스태킹 앙상블 기반 IP 카메라 침입 탐지 모델의 교차 도메인 간 강건성 평가 연구

이지수[1], 전상훈[2*]
[1]수원대학교 정보보호학과 학생, [2]수원대학교 정보보호학과 교수

Cross-Domain Robustness Evaluation of Stacking Ensemble-based IP Camera Intrusion Detection Models

Jisoo Lee[1], Sanghoon Jeon[2*]
[1]Student, Department of Information Security, The University of Suwon
[2]Professor, Department of Information Security, The University of Suwon

요 약 IP 카메라 보급 확대로 인한 보안 위협이 증가함에 따라 머신러닝 기반 이상 탐지 기술의 중요성이 커지고 있다. 그러나 기존 연구는 단일 데이터 세트에 편향되어 이기종 네트워크 환경에서의 일반화 성능 확보에 한계가 있다. 본 연구에서는 호스트 기반의 N-BaIoT와 네트워크 기반의 Kitsune 데이터 세트를 교차 활용하여 Stacking 앙상블 기반의 강건한 침입 탐지 모델을 제안한다. 특히 RF, NN, AdaBoost를 기저 학습기로 구성하고 Orange3 환경에서 하이퍼파라미터 최적화를 수행하였다. 실험 결과, 동일 도메인 환경에서는 AUC 0.999 이상의 높은 탐지 성능을 보였다. 반면 이종 도메인 전이 시나리오에서는 데이터 특성 차이로 인해 성능 저하가 발생하였다. Information Gain 기반 특징 분석 결과, N-BaIoT 데이터의 강한 이상 패턴과 Kitsune 데이터의 상대적으로 미묘한 이상 패턴 간의 비대칭성이 이러한 성능 저하의 주요 원인임을 확인하였다. 또한 제안된 Stacking 모델은 단일 머신러닝 모델 대비 보다 안정적인 성능을 보였으며, 복합 데이터가 혼재된 실제 보안 솔루션 환경에서도 활용 가능성을 시사한다.

주제어 : 침입 탐지 시스템, 스태킹 앙상블 학습, IoT 보안, 교차 도메인 침입 탐지, 이상 탐지

Abstract With the increasing security threats from the proliferation of IP cameras, the importance of machine learning-based anomaly detection is growing. However, existing studies are often biased toward single datasets, limiting their generalization performance in heterogeneous network environments. This study proposes a robust Stacking ensemble-based intrusion detection model by cross-utilizing the host-based N-BaIoT and network-based Kitsune datasets. Specifically, RF, NN, and AdaBoost were configured as base learners, and hyperparameter optimization was performed in the Orange3 environment. Experimental results show that the proposed model achieves an AUC greater than 0.999 under the intra-validation scenario. In inter-validation scenarios, performance degradation is observed due to differences in data characteristics. Information Gain analysis reveals that the asymmetric anomaly signals between the strong attack patterns in N-BaIoT and the relatively subtle anomalies in Kitsune are the primary cause of this performance degradation. Furthermore, the proposed stacking model demonstrates more robust performance than single machine learning models, suggesting its applicability to real-world security solutions where heterogeneous data sources coexist.

Key Words : Intrusion Detection System, Stacking Ensemble Learning, IoT Security, Cross-Domain Intrusion Detection, Anomaly Detection

본 연구는 원자력안전위원회의 재원으로 한국원자력안전재단의 지원을 받아 수행한 원자력안전연구사업의 연구결과입니다. (RS-2026-25524055)
*교신저자 : 전상훈(shjeon@suwon.ac.kr)
접수일 2026년 03월 06일 수정일 2026년 05월 11일 심사완료일 2026년 06월 05일

1. 서론

ICT(Information and Communications Technology) 기술의 발전에 따라 사물인터넷(Internet of Things, IoT)은 실생활에서 다방면으로 사용되고 있다. 특히 1인 가구와 반려동물 가구의 증가로 인해 실시간 모니터링이 가능한 IP 카메라의 수요가 급격히 증가하고 있다[1]. 여기에 맞벌이 가구의 증가로 영유아의 안전을 실시간으로 확인 및 기록하려는 홈캠 설치가 일반화되면서 IP 카메라는 사용자에게 편의성과 안전성을 제공하는 필수 가전으로 자리 잡았다[2].

하지만 수요가 증가함에 따라 보안이 충분히 고려되지 않은 기기들이 대량으로 보급되면서, 이러한 점을 악용하는 해킹 사례가 증가하고 있다. Yuhong Nan 등[3]의 연구에 따르면 분석 대상 IoT 앱의 약 31.8%가 사용자에게 명시적인 고지 없이 민감한 데이터를 노출하고 있으며, 이로 인해 프라이버시 침해 위협이 더욱 증가하였다. 이러한 보안 위협은 단순한 가능성을 넘어 대규모 실제 피해로 이어지고 있다. 최근 경찰청의 발표에 따르면 취약한 보안 설정을 악용한 해킹으로 약 12만 대의 IP 카메라 영상이 유출되어 해외 불법 사이트에서 거래된 사실이 확인되었다[4]. 조사 결과 유출된 영상의 62%가 해킹에 의한 것이었으며 대부분의 피해 기기가 단순한 비밀번호 조합을 사용하고 있어 보안 관리의 심각한 허점이 드러났다.

이러한 보안 위협을 해결하기 위해 머신러닝을 활용한 이상 탐지 기술에 관한 연구가 다양하게 진행되어 왔다. 특히 Random Forest, K-Nearest Neighbors 등의 알고리즘을 활용하여 IP 카메라 네트워크 내 침입을 탐지하려는 연구들이 높은 탐지 성능을 보여주고 있다[5-7]. 그러나 대다수의 선행 연구는 특정 데이터 세트 내에서의 정확도 향상에만 치우쳐 있어 데이터의 편향성으로 인한 일반화 성능의 한계를 지닌다. 실제 보안 현장에서는 제조사, 기기 사양, 네트워크 구성 등 환경적 변수가 매우 다양하므로 단일 환경에서 수집된 데이터에 의존하는 모델은 강건성을 보장하기 어렵다.

본 연구에서는 호스트 기반의 N-BaIoT 데이터 세트와 네트워크 트래픽 기반의 Kitsune 데이터 세트를 교차 활용하여, 단일 환경에 편향된 기존 모델의 한계를 극복하기 위한 Stacking 앙상블 기반의 강건한 침입 탐지 모델을 제안한다. 단순히 특정 지표의 향상을 넘어 Random Forest, Neural Network, AdaBoost의 하이퍼파라미터를 체계적으로 최적화하고, 이를 결합한 앙상블 모델이 이기종 네트워크 환경의 미학습 위협에 대해 어느 정도의 탐지 유효성을 유지하는지 실증적으로 검증한다. 특히 개별 기저 학습기들의 파라미터 변화(Trees, Layers, Estimators 등)에 따른 성능 추이를 정량적으로 분석함으로써, 데이터 도메인 변화에 강건한 모델의 일반화 성능을 분석한다. 최종적으로 대규모 IoT 트래픽 환경에서 적합한 최종 모델을 제시하는 것을 목표로 한다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 관련 연구와 본 연구에서 활용하는 두 종류의 데이터 세트의 구조와 특징을 다룬다. Ⅲ장에서는 데이터 전처리 방법 및 하이퍼파라미터 최적화를 포함한 Stacking 앙상블 모델의 설계 방법을 설명한다. Ⅳ장에서는 데이터 내(Intra) 및 데이터 간(Inter) 시나리오별 성능 검증 결과를 분석한다. 마지막으로 Ⅴ장에서는 본 연구의 결론을 도출하고 향후 연구 방향을 제시한다.

