J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

산업제어시스템을 위한 인공지능 보안 기술☆

AI-based Cybersecurity Solution for Industrial Control System

조 부 성[1] 김 문 석[1*]
Bu-Seong Jo Mun-Suk Kim

요 약

본 논문에서는 산업제어시스템(Industrial Control System, ICS)을 위한 보안기술의 동향과 인공지능 활용을 설명한다. 산업제어시스템은 대규모의 국가적 주요기반 및 산업분야에 적용되어 사용되기 때문에 사이버 공격에 의한 사소한 문제라도 막대한 경제적 비용을 발생 시킬 수 있다. 산업제어시스템은 기존의 IT (Information Technology) 시스템과는 다른 특성을 가지고 있기 때문에 최신 보안기술 개발을 위해서는 산업제어시스템에 맞는 보안위험요소를 고려해야한다. 산업제어시스템에서 최근에 발생되었던 사이버 공격의 방법, 기술을 분석한 프레임워크를 설명한다. 또한, 산업제어시스템의 보안을 위한 대표적인 기술인 침입탐지시스템의 동향을 알아보고 침입탐지시스템에 활용된 인공지능 기술을 분석한다. 구체적으로 침입탐지를 위한 인공지능 기술 적용을 위해 필요한 데이터 수집 및 분석, 인공지능 모델, 인공지능 모델 성능평가를 위한 기법을 설명한다.

☞ 주제어: 산업제어시스템, 인공지능, 침입탐지시스템, 데이터수집, 모델평가

ABSTRACT

This paper explains trends in security technologies for ICS. Since ICS is usually applied to large-scale national main infrastructures and industry fields, minor errors caused by cyberattack could generate enormous economic cost. ICS has different characteristic with commonly used IT systems, so considering security threats of ICS separately with IT is needed for developing modern security technology. This paper introduce framework for ICS that analyzes recent cyberattack tactics & techniques and find out trends in Intrusion Detection System (IDS) which is representative technology for ICS security, and analyzes AI technologies used for IDS. Specifically, this paper explains data collection and analysis for applying AI techniques, AI models, techniques for evaluating AI Model.

☞ keyword : ICS, AI, IDS, Data Collection, Model Evaluation

1. 서 론

산업제어시스템은 주로 대규모의 국가적 주요기반 및 산업분야에 적용되어 사용된다. 따라서, 사이버 공격에 의한 시스템 상의 사소한 동작의 오류는 막대한 경제적 피해를 가져올 수 있다 [1]. 산업제어시스템이 기반으로 하는 운영기술 (Operational Technology, OT)은 우리에게 익숙한 IT 시스템과 구별되는 하드웨어 및 소프트웨어 구성과 네트워크 통신에 관한 특성을 가지고 있기 때문에 IT 시스템에서 사용하는 범용 보안솔루션을 그대로 적용할 수 없다 [2]. IT 시스템은 데이터 기밀성 및 무결성 보장이 주요 관심사인 반면 운영기술은 안전한 작동을 보장하기 위한 장비의 가용성이 주요 관심사이다. 기술, 프로토콜 및 장비에 차이가 있어 IT 시스템의 관점에서 공격 표면은 컴퓨터, 서버, 프린터 및 네트워크 스위치 등이나 운영기술의 경우에는 프로그래머블 로직 컨트롤러 (Programmable logic controller, PLC), 원격 단말 장치 (Remote Terminal Unit, RTU), 휴먼 머신 인터페이스 (Human-Machine Interface, HMI), 원방감시제어 및 데이터 취득 시스템 (Supervisory Control And Data Acquisition, SCADA)등이 주요 공격

표면이다. 이러한 IT 시스템과 운영기술 간의 단절, 지식과 기술 격차의 악화는 운영기술로 하여금 공격에 더 취약하게 만든다 [3].

1 Dept. of Computer Science and Engineering, Sejong University, Seoul, 05006, Korea
* Corresponding author: Mun-Suk Kim (msk@sejong.ac.kr)
[Received 16 August 2022, Reviewed 19 August 2022(R2 17 October 2022), Accepted 28 October 2022]
☆ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. [IITP-2022-2021-0-01816, 메타버스 자율트윈 핵심기술 연구] 그리고, 2022년도 정부(과기정통부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2022-00165231).

(표 1) 기존 IT 시스템과 산업제어시스템의 비교[2]
(Table 1) Comparing IT System and ICS System[2]

항목	기존 IT 시스템	산업제어시스템
공격 표면	- 컴퓨터, 서버, 프린터, 스위치	- PLC, RTU, HMI, SCADA
하드웨어 및 소프트웨어	- 짧은 교체주기(3~5년) - 리눅스, 윈도우 & 범용 프로토콜 - 유지, 보수 용이	- 장기간 교체주기(15~20년) - 전용 어플리케이션 & 비공개 프로토콜 - 유지, 보수 어려움
네트워크 및 통신	- 인터넷 및 외부와 연결된 망구조 - 응답의 신뢰성이 중요하며 일부 통신 지연 허용	- 외부와 분리된 내부 통신망 구조 - 응답 시간이 중요하며통신 지연 불허
위험관리 목표	- 데이터의 기밀성 및 무결성 중요 - 일부 고장 및 장애 허용	- 인간의 안전 및 시스템 가용성 중요
사고영향	- 업무 불편 및 지연 등 미미한 경제적 피해 발생	- 산업현장 운영 중단으로 인한 피해 및 대규모 물리적, 경제적 피해 발생

2010년에 이란 핵시설에 대한 사이버 공격인 Stuxnet 사건이 발생하였고, 2014년에 개발된 트로이 목마 패밀리의 세번째 버전으로 알려진 BlackEnergy는 산업제어시스템 기반 시설을 대상으로 개발되었으며, 2015년에 키예브 배전 회사에서 3시간동안 30개의 변전소 단절을 유발하였다. 2017년에는 안전 목적으로 사용되는 프로그래머블 로직 컨트롤러 를 재프로그래밍해 실패 상태를 유발상태는 말웨어 TRITON 이 사우디아라비아 석유화학 공장을 예정없이 폐쇄시켜 버린 사고가 발생하였다. 이처럼 산업제어시스템에 대한 사이버 공격은 세계적으로 지속되고 있다. [4].

