의 정보만을 이용한 단순 강도 평가에 관한 연구를

* 서울과학기술대학교, 석사과정 (18101217@seoultech.ac.kr)
** 서울과학기술대학교, 박사과정 (sh.seo@seoultech.ac.kr)
*** 서울과학기술대학교, 교수, 교신저자 (chlee@seoultech.ac.kr)

소개한다. Ⅲ장에서는 패스워드 사전을 이용한 비밀번호 강도 측정 방법을 위한 사전 생성 방법에 관해 설명한다. Ⅳ장에서는 생성한 사전과 비교하는 비밀번호 강도 측정 모델에 관해 설명한다. 마지막으로 Ⅴ장에서는 연구의 결론 및 향후 연구 방향을 제시한다.

Ⅱ. 정량적 지표를 통한 비밀번호 강도 평가

비밀번호 강도 평가는 비밀번호의 보안성을 평가하고 개선하기 위해 필수적이다. 본 장에서는 단어의 복잡도와 출연 빈도를 통해 강도를 평가하는 방법을 소개한다.

2.1 엔트로피 측정

비밀번호 엔트로피는 비밀번호의 복잡성과 예측 불가능성을 수치화하는 지표로 사용된다. 엔트로피가 높을수록 비밀번호가 더 복잡하고, 추측하기 어렵다는 것을 의미한다. 비밀번호의 여러 가지 엔트로피 생성 방법 중 자주 사용되는 방법으로는 Shannon 엔트로피가 있다.

2.1.1 Shannon 엔트로피

Shannon 엔트로피는 정보 이론에서 도입된 개념으로, 비밀번호의 각 문자 선택이 독립적이고 균등하게 분포된다고 가정할 때 비밀번호의 복잡성을 측정한다. Shannon 엔트로피는 다음과 같은 수식으로 계산 된다:

$$H = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i)$$

여기서 $P(x_i)$는 비밀번호의 각 문자 x_i 가 선택될 확률을 의미한다. Shannon 엔트로피는 비밀번호의 길이와 문자 집합의 크기에 따라 증가하며, 이는 비밀번호가 더 많은 문자 집합에서 선택될 때 더 강력해진다는 것을 나타낸다[1].

2.1.2 엔트로피 계산 예시

예를 들어, 8자리 비밀번호가 있고 각 자리가 26개의 소문자, 26개의 대문자, 10개의 숫자, 그리고 10개의 특수문자로 이루어져 있다고 가정하자. 가능한 문자 집합의 크기는 72이다. 각 문자가 독립적으로 선택된다고 가정할 때, 비밀번호의 엔트로피 계산은 다음과 같다.

$$H = 8\log_2 72 \approx 8 \times 6.17 \approx 49.36 bits$$

이는 해당 비밀번호가 약 49.36-bit의 엔트로피를 가지며, 이는 약 $2^{49.36}$개의 가능한 조합을 가진다는 것을 의미한다. 따라서, 높은 엔트로피를 가진 비밀번호는 낮은 엔트로피의 비밀번호보다 추측하기 어려워 보안성이 높다.

엔트로피를 통해 비밀번호의 복잡성을 측정하는 것은 유용하지만 몇 가지 한계점이 존재한다. 먼저 엔트로피 계산은 비밀번호의 모든 문자가 독립적이고 동일한 확률로 선택된다는 가정을 기반으로 한다. 실제로는 사용자가 특정 패턴이나 습관에 따라 비밀번호를 생성하는 경우가 많아 이 가정이 항상 성립하지 않는다. 또한 엔트로피는 비밀번호의 길이와 문자 집합의 크기만을 고려하므로, 실제 공격 시나리오에서의 비밀번호 강도를 완벽하게 반영하지 못할 수 있다.

2.2 Zipf's Law

Zipf's Law는 자연어 처리, 검색 엔진 최적화, 생태학 등 다양한 분야에서 사용되는 법칙으로, 비밀번호 강도 평가에도 적용될

수 있다. Zipf's Law는 비밀번호 사용 빈도와 순위 사이의 관계를 설명하며, 일반적으로 빈도가 높은 비밀번호가 공격자에게 쉽게 추측될 수 있음을 나타낸다[2].

Zipf's Law는 단어의 빈도 f 와 해당 단어의 순위 r 사이에 다음과 같은 관계가 성립함을 보인다.

$$f \propto \frac{1}{r}$$

이는 순위가 r 인 단어의 빈도는 가장 빈도가 높은 단어의 빈도의 $1/r$에 해당한다는 것을 의미한다. 비밀번호의 경우, 가장 빈도가 높은 비밀번호부터 순위를 매겨 빈도와 순위의 관계를 분석할 수 있다.

Rockyou와 같은 대규모 비밀번호 데이터 세트에 Zipf's Law를 적용하여 비밀번호의 빈도와 순위를 분석할 수 있다. 데이터 세트에서 각 비밀번호의 출연 빈도를 계산하고 순위를 매겨 비밀번호의 강도 순위를 평가할 수 있다.

그러나 Zipf's Law를 통한 비밀번호 강도 평가는 다음과 같은 한계점이 존재한다.

일반화의 어려움: Zipf의 법칙은 데이터 세트에 따라 다르게 나타날 수 있으며, 모든 비밀번호 데이터 세트에 동일하게 적용되지 않을 수 있다.

비밀번호 패턴 미반영: 빈도와 순위만을 고려하므로, 사용자가 특정 패턴에 따라 생성한 비밀번호의 강도를 충분히 반영하지 못할 수 있다.

2.3 단순 비밀번호 강도 평가의 한계

엔트로피 측정과 Zipf's Law는 비밀번호 각각의 복잡도나 순위의 분석엔 유용하지만 실제 환경에서 비밀번호 강도로 적용하기에는 여러 한계가 존재한다.

- 복잡도 및 순위의 제한: 엔트로피는 비밀번호의 문자 조합의 복잡도를 측정하지만, 사용자의 실제 비밀번호 생성 습관이나 패턴을 충분히 반영하지 못한다. Zipf's Law는 비밀번호의 빈도와 순위를 분석하지만, 이는 단지 특정 데이터셋에 한정된 결과일 수 있다.
- 실제 공격 시나리오 반영 부족: 단순 비밀번호 강도 평가는 실제 공격 시나리오를 충분히 반영하지 못한다. 공격자는 다양한 추측 기법을 사용할 수 있으며, 단순히 복잡도나 빈도만으로 비밀번호 강도를 평가하는 것은 불완전할 수 있다.

이러한 한계로 인해, 비밀번호의 복잡도, Zipf's Law를 적용한 순위 등 단일 단어나 세트만을 사용해 생성한 정보는 주로 비밀번호 강도의 참고 자료로써 간접적으로 사용되고 실제 환경에서는 다양한 비밀번호들의 패턴을 반영한 비밀번호 사전과의 비교를 통해 비밀번호 강도를 보다 정확하고 포괄적으로 평가할 수 있는 사전 기반의 비밀번호 강도 측정 방법이 주로 사용된다.

Ⅲ. 비밀번호 사전 기반 상대적 평가 기법

사전 기반 비밀번호 강도 측정 방법은 실제 사용되는 비밀번호 목록을 저장한 비밀번호 사전과의 비교를 통해 사용자가 실제로 비밀번호를 생성하는 패턴을 반영하여 보다 정교한 강도 평가를 가능하게 한다. 이러한 방법은 비밀번호의 문자 조합뿐만 아니라, 사용자가 비밀번호를 생성할 때 따르는 특정 패턴을 분석하여 비밀번호의 예측 가능성을 평가한다.

사전 기반 비밀번호 강도 측정 방법은 확률 기반의 다양한 알고리즘으로 비밀번호 리스트를 생성하고 이들을 평가하는

과정으로 진행된다. 본 장에서는 우선 비밀번호의 확률에 기반하여 강도 평가의 비교 대상 비밀번호 사전을 생성하는 대표적인 기법들인 PCFG, Markov, n-gram 모델을 소개하고 생성한 사전과의 비교를 통해 상대적인 강도를 평가하는 다양한 기법들에 관해 설명한다.

3.1 PCFG 사전 생성 기법

PCFG (Probabilistic Context-Free Grammar Model) 모델은 비밀번호 생성 패턴을 모델링하여 학습하고 사전을 생성하는 기법이다. PCFG는 다음과 같이 동작한다.

- 비밀번호 템플릿 추출: CFG는 먼저 비밀번호를 템플릿으로 분해한다. 템플릿은 비밀번호의 문자 유형(알파벳, 숫자, 특수문자 등)을 나타낸다. 예를 들어, 비밀번호 "abc123"은 템플릿 "L3D3"로 표현된다. 여기서 "L3"는 3개의 알파벳 문자, "D3"는 3개의 숫자를 의미한다.
- 템플릿 학습: PCFG는 대규모 비밀번호 데이터 세트를 사용하여 다양한 템플릿과 각 템플릿의 구성 요소를 학습한다. 이 과정에서 각 템플릿의 빈도와 템플릿 내 구성 요소의 빈도를 계산한다. 이를 통해 비밀번호 생성에 사용되는 다양한 패턴을 파악한다.
- 사전 생성: 학습된 PCFG 모델을 사용하여 가능한 비밀번호의 사전을 생성한다. 이 사전은 다양한 템플릿과 그 템플릿에 맞는 구성 요소의 조합으로 이루어진다. 예를 들어, "L3D3" 템플릿에 따라 "abc123", "xyz789" 등의 비밀번호가 생성될 수 있다[3].

3.2 Markov 사전 생성 기법

Markov 모델은 비밀번호 세트의 현재 상태 문자 시퀀스를 기반으로 다음 문자를 예측하는 확률 모델이다. 이 모델은 비밀번호의 생성 과정을 연속된 상태 전이로 표현하여 비밀번호의 패턴을 학습하고 사전을 생성한다. Markov 모델은 다음과 같이 동작한다.

- 상태 전이 학습: Markov 모델은 비밀번호 데이터 세트에서 현재 상태의 각 문자 시퀀스의 전이 확률을 학습한다.
- $$P(c_i \mid c_{i-1}) = \frac{F(c_{i-1}, c_i)}{F(c_{i-1})}$$

여기서 $F(c_i)$는 c_i 로 시작하는 시퀀스의 빈도를 뜻하고 $F(c_{i-1}, c_i)$는 c_{i-1}문자 뒤에 c_i 문자로 이어지는 시퀀스의 빈도를 뜻한다.

예를 들어 비밀번호 데이터 세트에서 문자 'a' 뒤에 'b'가 나오는 빈도가 30회이고, 문자 'a'로 시작하는 모든 시퀀스의 빈도가 100회라면 P(b|a) = 30/100=0.3으로 계산된다.

- 비밀번호 전체 확률 계산: 비밀번호 전체의 생성 확률은 각 상태 전이 확률의 곱으로 계산된다. 예를 들어, 비밀번호 "abc"의 생성 확률은 다음과 같이 표현된다.
- $$P(abc) = P(a) \times P(b \mid a) \times P(c \mid b)$$

여기서 $P(a)$는 비밀번호가 a로 시작할 확률을 뜻한다.

- 사전 생성: 학습한 상태 전이 확률을 기반으로 초기 상태를 확률에 따라 설정하고 종료 조건(문자 길이)에 도달하기까지 다음 문자를 상태 전이 확률에 따라 무작위로 선택하여 단어를 생성한다[4].

3.3 n-gram 사전 생성 기법

n-gram 모델은 Markov 모델과 유사하나 현재 문자뿐 아니라 이전의 n개의 연속된 요소(문자 또는 단어)에 의존하여 다음 요소의 확률을 계산한다. n-gram의 상태 전이 확률은 다음과 같이 계산된다.

$$P(c_i \mid c_{i-n+1}, \cdots, c_{i-1}) = \frac{F(c_{i-n+1}, \cdots, c_{i-1}c_i)}{F(c_{i-n+1}, \cdots, c_{i-1})}$$

Markov와 마찬가지로 비밀번호 데이터 세트에서 "ab" 다음에 'c'가 나오는 빈도가 40회이고, "ab"로 시작하는 모든 시퀀스의 빈도가 200회라면 P(c|ab) = 40/200=0.2로 계산된다.

n-gram 또한 비밀번호 전체의 생성 확률을 계산하는데 예를 들어 3-gram 모델에서 "abcde"의 생성 확률은 다음과 같이 표현된다.

$$P(abcde) = P(a) \times P(b \mid a) \times P(c \mid ab) \times P(d \mid bc) \times P(e \mid cd)$$

n-gram 방식도 Markov 방식과 마찬가지로 비밀번호 사전을 생성하는 데 사용된다[5].

이러한 사전 생성 기법은 주로 실제 유출된 비밀번호 데이터 세트를 기반으로 학습하기에 현실적인 비밀번호 생성 패턴을 반영할 수 있다. 이는 비밀번호 강도 평가의 정확성을 높이며, 공격자들이 사용하는 비밀번호 추측 방식을 모델링하여 더욱 효과적인 보안 강화 방안을 마련하는 데 도움을 준다.

3.4 사전 비교 분석

사전 기반 비밀번호 강도 평가의 초기 접근 방식 중 하나는 사전 매칭 기법이다. 이 방법은 미리 생성한 사전에 있는 비밀번호와 사용자가 입력한 비밀번호를 비교하여 일치 여부를 판단한다.

- 단순 사전 비교: 사전에 포함된 비밀번호와 사용자가 입력한 비밀번호를 일대일로 비교하여 일치하는 경우 해당 비밀번호를 약한 비밀번호로 간주하는 기법으로 사전의 크기와 다양성이 비밀번호 강도 평가의 정확성에 큰 영향을 미치는 한계가 있다.
- 확장된 사전 비교: 단순 사전 매칭의 한계를 극복하기 위해 사전에 포함된 비밀번호에 다양한 변형을 적용하여 더 많은 비밀번호를 탐지할 수 있도록 한 기법이다.
- 규칙 사전 비교: 사전에 포함된 비밀번호뿐만 아니라, 특정 규칙을 사전에 사용하여 비밀번호를 생성하고 이를 사전과 매칭하는 방법이다. 이 방법은 비밀번호 생성 규칙을 기반으로 다양한 비밀번호를 생성하여 탐지 범위를 넓힌다.

사전 매칭 기법은 비밀번호 강도 평가의 초기 접근 방식으로 다양한 장단점이 존재한다. 먼저 추가적인 학습 과정 없이 생성한 사전과 단어를 단순 비교하기 때문에 자원 및 시간 효율성이 높다. 또한 구조가 단순하여 평가를 위한 파라미터 조정이 쉽고 직관적이며 배포 및 업데이트 등의 관리가 편리하다. 그러나 실제 환경에 사용되는 수많은 비밀번호 목록을 전부 적용하기엔 한계가 있어 실제 사용자의 수많은 패턴을 반영하지 못하는 단점이 있다.

3.5 Monte Carlo 기법

Monte Carlo 기법은 무작위 샘플링을 통해 비밀번호의 강도를 평가하는 방법이다. PCFG, Markov, n-gram 모델 등 다양한 방법으로 생성된 사전을 기반으로 비밀번호의 확률 분포를 기반으로 무작위 샘플링하고, 이 샘플의 빈도와 순위를 분석하여 비밀번호의 강도를 측정한다.

Monte Carlo 기법은 우선 비밀번호 a의

확률 P(a)과 a의 순위를 정의한다. 여기서 a의 순위$S_p(a)$는 모델 내에서 a보다 높은 확률을 가지는 비밀번호의 수를 의미한다.

$$S_p(\alpha) = \mid \beta \in \Gamma : p(\beta) > p(\alpha) \mid$$

그리고 n개의 비밀번호 샘플을 생성하고 샘플 Θ에서 각 비밀번호의 확률 p(B)를 계산하여 비밀번호 a보다 높은 확률을 가지는지 확인한다. 그 후 수식 $C_\triangle$를 사용하여 비밀번호 a의 상대적인 강도를 추정한다.(Δ : α보다 높은 확률을 가지는 비밀번호들의 집합)

$$C_\triangle = \sum_{\beta \in \Theta} \begin{cases} \dfrac{1}{p(\beta) \cdot n} & \text{if} p(\beta) > p(\alpha) \\ 0 & otherwise \end{cases}$$

$C_\triangle$ 값은 비밀번호 a의 강도를 나타내며 값이 작을수록 비밀번호의 강도가 높다는 것을 나타낸다[6].

Monte Carlo 기법은 비밀번호의 강도를 효율적으로 추정할 수 있다. 기존의 사전 비교 방법과는 달리 모든 비밀번호를 열거하지 않고도 비밀번호 추측 수를 추정할 수 있어 효율적이다. 또한 이 기법을 통해 도출한 비밀번호의 강도는 확률적 비밀번호 모델에 대해 일반적으로 적용할 수 있다. 즉, 다양한 비밀번호 모델에 대해 효율적으로 적용할 수 있다.

그러나 Monte Carlo 기법은 추정값에 대한 절대적인 정확성을 보장하지 못한다. 특히, 추정값의 정확성에 대한 신뢰 구간을 제공하지 않으며, 추정값이 부정확할 때 이를 식별하는 메커니즘이 부족한 단점이 있다. 또한 샘플 크기가 충분히 크지 않을 경우 추정값의 분산이 커질 수 있다. 이는 추정값의 정확성을 저하시킬 수 있다.

3.6 Confident Monte Carlo

Confident Monte Carlo는 Monte Carlo 기법의 한계를 개선하여 더욱 신뢰성 있는 비밀번호 강도 평가를 제공하는 기법이다. 이 기법은 Monte Carlo 기법의 장점을 유지하면서도, 추정값에 대한 신뢰 구간을 제공하여 추정값의 신뢰성을 높인다.

$$C_\triangle = \sum_{\beta \in \Theta} \begin{cases} \dfrac{1}{p(\beta) \cdot n} & \text{if} p(\beta) > p(\alpha) \\ 0 & otherwise \end{cases}$$

Confident Monte Carlo 기법은 Monte Carlo 기법과 마찬가지로 $C_\triangle$ 값을 통해 비밀번호의 강도를 평가하나 Hoeffding 부등식과 Chernoff 부등식을 사용하여 샘플링 기반 추정값에 대한 신뢰 구간을 계산한다. 이를 통해 비밀번호 강도 평가의 신뢰성을 보장한다[7].

신뢰 구간 : 추정된 비밀번호 강도 값이 실제 값에 얼마나 가까운지를 나타내는 범위. Confident Monte Carlo 기법은 두 부등식을 활용하여 신뢰 구간을 제시하고 추정값이 오차 범위 내 확률로 신뢰 구간 내에 있을 것임을 보장한다.