2. 관련 연구 및 데이터 세트

2.1 머신러닝 기반 네트워크 탐지 모델

IP 카메라에서 네트워크 침입 탐지 시스템의 성능을 향상시키기 위해 머신러닝을 활용한 다양한 연구가 진행되어 왔다.

최민영 등[5]의 연구는 IP 카메라 해킹으로 인한 사생활 침해 문제를 해결하기 위해 머신러닝 기반의 실시간 네트워크 공격 탐지 시스템 및 사용자 애플리케이션을 제안하였다. CIC IoT 2023 오픈 데이터 세트를 활용하여 네트워크 트래픽 패턴을 분석해 Dos, DDos, Spoofing 등 총 다섯 가지 공격 유형에 대한 다중 분류 모델을 구축하였다. 가장 우수한 성능을 보인 Random Forest 모델을 활용하여 실제 Tapo C200 IP 카메라를 대상으로 한 실험에서 100%의 DDoS 탐지율을 보였고 사용자에게 알림을 전송하기까지 평균 0.89초가 소요됨을 입증하였다. 애플리케이션을 통해 공격 탐지 시 즉각 알림 기능과 네트워크 보안 로그 확인 기능을 제공함으로써 사용자의 능동적인 대응을 제안하였다.

여승연 등[6]의 연구는 IoT 환경에서 빈번하게 발생하는 DoS 및 DRDoS 공격을 탐지하기 위해 시스템 및 네트워크 메트릭을 통합적으로 활용하는 머신러닝 기반 침입탐지 모델을 제안하였다. 라즈베리파이를 활용한 실제 테스트베드를 구축하고 패킷 크기와 전송 속도를 다양하게 하여 총 37종의 메트릭 데이터를 실시간으로 수집하였다. 실험 결과 이진 분류 및 다중 분류에 Random Forest

모델이 가장 좋은 성능을 보여 Random Forest 모델을 이용하여 공격을 탐지하는 방식을 제안하였다.

현미진[7]의 연구는 IoT 환경에서 봇넷 공격을 실시간으로 탐지하기 위해 머신러닝 알고리즘의 성능을 비교 분석하고 최적의 모델을 제안하였다. Bot-IoT 데이터 세트를 활용하여 Decision Tree, Naïve Bayes, Random Forest, KNN의 4가지 알고리즘을 대상으로 성능을 평가하였고 실험 결과 KNN 알고리즘이 DDos, DoS, Data theft 공격 탐지에서 99% 이상의 높은 정확도를 보였다. 하지만 데이터 불균형 문제로 Normal 트래픽과 Data theft 공격에 대한 탐지 성능은 상대적으로 낮게 나타났다.

하지만 이러한 선행 연구들은 단일 데이터 세트만을 활용하여 모델을 검증했다는 한계가 존재한다. 특정 환경에서 수집된 데이터로 학습된 모델은 제조사나 네트워크 환경이 다른 실제 환경에서 탐지 성능이 낮아지는 문제가 발생할 수 있다. 따라서 본 연구에서는 두 개의 오픈 데이터 세트를 교차 활용하여 미학습 환경에서의 탐지 유효성을 검증하고 이종 도메인 간의 성능 편차를 최소화할 수 있는 구성을 도출하고자 한다. 이를 통해 단일 모델의 한계를 극복하는 Stacking 기반의 강건한 보안 모델을 제안하고, 정밀한 파라미터 튜닝을 통해 실제 IP 카메라 보안 현장에서 요구되는 높은 일반화 성능과 탐지 신뢰성을 동시에 확보하는 것을 목표로 한다.

2.2 데이터 세트

Meidan 등[8]이 제안한 Network-based Detection of IoT Botnet Attacks (N-BaIoT)는 IoT 기기를 대상으로 하는 봇넷 공격을 탐지하기 위한 호스트 기반 데이터 세트이다. 실제 사용 IoT 기기 9종을 대상으로 수집되었으며 기기별 정상 트래픽과 대표적인 봇넷 공격인 Mirai 및 BASHLITE(Gafgyt)의 공격 트래픽이 포함되어 있다. 이 데이터 세트는 패킷의 흐름을 실시간으로 분석하기 위해 통계적 수치로 가공된 특징 기반의 CSV 파일로 구성되어 있다. 특징 추출 과정에서는 각 패킷이 도착할 때마다 소스 IP와 MAC 주소, 목적지 포트 등의 정보를 기반으로 행동 스냅샷을 캡처하며 총 115개의 네트워크 통계적 특징을 추출하였다. 특히 100ms부터 1min까지 5개의 서로 다른 시간 윈도우를 활용하여 데이터 전송의 급격한 변화를 다차원적으로 반영하고 있다. 본 연구에서는 IP 카메라 보안이라는 분석 목적에 집중하기 위해 전체 기기 중 베이비 모니터 1대, 보안 카메라 4대, 웹캠 1대를 포함한 총 6대의 기기를 분석 대상으로 최종 선정하였다.

Mirsky 등[9]이 제안한 Kitsune 데이터 세트는 실제 운영 중인 IP 카메라 비디오 감시 네트워크 환경에서 수집된 네트워크 기반 데이터 세트이다. 이 데이터 세트는 실제 네트워크 환경에 연결된 8대의 HD 카메라 트래픽을 기반으로 하며 네트워크 패킷의 통계적 흐름을 실시간으로 추적하여 총 115개의 특징을 추출한다.

본 연구에서 Kitsune 데이터 세트를 활용하는 주된 목적은 데이터의 환경 및 공격 시나리오의 다양성 확보에 있다. N-BaIoT가 기기 중심의 봇넷 공격에 집중한다면 Kitsune는 실제 IP 카메라에서 발생할 수 있는 9종의 보안 위협(ARP MitM(중간자 공격), SSDP Flood, SYN DoS, Video Injection 등)을 포함하고 있다. 특징 추출 측면에서는 N-BaIoT와 동일하게 특징 추출 알고리즘을 공유하고 있어 소스 IP, MAC 주소, 포트 정보 등을 기반으로 115개의 통계적 특징을 도출하므로 일반화 성능 분석을 수행하기에 최적의 조건을 갖추고 있다.

추출된 115개의 특징은 패킷의 흐름을 5개의 시간 윈도우(100ms, 500ms, 1.5s, 10s, 1min)로 나누어 분석하며 트래픽의 급격한 변화와 장기적인 이상 징후를 동시에 포착하도록 설계되었다. 구체적으로 MI(MAC-IP), H(Host-IP), HH_jit(Jitter) 범주는 데이터의 양과 평균 등을 나타내는 1차원(1D) 통계를 활용하고, HH(Host-Host)와 HpH(Host-Port-Host) 범주는 상관계수를 포함한 2차원(2D) 통계를 활용하여 기기 간의 복합적인 상호작용을 반영한다. Table 1은 본 연구에서 활용한 두 데이터 세트의 공통적인 네트워크 통계적 특징 구성을 나타낸다.