기존의 사이버 공격과 차별화된 지능형지속위협(Advanced persistent threat, APT) 등의 최신 공격기술들은 Stuxnet 사례와 같이 가능한 모든 방법을 동원하여 내부 네트워크에 침입한 뒤 장기간 대기하며 최적 시점에 맞추어 공격을 수행하는 등 매우 정교하고 지능적인 방법으로 사이버 공격을 수행하여 기존의 탐지 체계만으로는 대응하기 역부족이다. 이에 최근 대두되고 있는 인공지능 기술을 사이버보안에 접목시켜, 공격 데이터셋을 모으고 사이버공격을 탐지하는 기계학습 모델에 테스트베드들을 학습시키는 등 산업제어시스템에 적합한 인공지능 보안기술에 관한 연구들은 이미 진행되어오고 있으나 [12-18], 그 기간이나 규모가 그리 크지 않다. 따라서 산업제어시스템의 인공지능 보안 접목에 대한 연구 동향을 조사하고 요약하여 이 분야의 개론을 요약할 논문을 작성하고자 한다.

본 논문에서는 먼저 산업제어시스템의 개념과 기존의 IT 시스템과 구별되는 특성을 파악하고, 산업제어시스템만의 특성에 따른 보안 위협요소를 소개한다. 또한, 산업제어시스템을 대상으로 최근에 발생한 사이버 공격 기술을 체계적으로 분석할 기반이 되는 프레임 워크를

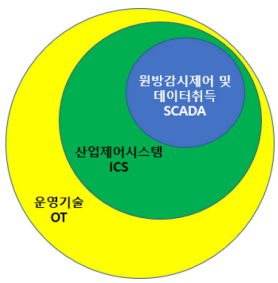
소개하고 특정 사례를 해당 프레임워크에 매핑하여 사이버 공격 방법 및 기술을 설명한다.

마지막으로 산업제어시스템의 보안을 위한 인공지능 기술을 활용한 침입탐지시스템의 연구를 설명하고, 침입탐지시스템의 인공지능 기술을 위한 데이터 수집 및 분석, 데이터 학습 및 분류 모델, 모델 평가 기법을 설명한다. 본 논문을 통해 산업제어시스템에서 보안기술의 동향과 인공지능 활용에 대한 이해를 높일 수 있다.

2. 산업제어시스템의 보안위협요소

산업제어시스템은 국가적 규모 주요기반 시설 및 산업분야에서 센서 측정값과 운용정보 수집, 정보의 처리/표시, 그리고 제어정보를 원거리에 위치한 장치로 전달하는 동작을 수행하는 시스템을 의미하고, 구체적인 예로는 원자력 및 화력 발전소, 수처리 시설, 발전, 중공업 및 배전 시스템과 같은 산업 부문 및 중요 기반 시설에서 흔히 볼 수 있다 [1][5].

그림 1과 같이 산업제어시스템은 원방감시제어 및 데이터 취득 시스템과 운영기술의 중간개념으로 말할 수 있다 [2].



(그림 1) 산업제어시스템의 개념
(Figure 1) Concept of ICS

(표 2) ICS의 주요 취약점 및 그에 따른 위협
(Table 2) Security weakness and threat factors in ICS

영역	취약점	위협
모니터링	- 생산 망 내 설비/서버 보안 위협 탐지 미흡 - 보안 솔루션 관리 미흡	- 조기탐지 및 대응 어려움
설비	- 외부 매체 통제 미흡 - 악성코드 방어 체계 미흡	- 악성코드 유입, 감염 및 확산
네트워크	- 망분리 부재 - 생산망 Segmentation 부재 - 무선 네트워크 접근에 대한 인증 부재	- 외부로부터의 악성코드 유입 - 감염 시 생산망 내 모든 구성요소에 영향 - 비인가자의 내부 네트워크 접속
서버	- 주요 어플리케이션에 대한 사용자 인증 부재 - 구성요소 간 인증/암호화 부재	- 비인가자의 PLC 접근 - 통신정보 유출/변조

원방감시제어 및 데이터취득 시스템은 지역적으로 분산되어 있는 장치들의 정보수집과 제어를 수행하는 시스템을 말하고, 시스템의 규모가 커지면서 산업제어시스템과 구분없이 ICS/SCADA로 통칭하여 사용되고 있다 [6].

산업제어시스템은 하드웨어 및 소프트웨어를 사용해 산업 장치와 공정 운영을 모니터링하고 설비를 직접 제어하는 모든 기술을 포괄하는 운영기술 방식을 이용하고, 분산 제어 시스템 (Distributed Control System, DCS), 프로그래머블 로직 컨트롤러, 센서, 엑추에이터와 같이 다른 산업장비, 자산, 프로세스 및 이벤트를 제어하는 시스템 구성을 포함한다 [4][5].

표 1과 같이 산업제어시스템은 기존 IT 시스템과 비교하여 여러 측면에서 다른 특성을 가진다. 이러한 차이점 때문에 산업제어시스템은 기존 IT 시스템과는 다르게, 표 2와 같은 취약점이 생기며 그에 상응하는 보안위협요소를 고려해야한다 [7].