3.6.1 Hoeffding 부등식

Hoeffding 부등식은 확률 변수의 합이 일정 범위를 벗어날 확률에 대한 상한을 제공한다.

$$P(\mid C_\triangle - E[C_\triangle] \mid \geq \epsilon) \leq 2\exp(-2n\epsilon^2)$$

※$E(C_\triangle)$: $C_\triangle$의 기댓값, ϵ: 허용 오차

3.6.2 Chernoff 부등식

Chernoff 부등식은 확률 변수의 곱이 일정 범위를 벗어날 확률에 대한 상한을 제공한다.

$$P(C_{\triangle} \geq (1+\delta)E[C_{\triangle}]) \leq \exp(-\frac{\delta^2 nE[C_{\triangle}]}{2+\delta})$$

※ δ: 상대적 허용 오차

Confident Monte Carlo 기법은 이 두 부등식을 활용하여 신뢰 구간의 상한과 하한을 계산한다.

신뢰 구간 하한 : Hoeffding 부등식을 사용하여 하한을 계산

$$L = E[C_{\triangle}] - \epsilon$$

$$※\epsilon = \sqrt{\frac{\ln(2/\delta)}{2n}}$$

신뢰 구간 상한: Chernoff 부등식을 사용하여 상한을 계산

$$U = (1+\delta)E[C_{\triangle}]$$

$$※\delta = \sqrt{\frac{2\ln(1/\delta)}{nE[C_{\triangle}]}}$$

3.6.3 Confident Monte Carlo 기법의 장단점

Confident Monte Carlo 기법은 기존 Monte Carlo 기법이 가진 비밀번호 강도 평가에서 신뢰성 문제를 개선하였다. 먼저 이 기법은 추정값에 대한 신뢰 구간을 제공하여, 추정값의 신뢰성을 높인다. 이를 통해 사용자는 추정값이 실제 값에 얼마나 가까운지를 판단할 수 있으며, 평가 결과의 신뢰도를 크게 향상할 수 있다. 신뢰 구간은 추정값의 범위를 제공하여, 추정값이 통계적으로 신뢰할 수 있는지 여부를 명확히 한다.

Confident Monte Carlo 기법은 Hoeffding 부등식과 Chernoff 부등식과 같은 엄격한 통계적 기법을 사용하여 신뢰 구간을 계산한다. 이러한 부등식들은 샘플링된 데이터의 특성을 기반으로 신뢰 구간을 수학적으로 보장하며, 추정값의 정확성과 신뢰성을 수학적으로 뒷받침한다. 이는 비밀번호 강도 평가의 신뢰성을 높이는 데 중요한 역할을 한다.

또한 이 기법은 Monte Carlo 기법의 장점을 유지하면서, 실제 유출된 비밀번호 데이터 세트를 기반으로 비밀번호 강도를 평가한다. 이는 실제 사용자의 비밀번호 생성 패턴을 반영하여 현실적인 비밀번호 강도 평가를 가능하게 한다.

그러나 Confident Monte Carlo 기법에는 몇 가지 단점도 존재한다. 먼저, 이 기법은 일반적인 Monte Carlo 기법보다 계산적으로 더 복잡하다. 신뢰 구간을 계산하기 위해 추가적인 샘플링과 계산이 필요하므로, 높은 계산 비용이 요구된다. 이는 실시간 비밀번호 강도 평가와 같은 응용에서 사용하기 어려울 수 있다.

또한 희귀한 비밀번호(확률이 매우 낮은 비밀번호)에 대해서는 여전히 추정값의 신뢰 구간이 넓을 수 있다. 이는 정확한 평가가 어려울 수 있음을 의미한다. 이러한 비밀번호는 데이터 세트에서 매우 드물게 나타나기 때문에, 충분한 샘플링을 통해서도 정확한 추정이 어렵다. 따라서, 희귀 이벤트에 대한 처리가 여전히 어려운 점이 단점으로 작용할 수 있다.

Confident Monte Carlo 기법은 기존 Monte Carlo 기법의 비밀번호 강도 평가에서 추정값의 신뢰성을 높이기 위해 신뢰 구간을 제공하는 기법이다. 이 기법은 통계적 보증을 통해 추정값의 정확성을 보장하지만, 계산 복잡성과 추가 샘플링 필요성 등의 단점도 존재한다. 이를 통해 사용자는 상황에 맞는 기법을 사용하여 비밀번호 강도를 평가할 수 있다.

3.7 비밀번호 강도의 평가 방법 비교

앞서 비밀번호 강도의 다양한 평가

방법들에 관해 설명하였다.

• 엔트로피 분석

장점: 비밀번호의 강도를 수치화하여 직관적으로 평가 가능

단점: 사용자의 비밀번호 생성 패턴을 반영하지 못함

주 사용 환경: 비밀번호 강도의 정량적 지표가 필요할 때

• Zipf's Law

장점: 비밀번호 사전의

단점: 특정 사용자의 정보를 포함한 저빈도의 비밀번호에 대한 강도 평가가 어려움

주 사용 환경: 비밀번호의 빈도 분포를 확인할 때

• 단순 사전 비교

장점: 구현이 간단하고 빠르며

단점: 사용되는 사전에 따라 강도의 정확성이 달라짐

주 사용 환경: 사전이 충분히 크거나 빠른 평가가 필요할 때

• 확장된 사전 비교

장점: 다양한 변형을 포함하여 더 많은 비밀번호를 탐지

단점: 사전의 크기에 따라 메모리 사용량과 계산 비용이 크게 증가함

주 사용 환경: 특정 정보를 포함한 비밀번호의 다양한 변조의 강도를 평가할때

• 규칙 사전 비교

장점: 특정 패던을 사용한 비밀번호의 평가하기에 적합함

단점: 복잡한 규칙이 많을수록 계산 비용이 증가하고, 모든 가능한 규칙을 포괄하기 어려움

주 사용 환경: 특정 사용자 패턴의 비밀번호를 평가할 때

• Monte Carlo Sampling

장점: 무작위 샘플링을 통해 효율적으로 비밀번호의 확률 분포를 반영

단점: 측정값의 절대적인 정확성을 보장하지 못하며 샘플 크기에 따라 추정값의 분산이 좌우됨

주 사용 환경: 확률 모델 기반의 비밀번호 강도 평가

• Confident Monte Carlo

장점: 신뢰 구간을 제공하여 추정값의 신뢰성을 높임

단점: 계산적으로 더 복잡하고 높은 계산 비용을 요구함

주 사용 환경: 더욱 정확한 강도 평가와 안전한 비밀번호의 신뢰 구간을 확인할 때

비밀번호 강도 측정 분석 결과를 고려할 때, 강도 측정을 위한 자원 및 연산 효율성과 강도 측정의 정확성은 서로 반비례 관계를 갖는다. 이에 따라 각 환경에 맞는 비밀번호 강도 평가 방법을 사용하여 적절한 강도의 비밀번호를 사용할 수 있다.

IV. 결론 및 향후 연구

본 논문에서는 단어 자체 정보만을 사용한 기법부터 단어 사전과 확률을 기반으로 비밀번호 강도를 측정하는 기법까지 다양한 비밀번호 강도 분석 기법들을 비교 분석하였고 각 기법의 장단점을 도출하였다.

또한 확률 모델 기반의 비밀번호 강도 평가 방법인 Monte Carlo 기법과 Confident Monte Carlo 기법과 비밀번호 강도 평가를 위한 사전 생성 기술인 PCFG, Markov, n-gram 모델을 분석하였는데 이러한 기법들은 실제 유출된 비밀번호 데이터셋을 기반으로 학습하여 현실적인 비밀번호 생성 패턴을 반영할 수 있다. 이는 비밀번호 강도 평가의 정확성을 높이는 데 기여하며, 공격자들이 사용하는 비밀번호 추측 방식을 모델링하여 보다 효과적인 보안 강화 방안을 마련하는 데 도움을 준다.

각 비밀번호 강도 측정 방법론은 앞서 기술한 장단점에 기반하여 비밀번호 강도

측정기를 적용하려는 시스템의 환경에 따라 그 적절성이 달라진다. 예를 들어 자원이 매우 제한적이고 유출 데이터를 확보할 수 없는 경우 확장된 사전 비교 기법을 통해 다양한 비밀번호 경우의 수를 확보하는 것이 유리하다.

앞서 말했듯 비밀번호의 강도 평가 과정에서 강도 측정을 위한 자원 및 연산 효율성과 강도 측정의 정확성은 서로 반비례하지만 최근 시스템에서 유출 사고를 유의하여 사용자의 비밀번호를 해시값 등의 난수 형태로 관리함에 따라, 각 시스템의 패스워드 정책에 적절한 학습 데이터가 충분하지 않을 가능성이 크다. 이에 따라, 학습할 수 있는 데이터의 양이 충분하지 않은 환경을 고려할 때, 효율적인 자원 활용 및 높은 정확성을 동시에 달성할 수 있는 연구가 필요하다.

참 고 문 헌

[1] S. Rass, S. König Password Security as a Game of Entropies entropy Apr, 2018
[2] D. Wang, G. Jian, X. Huang, P. Wang Zipf's Law in Passwords IEEE Transactions on Information Forensics and Security Nov, 2017.
[3] M. Weir, S. Aggarwal, M. Collins, H. Stem Testing metrics for password creation policies by attacking large sets of revealed passwords 17th ACM Conference on Computer and Communications Security. Oct, 2010.
[4] M. Heričko, B. Brumen Strength Analysis of Real-Life Passwords Using Markov Models Applied Sciences. Oct, 2021.
[5] Zhang, L., Luo, X., & Lin, Z. (2021). NEMO: Modeling Password Guessability Using Markov Models. GitHub. https://github.com/nemo
[6] M. Dell'Amico, M. Filippone. Monte Carlo Strength Evaluation:Fast and Reliable Password Checking, ACM Digital Library Oct, 2015.
[7] P. Liu, J. Blocki, W. Bai Confident Monte Carlo: Rigorous Analysis of Guessing Curves for Probabilistic Password Models 2023 IEEE Symposium on Security and Privacy (SP) May, 2023.

HMAC 기반 메시지인증을 위한 키관리 기법

*강윤희, **권태언

Secure key management for HMAC based message authentication

*Yunhee Kang, **Taeun Kwon

요 약

디지털 자료의 무결성 및 진본검증은 영상데이터를 안전하게 유지 및 관리하기 위한 정보기술의 핵심과제 중 하나이다. HMAC은 단순한 해시함수를 사용하여 전송 디지털 자료의 무결성 검증하는 이점을 갖지만 HMAC에 필요한 키관리를 요구한다. 본 논문에서는 이를 위해 Shamir의 비밀공유 스킴을 기반으로 HMAC의 메시지 인증코드의 생성 및 검증을 위한 비밀 키관리 기법을 제안한다.

Key words

Data Integrity, Authenticity Verification, HMAC, Shamir's Secure Sharing

I. 서 론

디지털 자료의 무결성 및 진본검증은 영상데이터를 안전하게 유지 및 관리하기 위한 정보기술의 핵심과제 중 하나이다. 최근 디지털 자료는 생산과 사용 과정에서 해당 콘텐츠의 무결성 검증은 주요한 절차로 고려되고 있다. 이를 위해 비대칭키를 사용하는 PKI(public key infrastructure) 기반의 서명 및 검증을 활용할 수 있으나 PKI 기법은 중앙기관을 통한 복잡한 키 생성과 유지의 관리 제약 및 서명 및 인증 과정에서의 성능제약을 갖는다. 여기서는 이를 해결하기 위해 간편한 해시함수를 사용하는 HMAC (keyed-hash message authentication code) 기법을 적용한다[1][4].

HMAC은 전송 디지털 자료의 무결성 검증하기 위해 해시함수를 사용하는 이점을 갖지만 HMAC 알고리즘 운영에 필요한 비밀키의 안전한 관리를 요구한다[1].

본 논문에서는 HMAC 알고리즘 운영에서 데이터 생성자와 소비자 양측이 공유하는 비밀키를 관리하기 위한 기법을 제안한다. 제안한 키 관리 기법은 키 생성이 필요한 비밀값을 다루기 위해 사용하며, 이를 위해

* 백석대학교, 교수 (yhkang@bu.ac.kr)
** ㈜하스퍼, 책임연구원, 교신저자 (peterkwon@harsper.co.kr)

지분을 나누어 비밀값을 재구성하는 Shamir의 비밀공유 스킴을 기반으로 한다.

Ⅱ. 설계된 키 교환 기법

HMAC는 암호화 해시 함수와 기밀 암호화 키를 수반하는 특정한 유형의 메시지 인증 코드이다[1]. 원본 메시지가 변하면 그 해시값도 변하는 해시함수의 특징을 활용하여 메시지의 데이터 무결성과 진본 확인을 동시에 수행하기 위해 사용한다.

HMAC는 해시코드의 생성 및 확인을 위해 대칭 키를 공유하는 방식으로 운영되며, 암호화 해시 함수와 공유 비밀 키를 결합과정에서 비밀키에 대한 안전한 관리가 요구된다. 여기서는 이와 같은 보안 취약성을 해결하기 위해 다수의 참여자에게 지분(share)를 분배한 후 키 생성이 필요할 경우 추가적인 지분을 확보하여 대칭키를 재구성하기 위한 Shamir 비밀공유[2,3] 기반으로 키 관리 스킴을 설계한다. 키교환을 통해 구성된 비밀값은 HMAC 알고리즘에 사용된다.

비밀공유는 비밀키의 소유자가 1명이 아닌 다수의 인증된(authorized) 참가자(participant) 들이 되고, 이 때 임의의 참가자는 비밀 자료의 일부 지분(share)을 소유하는 방식으로 임계치(threshold) 이상의 지분만 있으면 비밀정보를 복구할 수 있기 때문에 일부 조각이 유실되더라도 안전하다.

제안한 키 관리 스킴에서 지분 요청을 수행하는 참여자는 인증을 위해 PKI[4] 메커니즘을 사용하여 서명 및 검증을 수행하며, 지분 전달의 보안 강화를 위해 전달 내용에 XOR(exclusive OR)를 적용하여 자료의 보호를 강화한다. 설계된 키교환 기법은 그림 1과 같이 운영된다. 그림 1의 Logger와 Verifier은 키 교환 스킴의 단계 별 수행을 통해 메시지의 무결성과 인증을 보장하는 키 관리를 수행한다.

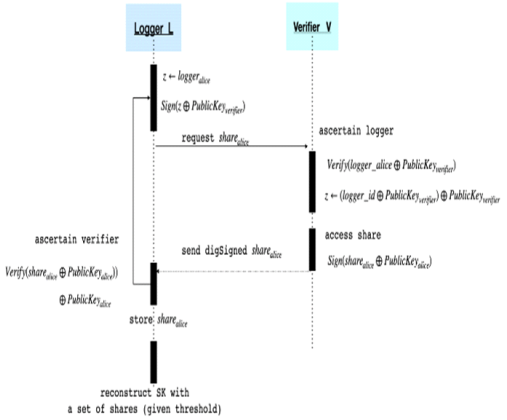


그림 1. 키 교환 스킴의 시스템 구성

그림 1에서 Logger_L은 참여자이며, Verifier_V는 모든 지분을 관리하는 관리자이다. 참여자와 관리자 모두 PKI 메커니즘을 사용하여 자신임을 증명할 수 있는 비대칭 키를 유지한다. Logger_L 의 지분 요청에 따라 Verifier_V는 지분 중 무작위 하나의 지분을 선택하고, 해당 지분을 Logger_L의 Public Key와 XOR를 수행하고 해당 내용에 자신의 Private Key로 서명을 작성하여 Logger_L에 전달한다. Logger_L은 앞선 내용과 같이 서명을 검증하고 XOR연산을 수행하여 지분을 추출한 뒤 대칭 키를 생성하기 위한 Threshold의 개수 만큼의 지분을 확보하기까지 Verifier_V에 반복적으로 지분을 요청한다.

Ⅲ. 실험 환경 및 성능평가

그림 2는 지분 생성 단계의 수행을 보인 것으로 주어진 비밀값, 지분갯수 n과 비밀값 생성을 위한 임계값 t으로 부터 지분 생성을

요청한다. 해당 생성 요청은 Logger에 의해 시작되며, 지분은 Verifier에서 생성하여 유지한다. Logger는 응답으로 비밀값을 얻기 위한 필요 지분 갯수인 임계값을 얻는다. 이 과정에서 비밀값을 얻기 위한 과정은 다음 두 단계로 이루어진다.

- 지분요청 및 비밀값 재구성
- 비밀값 검증

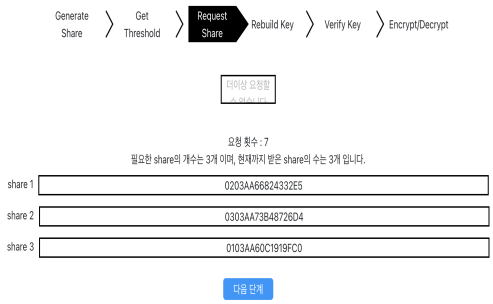


그림 2. 지분 요청 및 비밀값 재구성 과정 화면

IV 결론 및 향후 연구 방향

무결성 및 진본확인은 디지털데이터의 내용 변경이 없음을 검증하기 위한 목적으로 사용한다. 이 논문은 HMAC 메시지인증코드 생성을 위한 비밀키 관리를 위해 Shamir의 비밀공유를 사용하였다. 이를 위해 Shamir 지분으로부터 비밀값 재구성을 수행하기 위한 안전한 요청 및 응답 처리를 수행하였다. 이를 위해 설계된 키 관리 스킴은 생성영상의 무결성 검증을 위해 사용한다.

감사의 글

본 논문은 중소벤처중소벤처기업부(중소기업기술정보진흥원, RS-2023-00225234) 2023 년도 산학연 CollaboR&D 사업의 산업현장의 디지털 영상데이터의 AI 기반 무결성 및 검증 기술 개발과제의 지원을 받아 수행된 연구임

참 고 문 헌

[1] Pierre-Alain Fouque, David Pointcheval, and Sébastien Zimmer. 2008. HMAC is a randomness extractor and applications to TLS. In Proceedings of the 2008 ACM symposium on Information, computer and communications security (ASIACCS '08). Association for Computing Machinery, New York, NY, USA, 21–32.

[2] A. Shamir, "How to share a secret," Commun. ACM, vol. 22, no. 11, pp. 612-613, 1979.

[3] Dawson, E.; Donovan, D. (1994), "The breadth of Shamir's secret-sharing scheme", Computers & Security, 13: 69–78.