〈Table 1〉 Network Traffic Feature Specifications

Aggregation Level	Description	Statistics
MI	Traffic statistics based on Source MAC and IP addresses	weight, mean, variance
H	Traffic statistics from a specific Source IP	weight, mean, variance
HH	Traffic statistics between specific Source and Destination IPs	weight, mean, std, radius, magnitude, covariance, pcc
HH_jit	Inter-arrival time (Jitter) statistics between Source and Destination	weight, mean, variance
HpH	Traffic statistics based on IP and Destination Port	weight, mean, std, radius, magnitude, covariance, pcc

3. 제안 방법

3.1 데이터 세트 전처리

본 연구에서는 IP 카메라 환경의 침입 탐지 성능을 다각도로 검증하기 위해 호스트 기반의 N-BaIoT와 네트워크 트래픽 기반의 Kitsune 데이터 세트를 활용한다. 특정 환경에 편향되지 않은 모델의 강건성을 측정하기 위해 두 데이터 세트를 교차 활용하는 실험 환경을 구축하였다.

N-BaIoT 데이터 세트는 실제 IoT 기기 9종 중 본 연구의 목적에 부합하는 베이비 모니터, 보안 카메라 등 총 6대의 기기 데이터를 선정하여 사용한다. 해당 데이터는 정상 트래픽과 Mirai, BASHLITE 등 대표적인 봇넷 공격 트래픽을 포함하고 있으며, 실험의 효율성을 위해 파이썬(Python) 기반의 무작위 샘플링을 거쳐 총 90,000개의 데이터를 추출하였다.

Kitsune 데이터 세트는 실제 운영 중인 IP 카메라 감시 네트워크에서 수집된 데이터로, 초기 100만 개의 패킷이 정상 트래픽으로 구성된 후 공격이 발생하는 구조를 가진다. 모델의 정밀한 성능 측정을 위해 labels.csv 파일을 분석하여 공격(Label 1)이 발생하는 정확한 인덱스 구간을 식별하였으며, 해당 시점의 패킷만을 정밀하게 추출하는 과정을 거쳤다. 데이터 규모를 N-BaIoT와 유사한 수준으로 맞추기 위해 9가지 공격 유형별로 약 10,000개씩, 그리고 정상 트래픽 10,000개를 추출하여 총 97,037개의 데이터를 확보하였다. 이때 SYN DoS 공격과 같이 가용 데이터가 부족한 시나리오는 실제 환경을 반영하여 7,037개의 데이터를 전량 추출하여 사용하였다.

본 연구에서는 학습 데이터와 테스트 데이터 간의 특징 공간 불일치를 최소화하기 위해 Kitsune과 N-BaIoT 데이터 세트에서 공통적으로 제공되는 AfterImage 기반 115개 통계 특징을 사용하였다. 이를 위해 Kitsune 원본 데이터에 포함된 줄 번호 등 분석에 불필요한 인덱스 열을 제거하고, 모델 입력에 해당하는 순수 특징 값만을 분리하였다. 이후 N-BaIoT의 특징명을 Kitsune의 특징 체계와 대응되도록 재정렬 및 매핑함으로써 두 데이터 세트 간 입력 특징의 정합성을 확보하였다.

최종적으로 실제 보안 관제 환경의 이진 분류 성능을 평가하기 위해, 각 데이터 세트에 포함된 다양한 세부 공격 유형들을 하나의 공격 범주로 통합하였다. 이에 따라 본 논문에서는 모든 데이터를 정상(Normal)과 공격(Attack) 두 개의 클래스만 가지는 형태로 정의하여 사용하며, 구체적인 구성은 Table 2와 같다.

〈Table 2〉 Dataset Specifications for Machine Learning Training and Evaluation

Dataset	Category	Description
N-BaIoT	Data Source	Traffic from 6 real IoT devices
	Total Instances	90,000 samples
	Feature Count	115 statistical features
	Class	Normal, Attacks (Mirai, BASHLITE)
Kitsune	Data Source	Traffic from various network environments
	Total Instances	97,037 samples
	Feature Count	115 statistical features
	Class	Normal, Attacks (ARP MitM, Active Wiretap, Fuzzing, Mirai Botnet, OS Scan, SSDP Flood, SSL Renegotiation, SYN DoS, Video Injection)

3.2 머신러닝 모델 선정

본 연구의 모든 실험은 AMD Ryzen 7 5800X CPU(3.80 GHz, 8 cores), 32 GB RAM, Windows 11 Pro 23H2 환경에서 수행하였다. 데이터 전처리와 모델 학습 및 평가에는 Python 3.14.2와 Orange3 3.40.0을 사용하였다. 특히 Orange3의 Test and Score 위젯을 통해 10-fold 교차 검증을 실시함으로써 모델의 통계적 신뢰성을 확보하였다.

첫 번째 학습 모델로 Random Forest(RF)를 선정한 이유는 IoT 환경 특유의 고차원성과 데이터 노이즈를 효과적으로 제어하기 위함이다. RF는 독립적인 다수의 결정 나무를 학습시킨 뒤 그 결과를 평균화하는 배깅(Bagging, Bootstrap Aggregation) 기법을 기반으로 하며, 이는 모델의 분산을 줄여 과적합을 방지하는데 탁월한 효과가 있다. 특히 Adnan Rawashdeh 등[10]의 연구에 따르면 RF는 비선형 결정 경계(Non-linear decision boundaries)를 제공하며 수백만 개의 네트워크 플로우에 대한 확장성을 갖추고 있어 IoT 데이터 특성에 가장 적합한 선도적인 알고리즘 중 하나로 평가받는다. 또한 Dukka KarunKumar Reddy[11] 등은 RF가 네트워크의 정상 및 이상 동작을 통계적으로 모델링할 수 있는 가장 풍부한 분류기 중 하나임을 명시하고 배깅 접근 방식이 악의적인 네트워크 운영을 분류하고 처리하는데 매우 효과적임을 강조하였다.

두 번째 학습 모델로 신경망(Neural Network, NN)을 선정한 이유는 트리 기반 모델이 포착하기 어려운 비선형적이고 계층적인 특징 표현을 학습하여 지능화된 공

격 및 제로데이 위협에 대응하기 위함이다. 신경망은 입력 데이터(x_i)와 가중치(w_i)의 곱에 편향(b)를 더하고 비선형 활성화 함수(f)를 거쳐 결과를 도출하는 기본 원리를 가진다. 수식은 식 (1)과 같다.

$$y = f(\sum_{i=1}^{n} w_i x_i + b) \quad (1)$$

이러한 메커니즘을 통해 신경망은 고차원 데이터 내의 복잡한 상관관계를 추출하며 특히 정형 데이터뿐만 아니라 비정형적인 네트워크 트래픽 패턴에서도 우수한 성능을 발휘한다. Sangeetha V 등[12]은 LSTM(Long Short-Term Memory) 네트워크를 활용하여 트래픽의 시계열적 패턴을 학습한 결과, NF-BoT-IoT 데이터 세트에서 99%의 높은 탐지 정확도를 달성하며 신경망이 실시간 이상 탐지에서 강력한 해결책임을 입증하였다. 또한 Mei Liu와 Leon Yang[13]의 연구에 따르면 신경망 기반의 딥러닝 모델은 대규모의 동적인 IoT 데이터를 처리하는 데 있어 기존 통계적 방식보다 우수하며 특히 비지도 학습(Unsupervised learning) 기법과 결합될 경우 라벨링되지 않은 미지의 이상 징후를 탐지하는데 탁월한 효과가 있다.

