제로 트러스트 기반 보안모델 전개 전략에 관한 연구: 불법영상 탐지시스템 중심으로

이진용, 이용호, 민경호

한국정보통신기술협회

topjin55@tta.or.kr, abysskey@tta.or.kr, minkh0409@tta.or.kr

A Study on the Development Strategy of Security Model Based on Zero Trust: Focusing on Illegal Video Detection System

Jin Yong Lee, Yongho Lee, Kyung Ho Min

Telecommunications Technology Association

요 약

인공지능, 클라우드 및 쇼셜네트워크 플랫폼 기술의 발달로 고품질의 대용량 콘텐츠가 빠르게 유통될 수 있는 환경이 보편화되고 있다. 이와 같은 환경은 정보접근의 용이성, 풍부한 콘텐츠 제공 및 커뮤니케이션 증대 등 많은 긍정적 요소에 기여 하였으나, 불법 및 거짓 정보가 유통되는 부정적 측면 또한 무시할 수 없게 되었다.

이중 인공지능 기술과 결합된 불법 영상 촬영물, 딥페이크 및 딥보이스 등은 범죄 도구로 활용되어 금융자산 탈취, 사생활 침해, 사회적 혼란 등의 문제를 야기 시키고 있다. 이에 대한 대응을 위해 관련 업계에서는 불법 영상물 및 딥페이크 등에 대한 탐지기술 개발이 활발히 수행되고 있다. 그러나 이와 같은 탐지기술의 실제 운영을 위한 인프라 환경에 대한 위협분석 및 보안 설계 기준이 명확하지 않아 해커의 우회 공격수단으로 활용될 우려가 존재한다. 본 논문에서는 불법영상 탐지기술의 우회 및 신종 위협에 대응하기 위해 제로 트러스트 기반 보안 모델 전개 전략을 제언한다.

1. 서 론

인공지능, 클라우드 및 쇼셜네트워크 플랫폼 및 IoT 기술 발달 등으로 대용량의 고품질 콘텐츠가 언제 어디서나 다양한 형태로 제공되고 있다.

이와 같은 콘텐츠는 4차 산업혁명을 주도하며, 실생활을 윤택하게 하여 주는 긍적적 요소가 있는 반면 불법 및 거짓 정보를 기반으로 한 새로운 범죄의 도구로 이용되고 있다[1]. 콘텐츠를 이용한 범죄는 인공지능 기술과 결합하여 불법영상, 딥페이크 등으로 다양하게 활용되고 있으며, 이는 사회적 혼란부터 개인의 사생활 및 금전적 침해까지 다양한 형태로 나타나고 있다[2]. 이에 대응하기 위해 국가 및 관련 업계에서는 불법 영상 등을 탐지하기 위한 다양한 기술과 법·제도를 정비하고 있다[1, 2]. 그러나 탐지기술이 강화될수록 공격자는 새롭고 효과적인 취약점을 발견하기 위해 노력할 것이다. 불법 영상과 딥페이크 등은 사회적 혼란을 유도할 수 있고 여론을 조작할 수도 있다[2]. 특히, 여론 조작과 사회적 혼란은 전문 해커단체의 지능형·지속 공격을 수행될 수 있음을 예고한다[1, 3].

따라서 본 논문에서는 전문 해커단체 등이 불법영상 탐지기술을 우회하기 위해 수단으로 인프라 환경을 공격지점으로 활용하는 것에 대응하기 위한 수단을 강구하고자 하였다. 이때 기존의 경계 중심 보안모델의 한계를 개선하고 지능형·지속 공격에 효과적으로 방어할 수 있는 개념으로 최근 각광받고 있는 제로 트러스트 사상을 도입하는 것을 목표로 하였다. 이를 위해 불법영상 탐지시스템의 구조적 모델에서 파생되는 사이버 보안 위협을 제로 트러스트 기본원칙[4]에 의거하여 분석하고, 그 결과로 도출된 근거를 바탕으로 보안모델을 전개하는 전략을 제언하였다.

2. 관련 연구

2.1 제로 트러스트 사상

제로 트러스트는 모든 것을 신뢰하지 않는 다는 원칙을 기반으로 하는 보안 사상이다. NIST(미국 국립표준기술연구소)에서는 다음과 같이 제로 트러스트의 7개의 원칙을 제시하였다.