[4] "X.509: Information technology - Open Systems Interconnection - The Directory: Public-key and attribute certificate frameworks". ITU. Retrieved 6 November 2019

Session 2

논문발표

[반도체대학 301호]

- 좌장 -

김시호(연세대)

영-한 다국어 단어 임베딩을 통한 효율적인 문서 검색 시스템

*강어진, **유준

Implementation of a Synonym Search Service with English-Korean Multilingual Word Embedding

*Kang Eojin, **Joon Yoo

요 약

본 논문에서는 영어와 한국어 간 다국어 단어 임베딩을 이용한 동의어 검색 서비스를 구현하는 방법을 제시한다. IT 업계에서는 '머신러닝-기계학습-machine learning', '코스트-비용-cost'와 같이 동일한 의미를 지닌 단어들이 영어 및 한국어로 다양한 형태로 불리며 사용되고 있다. 이러한 단어들을 효율적으로 검색하고 확인할 수 있도록 돕기 위해, 본 논문에서는 FastText 기반의 Multilingual Word Embedding 모델인 MUSE를 활용하여 워드 벡터를 생성하였다.

본 연구에서는 위키피디아에서 제공하는 최신 영어 및 한국어 데이터 및 워드벡터를 사용하였으며, 추가적으로 영한 음역 데이터셋을 확보하여 임베딩 생성 시 학습에 포함시켰다. 이를 통해 영어와 한국어 간의 의미적 유사성을 효과적으로 반영한 임베딩을 구축하였다. 결과적으로, 제안된 시스템은 다양한 형태의 동의어를 정확하고 빠르게 검색할 수 있도록 하여 다국어 환경에서의 정보 검색 효율성을 크게 향상시킬 수 있음을 보인다.

본 논문에서는 모델 설계 및 구현 과정, 데이터 전처리 방법, 그리고 실험 결과와 함께 제안된 시스템의 유용성을 평가하고, 향후 연구 방향에 대해 논의한다.

Key words

Word Embedding, Multilingual word embedding, MUSE, FastText

I. 서 론

현대의 글로벌화된 사회에서 효율적인 정보 검색의 필요성은 날로 증가하고 있다. 특히 IT 업계에서는 다양한 개념이 새로이 등장하며 이에 대해 발음 시 대부분 원어를 그대로 옮겨 사용한다. 동일한 의미를 지닌 단어들이 영어 및 한국어로 다양한 형태로 사용되며, 원하는 정보를 얻기 위해 이러한 단어들을 효율적으로 검색하고 이해하는

* 가천대학교 AI·소프트웨어학부 학부생 (kaj1226@gachon.ac.kr)

** 가천대학교 AI·소프트웨어학부 부교수 (joon.yoo@gachon.ac.kr)

것은 중요한 과제이다. 그러나 기존의 단일 언어 기반 검색 시스템은 하나의 단어를 사용할 때만 검색이 가능하기에 동일한 의미의 단어를 여러 형태로 재검색하는 등 다국어 환경에서의 정보 검색 효율성을 저하시킬 수 있다.

이러한 문제를 해결하기 위해 본 연구에서는 다국어 단어 임베딩을 활용한 동의어 검색 서비스를 제안한다. 이를 통해 영어와 한국어 간의 의미적 유사성을 효과적으로 반영한 검색 시스템을 구축하고자 한다. 본 연구에서는 META에서 공개한 MUSE (Multilingual Unsupervised and Supervised Embeddings)를 활용하여 fastText 기반의 영어 및 한국어 워드 벡터를 결합한 영한 다국어 단어 임베딩을 생성하였으며, 위키피디아 및 AI Hub에서 제공하는 최신 영어 및 한국어 데이터셋과 영한 음역 데이터셋을 활용하여 워드벡터를 생성하였다.

다국어 단어 임베딩은 서로 다른 언어 간의 단어들을 같은 벡터 공간에 매핑함으로써, 의미적으로 유사한 단어들이 가까운 거리에 위치하도록 한다. 이를 통해 사용자는 영어와 한국어로 다양한 형태의 동의어 및 유사어를 한 번에 검색하고 확인할 수 있으며, 이는 다국어 환경에서의 정보 검색 효율성을 크게 향상시킬 수 있다.

본 논문에서는 제안된 시스템의 관련 연구 및 설계, 데이터 전처리 방법, 그리고 실험 과정과 결과를 상세히 설명한다. 또한, 제안된 시스템의 유용성을 평가하고, 향후 연구 방향에 대해 논의한다. 본 연구를 통해 다국어 정보 검색의 효율성을 향상시키고, 다양한 언어를 혼합하여 사용하는 IT 업계를 시작으로 많은 사용자들에게 보다 편리한 검색 환경을 제공하는 데 기여하고자 한다.

Ⅱ. 관련 연구

2.1 단일 언어 기반 단어 임베딩

초기 단어 임베딩 연구는 주로 단일 언어에 초점을 맞추었다. 대표적으로 Word2Vec (Mikolov et al., 2013) [1]은 단어의 의미를 벡터 공간에 표현하여 유사한 의미를 가진 단어들이 유사한 벡터 값을 가지도록 하였다. 이어서 GloVe (Pennington et al., 2014)와 fastText (Bojanowski et al., 2017) [2] 등 다양한 단일 언어 임베딩 기법들이 제안되었다. 그러나 이러한 모델들은 다국어 간의 의미적 유사성을 반영하지 못하는 한계가 있다.

2.2 다국어 임베딩 모델

다국어 임베딩 모델은 여러 언어의 단어들을 동일한 벡터 공간에 매핑함으로써, 언어 간의 의미적 유사성을 반영하고자 한다. 최근 연구들은 다국어 임베딩 생성 시 추가적인 정보를 활용하여 성능을 향상시키고자 하였다. Artetxe et al. (2018)은 언어 간 번역 쌍을 이용하여 임베딩을 학습하는 방법을 제안하였고, Lample et al. (2018)은 언어 모델을 통해 다국어 임베딩을 학습하는 방법을 제안하였다. 이러한 연구들은 다국어 임베딩의 성능을 크게 향상시켰으나, 여전히 특정 언어 쌍에 대해 제한적인 성능을 보이는 경우가 있다.

본 연구에서는 위키피디아에서 제공하는 최신 영어 및 한국어 워드벡터를 활용하고, 추가적으로 영한 음역 데이터셋을 확보하여 학습에 포함시킴으로써, 기존 연구들의 한계를 극복하고자 하였다. 이를 통해 영어와 한국어 간의 의미적 유사성을 더욱 효과적으로 반영한 임베딩을 생성하였으며, 이를 바탕으로 동의어 검색 서비스의 성능을 향상시켰다.

Ⅲ. 시스템 구조

현재 한 단어의 다양한 형태에 대한 검색을 위해서는 각각의 형태를 하나의 검색어로서 새로이 입력해야 한다. 이는 여러 번의 검색을 해야 한다는 수고로움이 있을 뿐더러 완벽하게 일치하지 않는 형태는 찾을 수 없다는 단점이 있다. 따라서 본 연구에서는 매번 다른 형태로 검색하지 않도록 효율적인 문서 검색 시스템을 구현하고자 한다. 이는 기존 검색 시스템이 검색어와 일치하는 부분을 문서 내에서 찾아낸다는 점을 이용하여, 이 때 입력으로 들어가는 검색어를 시스템 판단 하에 사용자 입력어와 유사한 의미의 여러 단어를 입력한다. ‘Machine Learning’이라는 단어는 한국어로 ‘머신러닝’ 또는 ‘기계학습’ 등으로 불리기에, 이러한 영-한 음역 및 번역어와 같이 외래어 중 여러 형태의 한국어로 존재하는 단어를 타깃으로 삼았으며, 영어 또는 한국어를 입력하였을 때 그와 동일하거나 유사한 뜻을 지닌 영/한 단어들을 출력하여 이 언어들을 활용한 검색 시스템을 구현한다.

이를 위해 한-영 다국어 워드 임베딩을 구현한다. 외래어에 대한 동의어 및 유사어를 찾기 위해서는 각 언어의 워드 벡터를 양방향으로 이용할 수 있어야 한다. 따라서 한국어 단어 입력 시 유사한 영단어 및 한국어가, 영단어 입력 시 유사한 한국어 및 영단어가 도출된다. 최종적으로는 단어뿐이 아닌 유사한 문장을 찾아내는 것을 목표로 하며, 이를 위한 단계로 우선 유사한 단어를 파악하여 검색어로써 활용하고자 한다.

Ⅳ. 실험 과정

4.1 데이터 수집 및 전처리

실험에 사용된 데이터는 위키피디아에서 제공하는 최신 영어 및 한국어 워드벡터를 기반으로 하였다. 추가적으로, 영어와 한국어 간의 음역 데이터를 확보하여 학습에 포함시켰다. 데이터 전처리 과정에서는 불필요한 기호와 중복 단어를 제거하고, 토큰화를 통해 단어를 분리하였다. 또한, fastText를 사용하여 각 단어의 벡터를 생성하였다.

4.2 모델 학습

앞서 생성된 데이터를 중심으로 MUSE (Multilingual Unsupervised and Supervised Embeddings)[3]를 활용한 다국어 워드 임베딩을 진행했다. MUSE는 META에서 배포한 다국어 임베딩 라이브러리로, 다국어 단어 임베딩 및 자연어 처리의 빠른 개발 및 성능 평가를 가능하게 하는 오픈소스이다. 지도학습 및 비지도학습을 모두 제공하며, 본 연구에서는 앞서 생성한 한국어 워드 임베딩 및 영어 워드 임베딩을 이용해 두 임베딩을 정렬하여 다국어 임베딩을 가능하게 하였다. 지도학습을 통해 진행하였으며, 반복 횟수를 3, 5, 7회로 나누어 실행하고 소스 벡터와 타깃 벡터를 영어와 한국어를 바꿔가며 진행하여 양방향 임베딩을 가능하게 하였다.

Ⅴ. 성능 평가

5.1 WordSim353 유사도 비교

동의어 검색 성능 평가 전, 각 언어 벡터의 정렬이 성공적으로 이루어졌는지 확인하기 위해 위해 Wordsim353 데이터셋을 활용하였다. 본 데이터는 단어 임베딩 중 두 단어의 유사도를 비교하는데 주로 사용되는 데이터셋으로, 단어 두 개와 그 두 단어의 유사도를 하나의 행으로 가지고 있다. 영어

벡터에서의 성능 평가를 위해 353개의 영단어 유사도 점수를 가지고 있으며, 한국어 벡터에서의 평가를 위해 이를 번역한 데이터를 활용하였다.

한-영 단어 임베딩에서의 성능 평가를 위해 한국어 Wordsim353과 한국어 벡터를 비교하였고, 영-한 임베딩에서의 성능 평가를 위해 영어 Wordsim353과 영어 벡터를 비교하였다. Wordsim353의 각 행에 위치한 두 단어에 대해 각 벡터에서의 유사도 점수를 계산하였고, 이를 데이터셋 내에서 제공된 유사도와 비교하여 일치율을 살폈다. 각 행별 일치율을 구한 뒤 이에 대한 평균을 계산하였으며, 그 결과는 한국어 벡터에서의 유사도 일치율 82.11%, 영어 벡터에서의 유사도 일치율 87.99%로 유의미한 일치율을 가져왔다.

```
word1,word2,wordSimScore,modelScore
가입,티켓,5.536,3.067
술,화학,4.125,2.2925
알루미늄,금속,6.625,4.334
```

그림 1. 한국어 WordSim353 유사도 비교

```
word1,word2,wordSimScore,modelScore
admission,ticket,5.536,4.4834
alcohol,chemistry,4.125,2.648
aluminum,metal,6.625,6.1241
```

그림 2. 영어 WordSim353 유사도 비교

```
Overall performance KO: 82.11%
Overall performance EN: 87.99%
```

그림 3. 각 WordSim353 유사도 비교 결과

타겟벡터와의 Wordsim353 유사도 비교 평가를 진행하지 않은 이유는, MUSE를 통한 다국어 워드 임베딩은 두 언어의 워드 벡터를 하나의 벡터로 합치는 것이 아닌, 각 벡터 스페이스 상에서 같은 의미의 단어는 유사한 벡터를 가지도록 정렬하는 방식이기 때문이다. 따라서 타겟벡터에는 소스벡터의 언어로 구성된 단어가 없기에 유사도 평가가 불가하다.

5.2 KNN(K-Nearest Neighbor) 평가

동의어 검색 성능을 평가하기 위해 K-nearest neighbors (KNN) 알고리즘을 사용하였다. 영-한 단어쌍으로 이루어진 음역 데이터셋을 활용하였다. 본 데이터셋에 존재하는 영단어에 대한 최근접 이웃을 확인한 후, 그 중 쌍으로 연결된 한국어 단어가 존재한다면 True, 존재하지 않는다면 False의 값을 매겼다. 음역 데이터셋은 IT 분야에 특화된 데이터셋으로 두산백과의 IT 외래어 사전을 참고하였으며, 전체 단어 쌍 중 True가 도출된 비율을 확인하였다. 영단어 자체가 존재하지 않을 경우에는 데이터 학습의 부족으로 인해 벡터 상 단어가 존재하지 않는 것이므로 이를 제외한 다른 단어들을 학습하였다.

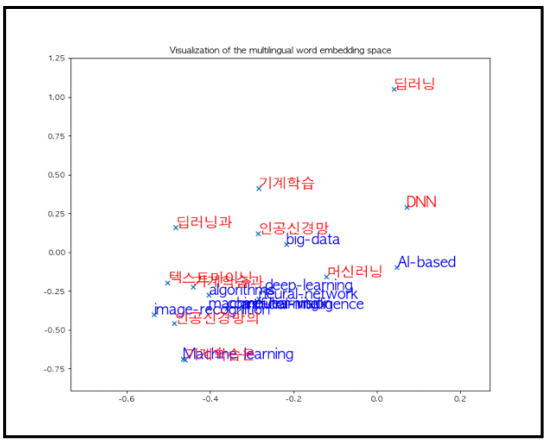


그림 4. 영-한 KNN 시각화

그림 4는 이러한 과정 중 'machine-learning'에 대한 영어 벡터에서의 최근접 단어어 10개와, 한국어 벡터에서의 최근접 단어 10개를 시각화한 것이다. 영어로는 'deep-learning', 'big-data' 등이 보이며, 한국어로는 '머신러닝', '기계학습', '기계학습의' 등의 단어를 볼 수 있다. 특히

‘Machine-learning’과 ‘기계학습은’의 벡터는 거의 일치한다고 볼 수 있으며, 이를 통해 영-한 및 한-영 워드 임베딩이 성공적으로 이루어졌음을 다시 한 번 확인할 수 있었다.

English Word	Korean Word	KNN Results	Match Found
Bit-Coin	비트코인	N/A	False
KakaoTalk	카카오톡	바이버, 카카오톡과, WeChat, 아이메시지, 카카오톡이…	True
Facebook	페이스북	페이스북, SNS, 페이스북은, 페이스북과, 뉴스피드의,…	True
Instagram	인스타그램	페이스북, SNS, Vyrl, 리그램, 인스타, 팔로우, 페…	False
Arduino	아두이노	아두이노, 아두이노의, 아두이노에, 아두이노를, 아두이…	True
protocol	프로토콜	프로토콜, OSPF, ppTp, TFTP, TCP의, 인증기술,…	True
YouTube	유튜브	유튜브, 유투브, 유투브의, 유트브, 유튜브나, 동영상,…	True
Ethernet	이더넷	이더넷, 40Gbps, DWDM, IGMP, 10Gbps, CPRI…	True
AlphaGo	알파고	알파고가, 알파고를, 기계학습, 인공지능, 딥마인드의,…	True

그림 5. IT 분야 음역 데이터셋 TF 결과

평가 결과 전체 739개의 단어쌍 중 True는 30.45%의 225개, False는 514개의 78.59%를 차지했다. 단, False의 경우 (1) 영단어가 애초에 존재하지 않는 경우와 (2) 영단어와 매칭되는 한국어가 없을 경우를 포함하는데, False에 해당하는 514개의 단어쌍 중 영단어가 애초에 존재하지 않는 경우는 321개를 차지했다. 따라서 영단어가 벡터 데이터 내부에 존재하는 경우에 한해서는 True는 53.82%를 차지한다. 이외에도 False에는 매칭되는 음역된 한국어 단어가 존재하지만 하이픈(-) 및 띄어쓰기, 대소문자 등으로 인해 불일치 판정을 받은 경우가 포함되므로 50% 이상의 True 판정은 다국어 단어 임베딩 벡터 상에서 음역 데이터 역시 유의미하게 찾아낼 수 있음을 의미한다.

VI. 향후 계획

본 실험을 통해 구현된 워드 벡터를 LLM과 연계하여 최종적으로 원하는 단어를 입력할 경우 문서 속에서 유사한 부분을 모두 찾아내는 것을 목표로 하고 있다. 이를 위해서는 첫째, 더 다양한 데이터셋에 대한 학습이 필요하다. 성능 평가 중 사용한 두산백과의 IT 외래어 사전을 활용할 때 영어 벡터에 영단어가 존재하지 않는 경우 또는 연결되는 한국어 단어가 존재하지 않는 경우 False로 판단하였는데, 이 경우가 IT 분야 데이터셋의 50%에 달한다. 이는 전반적으로 워드 벡터를 생성할 시 학습된 데이터셋의 부족에 원인이 있으므로, 초반 워드벡터 생성 시 더 많은 데이터셋을 확보하여 다양한 경우에 대한 단어를 가질 수 있어야 할 것이다. 이는 IT 분야는 물론 타 분야에 대한 데이터를 확보하여 더 높은 성능 향상을 위해 중요한 단계이다.

둘째, 문장 단위의 임베딩 및 LLM 연계가 필요하다. 현재 진행한 실험은 단어 단위의 다국어 임베딩으로, 유사한 단어만을 추출할 수 있다. 실제 LLM과 연계하여 ‘유사 검색 서비스’를 구현할 경우 사용자는 단어를 입력하겠으나 시스템은 단어뿐만 아닌 문장 단위의 유사 검색을 실시하는 것이 더욱 효율적일 것이다. 따라서 본 연구의 발전 계획은 우선 추가적인 데이터를 통해 학습한 단어 임베딩과 LLM을 연계하여 유사 단어 검색이 성공적으로 동작하는지 확인한 후 이를 문장 단위로 발전시키고자 한다.

Ⅶ. 결론

하나의 단어에 대해 단순 번역을 진행할 때에는 띄어쓰기 및 변형 등 여러 경우에 대한 번역 및 검색을 진행해야 한다. 하지만 본 실험과 같이 생성된 다국어 워드 임베딩을 통해 최근접 단어를 추출하여 유사어 검색에 활용한다면 같은 의미가 있는 여러 형태를 한 번에 얻을 수 있다. 한국어 벡터와 영어 벡터를 정렬하여 같은 의미의 단어 간에는 유사한 벡터를 지니도록 하고, 각각의 인풋 및 아웃풋으로 서로의 언어가 가능하게 하여 최종적으로는 원하는 언어의 단어를 입력하였을 때 그와 동일한 의미의 해당 언어

속 단어들은 물론 타 언어에 대한 유사어도 얻을 수 있었다. 이는 영한 이외의 다른 언어에도 적용될 수 있으며, 이를 실제 서비스와 연결하여 하나의 단어만 검색하여도 그와 유사한 의미를 지닌 단어 또는 문장을 찾을 수 있는 형태의 시스템 구현이 가능할 것이다.