마지막으로 세 번째 학습 모델로 AdaBoost를 선정한 이유는 IoT 네트워크 데이터의 고질적인 문제인 클래스 불균형을 해결하고 탐지하기 어려운 소수 클래스 공격에 대한 정교함을 높이기 위함이다. 부스팅 기법은 이전 단계에서 오분류된 샘플에 순차적으로 가중치를 부여하여 학습함으로써 모델의 편향을 효과적으로 감소시킨다. 이는 정상 트래픽이 공격 트래픽을 압도하는 IoT 환경에서 미세한 공격 패턴을 포착하는 데 필수적인 역량이다. Lamya Jaafouri 등[14]은 Gradient Boosting을 포함한 앙상블 프레임워크가 UNSW-NB15 데이터 세트에서 95.6%의 AUC와 88%의 F1 스코어를 달성하며 단일 모델 대비 복잡한 트래픽 패턴 포착 능력이 우수함을 입증하였다. 또한 Sujit Bebortta와 Suman Singh[15]의 연구에 따르면, AdaBoost를 활용한 알고리즘은 96.316%의 예측 정확도와 0.991의 매우 높은 AUC-ROC 값을 기록하며 대규모 이기종 네트워크의 침입 탐지 시스템(IDS) 설계에 적합한 선택임을 보여주었다. 더불어 Owais Bukhari 등[16]은 부스팅 기반 앙상블은 98.6%의 정확도를 기록하며 특히 UNSW-NB15 및 CICIDS2017 데이터 세트 내의 희귀 공격 탐지에서 기존 모델들을 압도하는 성능을 보였다.

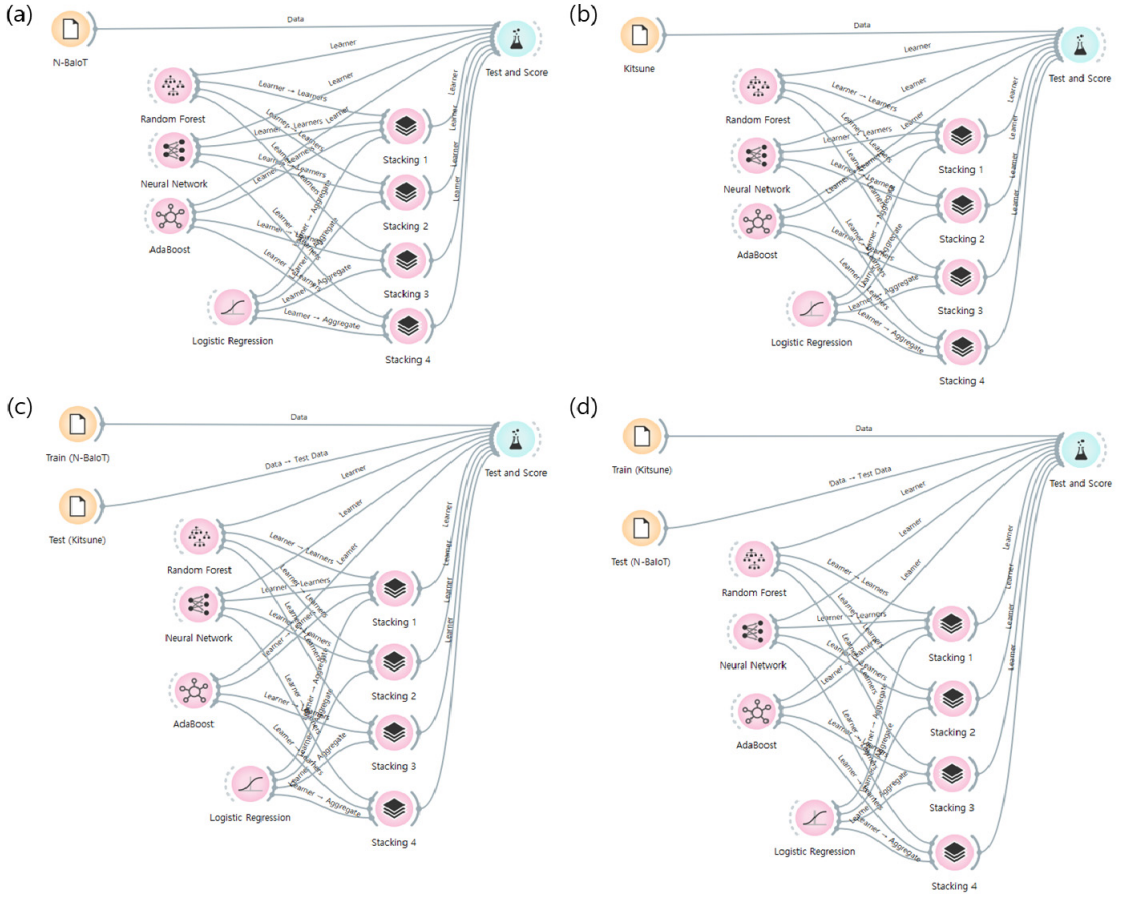
3.3 머신러닝 모델 하이퍼파라미터 최적화

본 연구에서는 Stacking 앙상블의 기초가 되는 개별 머신러닝 모델의 성능을 극대화하기 위해 Grid Search를 통한 하이퍼파라미터 최적화를 수행하였으며, 특히 네 가지 실험 시나리오(Case A, B, C, D)의 특성에 맞춰 각 모델의 파라미터를 정밀하게 조정하였다.

먼저 Random Forest 모델은 의사결정 나무의 개수를 [10, 100, 500, 1000] 범위에서 탐색하였다. 내부 검증 환경인 Case A(N-BaIoT)와 Case B(Kitsune)에서는 tree의 수가 100개일 때의 분류 정확도가 각각 0.999856과 0.99999로 최고점에 도달하여 100개를 공통 최적값으로 선정하였다. 강건성 연구의 핵심인 Case C(Train(N-BaIoT) to Test(Kitsune))에서는 나무의 개수가 10개(AUC 0.881)에서 100개(AUC 0.970)로 증가할 때 지표가 폭등하는 변곡점이 관찰되었으며 500개 이상부터는 오히려 성능이 하락하는 양상을 보여 100개를 최종 선정하였다. 반면 도메인 시프트가 극심한 Case D(Train(Kitsune) to Test(N-BaIoT))에서는 tree의 수가 많을수록 AUC가 0.314에서 0.407까지 점진적으로 상승하는 것을 확인하였고 이에 따라 복잡한 도메인 차이를 극복하기 위해 1000개의 tree를 최종 하이퍼파라미터로 선정하였다.

Neural Network 모델은 은닉층 구조를 [20, 100, (100, 100), (100, 100, 100)]으로 확장하며 최적점을 탐색하였다. 내부 검증인 Case A에서는 CA 0.999544로 가장 높은 정밀도를 보인 3개 층(100, 100, 100) 구조를 채택하였으나 Case B에서는 의외로 구조가 단순한 단일 층(20)에서 CA 0.999918의 최고 성능이 도출되어 과적합 없이 깔끔한 학습이 이루어진 해당 지점을 선택하였다. 이종 도메인 검증인 Case C에서는 AUC 수치가 높은 단일 층보다 CA(0.563)와 Recall(0.563)이 균형있게 도출된 2개 층(100, 100) 구조를 강건성 확보의 최적점으로 판단하였다. Case D 역시 AUC는 낮으나 실제 분류 정확도가 0.650으로 전 구간 중 가장 우수하게 나타난 2개 층(100, 100) 구조를 최종 선정하였다.