세부적인 보안위협요소 상황을 예시로 들면 다음과 같다. 먼저, 산업제어시스템의 설비들은 보통 규모가 크고 약 15년 이상의 긴 교체주기를 가지고 있기 때문에 기존 IT 시스템보다 최신 사이버 공격에 더 취약하다. 따라서, 산업제어시스템을 위한 보안기술 도입은 10년 이상의 미래를 고려하여 결정해야할 필요가 있다 [2]. 또한, 산업제어시스템은 많은 경우에 외부와 철저히 분리된 망에서 기존 IT 시스템에서 널리 사용되는 운영체제인 윈도우 또는 리눅스가 아닌 전용 실시간 운영체제를 이용하는 자체시스템을 구성한다. 또한, 산업제어시스템 내부간 통신은 IT 시스템에서 사용하는 프로토콜과 다른 Modbus, PI, ICCP 등과 같은 제어 프로토콜을 사용하기 때문에 기존 IT 시스템에서 사용되는 범용보안솔루션을 적용할 수 없다 [6]. 그리고, 산업제어시스템은 대규모 설비로 구성되는 경우가 많기 때문에 사이버 공격에 의해 발생되는 짧은 시간의 문제에도 막대한 피해를 당할 수 있다. 따라서, 산업제어시스템의 보안은 중단없이 시스템이 운용될 수 있는 가용성이 중요하다 [8].

MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) 프레임워크는 위협 시나리오 내에서 사전에 형성된 공격의 구조를 제공하는데 사용된다. 이 지식 기반은 대상 네트워크 내에서 추가 액세스, 침해 및 조작을 위한 적의 행동을 설명하는데 사용된다. 공격자의 목표, 작업 및 절차를 분류하기 위한 일관된 고려사항을 제공하므로, 시나리오 내에서 지정된 목적을 달성하는데 필요한 공격 절차 모델링에 적용될 수 있다 [9]. 실제 발생한 사이버 공격 사례를 기반으로 악위적 행위에 대해 공격방법 (Tactics) 와 공격기술(Techniques) 관점으로 매트릭스 형식으로 분류하여 엔터프라이즈, 모바일, 산업제어시스템을 대상으로 각각 제공한다. 산업제어시스템을 대상으로는 12개의 공격방법과, 89개의 공격기술이 존재하며, 각각의 공격기술에 대하여, 공격 완화 기법 (Migations), 공격단체조직 (Groups), 공격 소프트웨어 (Software), 탐지기법 (Detection) 또한 분류되어 있다 [10]. 산업제어시스템에 대한 공격 사례로 대표적인 2010년 이란의 원전시설에 발생한 Stuxnet 지능형지속위협 공격은 원전시설 내부 네트워크에 악성코드를 침투시킨 후 장기간 대기하며 최적 시점에 공격을 수행함으로써 공격 탐지를 피했다. 그림 2는 이란 원전시설에 발생한 Stuxnet 공격을 ATT&CK 프레임워크에 매핑한 것이다. 공격 방법 중에 탐색 (Discovery)에서는 네트워크 스니핑과 원격 시스템 정보검색 공격 기술이 사용되었고, 내부 확산 (Lateral Movement) 공격 방법에 대해서는 프로그램 다운

로드 공격기술, 그리고 수집 (Collection) 공격방법에 대해서는 I/O 이미지와 모니터 프로세스 상태의 공격기술이 사용되었다 [11].

Discovery	Lateral Movement	Collection
Network Connection Enumeration	Default Credentials	Automated Collection
Network Sniffing	Exploitation of Remote Services	Data from Information Repositories
Remote System Discovery	Lateral Tool Transfer	Detect Operating Mode
Remote System Information Discovery	Program Download	I/O Image
Wireless Sniffing	Remote Services	Man in the Middle
	Valid Accounts	Monitor Process State

(그림 2) Stuxnet사건의 ATT&CK 프레임워크 매핑
(Figure 2) Mapping Stuxnet incident on ATT&CK Framework

3. 산업제어시스템의 인공지능 보안

3.1 인공지능을 활용한 침입탐지시스템

산업제어시스템은 대규모의 국가적 주요기반 시설 또는 산업분야에 적용되기 때문에 사이버 공격에 의해 발생하는 사소한 오류 및 오작동에도 막대한 경제적 피해를 발생 시킬 수 있다 [1]. 따라서, 최근 산업제어시스템 내에 침투한 공격을 탐지하는 침입탐지시스템 (Intrusion Detection System, IDS)에 대한 연구가 진행되오고 있다 [6].

기존의 연구에서 개발된 침입탐지시스템은 탐지를 위해 모니터링하는 정보에 따라 크게 두가지 타입으로 분류할 수 있다. 첫번째 타입은 시스템의 장치들 사이의 주고받는 네트워크 트래픽을 모니터링하는 네트워크 기반 침입탐지시스템 (Network IDS, NIDS)이고 두번째 타입은 센서에 의해 측정된 데이터를 모니터링하는 측정 데이터 기반 침입탐지시스템 (Mesurement IDS, MIDS) 이다. 위의 두가지 타입의 침입탐지시스템 모두 공격탐지를 위한 네트워크 트래픽 또는 센서의 측정 데이터 분석을 위해 많은 경우에 인공지능 기술을 활용하고 있다 [12].