참 고 문 헌

[1] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean, Efficient Estimation of Word Representations in Vector Space, arXiv cs arXiv:1301.3781, pp.1-12, 2013, Retrieved from https://arxiv.org/abs/1301.3781 .

[2] Piotr Bojanowski, Edouard Grave, Armand Joulin, Tomas Mikolov, 'Enriching Word Vectors with Subword Information', arXiv cs arXiv:1607.04606, pp. 1-12, 2017, https://arxiv.org/abs/1607.04606 .

[3] Edouard Grave, Armand Joulin, Quentin Berthet, "Unsupervised Alignment of Embeddings with Wasserstein Procrustes", arXiv cs arXiv:1805.11222, pp. 1-11, 2018, https://arxiv.org/abs/1805.11222 .

휴머노이드 로봇 제어를 위한 ROS 노드 설계

*이동완, **안해은, ***지혜원, ****손애은, *****구본근

Design of ROS Nodes for controlling Humanoid Robot

*Dong-Wan Lee, **Hae-Eun Ahn, ***Hye-Won Ji, ****Ae-Eun Son and *****Bon-Gen Gu

요 약

휴머노이드 로봇은 관절들의 동작을 조합하여 걷기, 앉기, 중심 잡기 등 복잡한 동작을 수행한다. 로봇의 관절은 서보 모터 등을 이용하여 특정 범위 내의 회전 동작을 구현한다. 본 논문에서는 각 서보 모터의 동작을 제어하는 ROS 노드와 걷기 등 복잡한 동작을 수행하는 ROS 노드와 상호 작용 방법을 설계한다. 본 논문에서 제안하는 ROS 노드 사이의 상호 작용은 복잡한 동작을 기본적인 동작의 조합으로 표현할 수 있어 효과적인 동작 제어를 가능하게 한다.

Key words

ROS, humanoid, robot, node, ros topic, ros action

Ⅰ. 서 론

로봇 활용 분야의 확장은 로봇에 요구되는 기능 및 하드웨어 구성 요소를 증가시키고 있다[1-3]. 휴머노이드는 인간의 형태를 한 로봇을 의미하며, 인간의 동작을 모사하기 위해 서보 모터 등으로 구현한 다수의 관절을 가지고 있다. 인간이 여러 관절의 움직임을 이용하여 걷기, 앉기, 중심 잡기 등의 동작을 수행하므로, 이러한 인간의 동작을 모사하기 위한 휴머노이드 관절 즉, 각 모터의 정밀 제어가 필요하다.

교육용 휴머노이드 로봇의 동작 제어는 개발사에서 제공하는 참조 코드 또는 라이브러리를 이용하기 때문에 다양한 실험실 환경과 응용 분야에서 로봇의 동작을 적절하게 제어하기 어렵다.

본 논문에서는 특정 범위 내의 회전 동작을 통해 로봇 관절의 동작을 구현하는 모터를 제어하는 기본 동작 노드와 이들 노드와 상호 작용하여 걷기 등 복잡한 동작을 제어하는 노드 사이의 상호 작용 방법을 설계한다.

Ⅱ. ROS2 노드 상호작용 설계

그림 1은 휴머노이드의 관절 동작을 구현하는 서보 모터를 제어하는 기본 노드와

* 국립한국교통대학교, 학부생(slrtvi6600@gmail.com)
***** 국립한국교통대학교, 교수, 교신저자 (bggoo@ut.ac.kr)

이를 이용하여 복합적인 동작을 위한 노드 사이의 상호작용을 나타낸 것이다. 그림 1에 나타낸 교육용 휴머노이드 로봇은 열여섯 개의 서보 모터를 이용하여 로봇 관절을 구현하고 있다. 노란색 원으로 표시한 것은 각 서보 모터를 제어하는 ROS 노드를 나타낸 것으로 게시자(publisher)가 게시한 제어 정보를 수신하는 수신자(subscriber)이며, 수신한 제어 정보를 기반으로 서보 모터에 대한 기본적인 제어를 수행한다.

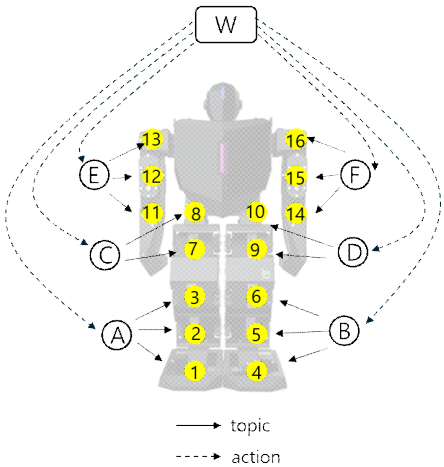


Fig. 1. Interaction between nodes

원내 문자 A에서 F는 휴머노이드 각 부분의 동작을 제어하며 ROS의 액션(action) 서버와 서보 모터 제어 정보를 게시하는 게시자 노드이다. 예를 들어, 노드 A는 걷기, 앉기, 중심 잡기 등의 복잡한 동작 구현을 위해 필요한 휴머노이드의 오른쪽 무릎 아래의 동작을 제어하는 것이며, 동작의 목표(goal)를 액션 클라이언트로부터 수신하여 목표 달성을 위해 필요한 서보 모터 동작을 위한 토픽을 게시한다. 둥근 사각형으로 표시한 것은 복잡한 동작을 결정하는 액션 클라이언트로 액션 서버에서 전송한 피드백 정보를 이용하여 현재의 상태를 모니터링하며 필요에 따라 새로운 목표를 전송한다.

Ⅲ. 결 론

본 논문에서는 관절의 움직임을 구현하기 위해 각 서보 모터 제어를 위한 단계를 계층화하고, 각각을 ROS 노드로 설계하는 것을 제안하였다. 휴머노이드 로봇 동작 제어를 위해 본 논문에서 제안한 방법은 복잡한 동작을 기본적인 동작의 조합으로 표현할 수 있어 환경 변화에 적응한 효과적인 동작 제어가 가능할 것으로 판단된다.

감사의 글

본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다. (2021RIS-001(1345370811))

참 고 문 헌

[1] Y. Kim, K. Kim, and Y. Bae, "An implementation of vector control of AC servo motor based on optical-EtherCAT network", The Journal of The Korea Institute of Electronic Communication Sciences, Vol.8, No.4, pp.583-587, 2013

[2] Y. Moon, S. Roh, K. Jo, and Y. Bae, "Design of robot joint structure using multiple motors", The Journal of The Korea Institute of Electronic Communication Sciences, Vol.7, No.2, pp.417-423, 2012

[3] Y. Moon, S. Roh, S. Lim, and Y. Bae, "An implementation of the Control System of the Mobile Robot using ROS," The Journal of The Korea Institute of Electronic Communication Sciences, Vol.8, No.11, pp.1713-1718, 2023

pix2pix-Swin: CGAN을 이용한 RGB-to-NIR 변환

*박인철, **진영완, ***김시호

pix2pix-Swin: RGB-to-NIR Translation with CGAN

*In-Cheol Park, **Young-Wan Jin and ***Shi-Ho Kim*

요 약

RGB 이미지를 근적외선(Near-Infrared, NIR) 이미지로 변환하는 기술은 자율주행 차량의 안정성과 안전성 향상에 기여할 수 있다. 본 연구에서는 악천후와 비정형 교통 상황에서 수집된 공개 데이터 세트인 IDD-AW를 활용하여 RGB 이미지를 NIR 이미지로 변환하는 방법을 제안한다. 이를 위해 pix2pixHD 모델을 기반으로 하는 pix2pix-Swin 모델을 사용하였다. pix2pix-Swin 모델은 Swin Transformer를 기존의 방법들과 비교하여 비교적 선명하고 디테일한 NIR 이미지를 생성할 수 있었으며, 이는 자율주행 차량의 안전한 주행을 위한 핵심 AI 소프트웨어 개발에 기여할 수 있을 것으로 기대된다.

Key words

Image generation, Image to Image translation, Autonomous Driving

Ⅰ. 서 론

최근 자율주행 차량 기술의 발전과 함께 다양한 환경에서의 안정적인 주행 능력이 중요해지고 있다. 특히 악천후나 비정형 교통 상황에서의 안전한 운행은 매우 중요하다[1]. 이러한 필요성에 따라 일반적인 RGB 이미지와 근적외선(Near-Infrared, NIR) 이미지 간의 변환 기술은 자율주행 시스템의 성능 향상에 중요한 역할을 할 수 있다. RGB 이미지는 가시광선 영역의 정보를 제공하지만, 악천후나 저조도 환경에서는 한계가 있다. 반면, NIR 이미지는 저조도 환경에서도 물체를 잘 감지하고, 안개나 비와 같은 악천후 조건에서도 우수한 투과성을 보이는 특징이 있다. 따라서 RGB 이미지를 NIR 이미지로 변환하는 기술은 자율주행 차량의 안정성과 안전성 향상에 기여할 수 있다.

본 연구에서는 악천후와 비정형 교통 상황에서 수집된 공개 데이터 세트인 IDD-AW를 활용하여 RGB 이미지를 NIR 이미지로 변환하는 방법을 제안한다. 이를

* 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (incheol97@yonsei.ac.kr)
** 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (tthatnn@yonsei.ac.kr)
*** 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (shiho@yonsei.ac.kr)

위해 고해상도 이미지 변환을 위한 GAN(Generative Adversarial Network)기반의 pix2pixHD[2] 모델을 기반으로 한 pix2pix-Swin 모델을 제안한다. 본 연구의 목적은 제안된 모델이 기존의 방법들과 비교하여 개선된 성능을 보임을 입증하고, 이를 통해 자율주행 차량의 안전한 운행에 기여하는 것이다.

Ⅱ. 본론

2.1 데이터 수집 및 전처리

자율주행 차량 연구를 위해 악천후와 비정형 교통 상황에서 수집된 공개 데이터 세트인 IDD-AW[4]를 사용한다. 이 데이터 세트는 비, 안개, 눈, 저조도 등 다양한 주행 환경에서 수집된 5000쌍의 RGB-NIR 이미지를 포함하고 있다. 학습과 테스트를 위해 이 데이터 세트를 카테고리 별로 8:2 비율로 분할하여 각각 학습과 검증에 사용한다.

2.2 pix2pix-Swin

pix2pixHD[2] 모델은 고해상도 이미지 변환을 위해 개발된 GAN(Generative Adversarial Network) 기반의 모델로, 기존의 pix2pix 모델을 확장하여 보다 선명하고 디테일한 이미지 변환을 가능하게 한다. 이 모델은 이미지의 다양한 크기에 대해 학습할 수 있도록 여러 스케일의 생성자와 판별자를 사용하여, 이미지의 전역적 및 지역적 특징을 모두 고려할 수 있다. 또한, 생성자가 판별자의 중간 레이어 특징을 모방하도록 유도하는 Feature matching loss과 생성된 이미지와 실제 이미지 간의 높은 수준의 유사성을 측정하기 위해 미리 학습된 VGG 네트워크를 활용하는 Perceptual loss을 사용하여, 생성된 이미지의 품질을 향상시키고 모델의 안정성을 높인다. pix2pixHD 모델은 고해상도 이미지 변환 작업에서 우수한 성능을 보여주었으며, 이를 기반으로 한 pix2pix-Swin 모델은 pix2pixHD와 Swin Transformer [3]를 결합하여 이미지 변환 성능을 더욱 개선시켰다.

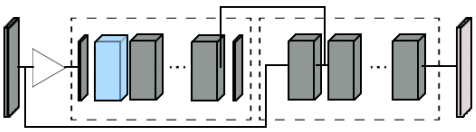


그림 1. pix2pix-Swin 개요도

Ⅲ. 결 론

본 연구에서는 RGB 이미지를 근적외선(NIR) 이미지로 변환하는 작업에 대해 다루었다. 이를 위해 pix2pixHD[2] 모델을 기반으로 하는 pix2pix-Swin모델을 제안하였다. RGB to NIR image translation 작업에 pix2pix-Swin 모델을 적용한 결과, 기존의 방법들과 비교하여 비교적 선명하고 디테일한 NIR 이미지를 생성할 수 있었으며, 제안된 모델은 RGB 이미지의 전반적인 구조를 잘 보존하면서 NIR 도메인의 정보를 반영할 수 있었다.

본 연구에서 제안된 pix2pix-Swin 모델은 RGB 이미지를 NIR 이미지로 변환하는 효과적인 방법을 제시하였다. 이 기술은 자율주행 차량의 안정적인 작동을 위한 핵심 AI 소프트웨어 개발에 기여할 수 있을 것으로 기대된다.

감사의 글

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의

지원을 받아 수행된 연구임 (RS-2023-00236245, 악천후/비정형 환경 변화에서의 Seamless 자율주행을 위한 인지/판단 AI SW 핵심기술 개발)

참 고 문 헌

[1] Yuxiao Zhang, Alexander Carballo, Hanting Yang, Kazuya Takeda."Perception and sensing for autonomous vehicles under adverse weather conditions: A survey."ISPRS Journal of Photogrammetry and Remote Sensing, 196, 2023.

[2] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro."High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs."IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[3] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows"IEEE/CVF International Conference on Computer Vision (ICCV), 2021.

[4] Shaik, Furqan, Malreddy, Abhishek, et al,"IDD-AW: A Benchmark for Safe and Robust Segmentation of Drive Scenes in Unstructured Traffic and Adverse Weather"IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024.

생성형 AI 기반 플랫폼 서비스에서의 UX 어포던스 및 지속사용의도간 관계연구

*박민혁, **구자준

A Study on the Relationship between UX Affordance and Continuous Use Intention in Generative AI-Based Platform Services

*Minhyuk Park, **JaJoon Koo

요 약

본 연구는 생성형 인공지능(AI)을 기반으로 하는 플랫폼 서비스에서의 사용자 경험(UX) 어포던스와 지속사용의도 간의 관계를 탐구하는 것을 목적으로 한다. 이를 위해, 생성형 AI 기반 플랫폼 서비스의 사용경험이 있는 사용자들을 대상으로 설문조사를 통해 데이터를 수집하고, 이를 바탕으로 실증분석을 진행하였다. 연구 결과, UX 어포던스는 사용자의 지속사용의도에 긍정적인 영향을 미치는 것으로 나타났다. 이와 같은 결과는 생성형 AI 기반 플랫폼 서비스에서 사용자의 지속사용의도를 향상시키기 위해 UX 디자인의 어포던스 요소가 가지는 중요성을 보여준다. 결론적으로, UX 어포던스를 고려하여 서비스를 디자인하는 경우, 사용자의 지속적인 플랫폼 서비스의 사용유도와 동시에 성공적인 운영에 기여할 수 있다는 시사점을 제공한다.

Key words

Generative AI, User Experience, UX Affordance, Continuous Usage Intention

I. 서 론

기술의 발전에 따라 등장한 인공지능(Artificial Intelligence, AI)은 일상생활 및 비즈니스환경 등 여러 상황에 있어서 큰 영향력을 펼치고 있다[1]. 지속적인 확장을 이루고 있는 인공지능은 최근 텍스트, 이미지 그리고 음성 등 다양한 형태의 콘텐츠를 생성과 동시에 사용자들에게 새로운 경험을 제공할 수 있는 생성형 인공지능(Generative AI)이라는 새로운 모습으로 큰 이슈로 자리잡게 되었다[2]. 생성형 AI가 발전함에 따라 이를 기반으로 한 다양한 형태의 플랫폼 서비스가 등장하고 있다.

* 성균관대학교 예술학협동과정, 박사과정(parer@naver.com)

** 성균관대학교 디자인학과, 교수, 교신저자(designer@skku.edu)

이와 같은 생성형 AI 기반의 플랫폼 서비스의 종류가 급속도로 증가함에 따라, 경쟁력과 차별화 제고를 위해 사용자 경험(User eXperience, UX) 요소 증진에 대한 필요성이 제기되기 시작되었다. UX는 사용자들이 서비스를 사용함에 있어서의 총체적 경험을 의미하며, 서비스의 지속적인 사용 여부를 결정하는 중요한 요소 중 하나라고 할 수 있다. 이를 기반으로 등장한 개념인 UX 어포던스는 사용자가 제품이나 서비스를 사용하면서 느끼는 행동 유도성을 의미하는데, 이는 사용자들이 서비스를 더욱 쉽게 이해하고 사용할 수 있다는 긍정적인 측면을 가지고 있다[3].

본 연구는 생성형 AI 기반 플랫폼 서비스에서의 UX 어포던스와 지속사용의도 간의 관계를 탐색하는 데 목적을 가지고 있다. 이를 위해, 생성형 AI 기반 플랫폼 서비스의 사용경험이 있는 사용자들을 대상으로 설문조사 기반의 데이터 수집 및 분석을 진행한 후, UX 어포던스와 지속사용의도 간의 관계를 파악하였다. 구체적으로, 본 연구는 UX 어포던스의 적용이 사용자들의 인지적, 정서적 반응을 비롯한 다양한 영역에 있어서 어떠한 영향을 미치는지 분석함으로써, 사용자들이 플랫폼 서비스를 지속적으로 사용하도록 유도할 수 있는 요인을 밝히고자 하며, 다음과 같은 주요 연구질문을 다루고자 하였다.

1. UX 어포던스가 지속사용의도에 미치는 영향은 어떠한가?

2. UX 어포던스와 지속사용의도 간의 관계 강화에 어떠한 전략이 필요한가?

본 연구의 결과는 생성형 AI 기반 플랫폼 서비스의 디자인 개선 및 플랫폼 서비스의 사용유도와 동시에 성공적인 운영 등에 있어서 유용한 시사점을 제공할 것이며, 나아가 관련 분야에 의미 있는 기초 자료를 제공할 수 있을 것으로 기대한다.

Ⅱ. 이론적 배경

2.1 생성형 인공지능

생성형 AI는 머신러닝(Machine Learning) 및 딥러닝(Deep Learning) 기반의 알고리즘을 이용하여 기존에 학습된 데이터를 기반으로 새로운 형태의 콘텐츠를 생성하는 기술이다[2].

생성형 AI는 사용자가 입력한 내용을 바탕으로 새로운 콘텐츠를 자동으로 만들어내는 특화점을 갖추고 있으며, 이와 같은 특화점은 기존에 자리매김하던 창작 및 개발의 범위를 확장하는 동시에 사용자들에게 새로운 경험을 제공하는 데 있어서 큰 영향력을 제공할 수 있다는 장점을 지닌다.