마지막으로 AdaBoost 모델은 반복 학습 횟수를 결정하는 Estimators 개수를 [10, 50, 100, 200] 구간에서 테스트하였으나, 모든 시나리오에서 파라미터 변화에 따른 성능 변동이 거의 발견되지 않았다. 구체적으로 Case A에서는 AUC 0.999542, Case B는 AUC 0.999844, Case C는 AUC 0.622, Case D는 AUC 0.351등으로 수치가 고정되는 양상을 보였기에, 모델의 구조적 한계



[Fig. 1] Overall Workflow showing Multi-Scenario Cross-Validation using Stacking Ensemble on N-BaIoT and Kitsune Datasets: (a) Intra-validation for N-BaIoT dataset (b) Intra-validation for Kitsune dataset (c) Inter-validation for cross-domain analysis (Train: N-BaIoT / Test: Kitsune) (d) Inter-validation for cross-domain analysis (Train: Kitsune / Test: N-BaIoT)

를 파라미터로 극복하기보다는 보안 솔루션의 실시간 탐지 속도와 모델의 경량성을 고려하여 안정적인 중간값인 50개를 모든 시나리오의 공통 파라미터로 확정하였다.

3.4 Stacking 앙상블 설계 및 실험 절차

본 연구의 전체적인 실험 절차는 Fig. 1과 같다. 파라미터 최적화가 완료된 개별 머신러닝 모델들을 바탕으로 최종적인 Stacking 앙상블 모델을 설계하였다. 기저 학습기로 RF, NN, AdaBoost의 3가지 모델을 선정하였다. 이러한 선정은 각 모델이 가진 서로 다른 학습 귀납 편향을 통해 IP Camera의 복잡한 데이터 특성을 상호 보완적으로 해결하기 위한 선택이다.

본 연구는 3개 모델의 결합(Stack 1: RF + NN + AdaBoost)이 가지는 당위성을 입증하기 위해 단일 모델뿐만 아니라 2개 모델로 구성된 다양한 앙상블 조합(Stacking 2: RF + NN, Stack 3: RF + AdaBoost, Stack 4: NN + AdaBoost)에 대한 비교 실험을 병행하였다.

최적화된 개별 모델들의 예측 결과를 입력 데이터로 활용하여 최종 판단을 내리는 메타 학습기로는 Logistic Regression을 선정하였다. 이는 개별 머신러닝 모델들이 가진 고유의 예측 편향을 메타 학습기가 가중치 학습을 통해 보완함으로써, 단일 모델만으로는 달성하기 어려운 이종 환경에서의 강건한 보안 성능을 도출하기 위함이다.

모델의 성능 검증은 N-BaIoT와 Kitsune 데이터 세트를 활용하여, 동일 환경 내의 탐지력(Intra-validation) 뿐만 아니라 학습되지 않은 이종 도메인으로의 전이 탐

지력(Inter-validation)을 측정하는 방식으로 설계하였다. 특히 N-BaIoT로 학습된 모델을 Kitsune 환경에 직접 적용하는 시나리오와 그 반대의 경우를 통해, 제안하는 3중 Stacking 모델이 단일 알고리즘 및 2중 결합 모델 대비 가지는 환경 적응력의 우위를 실증적으로 입증하고자 한다.

3.5 성능 평가 지표

본 연구에서는 IP 카메라를 대상으로 한 네트워크 공격 트래픽을 Positive 클래스로, 정상 트래픽을 Negative 클래스로 간주한다. 단순히 단일 환경의 수치를 제시하는 것을 넘어, 교차 데이터 세트 환경에서의 일반화 성능을 다각도로 분석하여 모델의 강건성을 평가한다.

머신러닝 모델의 성능 평가는 AUC(Area Under the Curve), CA(Classification Accuracy), Precision, Recall, F1-Score의 다섯 가지 주요 평가 지표를 활용하여 수행되었다. 각 지표를 산출하기 위한 모델의 예측 결과 유형은 다음과 같이 네 가지로 정의된다.

- True Positive(TP): 실제 공격 패킷을 모델이 공격으로 정확히 탐지한 경우
- True Negative(TN): 실제 정상 트래픽을 모델이 정상으로 정확히 식별한 경우
- False Positive(FP): 정상 트래픽을 공격으로 잘못 판단한 경우(오탐)
- False Negative(FN): 실제 공격 패킷을 정상으로 잘못 판단하여 놓친 경우(미탐)

위의 정의를 바탕으로 본 연구에서 채택한 성능 평가 지표의 산출 수식은 다음과 같다.

AUC는 ROC 곡선 아래의 면적을 나타내고 모델의 분류 성능을 종합적으로 나타내는 지표이다. 값이 1에 가까울수록 모델의 판별 능력이 우수함을 의미하며, 수식은 식 (2)과 같다.

$$AUC = \int_0^1 TPR(FPR)\,dFPR \tag{2}$$

CA(Classification Accuracy)는 전체 데이터 중 모델이 정상과 공격을 올바르게 예측한 비율을 나타낸다. 수식은 식 (3)와 같다.

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision은 모델이 공격(Positive)으로 예측한 결과 중 실제로 공격이었던 비율을 의미한다. 수식은 식 (4)과 같다.

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{4}$$

Recall은 실제 공격 패킷 중 모델이 공격으로 올바르게 탐지해낸 비율이다. 보안 시스템에서는 공격을 놓치지 않는 것이 매우 중요하므로 연구의 핵심 지표로 활용되며, 수식은 식 (5)와 같다.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

F1-Score(F1)는 Precision과 Recall의 조화 평균이고 두 지표 간의 균형 잡힌 성능을 평가하는 지표이며 식 (6)와 같다.

$$F1 - Score = 2 \times \frac{\mathrm{Precision} \times Recall}{\mathrm{Precision} + Recall} \tag{6}$$

4. 실험 결과 및 분석

본 장에서는 제안하는 Stacking 앙상블 모델과 개별 머신러닝 모델의 성능을 비교 분석한다. 실험은 동일한 네트워크 환경에서의 탐지력을 확인하는 데이터 세트 내 검증(Intra-validation)과 환경이 변화했을 때의 대응력을 확인하는 이기종 데이터 세트 간 검증(Inter-validation)으로 구분하여 수행되었다. 모든 성능 지표는 앞서 정의한 AUC, CA, F1-Score, Precision, Recall을 기준으로 측정되었다.

4.1 데이터 세트 내 검증(Intra-validation)

Table 3(a)은 동일한 도메인 내에서 학습과 테스트가 이루어진 N-BaIoT 데이터 세트 내 성능 평가 결과를 나타낸다. 이는 훈련된 머신러닝 모델이 N-BaIoT 환경 고유의 공격 특징을 얼마나 효과적으로 학습했는지를 보여주는 기초 지표로 활용된다.

실험 결과, 단일 머신러닝 모델(RF, NN, AdaBoost)과 네 가지 조합의 스태킹 앙상블 모델 모두 AUC 및 CA 지표에서 0.999 이상의 매우 높은 수치를 기록하며 안정적으로 수렴하는 양상을 보였다. 단일 모델 중에서는 Random Forest(RF)가 CA 0.999844로 가장 우수한 탐지 성능을 보였으며, 이는 트리 기반의 앙상블 기법이 N-BaIoT 데이터의 통계적 특징을 변별하는 데 최적화되어 있음을 보여준다.

스태킹 모델 분석을 살펴보면 세 가지 알고리즘을 모두 결합한 Stack 1(RF, NN, AdaBoost)은 CA 0.999833을 기록하여 단일 RF 모델에 근접하는 최상위 탐지력을

유지하였다. 이는 서로 다른 학습 메커니즘을 가진 모델들을 통합함에 있어 성능 저하 없이 안정적인 결합이 이루어졌음을 의미한다. 또한 RF와 NN을 결합한 Stack 2와 RF와 AdaBoost를 결합한 Stack 3 역시 각각 0.999789와 0.999711의 높은 CA를 기록하며 RF의 강력한 탐지 성능이 스태킹 구조 내에서도 유효하게 작용하고 있음을 보여주었다.