네트워크 기반 침입탐지시스템을 위해 [13]의 논문은 실제 산업제어시스템에서의 복잡한 네트워크 환경과 다양한 사이버 공격수단을 열린집합 (Open Set) 문제로 다룸으로써 알려진 공격 뿐아니라 아직 알려지지 않은 공격까지 인공신경망 (Artificial Neural Network, ANN) 기술을 통해 분류하는 시스템을 제안하였다. [14]의 논문 또한 ANN 기술을 이용하여 침입탐지시스템에서 고려하는 정보를 축소함으로써 전체적인 시스템의 계산 복잡성을 줄이는 방법을 제안한다. [15]의 논문은 지능적인 연결형 교통시스템의 사이버 공격에 대한 침입탐지시스템을 제안한다. [15]의 논문에서 제안하는 시스템의 침입탐지 동작은 네트워크 트래픽의 분석과 데이터 축소, 그리고 정상적인 서비스 요청과 비정상적인 서비스 요청을 구분하는 데이터 분류를 포함하는 총 세가지 단계를 수행한다. 이때, 데이터축소 단계는 심층신뢰신경망 (Deep Belief Network) 기술을 이용하고 데이터 분류를 위해서는 결정트리 (Decision Tree) 기술을 이용한다. [16]의 논문은 사이버 공격의 공격방법이 다양해

지고 계속 진화하면서 발전하는 특성을 고려하여 심층학습 (Deep Learning) 기술 기반으로 예측 가능한 공격 뿐아니라 예측이 힘든 공격까지 탐지 및 분류가 가능한 유연하고 효과적인 침입탐지시스템을 제안한다.

(표 3) 인공지능 기술을 활용한 침입탐지 시스템
(Table 3) IDS using AI technologies

IDS	참고논문	인공지능 모델	데이터셋
네트워크 기반 침입탐지시스템	[15]	심층신뢰망, 결정트리	KDD
	[13]	ANN	NF-BoT- IoT
	[16]	k-최근접 이웃, 결정트리, 서포트 벡터 머신	KDD
	[14]	ANN	
측정 데이터 기반 침입탐지시스템	[12]	k-최근접 이웃	HAI
	[17]	서포트 벡터 머신	시뮬레이터 구성
	[18]		

측정 데이터 기반 침입탐지시스템을 위해 [12]의 논문은 센서에 의해 측정된 데이터를 지도학습모델 (Supervised Learning Model)에 학습시켜 산업제어시스템의 정상적인 동작과 비정상적인 동작을 구분하는 방법을 제안한다. 이때, 학습 데이터는 정상 데이터와 사이버 공격에 의한 비정상 데이터를 균형있게 포함하도록 하여 침입탐지의 정확성을 높힌다. [17]의 논문은 연결형 자동차 내의 디지털 장치들 사이 통신 (Controller Area Network, CAN)을 위한 보안을 위해 전기적 신호의 전압 측정 데이터를 서포트 벡터 머신 (Support Vector Machine, SVM) 모델에 학습시켜 침입탐지를 수행하는 방법을 제안한다. [18]의 논문 또한 서포트 벡터 머신 모

델을 전력시스템의 모니터링 데이터로 학습시켜 사이버 공격탐지를 수행하는 기술을 제안한다. 표 3은 네트워크와 측정데이터 기반 침입탐지시스템에 관한 기존연구에서 사용한 인공지능 모델과 학습 및 테스트를 위한 데이터셋에 관한 정보를 나타낸다.

3.2 데이터 수집 및 분석

인공지능 모델을 학습하고 성능을 평가하기 위해서는 데이터셋 구축은 필수적인 과정이다. 데이터셋을 구축하기 위해서는 때로는 테스트 베드가 필요하다. 테스트 베드는 감독하에 실제 활동을 시뮬레이션하는 플랫폼으로, 주로 실험 활동과 그 결과를 확인하는데 사용되고, 이러한 활동 또한 복제할 수 있어야 한다. 현재 데이터 세트의 가용성이 낮다면 맞춤형 테스트베드를 생성해 다양한 산업 시나리오에서의 데이터를 생성해야 한다. 각 시나리오의 운영 내에서 개별 사건에 대한 테스트가 필요하지만 경제적 영향을 고려했을때, 개별 네트워크의 운영에서는 불가능하므로 실제 솔루션을 기반으로 사용자 지정 테스트 환경을 구축하고 이러한 환경에서 통신 및 다양한 사건을 테스트하여 데이터셋의 균형을 맞추는 것이 필수적이다. 따라서 테스트 환경은 산업제어시스템 대상 침입탐지시스템 보안 연구의 중요한 구성 요소이다 [19].

산업제어시스템 테스트베드를 구축하는데는 5가지의 필수 조건을 만족시켜야 한다. 첫번째는 다양성으로, 여러 제조업체의 다양한 장치와 소프트웨어를 사용해야 한다. 두번째는 확장성으로, 규모에 따라 여러 장치와 프로세스를 지원하면서도 시뮬레이션과 가상화를 통해 비용을 절감할 수 있어야 한다. 세번째는 복잡성으로, 데이터의 규모와 다양성이 증가할 때 관리자와 연구원들이 테스트베드를 관리하고 실험을 전개할 때 복잡해지는 것을 줄이기 위한 조치가 필요하고, 네번째는 데이터 수집으로, 시스템이 공격 혹은 시뮬레이션을 받고 있는 경우에 실험을 위한 적절한 데이터 수집이 가능해야 한다. 다섯번째는 안정성으로, 테스트베드는 연구원과 엔지니어의 안전 측면에서 위험을 최소화하도록 설계하고 테스트베드 자체도 외부 공격으로부터 안전해야 한다 [20].