이와 관련하여 등장한 생성형 AI 기반 플랫폼 서비스는 생성형 AI를 활용함으로써 사용자에게 다양한 도구와 기능을 제공한다. 이를 통해, 사용자들이 창작 및 개발활동 등을 보다 쉽게 하고, 새로운 상상력을 구현하는 데 있어 효율적인 도구의 모습을 보여준다.

2.3 UX 어포던스

어포던스(Affordance)는 사용자가 시스템을 어떻게 인식하고 사용할 수 있는지를 결정짓는 요소로, 이는 시스템의 디자인과 사용자의 경험 사이의 상호작용을 설명하는 데 중요한 역할을 한다. 그 중, UX 어포던스는 사용자 인터페이스(User Interface, UI) 디자인 측면에서 중요한 개념으로써, 사용자와 인터페이스가 서로 상호작용할 때 직관적으로 느끼는 사용 가능성을 의미한다[3].

감각적 어포던스, 기능적 어포던스, 물리적 어포던스, 인지적 어포던스로써 총 4가지로 구성되어 있는 어포던스 중 기능적 어포던스 및 인지적 어포던스가 UX 어포던스의 주된 요소라고 할 수 있다[4]. 이와 같은 어포던스 고려에 있어서 사용자가 직관적으로 이해하고 사용할 수 있도록 설계되어야 하는 동시에, 편의성, 접근성, 피드백의 적절성 등이 고려대상이라고 할 수 있다.

2.3 지속사용의도

지속사용의도(Continuous Use Intention)는 사용자가 제품이나 서비스를 지속적으로 사용하고자 하는 의지를 의미한다[5]. 지속사용의도는 UX 및 만족도 등이 높을수록 강화되며, 시스템의 성공과 함께 장기적인 사용자 유지에 중요한 지표가 된다. 지속사용의도는 주로 정보시스템(Information System, IS) 연구에서 많이 다루어지는 측면이 있다[6].

기술수용모델(Technology Acceptance Model, TAM), 기대일치이론(Expectation Confirmation Theory, ECT), 자기효능감이론(Self Efficacy, SE), 플로우이론(Flow Theory) 등의 이론을 바탕으로 연구되는 횟수가 많다. 지속사용의도는 사용자의 만족도와 사용용이성, 지각된 유용성 등에 의해 큰 영향을 받으며, 시스템을 통해 자신의 목표를 효과적으로 달성할 수 있다고 믿는 정도가 높아지는 경우 지속사용의도에 긍정적인 영향을 가져온다고 할 수 있다.

Ⅲ. 연구방법

본 연구의 대상은 생성형 AI 기반 플랫폼 서비스 사용경험이 있는 10대부터 50대 이상까지 다양한 연령층의 사용자들로 한정하였고, 사용자들을 대상으로 설문조사를 실시하였다. 구글 폼을 활용한 온라인 설문조사를 통해 데이터를 수집하였고, 설문조사 기반의 데이터 수집 후, SPSS를 활용하여 설문조사 결과를 분석하였다. 상관관계 분석을 통해 UX 어포던스와 지속사용 의도 간의 관계를 파악하였다.

설문조사지는 기능적 어포던스 5문항, 인지적 어포던스 6문항, 지속사용의도 6문항을 포함하여 총 16문항으로 구성되었으며, 응답 시간은 약 7분 내외였다. 설문조사는 총 1주 동안 진행되었으며, 총 108명의 응답이 수집되었다. 이 중 불완전하거나 일관성이 없는 응답을 제외하고 최종적으로 100명의 응답을 분석에 사용하였다.

표 1. 신뢰도 통계량

Cronbach의 알파	표준화된 항목의 Cronbach의 알파
.834	.836

표 2. UX 어포던스와 지속사용의도 간 상관관계

	기능적 어포던스	인지적 어포던스	지속 사용의도
기능적 어포던스	1		
인지적 어포던스	.728**	1	
지속 사용의도	.537**	.626**	1

**p<0.01

Ⅳ. 결 론

본 연구는 생성형 AI 기반 플랫폼 서비스의 UX 어포던스와 지속사용의도 간의 관계를 탐색하였다. 연구결과를 기반으로 하여 앞서 언급한 연구질문에 대해 살펴보았을 때,

도출할 수 있는 내용은 다음과 같다. UX 어포던스는 지속사용의도에 대체적으로 긍정적인 영향을 미치는 것을 확인할 수 있었다. 또한, UX 어포던스와 지속사용의도 간의 관계 강화를 위해 사용자 데이터 수집 및 분석과 피드백 제공, 지속적인 업데이트 등의 요소가 추가적으로 필요할 것으로 보였다.

본 연구는 사용자의 어포던스를 고려하여 서비스를 디자인하면 사용자의 지속 사용을 유도할 수 있으며 서비스의 성공적인 운영에 기여할 수 있다는 시사점을 제공한다. 하지만, 본 연구는 다음과 같은 한계점이 발견되었다. 표본의 크기가 작아 연구 결과의 일반화에 한계가 있었으며, 생성형 AI 기반 플랫폼 서비스의 종류가 다양하기 때문에, 범주에 따라 연구 결과가 달라질 수 있다.

향후 연구에서는 표본의 크기를 확대하고 다양한 생성형 AI 기반 플랫폼 서비스를 대상으로 UX 어포던스와 지속사용의도 간의 영향 관계를 더욱 구체적으로 파악하기 위해 다양한 변수를 고려한 연구가 필요할 것으로 생각된다. 이에 대한 보완점이 마련된다면, 생성형 AI 기반 플랫폼 서비스의 활성화에 대한 의미있는 인사이트를 제공할 수 있을 것으로 기대된다.

참 고 문 헌

[1] 국경완. (2019). 인공지능 기술 및 산업 분야별 적용 사례.정보통신기획평가원 주간기술동향,1888, 15-27.

[2] 이수환, & 송기상. (2023). 생성형 인공지능의 교육적 활용에 대한 국내 연구 동향 탐색. 컴퓨터교육학회 논문지, 26(6), 15-27.

[3] 노주희. (2021). 모바일 간편 결제 사용자를 위한 UX 인지적 어포던스 관점의 선호도 연구-국내 3 대 모바일 간편결제서비스를 중심으로. 조형미디어학, 24(1), 202-210.

[4] 안진호. (2015). 어포던스 (affordance) 중심의 정보시스템 감리에서의 UI/UX 감리방안 연구. 한국 IT 서비스학회 학술대회 논문집, 2015(3), 363-366.

[5] 정용국, & 장위. (2020). 구독형 OTT 서비스 특성이 이용자 만족과 지속 사용 의도에 미치는 영향: 넷플릭스 이용자를 대상으로. 한국콘텐츠학회논문지, 20(12), 123-135.

[6] 서필수, 이용기, & 정남호. (2013). 기대일치모형을 이용한 항공예약시스템의 지속사용의도 영향요인 고찰. 호텔경영학연구, 22(3), 249-264.

Windows VBS 악성코드 공격 동향 및 대응 방안 연구

*전규현, **이주현, ***서정택

A Study on Attack Trends and Countermeasures to VBS Malware Attacks in Windows

*GyuHyun Jeon, **JuHyeon Lee and ***Jung Taek Seo

요 약

Visual Basic Script(VBS) 기반 악성코드는 정상적인 파일로 위장하거나 취약점을 악용하여 최근까지도 멀웨어 및 RAT 계열 악성코드를 통한 피해가 지속적으로 발생하고 있다. 이에, VBS 악성코드 공격 탐지 및 대응 가능한 효과적인 방어 기법에 대한 연구가 필요하다. 이에 본 논문에서는 Windows 환경에서의 VBS 악성코드 공격 사례를 분석하고 사용된 공격 기법을 분류하여 정형화한다. 이후, Python Watchdog 기반 이중 확장자 파일 실시간 모니터링, 차단, 삭제 자동화를 구현한다. 또한, Windows 이벤트 뷰어 내 이벤트 ID 및 XML 기반 Windows PowerShell 로그 분석을 통한 VBS 악성코드 탐지, 네트워크 구성 변경 여부 확인을 통한 프로세스 차단 및 삭제를 제안하였다.

Key words

VBS, Script, Malware, Case Analysis, Windows

Ⅰ. 서 론

Visual Basic Script(VBS) 기반 악성코드는 문서형, 실행 파일, LNK 파일 등 다양한 형태로 존재하며, 정상적인 파일로 위장하거나 취약점을 악용하여 사용자가 파일을 실행하도록 유도한 후, 실행 시 악성행위가 수행된다[1]. 최근까지도 VBS를 악용한 다크게이트(DarkGate), 이모텟(Emotet), 칵봇(QakBot) 등 멀웨어 및 Remote Access Trojan(RAT) 계열의 악성코드 공격이 발생하고 있다[2]. 이로 인해 공격 대상의 시스템과 네트워크 등 공격 표면(Attack Surface)에 초기 접근한 후, 키로깅 등 개인 정보 탈취 및 로더(Loader)를 통한 악성 페이로드 다운로드 등 추가적인 피해가 지속적으로 발생하고 있다. 또한, 악성코드 탐지 회피를 위한 VBS 난독화(Obfuscate) 등 다양한 공격 기법이 개발되고 있다. 따라서 VBS 악성코드 공격 탐지

* 가천대학교 정보보호학과 석사과정 (pengchan88@gachon.ac.kr)
** 가천대학교 정보보호학과 박사과정 (202240226@gachon.ac.kr)
*** 가천대학교 컴퓨터공학부 스마트보안전공 교수 (seojt@gachon.ac.kr)

및 대응 가능한 효과적인 방어 기법에 대한 연구가 필요하다.

이에 본 논문에서는 Windows 환경에서의 VBS 악성코드 공격 대응 방안에 대한 연구를 진행한다. 먼저 실제 Windows VBS 악성코드 공격 사례를 분석하여 사용된 공격 기법을 도출한다. 이를 기반으로 Python Watchdog 기반 이중 확장자 파일에 대한 실시간 모니터링, 차단, 삭제 자동화를 구현한다. 또한, Windows 이벤트 뷰어 내 이벤트 ID 및 XML을 기반으로 Windows PowerShell 로그 분석을 수행하여 VBS 악성코드를 탐지한다. 이후, 네트워크 구성 변경 여부 확인을 통한 프로세스 차단 및 삭제를 제안하였다.

2장에서는 VBS 개요, 3장에서는 VBS 악성코드 공격 동향 분석, 4장에서는 VBS 악성코드 공격 대응방안, 5장에서는 결론 및 향후 연구방향에 대해 기술하였다.

Ⅱ. VBS 개요 및 취약점

2.1 VBS 개요

VBS는 MicroSoft社에서 개발한 Visual Basic 기반의 경량 액티브 스크립트 언어이다[3]. VBS는 기본적으로 Function/End Function 구성으로 작성되어 있으며, 실행 중인 환경의 요소에 접근하기 위해 Microsoft Component Object Model(COM) 개체를 사용하는 스크립팅 기법을 사용한다.

[표 1] VBS 주요 특징

VBS(*.vbs)	설명
목적	• 작업 자동화
실행 환경	• Windows(IE, IIS, WSH 등)
기능	• 네트워크 구성 변경 • 시스템 백업 및 계정 관리 • 문서 프로그램 매크로 등
정찰	• 주로 합법적인 이메일 또는 다운로드 파일로 위장

[표 1]은 VBS의 사용 목적, 실행 환경, 공격 표면 접근을 위한 정찰(Reconaissance) 방법, 주요 기능 등 주요 특징을 나타낸 것이다. VBS 기능은 Windows 98 이후의 버전에 기본적으로 내장되어 있으며 주로 Internet Explorer(IE), Internet Information Services(IIS), Windows Scripting Host(WSH) 등 Windows 호스트 환경에서 실행 가능하다[4].

해당 환경에서 네트워크 구성 변경 시, 시스템 백업, 계정 관리 등 시스템 관리 영역에서 반복적인 작업의 자동화를 위해 사용되고 있다. 최근에는 HWP, PDF 파일 및 MS Office의 Word 및 Excel 등 문서화 작업 기능을 지원하는 응용 프로그램에서 매크로로 자주 사용된다[5].

2.2 VBS 취약점

[표 2] VBS 주요 취약점 및 난독화 기법

VBS(*.vbs)	설명
취약점	• 원격/임의 코드 실행 허용 • VBS 코드 주입 • 개인 정보 탈취 등
난독화	• Wscript.Shell, Base64, XOR 등

[표 2]는 VBS 주요 기능의 취약점을 악용하는 공격을 수행하기 위해 자주 사용되는 취약점 및 난독화 기법을 나타낸 것이다. 먼저, 공격 정찰 단계에서는 악성 VBS가 포함된 파일을 합법적인 이메일 또는 다운로드 파일로 위장한 후 배포하는 방법이 자주 사용된다[6]. 이후, 공격 목표 달성을 위한 다양한 VBS 취약점을 악용한다. 현재 공개적으로 알려진 보안 취약점 목록인 Common Vulnerabilities and Exposures (CVE)를 기준으로 102개의 VBS 취약점이 존재한다[7]. 공격에 악용되는 주요 취약점에는 원격 코드 및 임의 코드 실행 허용, VBS 코드 주입, 개인 정보 탈취가 존재한다.

해당 취약점을 포함한 공격 코드를 작성할 시, 공격 탐지를 우회하기 위한 Wscript, Shell, Base64, XOR 인코딩 등 다양한 난독화 기법이 사용되며 여러 번 난독화도 가능하다[8].

Ⅲ. VBS 악성코드 공격 동향 분석

3.1 Konni APT

2024년 03월, 북한 Konni로 추정되는 APT 그룹의 공격을 통해 국내 기업을 대상으로 '첨부.zip' 악성파일이 배포되었다[9]. 해당 파일은 '첨부1_성명_개인정보수집이용동의서.docx.lnk', '첨부2_*** 메일 내용(참고).pdf' 총 2개의 하위파일로 구성되어 있으며, LNK(바로가기) 파일이 실제 공격을 수행하는 악성 파일이다. 이는 이중 확장자 취약점을 악용한 것으로, MicroSoft社의 Word DOCX 문서로 위장한다. 해당 악성파일은 난독화된 PowerShell 명령을 포함하고 있으며, 실행 시 바로가기 전체 길이 값인 0x16EF7F1A를 확인한다. DOCX 파일은 원본과 같은 경로에 동일한 이름으로 생성 및 공용(Public) 폴더 경로에 'UHCYbG.cab' 파일명의 Windows CAB 파일을 생성한다.

[표 3] start.vbs 코드

행	VBS 코드
1	set obj = GetObject("new:9BA05972-F6A8-11CF-A442-00A0C90A8F39")
2	set itemObj = obj.Item()
3	jHxescZHDcebgaNZ = Left(WScript.ScriptFullName, InstrRev(WScript.ScriptFullName, "₩") - 1)
4	itemObj.Document.Application.ShellExecute jHxescZHDcebgaNZ & "₩" & "09402649" & ".b" & "at", Null, jHxescZHDcebgaNZ, Null, 0
5	set obj = Nothing

CAB 파일의 압축이 해제되면 [표 3]과 같이 'start.vbs' 및 여러 BAT 파일이 생성되며, 'start.vbs' 스크립트 내 지정된 BAT 파일이 실행되면서 사용자 정보 수집 및 유출, 추가 악성파일 설치, 레지스트리 등록을 통한 지속성 획득 등 악성 작업을 수행한다.

3.2 Ande Loader

2024년 04월, 북미 제조업을 대상으로 Ande Loader 악성코드가 유포되었다[10]. 해당 악성코드는 Base64로 인코딩된 PE 파일을 포함한 VBS 이며, PowerShell로 디코딩하여 함수 호출, C&C 서버 주소 정보가 포함된 파라미터 전달, 그리고 PE 파일을 실행할 수 있는 'VAI' 메서드를 직접 호출한다. 이후, C&C 서버와 연결되어 Base64로 인코딩된 최종 페이로드인 'NjRAT' 악성코드를 다운로드 및 디코딩한다. 'NjRAT' 악성코드를 실행하기 위해 이미지 스위칭(Image Switching) 기법을 사용한다. [표 4]는 이미지 스위칭 기법에서 사용된 프로세스 관련 함수를 나타낸 것이다.

[표 4] 사용된 프로세스 관련 함수

함수명	설명
Create Process	• 프로세스 생성
Write Process Memory	• 현재 프로세스의 지정된 버퍼에서 지정된 프로세스의 주소 범위로 데이터를 복사(쓰기)
SetThread Context	• 지정된 스레드의 Context를 설정 • 실행 주소에 대한 Context를 불러온 후, 수정 및 설정함
Resume Thread	• 스레드 시작 또는 재시작

먼저 .NET 프레임워크를 구성하는 프로세스를 무작위로 선정한다. 이후, 프로세스를 생성 및 정상 프로세스의 코드를 메모리상에서 매핑 해제하는 'ZwUnmapViewOfSection' API를 사용한다. 해제된 영역에 교체할 악성 PE

이미지를 삽입하여 스레드의 실행 주소를 삽입된 PE 이미지의 Entry Point로 설정한 후, 스레드를 시작한다. 이미지 스위칭 이후, 'NjRAT' 악성코드는 앞서 무작위로 선정한 정상 프로세스명으로 위장하여 실행되며 시스템의 키로깅 데이터를 수집하여 레지스트리에 저장한 후, C&C 서버에 업로드한다.

3.3 Remcos RAT

2024년 06월, Unix-to-Unix Encode (UUE)로 인코딩된 VBS 난독화 파일을 통한 Remcos RAT 악성코드 다운로더가 견적서 등으로 위조한 피싱 메일 형태로 유포되었다[11]. UUE는 binary 데이터를 ASCII 텍스트 형식으로 인코딩하는 방식이며, Anti-Virus 프로그램의 탐지를 우회하기 위해 적용한 것으로 추정된다. UUE 파일을 디코딩하면 난독화된 VBS를 확인 가능하다.

해당 VBS 스크립트는 "C:\Users\사용자\AppData\Local\Temp" 경로에 'Talehmmedes.txt' 파일명으로 Power Shell를 저장한 후 실행한다. 해당 TXT파일은 "C:\Users\사용자\AppData" 경로에 PowerShell 스크립트 실행 기능을 수행하는 'Haartoppens.Eft' 파일과 추가 PowerShell 스크립트를 다운받기 위한 C&C 서버의 문자열이 포함되어 있다. 추가 PowerShell 스크립트는 지속성 획득을 위한 레지스트리를 등록하며 최종적으로 Remcos RAT 악성코드를 다운 및 실행하기 위한 또 다른 C&C 서버에 접속한다. Remcos RAT 악성코드는 시스템의 키로깅 데이터를 수집하여 "C:\Users\사용자\AppData" 경로에 저장한 후, C&C 서버에 업로드한다.