① 자원에 접근할 수 있는 모든 데이터와 서비스는 리소스로 식별한다.
② 네트워크는 위치와 무관하게 모든 통신은 보호된다.
③ 자원의 접근은 세션단위로 허가한다.
④ 자원의 접근은 동적 정책으로 결정된다.
⑤ 모든 자산의 무결성과 보안상태는 감시하여 관리된다.
⑥ 자원의 접근에 대한 인증/인가는 동적으로 수행된다.
⑦ 자산, 네트워크 등 자원으로 접근하는 모든 시나리오에 대해 가능한 많은 정보를 수집한다

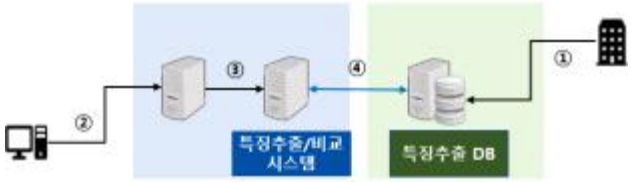
이와 같은 사상을 구현하기 위해 NIST에서는 정책을 수립하는 정책 결정포인트(PDP: Policy Decision Point)와 수립된 정책을 집행하는 정책집행포인트(PEP: Policy Enfocement Pont)를 기반으로 제로 트러스트의 아키텍처를 구성하였다[3, 4].

2.2 불법 영상탐지 시스템

(그림 1)은 불법영상 탐지시스템의 개념도로 국가 및 관련 기관으로부터 수집된 불법 영상의 특징이 특징추출 DB에 저장되어 이용자가 업로드한 불법 영상과 상호 검증될 수 있도록 불법영상 특징추출/비교 시스템을 구성하도록 되어 있다[5].

3. 제로 트러스트 보안모델 전개전략

본 논문에서는 불법영상 탐지시스템의 인프라 및 서비스 구성의 특성을 기반으로 사이버 보안 위협을 분석하고 제로 트러스트 기반 보안모델 전개 전략을 제시한다.



(그림 1) 불법영상 탐지시스템 개념도

3.1 불법영상 탐지시스템의 사이버 보안 위협분석

(그림 1)에서 각 구간별 특징에 따라 다음과 같은 사이버 보안 위협을 식별할 수 있다.

① 구간은 불법영상 특징추출 DB가 수집·저장되는 구간으로 관리주체가 악성코드에 감염될 경우로 탐지 결과가 왜곡될 수 있는 위협을 내포하고 있다.

② 구간은 영상이 업로드되는 구간으로 이용자가 비정상 파일 및 악성코드를 고의·비고의적으로 업로드할 경우 추가 공격으로 활용될 수 있다.

③ ~ ④ 구간은 이용자로부터 받은 영상정보를 특징추출/비교 시스템으로 전송하고, 특징추출 DB와 비교하여 최종 결과를 도출하는 구간으로 파라미터 변조를 통한 결괏값 왜곡, 내부자 일탈행위 등으로 의도하지 않은 결과를 발생시킬 수 있다.

<표 1>은 위에서 언급된 불법영상 탐지시스템 구성요소의 각 구간별 발생할 수 있는 위협을 NIST의 제로 트러스트 7개 원칙을 기반으로 재구성한 내용이다.

<표 1> 제로 트러스트 원칙에 기반한 불법영상 탐지시스템의 사이버 위협분석

원칙	위협구간	위협항목	위협분석
(1)리소스	①, ②	악성코드	악성코드 감염에 따른 시스템 위·변조
(2)네트워크	①~④	위·변조	결괏값 전송의 파라미터 변조
(3)접근통제	③, ④	과다허용	비인가자의 접근
(4)정책관리	④	내부위반	내부자의 일탈행위
(5)모니터링	①~④	운영미흡	로그 기록·분석 누락·미흡
(6)인증/인가	①	정적인가	계정 탈취 및 비인가자의 접근
(7)정보수집	①~④	수집범위	핵심 수집 대상·항목 누락·미흡

3.2 제로 트러스트 기반 보안모델 구축 전략

(그림 2)는 3.1 불법영상 탐지시스템의 사이버 보안 위협분석에 기반하여 구성한 제로 트러스트 기반 보안모델의 개념도이다.

데이터 영역은 실제 서비스가 수행되는 곳을 통제하기 위한 영역이며, 정책집행포인트가 구성된다. 컨트롤 영역은 데이터 영역을 통제하기 위한 정책을 문맥적 분석을 통해 동적으로 생성하는 영역으로 정책결정포인트가 구성된다. 불법영상 탐지시스템의 제로 트러스트 전개 전략은 <표1>의 7개의 원칙과 (그림 2)의 각 주체별 항목에 따라 다음과 같이 구성된다.