실제 산업제어시스템이 설치되어 있는 환경에서 다양한 사이버 공격에대한 데이터 수집을 수행한다면 수행과정 동안에 오류로 인한 위험성이 크다는 문제가 있다. 예를 들어, 원전계측제어시스템에서 다양한 사이버 공격에 대한 데이터를 수집한다고 가정한다면 실험과정 동안 예측하지 못한 오류로 인해 원자력 발전소의 동작이 멈추는 등의 큰 피해가 발생할 수 있다. 이런 위험성으로 인해 정상 데이터를 수집하는 것보다 사이버 공격으로 인한 비정상 데이터 수집이 어렵다는 단점이 있다. 3.1장에서 소개한 침입탐지시스템 연구에서는 데이터 수집을 위해 공개된 데이터셋을 이용하거나 해당 시스템의 테스트베드 또는 시뮬레이터를 이용한다. 이탈리아 파도바 대학의 SPRITZ (Security & PrivacyResearch Group) 그룹은 공개된 산업제어시스템 테스트베드를 조사하여 기관, 범위, 범주, 라이선스에 따라 다음 웹사이트 문헌에 분류해 놓았다 [21].

구체적으로 네트워크 기반 침입탐지시스템을 살펴보면, [13]의 논문에서는 ANN 모델을 학습하고 테스트하기 위해 공개된 NF-BoT-IoT 데이터셋과 가스파이프라인 데이터셋을 이용한다. NF-BoT-IoT 데이터셋은 서비스거부 및 분산서비스거부 공격 (Distributed Denial Of Service, DDoS), 정찰 (Reconnaissance), 도용 (Theft)의 네가지 공격을 포함한다. 가스파이프라인 데이터셋은 서비스거부 공격과 정찰 공격을 포함하는 7개의 공격을 포함한다. [14], [15], [16]의 논문들은 KDD (Knowledge Discovery in Data mining) 데이터셋을 이용한다. KDD는 서비스거부와 프루브 (Probe)를 포함하는 총 다섯가지 공격 클래스를 가지고 있고 22가지의 세부 공격을 포함한다.

측정 데이터 기반 침입탐지시스템의 데이터 수집에 대해 설명하면, [12]의 논문의 경우는 국가보안기술연구소에서 제공하는 HAI (Hardware-in-the-loopbased Augmented ICS) 데이터셋을 이용한다. 이는 HIL(Hardware-in-the-loop) 테스트베드를 통하여 수집되었으며, 터빈, 보일러, 수처리 시스템, 냉각설비 등에 대한 데이터로 구성되며, 이 데이터셋을 통해 복잡한 프로세스 설정이 가능하고 다양한 공격도 반복적으로 수행이 가능하다. [17]와 [18] 논문의 경우는 차량내 통신을 위한 시뮬레이터와 전력시스템을 위한 시뮬레이터를 각각 구성하여 측정 데이터를 수집하였고 인공지능모델의 학습과 테스트를 위해 사용하였다.

그 밖에도 데이터수집을 위해 [10]의 논문에서는 시스템을 가상화하는 디지털 트윈 (Digital Twin) 기법을 활용하여 실제 시스템과 동일한 가상환경에서 다양한 설명 및 공격 수행을 통한 데이터 수집을 설명한다. [6]의 논문에서는 시스템의 다양한 보안 로그와 이벤트, 각종 운용 정보를 통합하여 상관관계를 분석하는 SIEM (Security Information & Event Management) 기술을 소개한다.

수집한 데이터에 대해서 인공지능 모델을 학습하기 위해서는 정상 데이터와 사이버 공격에 의한 비정상 데이터를 모두 사용하는 것이 필요하다. 이때, 데이터셋에

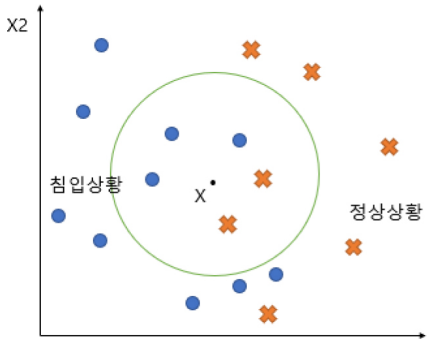
정상 데이터의 비율이 비정상 데이터의 비율보다 큰 차이로 많게 되면 인공지능 모델의 침입탐지 정확성이 저하되는 문제가 발생한다 [22]. 이 문제를 해결하기 위해 [23]의 논문에서 소개하는 SMOTE (Synthetic Minority Over-sampling Technique) 방법을 사용할 수 있다. SMOTE 방법은 희귀 클래스의 기존 데이터를 사용하여 새 인스턴스를 생성하는 무작위 오버샘플링 접근 방식이다. 기존에 있는 데이터를 복제하는 대신 소수 범주의 데이터들을 서로 보간하여 새로운 인공적인 데이터를 합성한다. 먼저 k-최근접 이웃 (K-Nearest Neighbors, KNN) 알고리즘을 사용하여 소수 범주의 데이터들과 가장 가까운 데이터들을 찾은 뒤 새로 합성된 데이터가 그 성향을 반영한다. 이 방법으로 인공지능 모델 학습에 있어 과적합 문제의 위험을 피할 수 있다 [24].