3.4 공격 기법 도출 및 정형화

[표 5]는 앞서 분석한 공격 사례를 기반으로 각 공격 사례에서 사용한 공격 기법을 도출한 후, MITRE ATT&CK 프레임워크의 공격 기법에 매핑하여 정형화한 것이다. MITRE ATT&CK 프레임워크는 MITRE Corporation社에서 개발한 다양한 공격 기법에 대한 정보를 분류하는 보안 프레임

[표 5] 공격 사례별 사용된 공격 기법 및 정형화

악성코드	공격 기법	MITRE ATT&CK Techniques
Konni APT	• 스피어 피싱 • 사용자의 이중 확장자 파일 실행 • PowerShell 및 VBS 실행 • PowerShell 난독화 • 레지스트리 등록을 통한 지속성 확보 • 키로깅을 통한 사용자 정보 수집 • C2 서버 통신(악성 페이로드 다운로드 및 키로깅 정보 업로드)	• T1566(Phishing) • T1204(User Execution) • T1059(Command and Scripting Interpreter) • T1027(Obfuscated Files or Information) • T1547(Boot or Logon Autostart Execution) • T1056(Input Capture) • T1071(Application Layer Protocol) 및 T1041(Exfiltration Over C2 Channel)
Ande Loader	• PowerShell 및 VBS 실행 • 'NjRAT' 악성코드의 Base64 난독화 • 'ZwUnmapViewOfSection' API 사용(이미지 스위칭) • 정상 프로세스명으로 위장한 후, 악성코드 실행 • 키로깅을 통한 사용자 정보 수집 • C2 서버 통신(악성 페이로드 다운로드 및 키로깅 정보 업로드)	• T1059(Command and Scripting Interpreter) • T1027(Obfuscated Files or Information) • T1055(Process Injection) • T1574.002(DLL Search Order Hijacking) • T1056(Input Capture) • T1071(Application Layer Protocol) 및 T1041(Exfiltration Over C2 Channel)
Remcos RAT	• 견적서 등으로 위조한 피싱 메일 • 사용자가 위조된 파일을 실행 • VBS 및 PowerShell 스크립트 실행 • UUE 인코딩을 통한 난독화 • 레지스트리 등록을 통한 지속성 확보 • 키로깅을 통한 사용자 정보 수집 • C2 서버 통신(악성 페이로드 다운로드 및 키로깅 정보 업로드)	• T1566(Phishing) • T1204(User Execution) • T1059(Command and Scripting Interpreter) • T1027(Obfuscated Files or Information) • T1547(Boot or Logon Autostart Execution) • T1056(Input Capture) • T1071(Application Layer Protocol) 및 T1041(Exfiltration Over C2 Channel)

워크이다[12]. 실제 사이버 공격 사례를 기반으로 공격자의 행위를 여러 가지 전술과 기법으로 분류하여 정형화할 수 있다.

도출된 공격 기법을 정형화한 결과, Konni APT 및 Remcos RAT 악성코드에서 사용한 공격 기법이 서로 유사함을 확인하였다. Ande Loader 악성코드의 경우, 프로세스의 이미지 스위칭 기법을 사용하는 등 프로세스 관련 공격 기법의 비중이 높은 것을 알 수 있다.

IV. VBS 악성코드 공격 대응방안

4.1 VBS 악성코드 공격 대응방안 개요

VBS 악성코드 파일을 탐지하여 생성 또는 실행되기 전 자동적으로 중지시키거나 삭제해야한다. 앞선 공격 사례에서는 이중 확장자 취약점을 악용하여 LNK 파일에 대한 사용자의 직접적인 실행을 유도하였다. 따라서 이중 확장자 파일 생성에 대한 파일 탐지 및 삭제가 필요하다. [표 6]은 해당 기능을 구현 및 악성코드 정상 작동의 차단을 구현 및 검증하기 위한 환경을 나타낸 것이다. VM 가상머신 내 Windows 7 및 10 운영체제를 설치하여 Python 패키지를 설치한 후, 소스코드를 작성하였다.

[표 6] 구현 및 검증 환경

유형	설명
가상 머신	• VMWare Workstation 16 Pro
운영체제	• Windows 7, 10
프로그래밍언어	• Python
소스 코드 편집 도구	• Visual Studio Code

4.2 VBS 악성코드 공격 대응방안 구현 및 검증

본 절에서는 앞서 분석한 실제 VBS 악성코드를 구축한 검증 환경에서 구동하여 제안한 기법이 VBS 악성코드를 정상적으로 탐지 및 차단하는지 검증한다.

Python Watchdog은 지정한 파일 경로에서 파일 시스템 이벤트에 대한 모니터링을 수행하는 파일 모니터링 모듈이다. [표 7]은 Python Watchdog을 사용하여 이중 확장자 파일 탐지를 위한 모니터링을 수행하는 코드를 구현한 것이다.

[표 7] 실시간 탐지 및 삭제 관련 주요 Python 코드

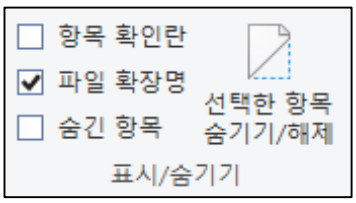
Class	주요 코드
Watcher	DIRECTORY_TO_WATCH = "파일 경로" def __init__(self): self.observer = Observer()

Handler	parts = file_name.split('.')
	if len(parts) > 2:
	try:
	os.remove(file_path)
	print(f"삭제 완료: {file_path}")
	except Exception as e:
	print(f"삭제 실패: {file_path}")

Watchdog는 Watcher 및 Handler로 구분되며 Watcher 에서는 모니터링할 파일 경로를 지정하고 Observer 객체를 생성하여 실시간 모니터링을 수행할 수 있도록 한다. Handler 에서는 이중 확장자 문자열 탐지 조건을 설정한 후, 탐지된 파일을 자동으로 삭제하였다. 작성한 코드를 실행하여 모니터링 프로세스를 추가한 후, Konni APT 악성코드 파일을 생성한 결과, '.' 문자열이 2개 이상인 이중 확장자 파일을 성공적으로 탐지 및 자동 삭제하였다. 한편, 상당수의 문서형 악성코드는 %temp%, %AppData%, %Public% 경로에 악성 VBS 파일을 생성한다. 해당 파일 경로들은 악성파일을 저장하는 경로로 자주 악용되므로 Watcher 내 경로 설정 및 추가적인 VBS 및 BAT 확장자 파일 필터링을 통해 삭제 및 실행을 차단해야 한다.

또한, 사용자는 [그림 1]과 같이 파일 확장명 설정을 활성화 하여 파일 확장자를 확인해야 한다. LNK 확장자는 바로가기 파일

유형이며 [그림 2]와 같이 파일 아이콘의 화살표 이미지가 표시되어 있으므로 실행 시 유의해야한다.



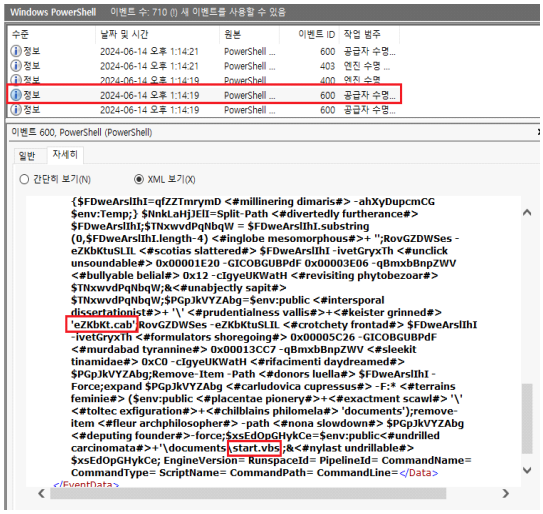
[그림 1] 파일 확장자 활성화

[그림 2] LNK 파일 유형 및 아이콘

시스템에서 VBS 생성 및 실행 시, 발생하는 이벤트 로그를 확인할 수 있는 이벤트 뷰어(eventvwr.msc)기능을 이용하여 로그 확인 및 감사·추적하여 이상 행위가 탐지될 경우 차단하여야 한다. 또한, 윈도우 이벤트 뷰어에서 객체 액세스 감사 기능을 실행하게 되면 보안 로그가 저장되는데, 발생한 이벤트 내용에 따라 저장되는 ID가 다르므로 각 이벤트 ID에 해당하는 내용들을 분석하여 공격 탐지 및 차단해야 한다. [그림 3]은 Konni APT 악성코드에 대한 이벤트 뷰어 내 로그를 나타낸 것이다.

Konni APT 악성코드는 PowerShell을 사용하는 악성코드이므로 Windows PowerShell 이벤트를 확인하여 분석하였다, 먼저 이벤트 ID 400은 PowerShell 시작, ID 403은 PowerShell 중지, ID 600은 PowerShell 코드 실행 이벤트를 나타낸다. XML 형태로 분석한 결과, 악성코드 생성 시 'UHCYbG.cab' 및 'start.vbs' 파일을 생성하는 PowerShell 명령 인자 값이 탐지되었다.

또한, 공격자는 추가 페이로드를 다운로드 받기 위해 C&C 서버 연결을 이용하기 때문에



[그림 3] Windows 이벤트 뷰어, Konni APT 악성코드 분석 결과

네트워크가 새로 구성되므로 네트워크 구성이 변경되는지 확인하여야 하며, 일반적으로 실행되지 않거나 인식되지 않는 프로세스가 생성되는지 모니터링한 후, 차단 및 삭제하여야 한다.

V. 결론 및 향후 연구방향

본 논문에서는 Windows 환경에서의 VBS 악성코드 공격 대응 방안에 대한 연구를 수행하였다. 그 결과, 최근 3개월 이내의 공격 사례 3건을 확인 및 분석하였다. Python Watchdog을 기반으로 이중 확장자 파일 실시간 모니터링, 차단, 삭제 자동화를 구현하였다. 또한, VBS 공격에 자주 사용되는 파일 경로를 분석 및 VBS 파일 필터링을 제안하였다. 그리고 Windows 이벤트 뷰어 내 이벤트 ID 및 XML 기반 Windows PowerShell 로그 분석을 통한 VBS 악성코드 탐지 및 네트워크 구성 변경 여부 확인을 통한 프로세스 차단 및 삭제를 제안하였다.

향후 연구로는, 본 제안 방안 이외의 정적 및 동적 탐지 기법을 조사하고, 실제 공격 사례에 해당 기법들을 적용하여 도출된

결과를 기반으로 비교 분석 연구를 진행할 것이다.

Acknowledgement

이 논문은 2024년도 정부(과학기술 정보통신부)의 재원으로 정보통신기획 평가원의 지원을 받아 수행된 연구 (No.2021-0-00493, 5G Massive 차세대 사이버공격 기만기술 개발, 50%)이며, 2024년도 정부 (과학기술 정보통신부)의 재원으로 정보통신기획 평가원의 지원을 받아 수행된 연구임 (RS-2024-00354169, 위협모델/XAI 기반 네트워크 이상행위 탐지·대응 및 사이버위협 예측 기술 개발, 50%)

참 고 문 헌

[1] KHUSHALI, Vala. A Review on Fileless Malware Analysis Techniques. International Journal of Engineering Research & Technology (IJERT), 9.05, 2020
[2] 보안뉴스, MS 오피스 매크로 대신 공격자들이 사용해 오던 VB스크립트, 조만간 사라진다, https://m.boannews.com/html/detail.html?idx=122647, 2023.10 [last access 2024/06/24]
[3] Bandlish, M., & Jain, N., Automation Testing, 2021.
[4] tutorialspoint, VBScript – Overview, https://www.tutorialspoint.com/vbscript/vbscript_overview.htm, [last access 2024/06/24]
[5] RoleCatcher, VBScript: The Complete Skill Guide, https://rolecatcher.com/en/skills/knowledge/information-and-communication-technologies/software-and-applications-development-and-analysis/vbscript/, 2023.12 [last access 2024/06/24]
[6] MITRE, Command and Scripting Interpreter: Visual Basic, https://attack.mitre.org/techniques/T1059/005/, [last access 2024/06/24]
[7] CVE, https://cve.mitre.org/index.html, [last access 2024/06/24]
[8] Koutsokostas, Vasilios, et al., Invoice# 31415 attached: Automated analysis of malicious Microsoft Office documents, Computers & Security, 114, 102582, 2022
[9] Genians, 비트코인 시세 급등에 따른 해킹 피해 주의보, https://www.genians.co.kr/blog/threat_intelligence/bitcoin, 2024.03 [last access 2024/06/24]
[10] 잉카인터넷, 이미지 스위칭을 이용하는 Ande Loader, https://isarc.tachyonlab.com/5628, 2024.04 [last access 2024/06/24]
[11] ASEC, Remcos RAT Distributed as UUEncoding (UUE) File, https://asec.ahnlab.com/en/66463/, 2024.06 [last access 2024/06/24]
[12] MITRE, MITRE ATT&CK, https://attack.mitre.org/, [last access 2024/06/24]

생성형 대형 언어 모델(LLM) 활용 영상의 날씨 조건 자동 인식 및 분류 방법

*주형진, **송한빈, ***김시호

Generative LLM-based automatic Classification and annotation of Weather environment in Image dataset

*Hyeongjin Ju, **Hanbin Song and ***Shiho Kim

요 약

최근 기후 변화로 인한 악천후의 빈도와 강도가 증가하면서, 특히 자율주행 상황에서의 악천후 인식의 중요성이 더욱 부각되고 있다. 정확한 날씨 예측과 분류는 안정적인 자율주행을 위해 필수적이다. 본 연구는 자율주행 차량이 악천후 조건을 효과적으로 인식하고 대응할 수 있도록, 대형 언어 모델(LLM)을 활용한 이미지-텍스트 쌍 데이터 생성 방법을 제안한다. 다양한 악천후 상황을 나타내는 이미지를 수집하고, 이에 상응하는 텍스트 설명을 자동 생성하여 이미지와 텍스트 간의 연관성을 강화하였다. 이를 통해 기계 학습 모델이 악천후 조건을 보다 정확하게 분류할 수 있도록 지원한다. 본 연구에서 제시하는 방법은 기상 데이터의 품질을 높이고, 기상 예측 시스템의 성능 향상에 실질적인 기여를 할 뿐만 아니라 자율주행 기술의 안전성과 신뢰성을 크게 향상시킬 것으로 기대된다.

Key words

Adverse Weather environment, Autonomous Driving, Lager Language Model

I. 서 론

최근 기후 변화로 인한 악천후의 빈도와 강도가 증가하면서, 안전하고 효율적인 자율주행 시스템 개발의 필요성이 더욱 강조되고 있다. 자율주행 차량은 다양한 주행환경에서 안정적으로 동작해야 하며, 특히 악천후 상황에서의 인식과 대응 능력은 차량의 안전성과 직결된다[1]. 눈, 비, 안개, 폭풍 등 다양한 악천후 조건은 자율주행 시스템의 센서와 알고리즘에 큰 도전 과제를 제시한다. 이러한 도전 과제를 극복하기 위해서는 정확한 날씨 예측 및 분류가 필수적이다.

기존의 자율주행 시스템은 주로 특정 기상 조건에서 훈련된 데이터 세트를 기반으로

* 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (wngudwls000@yonsei.ac.kr)

** 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (thgksqls369@yonsei.ac.kr)

*** 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (shiho@yonsei.ac.kr)

동작하였으나, 악천후 상황에서의 성능은 여전히 제한적이다. 이는 악천후 조건을 충분히 반영하지 못한 데이터 세트와 기계 학습 모델의 한계로 인해 발생한다. 따라서 자율주행 시스템이 다양한 악천후 상황을 효과적으로 인식하고 대응할 수 있도록 하는 방법론의 개발이 절실히 요구된다.

본 연구에서는 대형 언어 모델(LLM[2])을 활용하여 이미지-텍스트 쌍 데이터 생성 방법을 제안한다. 이 방법은 다양한 악천후 상황을 나타내는 이미지를 수집하고, 이에 상응하는 텍스트 설명을 자동으로 생성함으로써 이미지와 텍스트 간의 연관성을 강화하는 것을 목표로 한다. 이를 통해 기계 학습 모델이 악천후 조건을 정확하게 분류하고 예측할 수 있도록 지원한다.

Ⅱ. 본 론

2.1 데이터 수집 및 전처리

주행 상황에서 악천후 이미지-텍스트 쌍 데이터 생성을 위해 ACDC 데이터 세트[3]를 활용하였다. ACDC 데이터 세트는 자율주행 차량의 다양한 주행환경에서 수집된 고해상도 이미지와 풍부한 주석 데이터를 포함하고 있어, 악천후 상황을 효과적으로 반영할 수 있는 이상적인 데이터 세트이다. 특히, ACDC 데이터 세트는 주행 상황에서의 악천후 관련 벤치마크로 널리 사용되고 있어, 다양한 기상 조건에서 자율주행 시스템의 성능을 평가하는 데 중요한 역할을 한다.

ACDC 데이터 세트는 맑음, 눈, 비, 안개 등의 다양한 기상 조건을 포함하며, 각 이미지에 대한 상세한 주석 정보를 제공한다. 본 연구에서는 이 데이터를 기반으로 대형 언어 모델을 활용하여 이미지-텍스트 쌍 데이터를 생성했다.

2.2 대형 언어 모델

대형 언어 모델(LLM)은 대규모 데이터 세트에서 학습하여 자연어 처리 및 생성 능력을 갖춘 모델로, 최근 다양한 분야에서 혁신적인 성과를 보여주고 있다. 최근 발표된 대표적인 LLM으로는 OpenAI의 GPT-4[4]와 같은 모델이 있으며, 이들은 수십억 개의 매개변수를 사용하여 복잡한 언어 패턴을 이해하고 생성할 수 있다. LLM은 언어의 문법과 의미를 이해하고, 텍스트의 맥락을 파악하여 자연스럽고 유창한 문장을 생성할 수 있다. 또한, LLM은 새로운 데이터에 대해 빠르게 적응할 수 있는 능력을 갖추고 있어, 몇 가지 예시만으로도 특정 작업을 수행하는 데 필요한 지식을 학습할 수 있다.