반면 RF가 제외된 NN과 AdaBoost의 조합인 Stack 4는 CA 0.999589로 모든 스태킹 조합 중 상대적으로 낮은 수치를 기록하였는데, 이는 N-BaIoT 환경의 공격 탐지에 있어 RF의 배깅 기법이 제공하는 안정성이 핵심적인 요인으로 작용하고 있음을 반증한다. 그럼에도 불구하고 Stack 4 역시 0.999 이상의 높은 AUC를 유지하고 있어, 스태킹 앙상블 기법이 개별 학습기의 한계를 보완하여 전반적인 탐지 신뢰도를 높이는 데 기여하고 있음을 확인할 수 있다.

Table 3(b)은 동일한 도메인 내에서 학습과 테스트가 이루어진 Kitsune 데이터 세트 내 성능 평가 결과를 나타낸다. 이는 훈련된 머신러닝 모델이 Kitsune 네트워크 환경 고유의 공격 특징을 얼마나 효과적으로 학습했는지를 보여주는 기초 지표로 활용된다.

실험 결과, Kitsune 환경에서도 모든 개별 모델 및 네 가지 스태킹 조합이 AUC 0.999 이상의 압도적인 수치를 기록하며 데이터 세트의 복잡한 네트워크 흐름 특징을 완벽하게 학습했음을 입증하였다. 개별 모델 중에서는 Random Forest(RF)가 CA 0.99999를 기록하여 사실상 완벽한 탐지 성능을 보였으며, AdaBoost 역시 CA 0.999959로 이에 근접한 성과를 거두어 트리 기반 앙상블 기법이 Kitsune 데이터의 변별력을 확보하는 데 매우 효과적임을 확인하였다.

〈Table 3〉 Performance Metrics of Intra-validation

Dataset	Model	AUC	CA	F1	Prec	Recall
(a) N-BaIoT	RF	0.9999	0.9998	0.9998	0.9998	0.9998
	NN	0.9998	0.9995	0.9995	0.9995	0.9995
	AdaBoost	0.9995	0.9996	0.9996	0.9996	0.9996
	Stack1	0.9999	0.9998	0.9998	0.9998	0.9998
	Stack2	0.9999	0.9997	0.9997	0.9997	0.9997
	Stack3	0.9999	0.9997	0.9997	0.9997	0.9997
	Stack4	0.9999	0.9995	0.9995	0.9995	0.9995
(b) Kitsune	RF	0.9999	0.9999	0.9999	0.9999	0.9999
	NN	0.9999	0.9997	0.9997	0.9997	0.9997
	AdaBoost	0.9998	0.9999	0.9999	0.9999	0.9999
	Stack1	0.9999	0.9999	0.9999	0.9999	0.9999
	Stack2	0.9999	0.9999	0.9999	0.9999	0.9999
	Stack3	0.9999	0.9999	0.9999	0.9999	0.9999
	Stack4	0.9999	0.9999	0.9999	0.9999	0.9999

Stack 1(RF, NN, AdaBoost)은 CA 0.999969를 기록하며 모든 스태킹 조합 중 가장 높은 정확도를 달성하였다. RF를 포함한 Stack 2(RF, NN)와 Stack 3(RF, AdaBoost)은 동일하게 CA 0.999959를 기록하며 안정적인 탐지력을 유지하였고, RF가 제외된 조합인 Stack 4(NN, AdaBoost) 또한 CA 0.999959라는 높은 정확도를 기록하며 모든 스태킹 구조가 개별 모델의 성능 하한선을 견고하게 지지하고 있음을 보여주었다.

4.2 데이터 세트 간 검증(Inter-validation)

Table 4(a)는 학습 도메인(N-BaIoT)과 상이한 이기종 네트워크 도메인(Kitsune) 환경에서 탐지 모델을 평가한 교차 도메인 간 강건성 평가 결과를 나타낸다.

실험 결과, 도메인 시프트 현상으로 인해 모든 모델의 성능이 데이터 세트 내 검증 대비 전반적으로 하락하였으나, 본 연구에서 제안하는 Stack 1(RF, NN, AdaBoost)은 CA 0.90619 및 Recall 0.975287을 기록하며 모든 비교군 중 가장 압도적인 강건성을 입증하였다. 특히 단일 모델 중 Neural Network(NN)의 Recall이 0.513994로 급락하며 공격 탐지에 심각한 결함을 보인 것과 대조적으로, Stack 1은 이러한 단일 모델의 편향성을 메타 학습 과정을 통해 효과적으로 보정하여 최상의 탐지 성능을 보여주는 것을 확인하였다.

스태킹 조합별 연구 결과를 살펴보면, 3개 모델을 모두 결합한 Stack 1이 AUC 0.967926로 가장 높은 성능을 보였다. RF와 NN만을 결합한 Stack 2는 CA 0.572627, Recall 0.523679라는 저조한 성능을 보였는데, 이는 도메인 변화에 취약한 NN의 예측 오류가 전체 구조에 부정적인 영향을 미쳤음을 시사한다. 또한 RF와 AdaBoost의 조합인 Stack 3(CA 0.897926, Recall 0.96598)과 NN과 AdaBoost의 조합인 Stack 4(CA 0.874584, Recall 0.939957) 모두 준수한 성능을 보였으나, 제안 모델인 Stack 1의 수치에는 미치지 못하였다. 이는 트리 기반의 안정성(RF), 비선형 패턴 추출(NN), 그리고 오차 보정(AdaBoost)이라는 세 가지 이질적 메커니즘이 모두 통합되었을 때 비로소 미지의 네트워크 위협에 대응할 수 있는 최적의 시너지가 발생함을 실증적으로 증명하는 결과이다.

〈Table 4〉 Performance Metrics of Inter-validation

Dataset	Model	AUC	CA	F1	Prec	Recall
(a) Train (N-BaIoT) → Test (Kitsune)	RF	0.9707	0.8916	0.9415	0.9115	0.9737
	NN	0.8846	0.5639	0.6789	0.9996	0.5139
	AdaBoost	0.6227	0.8745	0.9307	0.9217	0.9399
	Stack1	0.9679	0.9061	0.9491	0.9243	0.9752
	Stack2	0.9576	0.5726	0.6873	0.9997	0.5236
	Stack3	0.9372	0.8979	0.9443	0.9237	0.9659
	Stack4	0.8183	0.8745	0.9307	0.9217	0.9399
(b) Train (Kitsune) → Test (N-BaIoT)	RF	0.3491	0.4603	0.4211	0.3882	0.4603
	NN	0.1662	0.6506	0.5270	0.4576	0.6506
	AdaBoost	0.3512	0.4638	0.4264	0.3956	0.4638
	Stack1	0.3737	0.4600	0.4206	0.3875	0.4600
	Stack2	0.3750	0.4599	0.4204	0.3872	0.4599
	Stack3	0.3991	0.4638	0.4264	0.3955	0.4638
	Stack4	0.1709	0.4598	0.4204	0.3872	0.4598

Table 4(b)는 학습 도메인(Kitsune)과 상이한 이기종 네트워크 도메인(N-BaIoT) 환경에서 모델을 평가한 교차 도메인 간 강건성(Inter-validation) 평가 결과를 나타낸다. 이는 훈련된 모델이 데이터 특성이 완전히 다른 환경으로 전이되었을 때 발생하는 성능 저하를 측정하고, 그 속에서도 최소한의 탐지 가용성을 유지하는지 확인하는 실전적 지표로 활용된다.