3.3 인공지능 모델

침입탐지시스템은 사이버 공격 탐지를 위해 정상 데이터와 비정상 데이터를 분류하고, 이를 위해 다양한 인공지능 모델을 사용한다. 이번 장에서는 4.1에서 설명한 네트워크 및 측정 데이터 기반 침입탐지시스템에서 사용한 k-최근접 이웃, 심층 신뢰망, 의사 결정 트리, ANN, 서포트 벡터 머신에 대해 설명한다.

k-최근접 이웃 기술은 이미 알려진 개체들을 훈련집합 형태로 메모리에 기억시킨 다음 그 중 유사한 개체를 선택하여 선택된 개체의 값에 따라 새로운 개체의 값을 예측하는 방식의 분류 알고리즘이다 [25]. 기계 학습에 적용되는 전형적인 비모수 분류기이며, 이러한 기법의 사고 방식은 레이블이 없는 데이터 샘플을 고려할 이웃 수를 정의하는 정수인 k개의 가장 가까운 이웃 클래스에 이름을 지정하는 것이다. 그림 3은 k = 5 인 k-최근접 이웃 분류기를 보여준다. 점 X는 분류해야 하는 레이블이 없는 날짜의 인스턴스를 나타낸다. X의 5개의 가장 가까운 이웃에는, 침입상황 클래스로부터의 3개의 유사 패턴과 정상상황 클래스로부터의 2개의 유사 패턴이 있다. 과반수 투표를 하면 X를 침입상황 클래스에 할당할 수 있다 [26].

심층 신뢰망은 연결된 레이어의 각 쌍이 제한된 볼츠만 머신 (Restricted Boltzmann Machine, RBM)인 다중 레이어 네트워크이다 [27]. 제한된 볼츠만 머신은 2계층 네트워크로, 계층 간 각 장치는 양방향 연결을 가지며 동일한 계층 내의 장치는 연결되지 않는다. 심층 신뢰망은 이전 제한된 볼츠만 머신의 출력이 후속 제한된 볼츠만 머신의 입력이 되도록 겹쳐 쌓아 형성된 심층 신경망 모델



(그림 3) k = 5 인 경우 k-최근접 이웃 분류기의 예
(Figure 3) Example of KNN for k = 5

이며 데이터는 하단 제한된 볼츠만 머신을 통해 네트워크에 입력되고, 상단 2계층은 양방향 연결, 나머지 계층은 단방향 연결이다. 심층 신뢰망의 학습은 전처리과정과 미세 조정 두 부분으로 나뉜다. 전처리과정은 첫번째 제한된 볼츠만 머신 입력 데이터에서 시작하여 첫번째 제한된 볼츠만 머신의 출력을 두번째 제한된 볼츠만 머신의 입력으로 시작하여 모든 계층이 완료될 때까지 계층별로 훈련하는 비지도 학습 프로세스이다. 전처리과정 후 가중치와 편향을 포함하는 각 계층의 매개변수는 네트워크 매개변수의 초기 값으로, BP (BackPropagation) 알고리즘을 사용하여 전체 네트워크의 매개변수를 조정하는 것은 지도 학습 과정이다 [28].

의사 결정 트리는 설명변수에 대한 목표변수 값을 예측하는 모형으로, 가장 설명력이 높은 독립 변수 순으로 분류하여 If-Then 규칙을 생성해 나가는 분류나무 모델이고, 테스트 속성을 식별하는 데 사용되는 결정 노드와, 테스트 속성 값을 기반으로 가능한 결정을 나타내는 분기, 그리고 인스턴스가 속한 클래스를 구성하는 잎 세가지 기본 요소로 구성된다. 트리의 깊이에 따른 과적합 문제는 원래의 데이터에서 일부 데이터를 선별한 결정트리를 여러 개 만들고 각 의사 결정 트리의 결론을 통합해 최종 결과치를 내는 방식으로 해결할 수 있다 [26][29] .

ANN은 인간의 두뇌가 작동하는 것처럼 모방하는 처리에 의해 주어진 특정 데이터 세트에서 패턴을 식별하는데 사용되는 알고리즘의 조합 또는 세트이다. 이 메커니즘의 중요성은 변화하는 입력에 적응하고 출력을 위해 네트워크를 다시 설계할 필요가 없다는 것이다. 입력 계층, 하나 이상의 은닉 계층 및 출력 계층을 포함하는 노드 계층으로 구성되며, 각 노드나 인공 신경망은 다른 노드에 연결되고 연관된 가중치 및 임계값이 있다. 개별 노

드의 출력이 지정된 임계값보다 높으면 해당 노드가 활성화되어 네트워크의 다음 계층으로 데이터를 보내고, 그렇지 않으면 데이터가 전달되지 않는다 [30][31].

서포트 벡터 머신은 분할 초평면에 의해 정의되는 식별 분류기로, 커널 함수를 사용하여 학습 데이터를 고차원 공간에 매핑하여 침입상황이 선형으로 분류되도록 한다. 데이터를 분류하거나 결과를 정확하게 예측하기 위해 데이터 마이닝 시 분류와 회귀라는 두가지 유형을 사용하여 알고리즘을 학습한다. 분류는 알고리즘을 사용하여 테스트 데이터를 특정 범주로 정확하게 할당하는 것이고 회귀는 종속 변수와 독립 변수 간의 관계를 이해하는데 사용된다. 속성 수가 많고 데이터 포인트 수가 적을 때 주로 사용되는 모델이다 [26][32].

3.4 인공지능모델 평가기법

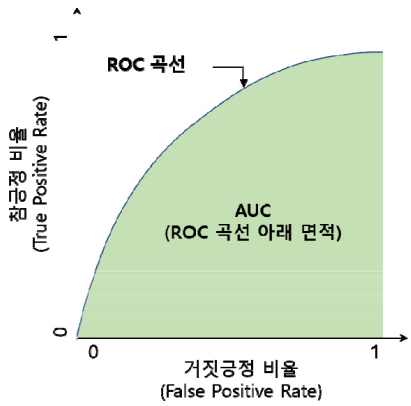
침입탐지시스템은 앞에서 설명한 것처럼 다양한 인공지능 모델을 사용할 수 있다. 각 시스템 환경에서 적합한 인공지능 모델을 선택하기 위해서는 인공지능 모델의 침입탐지 정확도 성능평가를 위한 기법이 필요하다.