2.3 데이터 생성

본 연구에서 수행한 이미지-텍스트 쌍 데이터 생성 과정은 다음과 같다. 먼저 LLM에 텍스트를 생성할 이미지와 함께 입력할 질문을 작성한다. 이때, 입력하는 질문의 품질에 따라 생성되는 텍스트의 정확도가 달라지기 때문에 적절한 질문을 제공하는 것이 중요하다. 질문을 작성하고 이미지 데이터와 함께 LLM에 입력하여 이미지-텍스트 쌍 데이터를 생성한다. 이후 생성한 텍스트의 정확도를 측정하기 위해 정답 클래스와 비교한다. 본 연구에서는 데이터 생성을 위한 LLM으로 GPT-4o를 사용하였다. 생성한 데이터의 정확도는 아래 표와 같다.

표 1. GPT-4o 정확도

		Prediction					Recall
		Clear	Fog	Rain	Snow	Cloudy	
Ground Truth	Clear	2978	18	2	0	2	0.993
	Fog	0	1000	0	0	0	1
	Rain	0	0	1000	0	0	1
	Snow	0	0	21	979	0	0.979
Precision		1	0.982	0.978	1	0	

Ⅲ. 결 론

본 연구에서는 자율주행 차량이 다양한 악천후 상황을 효과적으로 인식할 수 있도록 하는 이미지-텍스트 쌍 데이터 생성 방법을 탐구하였다. 먼저, 악천후 관련 이미지 데이터 세트를 수집하고, 이를 전처리하였다. 데이터는 네 가지 주요 클래스(맑음, 비, 눈, 안개)로 구분되어 정확한 출력을 위한 질문과 함께 입력 데이터로 사용된다. 이후 대형 언어 모델을 활용하여 텍스트 설명을 생성하고, 이를 이미지와 결합하여 이미지-텍스트 쌍 데이터를 구성하였다.

본 연구에서 제안된 이미지-텍스트 쌍 데이터 생성 방법은 자율주행 기술의 안전성과 신뢰성을 높이는 데 중요한 기여를 할 것으로 기대된다. 나아가, 이 방법은 자율주행 상용화를 위한 중요한 기반을 마련하는 데 있어 AI 소프트웨어의 핵심 기술로 자리매김할 것이다.

감사의 글

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00236245, 악천후/비정형 환경 변화에서의 Seamless 자율주행을 위한 인지/판단 AI SW 핵심기술 개발)

참 고 문 헌

[1] ZHANG, Yuxiao, et al. Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 196: 146-177.

[2] ZHAO, Wayne Xin, et al. A survey of large language models. arXiv preprint arXiv:2303.18223, 2023.

[3] SAKARIDIS, Christos; DAI, Dengxin; VAN GOOL, Luc. ACDC: The adverse conditions dataset with correspondences for semantic driving scene understanding. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. p. 10765-10775.

[4] ACHIAM, Josh, et al. Gpt-4 technical report. arXiv preprint arXiv:2303.08774, 2023.

PNC 모델을 활용한 하천수위 예측 딥러닝 네트워크 개발

*이은서, **박귀만, ***배영철

Development of a Deep Learning Network for River Level Prediction Using the PNC Model

*Eunseo Lee, **Gwiman Bak and ***Youngchul Bae

요 약

본 연구에서는 PNC 모델을 이용하여 서울시 한강대교의 하천 수위를 예측하는 딥러닝 네트워크를 개발하였다. PNC+MLP와 LSTM+MLP 네트워크를 비교 평가하였으며, 입력 데이터로는 강수량, 습도, 온도를 사용하였다. 두 모델 모두 학습 데이터에서 우수한 성능을 보였으나 시험 데이터에서는 과적합 문제로 성능이 저하되었다. 본 연구는 다양한 딥러닝 모델의 하천 수위 예측 가능성을 확인하였다.

Key words

PNC, River flow level, Prediction, LSTM

Ⅰ. 서 론

최근 기후 변화와 도시화로 인해 홍수와 같은 극한 기상 현상의 발생 빈도가 증가하고 있다. 이러한 현상은 특히 도시 지역의 강이나 하천 수위 변화에 큰 영향을 미쳐, 효율적인 수위 예측 및 관리의 중요성이 부각되고 있다[1].

본 연구는 Positive and negative perceptron convolution (PNC) 모델을 이용하여 딥러닝 네트워크 구축하고 서울시 한강대교의 하천수위를 예측한다. PNC 모델은 PNP 모델에 CNN을 적용하여 개발한 딥러닝 모델이다[2]. 성능 비교를 위해 LSTM 모델을 이용하여 딥러닝 네트워크를 구축하여 예측 성능을 RMSE와 SMAPE로 비교한다.

Ⅱ. 본론

2.1 딥러닝 네트워크 구성도

본 연구에서는 하천수위를 예측할 수 있는 2개의 딥러닝 네트워크를 구축하였다.

* 전남대학교 전기및반도체공학과 일반대학원 (cielo678@naver.com)
** 전남대학교 전기및반도체공학과 일반대학원 (qkrrlend@naver.com)
*** 전남대학교 전기컴퓨터공학부 교수 (ycbae@jnu.com)

각각의 네트워크는 PNC+MLP, LSTM+MLP로 구성되어 있다. 입력 데이터는 강수량과, 습도, 온도를 사용하였으며 각 모델의 은닉층은 100개로 지정하였다.

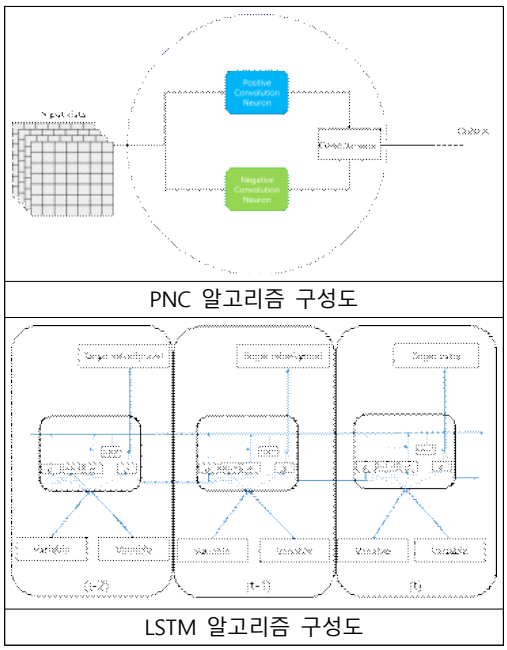


그림 1. 알고리즘의 구성도

2.1 성능 측정

표 1. 각 네트워크의 성능 비교

	PNP+MLP		LSTM+MLP	
	Train	Test	Train	Test
RMSE	40.35	1243.56	2.42	1271.30
sMAPE	14.26	48.74	0.81	47.52

표 1은 PNC+MLP 네트워크와 LSTM+MLP 네트워크의 학습 데이터와 시험데이터로 예측한 결과를 RMSE와 SMAPE로 측정한 결과를 나타낸다. PNC+MLP 네트워크는 학습 데이터에서 RMSE 40.35, sMAPE 14.26을 기록하였으나, 시험 데이터에서는 RMSE 1243.56, sMAPE 48.74로 성능이 급격히 저하되었다. 반면, LSTM+MLP 네트워크는 학습 데이터에서 RMSE 2.42, sMAPE 0.81로 매우 우수한 성능을 보였으나, 시험 데이터에서는 RMSE 1271.30, sMAPE 47.52로 마찬가지로 성능이 저하되었다.

Ⅲ. 결 론

본 연구에서는 PNC+MLP와 LSTM+MLP 네트워크를 활용하여 서울시 한강대교의 하천 수위를 예측하였다. 두 모델 모두 학습 데이터에서는 우수한 성능을 보였으나, 시험 데이터에서는 성능이 저하되는 과적합 문제를 나타냈다. 특히, LSTM+MLP 모델은 학습 데이터에서 매우 낮은 RMSE와 sMAPE를 기록하여 시계열 데이터 처리에 강점을 보였으나, 시험 데이터에서의 일반화 성능이 부족한 모습을 보였다. 결과적으로, 하천 수위 예측에 있어 다양한 딥러닝 모델의 적용 가능성을 확인하였으며, 향후 연구에서는 데이터의 다양성 및 모델의 일반화 성능을 향상시키는 것이 필요하다.

감사의 글

본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-002)

참 고 문 헌

[1] Yang, Q., Zheng, X., Jin, L., Lei, X., Shao, B., & Chen, Y., Research progress of urban floods under climate change and urbanization: a scien tometric analysis. Buildings, Vol. 11, No. 12, 2021, 628.

[2] Bak, G., & Bae, Y., Deep learning algorithm development for river flow prediction: PNP algo

rithm, Soft Computing, Vol. 27, No. 18, 2023, papers 13487-13515, doi: 10.1007/s00500-023-12345-6.

Jester를 활용한 핸드제스처 감지 방법 비교 연구

*이창용, **이종윤, ***권소영, ****이용환

Comparative Study of Hand Gesture Detection Methods using Jester

*Chang-Yong Lee, **Jong-Youn Lee, ***So-Young Kwon and ****Yong-Hwan Lee

요 약

사람과 기계간의 소통을 하는 방법 중 하나인 핸드 제스처 인식의 방법은 크게 정적 지향 방법과 동적지향 방법으로 나뉘며 이를 다시 여러 분류로 분류가 가능하다. 이 논문에서는 카메라를 사용하여 이미지를 찍은 데이터셋을 활용한 방법들에 대해 연구해보고, 이를 추후 FGPA로 구현해보는 목적을 가진다. 스킵 커넥션을 활용한 MKTB와 GRB를 사용하는 방법이 TSN을 사용한 방식보다 1.12%의 더 높은 정확도를 보인다.

Key words

Hand Gesture, Image Processing, Convolution Neural Network, Skip Connection, Jester

I. 서 론

핸드 제스처란 손의 움직임을 통하여 사람과 사람간의 소통을 하는 방법이거나, 글러브나, 카메라 등을 통하여 사람과 기계가 서로 소통하는 방법을 말한다. 이 논문에서는 이미지셋을 이용한 제스처 분류 방법에 대해 논한다. 제스처 분류 방법은 정적 제스처 지향 및 동적 제스처 지향 방법이 포함된다[1]. 정적 제스처는 정지 영상의 자세를 인식하려면 랜덤 포리스트[2]나 템플릿 매칭 방법[3]과 같은 일반 분류기로 사용이 가능하다. 동적 제스처 인식은 시간적 측면을 가지므로 비디오에서 동작을 시연하기 위해 더 많은 노력이 필요하다. FSM(Finite State Machine)[4]가 일반적으로 사용되는데, 이 경우 상태는 자세를 나타낼 수 있는 반면, 전환은 동작 정보를 나타내기 위해 사용된다. 현재는 CNN(Convolution Neural Network)을 활용한 핸드 제스처 방법이 많이 사용되고 있으며, 1DCNN, 2DCNN[5], ConvLSTM

* 금오공과대학교 일반대학원 전자공학과 박사과정 (lcy5200@kumoh.ac.kr)
** 금오공과대학교 일반대학원 전자공학과 석사과정 (ljyun0807@naver.com)
*** 금오공과대학교 일반대학원 전자공학과 박사과정 (papaya4040@kumoh.ac.kr)
**** 금오공과대학교 전자공학부 교수, 교신저자 (yhlee@kumoh.ac.kr)

[6] 등이 사용된다.

Ⅱ. 방법 비교

2.1 TSN을 사용하는 방법

TSN(Temporal Segment Network)을 사용한 방법[7]은 검출기와 분류기로 2가지로 구성된다. 검출기의 경우 일련의 이미지에서 검출기 큐 마스크를 실행하여 제스처 클래스와 제스처 클래스를 구분한다. 분류기 모델의 스위치 역할을 하는 것으로 제스처를 감지하면 분류기가 대기열 분류기의 프레임에 의해 활성화되고 공급되는 시스템이다. 분류기는 모델의 크기나 복잡성에 제한이 없어 우수한 구조를 분류기로 선택하여 사용한다.

2.2 MKTB와 GRB를 사용하는 방법

이번에 제안된 방법[8]은 기존에 사용된 3DCNN이나 ConvLSTM의 방법보다 계산 시간의 감소와 높은 정확도를 얻기 위하여 MKTB(Multi-Kernel Temporal Block)와 GRB(Global Refinement Block)를 사용한다. 여기서 제안된 MKTB는 기존의 공간 및 시간 차원에 대한 컨볼루션 연산이 아니라 시간 정보 학습에 중점을 두는 구조로 설계된다. MKTB의 효율성을 더욱 유지하기 위해 시간 컨볼루션을 사용하여 각 채널에 대해 독립적으로 계산을 수행한다. 이는 MKTB 블록이 각 채널의 시간 정보 모델링에 중점을 두는 구조를 말하고, 피라미드 형상은 요소별 합산에 의해 융합된 후, 재구성 작업이 진행된다. 다른 1x1 컨볼루션이 끝에 연결되어 출력이 입력과 동일한 수의 채널을 갖도록 한며, 모델 교육을 용이하게 하기 위해 스킵 연결이 추가한다. GRB는 모든 위치에서 가중된 합을 말하며 fully-connected layer 이전단계에서 사용되는 구조를 가진다.

Ⅲ. Dataset

사용된 데이터셋은 Jester로, 작업자들이 랩탑 카메라 및 웹캠 앞에서 수행한 148,092개의 제스처 비디오가 있는 데이터셋이다. 클래스는 27개가 있으며, 각 클래스는 평균 5,000개 이상의 인스턴스를 가지고 있는 것이 특징이다.

Ⅳ. 결 론

TSN을 사용하는 방법의 정확도도 높은 편이지만 MKTB와 GRB를 사용하는 방법이 1.12% 가량 높은 정확도를 가진다. MKTB와 GRB를 적용하여 1DCNN을 중간에 피라미드 구조로 삽입하는 스킵 커넥션 구조를 통해 더 좋은 결과를 얻었다. 추후 FPGA를 사용하여 구현할 때에도 모든 앞에서 언급한 방법들을 직접 구현하여 데이터와 비교할 예정이다. 또한 데이터셋을 자체적으로 만들지 않고 사용이 가능하도록 시스템을 수성하는 방안을 향후 과제로 한다.

감사의 글

이 논문은 2024년도 정부(산업통상 자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0017011, 2024년 산업혁신인재성장지원사업)

참 고 문 헌

[1] S. S. Rautaray and A. Agrawal, Vision based hand gesture recognition for human computer interaction: a survey, Artificial Intelligence Review, Vol. 43, 2015, pages 1–54.

[2] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, Real-time human pose recognition in parts fro m single depth images, Communications of the ACM, Vol. 56, 2013, pages 116–124.
[3] K. Oka, Y. Sato, and H. Koike, Real-time fingertip tracking and gesture recognition, IEEE Comput er graphics and Applications, Vol.22, Issues 6, 2002, pages 64–71.
[4] P. Hong, M. Turk, and T. S. Huang, Gesture modeling and recognition using finite state machines, In IEEE International Conference on Automatic Face and Gesture Recognition, Cat. No. PR00580, 2000, pages 410–415.
[4] Liang Zhang, Guangming Zhu, Peiyi Shen, Juan Song, Syed Afaq Shah, and Mohammed Benna moun, Learning spatiotemporal features using 3dcnn and convolutional lstm for gesture recog nition, In 2017 IEEE International Conference on Computer Vision Workshops(ICCV), 2017, pages 3120–3128.
[5] Pradyumna Narayana, Ross Beveridge, and Bru ce A Draper, Gesture recognition: Focus on the hands, In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 201 8, pages 5235–5244.
[6] Okan Köpüklü, Ahmet Gunduz, Neslihan Kose, Gerhard Rigoll, Real-time Hand Gesture Detecti on and Classification Using Convolutional Neur al Networks, In 2019 IEEE 20th International Conference on Information Reuse and Integrati on for Data Science, 2019. pages 1-8.
[7] Yang Yi1, Feng Ni, Yuexin Ma, Xinge Zhu, Yuank ai Qi, Riming Qiu, Shijie Zhao, Feng Li and Yongtao Wang, High Performance Gesture Rec ognition via Effective and Efficient Temporal Modeling, In IJCAI, 2019, pages 1003-1009

IoT 환경 암호화 트래픽 대상 사이버공격탐지 시스템 제안

*지일환, **이주현, ***서정택

Anomaly Detection System for Encrypted Traffic in IoT Environments

*Ilhwan Ji, **Juhyeon Lee and ***Jung Taek Seo

요 약

과거 IoT 환경에서의 평문통신 취약점을 활용한 다양한 사이버공격이 발생하였으며 금전적 피해 및 인명 피해를 발생시켰다. 이러한 이유로 IoT에 대한 암호화통신의 적용이 요구되어왔으며, 현재 대다수의 IoT 환경에 암호화 통신이 적용되었다. 암호화 통신의 적용으로 평문통신의 취약점으로 인한 사이버공격에 대한 대응은 기능해졌지만, 공격자들 또한 암호화 통신을 통해 사이버공격을 감행하고 있다. 이러한 이유로 암호화 트래픽에 대한 사이버공격 탐지가 수행되어야 한다. 하지만 암호화 트래픽은 payload를 포함하는 중요정보가 암호화되어있어 네트워크 패킷의 주요 정보를 기반으로 공격의 증거를 추출하는 접근 방식은 더 이상 유효하지 않다. 이러한 이유로, 본 논문에서는 IoT 환경에서 발생하는 암호화 트래픽 대상 사이버공격탐지 시스템을 제안한다. 본 연구에서 제안하는 IoT 환경 암호화 트래픽 대상 사이버공격탐지 시스템의 성능 평가를 위하여 IoT 환경에서 발생하는 정상 및 7개 범주의 사이버공격에 의하여 발생한 암호화 트래픽을 포함하는 dataset인 CICIoT2023를 사용하였으며, accuracy 0.99739, precision 0.99154, recall 1.0, f1 score 0.99575, roc_auc 0.99812와 같이 높은 성능을 도출하였다.