실험 결과, Kitsune 데이터로 학습된 모델들이 N-BaIoT 도메인에 적용되었을 때 모든 모델의 AUC가 0.4 이하로 급락하는 등 극심한 성능 저하가 관찰되었다. 특히 Neural Network(NN)의 경우 CA와 Recall은 0.650622로 가장 높게 나타났으나 AUC가 0.166235로 매우 낮게 기록되었다. 반면 Random Forest(RF)와 AdaBoost는 각각 0.349159와 0.351275의 낮은 AUC 성능을 보였다.

스태킹 조합 간의 비교를 살펴보면, 제안 모델인 Stack 1(RF, NN, AdaBoost)은 AUC 0.373704를 기록하며 단일 RF나 NN보다 안정적인 수치를 보여주었다. 또한 RF와 AdaBoost를 결합한 Stack 3가 AUC 0.399186으로 모든 조합 중 가장 높은 수치를 기록했는데, 이는 극단적인 도메인 시프트 환경에서는 NN의 개입보다 트리 기반 앙상블 모델들의 견고한 결합이 오탐을 줄이는 데 더 효과적일 수 있음을 보여준다. Stack 2(RF, NN)는 AUC 0.375036, Stack 4(NN, AdaBoost)는 AUC 0.170916을 기록하며, 전반적으로 RF가 포함된 조합들이 NN 단독 오차를 보정하며 성능 하락을 방어하고 있음을 확인하였다.

종합적으로 데이터 세트 간 검증 결과를 분석한 결과, 도메인 간의 환경 차이가 클수록 모든 모델에서 성능 하락이 관찰되었으나 모델 조합에 따라 방어 성능에 차이가 있음을 확인하였다. 전체 시나리오를 아우르는 안정성 측면에서는 Stack 1이 가장 균형 잡힌 지표를 유지하였다.

4.3 도메인 비대칭성에 따른 성능 저하 원인 분석

이러한 성능 저하는 두 데이터 세트의 트래픽 수집 환경 차이에서 기인한다. N-BaIoT는 기기별로 격리된 환경에서 Mirai·BASHLITE 봇넷이 생성하는 대용량 스캔 트래픽을 수집하였으므로, 정상·공격 간 통계적 분리도가 매우 높다. 반면 Kitsune는 실제 운영 네트워크에 8대의 카메라가 혼재된 환경에서 ARP MitM, Video Injection 등 정상 트래픽과 혼재되기 쉬운 9종의 공격을 포함하므로, 개별 특징의 이상 신호 강도가 상대적으로 낮다. 이 때문에 N-BaIoT 기반 모델은 강한 신호에 최적화된 결정 경계를 형성하여 Kitsune의 미묘한 패턴을 정상으로 오인(High FN)하는 경향이 나타난다. 반대로 Kitsune 기반 모델은 미세한 통계적 변화에 과도하게 최적화되어, N-BaIoT의 거대한 트래픽 변동을 도메인 노이즈나 미학습 패턴으로 인식하면서 오탐을 발생시킨다. 결과적으로 이러한 데이터 세트 간의 통계적 비대칭성이 양방향 교차 검증 모두에서 급격한 성능 저하를 야기하는 본질적인 원인이 된다.

본 실험의 교차 도메인 검증(Inter-Validation from Train(Kitsune) to Test(N-BaIoT))에서 관찰된 급격한 성능 저하의 원인을 규명하기 위해 각 도메인의 특징별 Information Gain(IG)을 분석하였다. IG는 특정 특징이 공격과 정상 트래픽을 분류하는 데 기여하는 정보의 양을 수치화한 지표로, 값이 클수록 도메인의 이상 신호가 뚜렷함을 의미한다.

학습 데이터인 Kitsune와 평가 데이터인 N-BaIoT의 특징별 IG 순위는 Fig. 2와 같다. 분석 결과, Fig. 2(a)의 Kitsune 도메인은 상위 특징들의 IG가 최대 0.234 수준으로 나타났다. 이는 공격 트래픽이 정상 범위와 매우 흡사하게 섞여 있어 N-BaIoT 대비 상대적으로 미묘하고 감지하기 어려운 약한 이상 신호(Subtle/Weak Anomaly)의 특성을 가짐을 시사한다. 반면 Fig. 2(b)의 N-BaIoT 도메인은 상위 특징들의 IG가 최대 0.627을 기록하며 Kitsune와 비교했을 때 상대적으로 매우 뚜렷하고 강한 이상 신호(Strong Anomaly)의 특성을 나타냈다. N-BaIoT의 정보 밀도는 Kitsune 대비 약 2.7배 높게

		#	Info. gain
1	MI_dir_L0.01_weight		0.234
2	H_L0.01_weight		0.234
3	HH_jit_L3_mean		0.234
4	HH_jit_L1_mean		0.234
5	HH_jit_L0.1_mean		0.234
6	HH_jit_L0.01_mean		0.234
7	HH_jit_L0.1_variance		0.234
8	HH_jit_L3_variance		0.233
9	HH_jit_L1_variance		0.233
10	HH_jit_L5_variance		0.233

(a)

		#	Info. gain
1	HH_L0.01_weight		0.627
2	HH_jit_L0.01_weight		0.627
3	HH_L0.1_weight		0.620
4	HH_jit_L0.1_weight		0.620
5	HH_jit_L0.01_mean		0.606
6	HH_jit_L0.01_variance		0.592
7	MI_dir_L1_weight		0.590
8	H_L1_weight		0.590
9	MI_dir_L0.1_weight		0.590
10	H_L0.1_weight		0.590

(b)

[Fig. 2] Comparison of Information Gain by Feature: (1) Kitsune Information Gain (2) N-BaIoT Information Gain

측정되었으며 이는 해당 도메인의 공격 트래픽이 통계적으로 매우 명확한 변별력을 가짐을 수학적으로 보여준다.

단순히 최댓값 비교를 넘어, 상위 N개 특징의 Information Gain 누적 분포를 비교하면 도메인 비대칭성이 더욱 뚜렷이 드러난다. N-BaIoT는 상위 10개 특징만으로도 전체 분류 정보의 대부분을 포괄하는 반면, Kitsune는 유효 정보가 더 많은 특징에 분산되어 있어 모델이 소수의 지배적 특징에 과도 의존하는 단일 모델(RF, AdaBoost)에서 전이 실패가 두드러진다. 반면 Stacking 앙상블은 RF가 고-IG 특징을, NN이 다차원 잠재 표현을, AdaBoost가 저-IG의 미묘한 신호를 각각 담당하는 역할 분담이 이루어지므로, 단일 모델 대비 특정 특징 분포에 대한 의존도가 낮아 도메인 시프트에 더 강건하다.

이러한 도메인 간 정보량의 상대적 비대칭성은 전이 학습 시 성능 하락의 결정적 요인이 된다. 상대적으로 미묘한 신호를 포착하기 위해 미세한 통계적 변화에 과도하게 최적화된 모델이, 강한 신호 기반의 N-BaIoT 도메인에 적용될 경우 두 도메인 간의 거대한 통계적 편차를 도메인 노이즈로 오인하게 된다. 이로 인해 발생하는 Negative Transfer 현상이 Case D 시나리오에서 관찰된 AUC 급락의 본질적인 원인임을 확인하였다.