인공지능 모델을 그림 4과 같은 혼동행렬을 통해 평가할 수 있다. 인공지능 모델을 활용하여 실제 (Actual Condition)와 예측 (Predicted Condition)의 값을 측정하여 혼동행렬의 참긍정 (True Positive), 거짓부정 (False Negative), 거짓긍정 (False Positive), 참부정 (True Negative)를 측정한다.

		예측 (Predicted Condition)	
		예측긍정 (Predicted Positive)	예측부정 (Predicted Negative)
실제 (Actual Condition)	긍정 (Positive)	참긍정 (True Positive)	거짓부정 (False Negative)
	부정 (Negative)	거짓긍정 (False Positive)	참부정 (True Negative)

(그림 4) 모델 평가를 위한 혼동 행렬
(Figure 4) Confusion Matrix for model evaluation

그림 5는 참긍정 비율 (True Positive Rate)과 거짓긍정 비율 (False Positive Rate)을 나타내어 만들어진 ROC (Receiver operator characteristic) 곡선이다. 그림 5와 같이 ROC 곡선 아래 면적을 AUC (Area under the ROC Curve)라 하고, AUC의 값이 클수록 인공지능 모델의 예측 정확도가 높다고 할 수 있다 [33].



(그림 5) 모델 성능을 의미하는 AUC
(Figure 5) AUC that means performance of model

4. 결 론

산업제어시스템은 대규모의 국가적 기반시설 및 산업분야에 적용되는 시스템이기 때문에 사소한 보안문제도 막대한 피해를 발생시킬 수 있다. 또한, 기존의 IT 시스템과 다른 장치구성 및 통신 특성을 가지고 있기 때문에 산업제어시스템에 적합한 보안기술의 연구가 필요하다. 보안 안전에 치명적인 산업제어시스템의 특성상, 산업제어시스템 보안의 궁극적인 목표는 어떠한 상황에서도 산업제어시스템의 안정성에 영향이 없게 하는 것이다. 산업제어시스템 내부에 잠입한 사이버 공격이 발생시킬 수 있는 산업제어시스템 안전을 저해하는 경우들을 분석하고 이를 탐지하여 그 영향을 원천적으로 방지하는 방법이 필요하다. 지능형지속위협과 같이 발전되고있는 최근의 공격 기술들에 대응하여, 최근 대두되는 인공지능 기술을 사이버보안에 활용한다면 이 분야에 상당한 좋은 영향을 미칠 것으로 예상된다. 본 논문에서는 이에 맞추어 산업제어시스템의 보안위협요소를 분석하였고 최근에 산업제어시스템에서 발생한 사이버 공격 방법 및 기술을 설명하였다. 그리고, 산업제어시스템의 보안을 위해 연구되고 있는 대표적인 기술인 침입탐지시스템에서의 인공지능의 활용 방법 및 동향을 설명하였고, 이러한 정보를 제공함으로써, 연구자들이 기존의 IT시스템보다 접하기 어려운 산업제어시스템 분야에서, 보안 솔루션을 개발하는 방법을 더 쉽게 접하며 익히고, 지능형지속위협 등의 최신 공격 기술에 대항할 인공지능 기술을 접목하는 논문들도 소개하여, 산업제어시스템 보안 향상에 전반적으로 기여하기 쉽게 될것으로 기대한다.

References

[1] Hyun-Seok Kim, Dong-Gue Park, "Implementation of abnormal behavior detection system based packet analysis for industrial control system security", Journal of the Korea Academia-Industrial cooperation Society, Vol. 19, Issue 4, 47-56, 2018.
https://doi.org/10.5762/KAIS.2018.19.4.47

[2] https://www.epnc.co.kr/news/articleView.html?idxno=92835

[3] Shaharyar K., Alberto V., Geet K., Jonathan E., Tommaso P., Sabino C., Micheal S., "Cyber Range for Industrial Control Systems (CR-ICS) for Simulating Attack Scenarios", Italian Conference on Cybersecurity, 2021.

[4] Mauro C., Denis D., Federico T., "A Survey on Industrial Control System Testbeds and Datasets for Security Research", IEEE Communications Surveys& Tutorials, 2021.
https://doi.org/10.1109/comst.2021.3094360

[5] David B., Maede Z., Aiman E., Raj J., Khaled K., Nader M., "Cybersecurity for Industrial Control Systems: A Survey", Industrial Control Systems Security, 2019.
https://doi.org/10.48550/arXiv.2002.04124

[6] Dong-Gue Park, Deok-Jo Jeon, "A Model of Monitoring for Security of Industrial Control System", Journal of Korean Institute of Information Technology, vol.13, Issue 7, 1-16, 2015.
http://doi.org/10.14801/jkiit.2015.13.7.1

[7] https://blog.lgcns.com/1741

[8] Yong-Hee Jeon, "산업제어시스템 보안을 위한 네트워크 설계 및 구조", Korea Institute of Information Security & Cryptology, v.19, Issue 5, 2009.

[9] Austris U., Bernhards B., "Industrial and Automation Control System Cyber Range Prototype for Offensive Capability Development", 8th International Conference on Information Systems Security and Privacy, 2022.
https://doi.org/10.5220/0010879500003120

[10] Myung-Kil Ahn, Jung-Ryun Lee, "Research on Threat Analysis methodology based on ICS-ATT&CK for ICS/SCADA system", 2020.