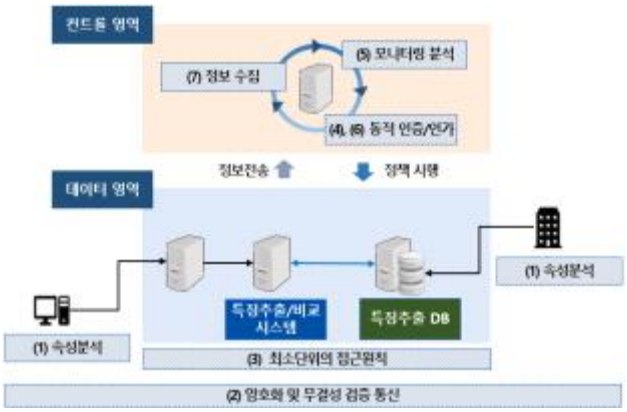
(1) 이용자 및 특징추출 DB를 구현하는 접근주체이다 접근대상이 되며, 주체의 보안설정을 검증하는 정적 속성과 위반적 행위를 식별할 수 있는 동적 속성에 기반한 속성분석 환경을 구성한다.

(2) 통신구간은 암호화 전송을 원칙으로 하되, 복호화 구간의 파라미터 위·변조에 대응하기 위해 무결성을 검증할 수 있는 파라미터 검증 값(해쉬 값, 타임스탬프값 등)을 포함하여 통신할 수 있어야 한다.

(3) 각 시스템의 접근구간은 세션 값을 바탕으로 최소접근 원칙을 준수할 수 있도록 구성한다.

(4), (6) 내부자의 일탈행위와 계정 탈취 등을 고려하여 (1)에서 명시된 속성분석에 기반한 동적 접근 및 인가 정책을 적용한다.

(5), (7) 정보는 정형·비정형의 수집되어야 하며, 분석은 시나리오 기반의 문맥적 분석을 수행하여야 하며, 이와 같이 분석된 결과가 동적 인증/인가 정책으로 데이터 영역에서 실행될 수 있어야 한다.



(그림 2) 제로 트러스트 기반 불법영상 탐지시스템 보안모델 전개 적략 개념도

4. 결론

불법영상 촬영물, 딥페이크 등이 범죄의 도구로 활용될 경우 국가, 경제, 개인에 발생하는 피해의 파급력은 넓고 깊다. 이에 따라 불법 영상을 탐지하기 위한 핵심기술들이 활발히 개발되고 있다. 그러나 이와 같은 방어기술이 진보될수록 공격자는 새로운 우회경로를 찾기 위해 노력할 것이다.

본 논문에서는 사회적 혼란 등을 목표로하는 고도의 해커단체가 불법영상 탐지시스템의 인프라 구간별 존재하는 각종 취약점을 탐지기술을 우회하기 위한 수단으로 활용할 것에 중점을 두었다. 따라서, 전문 해커단체의 지능형·지속공격에 대응하기 위해 제로 트러스트의 사상적 개념을 실제 시스템 아키텍처로 전개하기 위한 연구를 수행하였다. 이를 위해 제로 트러스트 기본원칙이 입각한 사이버 보안 위협을 도출하고 이를 방어하기 위한 전략적 개념의 모델을 제언하였다. 향후 불법영상 탐지시스템의 다양한 시나리오의 추가 위협을 분석하고, 이에 대한 타당성 및 신뢰성을 검증할 경우 보다 실효성 있는 모델로 구체화될 수 있을 것으로 기대된다.

참 고 문 헌

[1] J. H. Kim, “A Study on the Dangers of Illegal Filming and Counterplan”, The Institute of Legal Studies, Konkuk Univ, Vol.48, pp.439-464, 2021

[2] T. R. M, M. Arunachalam, R. R, K. Rammaraj, M. Arunachalam, K. P, “A Review on the Detection of Deep Fake and Propaganda Videos and Images-based Voice and Facial Manipulation using AI Techniques”, IEEE ICACRS-2023, pp.1083-1087, 2023

[3] J. Y. Lee, B. H. Choi, N. Koh, S. Chun, "A Study on How to Build a Zero Trust Security Model," KIPS Transactions on Computer and Communication Systems, Vol.12, No.6, pp.189-196, 2023.

[4] Scott Rose, Oliver Borchert, Stu Mitchell, Sean Connelly, “NIST Special Publication 800-207, Zero Trust Architecture”, National Institute of Standards and Technology, 2020.

[5] etnews, https://www.etnews.com/20210707000089, retrieved at May 6, 2024