Case D의 극단적 성능 하락은 Negative Transfer로 설명된다. Kitsune의 낮은 IG 기반으로 학습된 결정 경계가 N-BaIoT의 강한 이상 신호를 오히려 아웃라이어(노이즈)로 오인하기 때문이다. 이 상황에서 단일 모델들은 각자의 귀납적 편향이 그대로 전이 실패로 이어지는 반면, 본 연구의 Stacking 구조에서는 메타 학습기(Logistic Regression)가 세 기저 학습기의 예측 확률을 입력으로 받아 '어떤 학습기의 판단을 얼마나 신뢰할 것인가'를 재학습하게 된다. 따라서 특정 학습기의 예측이 극단적으로 편향되더라도 나머지 학습기들의 보정 효과로 인해 AUC 하락 폭이 단일 모델 대비 상대적으로 완화된다.

5. 결론

본 연구는 이기종 IP Camera 네트워크 환경에서의 일반화 성능 향상을 위해 Stacking 앙상블 기반 침입 탐지 모델을 제안하였다. N-BaIoT와 Kitsune 데이터 세트를 교차 활용하여 도메인 간 성능 편차를 분석하고, 최적의 하이퍼파라미터 조합을 도출하였다.

실험 결과, Stack 1(RF, NN, AdaBoost) 모델은 동일 도메인 검증에서 0.999 이상의 높은 성능을 보였고, Case C에서도 Recall 0.975를 기록하여 단일 모델 대비 우수한 강건성을 나타냈다. 그러나 Case D에서는 성능 저하가 발생하였으며, Information Gain 분석 결과 N-BaIoT와 Kitsune 간 이상 신호 강도 차이로 인한 도메인 비대칭성이 주요 원인으로 확인되었다.

본 연구는 통계적 특성이 다른 두 데이터 세트를 교차 분석하여 단일 데이터 세트 의존성을 완화하고, Stacking 구조가 미학습 도메인에서 개별 모델의 편향을 보정할 수 있음을 확인하였다. 또한 IP 카메라 보안 환경에서는 기기 및 트래픽 특성에 따른 유연한 모델 구성과 임계값 설정이 필요함을 보여주었다.

다만, 미묘한 이상 신호 기반 학습 모델이 강한 이상 신호 환경으로 전이될 때 성능이 저하되는 한계가 확인되었다. 향후에는 도메인 적응 기법과 XAI 분석을 적용하여 도메인 간 성능 격차를 완화하고 탐지 결과의 신뢰성과 활용성을 높일 계획이다.

REFERENCES

[1] D.S.Kim, While single-person households are away from home: How far has home security come?[Internet], https://www.dailypop.kr/news/articleView.html?idxno=65942.

[2] B.C.Won, Protecting my life and home on my own: The home CCTV boom[Internet], https://www.boannews.com/media/view.asp?idx=45655.

[3] Y.Nan, X.Wang, L.Xing, X.Liao, R.Wu, J.Wu, Y.Zhang and X.Wang, "Are You Spying on Me? Large-Scale Analysis on IoT Data Exposure through Companion Apps," in Proceedings of the 32nd USENIX Security

Symposium, Anaheim, CA, USA, 2023, pp.6664-6682.

[4] Korean National Police Agency, "Arrest of 4 suspects for hacking 120,000 IP cameras and selling stolen videos," Press Release, 2025.

[5] M.Choi, A.Cho and S.Lee, "Real-Time Network Attack Detection and APP Implementation for Wi-Fi Based IP Cameras Using Machine Learning," The Journal of Korean Institute of Communications and Information Sciences, Vol.50, No.8, pp.1279-1289, 2025.

[6] S.Y.Yeo, S.Y.Jo and J.Kim, "Machine Learning-based Detection of DoS and DRDoS Attacks in IoT Networks," Journal of The Korea Society of Computer and Information, Vol.27, No.7, pp.101-108, 2022.

[7] M.J.Hyun, "A comparative study of the performance of machine learning algorithms to detect malicious traffic in IoT networks," Journal of Digital Convergence, Vol.19, No.9, pp.463-468, 2021.

[8] Y.Meidan, M.Bohadana, Y.Mathov, Y.Mirsky, D.Breitenbacher, A.Shabtai and Y.Elovici, "N-BaIoT: Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders," IEEE PERVASIVE COMPUTING, Vol.17, No.3, pp.12-22, 2018.

[9] Y.Mirsky, T.Doitshman, Y.Elovici and A.Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection," in Proceedings of the 2018 Network and Distributed System Security Symposium (NDSS), San Diego, CA, USA, 2018.

[10] A.Rawashdeh, M.Alkasassbeh, M.Alauthman and M.Almseidin, "A stacked ensemble approach to identify internet of things network attacks through traffic analysis," Bulletin of Electrical Engineering and Informatics, Vol.13, No.6, pp.4316-4326, 2024.

[11] D.K.K.Reddy, H.Behera, G.M.S.Pratyusha and R.Karri, "Ensemble Bagging Approach for IoT Sensor Based Anomaly Detection," in Proceedings of the First International Conference on Intelligent Computing in Control and Communication, Singapore: Springer, 2021, pp.647-665.

[12] V.Sangeetha, R.C.A.Naidu, A.Bhat and P.Kulkarni, "Integrating Deep Learning with Ensemble Approach for Anomaly Detection in Network Traffic," in Proceedings of the 2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkuru, India, 2024, pp.1-5.

[13] M.Liu and L.Yang, "IoT Network Traffic Analysis with Deep Learning," in Proceedings of the 2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Biarritz, France, 2024, pp.184-189.

[14] L.Jaafouri, M.Et-Tolba, C.Hanin and I.Lahsen-Cherif, "IoT Network Security: Ensemble-Based Approaches for Anomaly Detection," in Proceedings of the 2025 International Conference on Circuit, Systems and Communication (ICCSC), Fez, Morocco, 2025, pp.1-6.

[15] S.Bebortta and S.Singh, "An Opportunistic Ensemble Learning Framework for Network Traffic Classification in IoT Environments," in Proceedings of the Seventh International Conference on Mathematics and Computing, Singapore: Springer, 2022, pp.473-484.

[16] O.Bukhari, P.Agarwal, D.Koundal and S.Zafar, "Anomaly detection using ensemble techniques for boosting the security of intrusion detection system," Procedia Computer Science, Vol.218, pp.1003-1013, 2023.

이 지 수(Jisoo Lee) [준회원]

- 2024년 3월 ~ 현재 : 수원대학교 정보보호학과 학사과정

〈관심분야〉
정보보호, AI 보안

전 상 훈 (Sanghoon Jeon) [정회원]

- 2012년 2월 : 경북대학교 IT대학 심화 전자공학 공학사
- 2014년 2월 : 대구경북과학기술원 정보통신융합공학전공 공학석사
- 2020년 8월 : 대구경북과학기술원 정보통신융합전공 공학박사
- 2020년 3월 ~ 2020년 8월 : 한양대학교 산학협력단 선임연구원
- 2020년 9월 ~ 2022년 9월 : 한양대학교 의과대학 응급의학과 포닥연구원
- 2022년 10월 ~ 2023년 9월 : 한양대학교 의과대학 응급의학과 연구조교수
- 2023년 10월 ~ 현재 : 수원대학교 지능형SW융합대학 정보보호학과 조교수

〈관심분야〉
웨어러블컴퓨팅, 의료인공지능, CPS보안