Key words

IoT cybersecurity, Ecrypted Traffic, Cyberattack Detection

Ⅰ. 서 론

IoT(Internet of Things)는 물리적 객체들이 인터넷을 통하여 서로 또는 중앙서버와 연결되어 정보를 교환하고 상호작용할 수 있게 하는 기술을 의미한다[1]. IoT는 스마트시티, 스마트 팩토리, 헬스케어, 항공/우주, 해양선박, 국방, 건설 분야 등 여러 중요 분야에 사람의 개입 없이 운영의 자동화 및 지능화를 목적으로 도입되어 활용되고 있다[2]. IoT 도입 초기에는 데이터를 암호화하고 복호화하기 위한 CPU, 배터리 등 각 구성 요소의 성능 부족으로 암호화 통신을 도입하기 어렵기 때문에 대부분 IoT 환경은 평문기반 통신을

* 가천대학교 정보보호학과 일반대학원 석사과정 (ilhwan1013@gachon.ac.kr)

** 가가천대학교 정보보호학과 일반대학원 석사과정 (02240222@gachon.ac.kr)

*** 가천대학교 컴퓨터공학부 스마트보안전공 교수 (seojt@gachon.ac.kr)

사용하였다. 평문기반의 통신은 그 자체로 보안취약점으로 작용하여 reply attack, sniffing, snooping 및 spoofing 등 다양한 사이버공격에 매우 취약하다. IoT는 여러 중요 분야에 적용되어 활용되고 있기 때문에, IoT에 대한 사이버공격이 발생할 경우 천문학적인 금전적 피해와 인명피해를 초래할 수 있다[3]. 이러한 이유로 IoT에 대한 암호화통신의 적용이 요구되어왔으며, IoT device의 컴퓨팅 성능향상과 경량암호의 발전으로 인하여 IoT에 대한 암호화 통신 적용이 가능해졌다. 암호화 통신의 적용으로 평문통신의 취약점으로 인한 사이버공격에 대한 대응이 가능해졌지만, 공격자들 또한 암호화된 통신 채널을 통해 사이버공격을 수행하고 있다. IoT를 구성하고 암호화 통신을 사용하는 장치에 대한 권한탈취 및 악성코드를 통해 제어권을 가진 뒤 해당 device를 매개로 사이버공격을 수행하게 된다면, 대부분의 사이버공격이 암호화된 채널을 통하여 발생 가능하다. IoT의 특성상 많은 device에 연결 되어있지만, 모든 device에 대한 보안관리가 매우 어려운 실정이기 때문에 여러 유형의 사이버 공격이 발생할 가능성이 높다. 하지만, 암호화 통신에 의하여 네트워크 패킷 내 payload등의 중요 정보가 암호화되어있기 때문에 네트워크 패킷의 주요 정보를 기반으로 사이버 공격을 탐지하는 방법은 암호화 트래픽에 대한 적용이 불가한다[4]. 이러한 이유로, 본 논문에서는 기존 IoT 대상 사이버공격 방식의 취약점을 보완하고 암호화 네트워크 트래픽에 대하여 사이버공격을 탐지할 수 있는 사이버공격 탐지시스템을 제안한다. 또한, 제안하는 사이버공격 탐지시스템 검증을 위하여 IoT 환경에서 수집한 암호화 트래픽을 포함하는 “CICIoT2023[5]”를 사용하였다. 실험결과 Accuracy 99.25%, Precision 97.63%, Recall 100%, F-1 Score 98.83%를 도출하였다.

본 논문의 구성은 다음과 같다. 2장에서는 IoT 대상 이상탐지 관련 연구에 대해 분석하고, 3장에서는 IoT 발생 암호화트래픽 대상 이상탐지시스템을 제안한다. 이후 4장에서는 제안하는 이상탐지시스템의 검증을 수행하여 5장에서는 결론 및 향후 연구 방향을 제시한다.

Ⅱ. 관련연구

Liu 등 6명[6]은 IoT에 대한 사이버공격 탐지를 위하여 IoT 에지 장치에 배포할 수 있는 페이로드 기반 이상 감지 프레임워크를 제안하였다. 해당 연구는 네트워크 패킷 내 payload를 식별하고 이를 머신러닝 및 딥러닝 알고리즘에 학습할 수 있는 형태로 전처리하여 이를 feature로 사용하였다. 해당 연구에서는 CNN-LSTM 기반의 사이버공격 탐지 모델을 설계하였으며, CICIDS 2017, ISCX 2012 dataset를 이용한 사이버공격 탐지 성능 검증 결과 f1 score 0.9732의 높은 성능을 도출하였다.

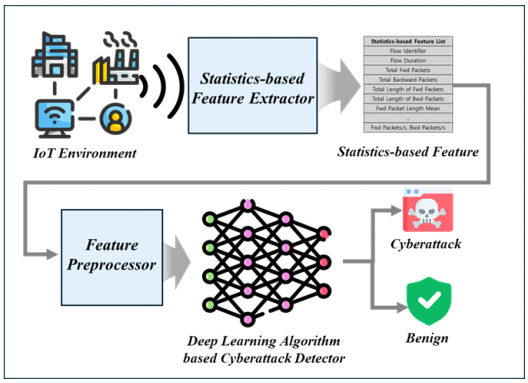
Cai 등 5명[7]은 Industrial Internet of Things (IIoT)에 대한 사이버공격 탐지를 위하여 네트워크 패킷 내 페이로드를 기반으로하는 CapBad이상탐지기를 제안하였다. CapBad은 산업용 제어 프로토콜 패킷을 모델링하고 패킷의 페이로드 특성을 자동으로 학습하고 학습한 정보를 기반으로 이상을 탐지한다. CapBad의 성능 평가 결과, roc_auc 0.974의 높은 성능을 도출하였다.

Kim 등 3명[8]은 IIoT에 대한 이상탐지를 위하여 autoencoder based payload anomaly detection (APAD)을 제안하였다. 해당 프레임워크는 네트워크 수집 후 payload를 포함한 식별할 수 있는 정보를 feature로 도출하고 이에 대한 preprocessing을 진행하였다. 이후

autoencoder 기반의 이상탐지기를 통하여 사이버공격을 탐지한다. APAD의 성능평가 결과, accuracy 0.944, recall 0.983의 높은 성능을 도출하였다.

앞서 제시한 IoT대상 사이버공격 탐지 연구가 네트워크 패킷에 대해 deep packet inspection(DPI)를 수행하여 주요정보를 식별하고 이를 feature로 사용하거나 payload의 직간접적인 통계 정보를 feature로 사용하는 방식의 연구가 대다수였다. 하지만 IoT 환경에 대한 암호화 통신 방식 적용으로 인하여 IoT 환경 발생 암호화 트래픽 대상 이상탐지 연구수행이 필요하다.

Ⅲ. IoT 환경 암호화 트래픽 대상 사이버 공격 탐지시스템 제안



[그림 1] IoT 환경 암호화 트래픽 대상 이상탐지시스템 구성도

본 논문에서 제안하는 IoT 환경 암호화 트래픽 대상 이상탐지시스템 구성도는 [그림 1]과 같다. IoT 대상 사이버공격 탐지시스템은 IoT 환경에서 발생하는 네트워크 트래픽을 수집하고 해당 데이터를 기반으로 사이버공격을 탐지한다.

IoT 환경 암호화 트래픽 대상 이상탐지시스템은 Statics-based Feature Extractor, Feature Preprocessor 및 AI based Cyberattack Detector로 구성된다. 평문 트래픽과 달리 암호화된 트래픽에는 복호화 없이 직관적으로 식별할 수 있는 정보가 포함되어 있지 않다. **(1)Statics-based Feature Extractor**는 암호화 패킷의 중요 정보가 암호화 되어 있음에 따라, 패킷의 단일 정보가 아닌 식별 가능한 정보를 기반으로 통계적 정보를 도출한다. 이렇게 추출된 통계적 특징은 암호화된 트래픽 흐름의 행동 패턴을 도출하여 사이버공격 탐지를 위한 모델 학습 및 검증에 활용가능하다. **(2)Feature Preprocessor**는 AI 기반 이상탐지 모델의 학습 및 이상탐지 효율을 높이기 위하여 Statics-based Feature Extractor에서 추출된 Feature를 학습 및 이상탐지에 유리한 형태로 변환한다. Feature Preprocessing 방식에는 결측치 제거, Feature Selection, Normalization등이 존재한다. **(3)AI based Cyberattack Detector**는 암호화 트래픽에서 식별가능한 정보 및 이에 대한 통계 분석 정보를 기반으로 사이버공격을 탐지한다. 사이버공격 탐지기는 정상 암호화 트래픽에서 추출한 Statistics-based Feature만을 학습하며, 정상 암호화 트래픽 정보를 기반으로 사이버공격을 탐지한다. 지도학습 기반 사이버공격 탐지방식은 비교적 높은 이상탐지율과 공격 내 공격유형을 분류할 수 있다는 장점을 가진다. 하지만 지도학습 기반 사이버공격 탐지시스템은 학습된 사이버공격 유형 이외의 새로운 사이버공격은 탐지 및 분류하지 못하는 단점이 존재한다[9]. 이러한 이유로, 본 논문에서는 IoT 환경에서 수집된 정상 데이터만을 모델에 학습하는 방식을 사용한다.

IV. IoT 환경 암호화 트래픽 대상 사이버 공격 탐지시스템 검증

본 장에서는 본 논문에서 제안하는 IoT 환경 암호화 트래픽 대상 사이버 공격 탐지시스템의 유효성을 실험을 통해 입증한다.

4.1 Dataset

본 논문의 실험에 사용한 Dataset인 “CICIoT2023”은 Z-wave, Zigbee, WiFi를 기반으로 통신하는 105개의 IoT 기기로 구성된 IoT 환경에 대하여 정상상태에서 수집한 TLS 1.2 기반 암호화 네트워크 packet 및 DDoS, DoS, Recon 대상 공격, Web 대상 공격, Brute Force, Spoofing 및 Mirai 등 7가지 범주의 사이버공격으로 인하여 발생한 TLS 1.2 기반 암호화 네트워크 packet으로 구성된다. 본 논문에서 제시하는 사이버공격 탐지시스템의 학습 및 검증 실험에는 네트워크 패킷 파일에서 Statics-based Feature Extractor를 사용하여 도출된 19,331개의 훈련 데이터(정상데이터 : 19,331) 및 테스트 데이터 14,578개(정상: 10,119개, backdoor_Malware: 498개, browser hijacking: 500개, command injection: 500개, DDoS-SlowLoris: 500개, Dictionaty Brute Force: 500개, DNS_Spoofing: 500개, Uploading_attack: 500개, Sql injection:501개, XSS: 460개)를 사용하였다. 사이버공격 탐지시스템의 검증 및 테스트를 위하여 정상데이터 구간 내 일부 구간을 선별하여 사이버공격 데이터를 삽입하였다.

4.2 실험설정

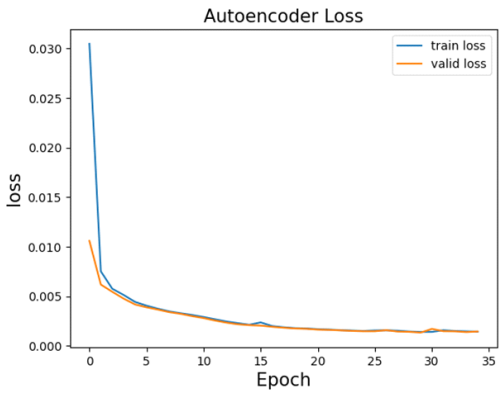
1) Statics-based Feature Extractor : 본 논문에서는 IoT 환경에서 발생한 암호화 네트워크 패킷에 대하여 통계 기반 feature 추출을 위하여 CICFlowmeter를 사용하였다. CICFlowMeter는 원시 패킷 데이터에서 통계 기반 feature를 추출하기 위해 설계된 오픈 소스이다. CICFlowMeter는 패킷 수, 바이트 수, 지속 시간, 패킷 전송 시간 통계 등 다양한 트래픽 흐름에 대한 통계 정보를 도출한다. CICFlowmeter는 암호화 트래픽에서 식별 가능한 정보에 대한 평균, 중앙값, 표준편차와 같은 통계를 계산하여 암호화된 트래픽 내의 데이터 크기 분포를 도출한다. 또한, 패킷 길이의 분포를 고려하여 평균 패킷 길이, 패킷 길이의 변화, 엔트로피 등 통계 기반 feature를 생성한다.

2) Feature Preprocessor : CICFlowmeter에 의해 생성된 통계 기반 Feature는 총 82개이다. 이러한 Feature에는 통계적 Feature 및 메타데이터 기반의 Feature 외 IP Address, MAC Address, Port number등이 포함된다. 이러한 Feature는 데이터셋을 추출한 환경에 한정되어, 해당 시스템의 일반화 측면에서는 부정적인 측면을 가지기 때문에 제외하였다. 이후, Feature 간 상관관계가 없는 Feature는 제외하여 총 66개의 Feature를 도출하였다. 또한, Feature간 값의 범 위 차이로 인한 학습 및 이상탐지 효율저하를 막기 위하여 Max-Abs scaling 방식을 적용하여 데이터 정규화를 진행하였다.

3) AI based Cyberattack Detector : 본 논문에서 제시하는 AI based Cyberattack Detector는 3장에서 제시한 정상 암호화 트래픽 데이터만을 학습하여 사이버공격을 탐지하는 방식 사용을 위하여 Autoencoder 방식을 적용하였다. Autoencoder 기반 사이버공격 탐지기는 각각 66, 64, 32, 16, 8개의 노드를 가지는

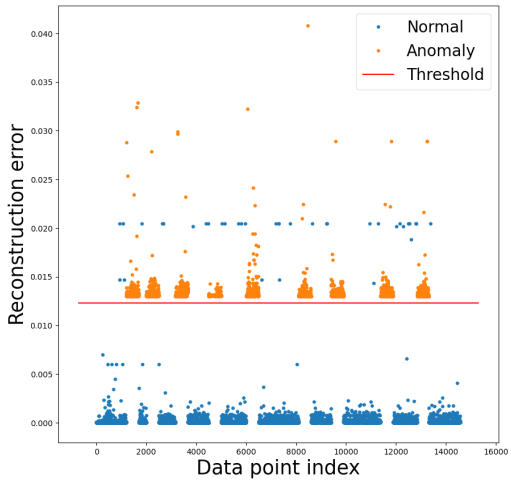
Dense Layer로 구성된 Encoder와 Decoder로 설계하였다. Autoencoder는 각 layer의 activation function으로 ReLU를 사용하였다. optimizer는 Adam을 사용하였으며, learning rate는 0.001을 적용하였다. Autoencoder는 Encoder를 이용하여 정상 데이터를 압축 후 Decoder를 이용하여 복원하여 생성한 데이터와 정상 데이터의 차이인 복원오차를 최대한 작게 도출하게 학습한다. 정상 데이터의 특성만을 학습한 Autoencoder 기반 이상 탐지기에 사이버공격 데이터를 입력할 경우 정상데이터에 비해 큰 복원오차가 출력되고, 출력된 복원오차 값이 설정한 threshold 보다 크다면 해당 데이터를 사이버공격 데이터로 탐지한다.

4.3 실험결과



[그림 2] Autoencoder기반 이상탐지 모델 Loss

[그림 2]는 Autoencoder 기반 이상탐지 모델이 의 학습 및 검증 Loss를 나타낸다. [그림 2]에서 파란색 선은 train loss를 의미하며, 주황색선은 validation loss를 의미한다. [그림 2]에서 볼 수 있듯이 Autoencoder 기반 이상탐지 모델은 validation loss가 0에 가깝게 학습되었으며, validation loss 0.0000271를 도출하였다.



[그림 3] 암호화 트래픽 대상 이상탐지 결과

[그림 3]은 Autoencoder 기반 이상탐지 모델의 암호화 트래픽을 포함하는 검증 dataset 대상 사이버공격 탐지 결과를 나타낸다. [그림 3]에서의 파란색 점은 정상데이터를 의미하며, 주황색 점은 이상데이터를 의미한다. 빨간색 직선은 이상행위 탐지를 위한 임계치를 나타낸다. 임계치 값은 이상탐지 모델의 precision과 recall이 같을 때의 값으로 설정하였다. 실험 결과 본 논문에서 제시한 Autoencoder 기반 이상탐지 모델은 정확도 99.25%, Precision 97.63%, Recall 100%, F-1 Score 98.83%의 높은 성능을 도출하였다.

V. 결론 및 향후 연구 방향

본 논문에서는 IoT 환경에 암호화 통신방식이 적용됨에 따라 발생하고 있는 은닉되어 수행되는 사이버 공격을 탐지 하기 위한 IoT 환경 발생 암호화 트래픽 대상 이상탐지 시스템을 제안하였다. 제안하는 이상탐지시스템 검증을 위하여 IoT 환경에서 수집한 암호화 트래픽을 포함하는 “CICIoT2023" 을 이용하여 모델 학습 및 검증을 통하여 Accuracy 99.25%,

Precision 97.63%, Recall 100%, F-1 Score 98.83%의 높은 성능을 도출하였다. 해당 실험결과 암호화 통신을 수행하는 IoT 환경에 본 논문에서 제시하는 이상탐지 시스템 적용 시 대부분의 사이버 공격에 대한 탐지가 가능할 것으로 판단한다.

향후 연구에서는 암호화 통신 IoT에 대한 실시간 발생 네트워크 트래픽 수집 및 실시간 수집 암호화 트래픽에 대한 사이버공격 탐지 연구를 수행함으로써 고도화 및 현장 적용을 위한 필요기술을 개발할 것이다.

Acknowledgement

이 논문은 2024년도 정부(과학기술 정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구 (No.2021-0-00493, 5G Massive 차세대 사이버공격 기만기술 개발, 50%)와 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(RS-2023-00241376, 해양 선박 공공 서비스·인프라의 암호화 사이버위협에 대한 네트워크 행위기반 보안관제 기술 개발, 50%).

참 고 문 헌

[1] Domínguez-Bolaño T, Campos O, Barral V, Escudero CJ, García-Naya JA. An overview of IoT architectures, technologies, and existing open-source projects. Internet of Things. 2022;20:100626.

[2] Dargaoui S, Azrour M, El Allaoui A, Amounas F, Guezzaz A, Attou H, et al. An overview of the security challenges in IoT environment. Advanced Technology for Smart Environment and Energy. 2023:151-60.

[3] Djenna A, Harous S, Saidouni DE. Internet of things meet internet of threats: New concern cyber security issues of critical cyber infrastructure. Applied Sciences. 2021;11(10):4580.

[4] Wang W, Zhu M, Zeng X, Ye X, Sheng Y, editors. Malware traffic classification using convolutional neural network for representation learning. 2017 International conference on information networking (ICOIN); 2017: IEEE.

[5] Neto ECP, Dadkhah S, Ferreira R, Zohourian A, Lu R, Ghorbani AA. CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. Sensors. 2023;23(13):5941.

[6] Liu J, Song X, Zhou Y, Peng X, Zhang Y, Liu P, et al. Deep anomaly detection in packet payload. Neurocomputing. 2022;485:205-18.

[7] Cai J, Wang Q, Luo J, Liu Y, Liao L. Capbad: Content-agnostic, payload-based anomaly detector for industrial control protocols. IEEE Internet of Things Journal. 2021;9(14):12542-54.

[5] Kim S, Jo W, Shon T. APAD: Autoencoder-based payload anomaly detection for industrial IoE. Applied Soft Computing. 2020;88:106017.

[8] oppi T, Ceccarelli A, Puccetti T, Bondavalli A. Which algorithm can detect unknown attacks? Comparison of supervised, unsupervised and meta-learning algorithms for intrusion detection. Computers & Security. 2023;127:103107.