[11] https://attack.mitre.org/software/S0603

[12] Sohrab M., Alireza A., Kang K. Y., Arman S., "A Machine Learning Approach for Anomaly Detection in Industrial Control Systems Based on Measurement Data", Electronics, 2021.
https://doi.org/10.3390/electronics10040407

[13] Chao W., Bailing W., Yunxiao S., Yuliang W., Kai W., Hui Z., Hngri L., "Intrusion Detection for Industrial Control Systems Based on Open Set Artificial Neural Network", Security and Communication Networks, 2021.
https://doi.org/10.1155/2021/4027900

[14] Akashdeep S.,Ishfaq. M., Neeraj K., "A feature reduced intrusion detection system using ANN classifier", Expert Systems with Applications, 2017.
https://doi.org/10.1016/j.eswa.2017.07.005

[15] Aloqaily M., Otoum S., Al Ridhawi I., Jararweh Y., "An intrusion detection system for connected vehicles in smart cities", Ad Hoc Networks, 2019.
https://doi.org/10.1016/j.adhoc.2019.02.001

[16] R. Vinayakumar, Mamoun A., K. P. Soman, Prabaharan P., Ameer A. N, "Deep Learning Approach for Intelligent Intrusion Detection System", IEEE Access, 2019.
https://doi.org/10.1109/access.2019.2895334

[17] Won-Suk Choi, Kyung-Ho Joo, Hyo-Jin Jo, Moon-Chan Park, Dong-Hoon Lee, "VoltageIDS: Low-Level Communication Characteristics for Automotive Intrusion Detection System", IEEE Transactions on Information Forensics and Security, 2018.
https://doi.org/10.1109/tifs.2018.2812149

[18] Mete O., Inaki E.,Fatos T. Y. V., Sanjeev R. U., H. Vincent P, "Machine learning methods for attack detection in the smart grid", IEEE Transactions on Neural Networks and Learning Systems, 2015.
https://doi.org/10.1109/tnnls.2015.2404803

[19] Ondrej P., Petr B., Karel K., Radek F., Jiri M, "Application Perspective on CyberSecurity Testbed for Industrial Control Systems", Sensors, 2021.
https://doi.org/10.3390/s21238119

[20] Gardiner J., Craggs B., Green B., Rashid A., "Oops I Did it Again: Further Adventures in the Land of ICS Security Testbeds", the ACM Workshop, 2019.
https://doi.org/10.1145/3338499.3357355

[21] https://spritz.math.unipd.it/projects/ics_survey

[22] Jae-Hyun Seo, "A Comparative Study on the Classification of the Imbalanced Intrusion Detection

Dataset Based on Deep Learning", Korean Institute of Intelligent Systems, vol.28, Issue 2, 152-159, 2018. https://doi.org/10.5391/jkiis.2018.28.2.152

[23] Nitesh V. C., Kevin W. B., Lawrence O. H., W. Philip. K., "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research, 2002. https://doi.org/10.1613/jair.953

[24] Maede Z., Marcio A. T., Raj J., "Effect of Imbalanced Datasets on Security of Industrial IoT Using Machine Learning", IEEE International Conference on Itelligence and Security Informatics, 2018. https://doi.org/10.1109/isi.2018.8587389

[25] Chang-Hwan Lee, "Calculating Attribute Weights in K-Nearest Neighbor Algorithms using Information Theory", Journal of KIISE: Software and Applications, vol.32, Issue 9, 920-926, 2005.

[26] Ansam K., Iqbal G., Peter V., Joarder K., "Survey of intrusion detection systems: techniques, datasets and challenges", Cybersecurity, 2019. https://doi.org/10.1186/s42400-019-0038-7

[27] https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures

[28] Guangzhen Z., Cuixiao Z., Lijuan Z., "Intrusion Detection using Deep Belief Network and Probabilistic Neural Network", IEEE International Conference on Computational Science and Engineering, 2017. https://doi.org/10.1109/cse-euc.2017.119

[29] Pil-Sung Jang, "의사결정트리 기반 혁신기업 특성 분석", Korea Technology Innovaton Society, 2019.

[30] https://www.ibm.com/cloud/learn/neural-networks

[31] Rao F. A., Amagd M.,Ebrahim A. A. G., Ammar A. A, "Survey on Cyber Security for Industrial Control Systems", International Conference on Data Analytics for Business and Industry, 2021. https://doi.org/10.1109/ICDABI53623.2021.9655902

[32] https://www.ibm.com/cloud/learn/supervised-learning

[33] Marzban C., "The ROC Curve and the Area under It as Performance Measures", Weather and Forecasting, 2004. https://doi.org/10.1175/825.1

◐ 저 자 소 개 ◑

조 부 성(Bu-Seong Jo)
2022년 세종대학교 물리천문학과(이학사)
2022년 세종대학교 컴퓨터공학과(공학사)
2022년 세종대학교 일반대학원 컴퓨터공학과(재학)
관심분야 : 인공지능, 사이버보안, etc.
E-mail : seoulrhdgns1@gmail.com

김 문 석(Mun-Suk Kim)
2007년 연세대학교 컴퓨터과학과(공학사)
2009년 연세대학교 대학원 컴퓨터과학과(공학석사)
2016년 연세대학교 대학원 컴퓨터과학과(공학박사)
2016년~2021년 미국국립표준기술연구소 Guesst Researcher
2021년~현재 세종대학교 컴퓨터공학과 조교수
관심분야 : 무선 네트워크, 인공지능, 무선 센싱, etc.
E-mail : msk@sejong.ac.kr