JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 31 No. 11, pp. 299-309, November 2026
https://doi.org/10.9708/jksci.2026.31.11.299

A Structure for Identifying and Controlling Unapproved External Connections of AI Development Tools in the Financial Sector

Seungman Jin*, Deukjo Hong**

*Ph.D. course, Div. of Electronics and Information Eng. (Computer Eng.), Jeonbuk National Univ., Jeonju, Korea
**Professor, Dept. of Computer Science and Artificial Intelligence, Jeonbuk National Univ., Jeonju, Korea

[Abstract]

In this paper, we propose an observation and decision structure for identifying unapproved external connections of AI development tools and tool-integration protocols such as the Model Context Protocol (MCP) in financial-sector development environments without SSL/TLS decryption, combining endpoint process attribution with a host-specific approval assessment outside the endpoint for each connection. The structure collects information on the host and child processes that created a connection at the endpoint, assesses destination approval at a policy point outside the endpoint, and joins the two records to determine, using a non-compensatory decision matrix in which approval conditions take precedence, whether the destination falls within the host's approved scope. Thus, even a connection created by a registered executable image is identified as outside the approval scope if its destination is not approved under the applicable policy, and each decision is retained with the applied policy version and raw records so that it can be reconstructed for audit, ensuring audit traceability. In addition, a five-step extension procedure allows the structure to be updated as new threats emerge and observation conditions change.

▸ **Key words**: Financial AI Security, Model Context Protocol, Unapproved Connection, Process Attribution, Host-Specific Approval

[요 약]

본 논문에서는 SSL/TLS 복호화를 적용하지 않는 금융권 개발 환경에서 AI 개발도구와 도구 연동 프로토콜(MCP 등)의 미승인 외부 연결을 식별하기 위해, 단말의 프로세스 식별 정보와 단말 외부의 호스트별 승인 대조를 동일 연결 단위로 결합하는 관측 및 판정 구조를 제안하였다. 본 제안 구조는 연결을 만든 호스트와 하위 프로세스의 정보를 단말에서 수집하고, 목적지의 승인 여부는 단말 외부의 정책 지점에서 대조한 뒤 두 기록을 결합하여 승인 조건이 우선하는 비보상적 판정표를 이용하여 목적지가 해당 호스트의 승인 범위에 포함되는지를 판단하는 것으로, 등록된 실행 이미지가 만든 연결이라도 목적지가 적용 정책의 승인 범위 밖이면 승인 범위 밖 연결로 식별되며 각 판정은 적용 정책 버전과 원시 기록을 함께 남겨 감사 시 재구성할 수 있도록 하여 감사 추적성을 확보하고 5단계 확장 절차를 통해 신규 위협과 관측 조건의 변화에 맞추어 구조를 갱신할 수 있도록 하였다.

▸ **주제어**: 금융 AI 보안, MCP, 미승인 연결, 프로세스 식별, 호스트별 승인

• First Author: Seungman Jin, Corresponding Author: Deukjo Hong
*Seungman Jin (smjinking@naver.com), Div. of Electronics and Information Eng. (Computer Eng.), Jeonbuk National Univ.
**Deukjo Hong (deukjo.hong@jbnu.ac.kr), Dept. of Computer Science and Artificial Intelligence, Jeonbuk National Univ.
• Received: 2026. 10. 06, Revised: 2026. 10. 19, Accepted: 2026. 10. 29.

Copyright © 2026 The Korea Society of Computer and Information
http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945
www.kci.go.kr

I. Introduction

생성형 AI 개발도구는 코드의 작성, 수정 및 리뷰를 지원하고, 도구 연동 프로토콜을 통해 파일시스템, 데이터베이스, 이슈트래커 등의 기능을 제공하는 서버와 연결된다[1]. 개발 단말에서 이러한 기능을 이용할 때 조직이 관리할 대상은 도구의 설치 여부만이 아니라 실행 중 발생하는 외부 연결까지 넓어진다. 등록된 개발도구나 하위 프로그램이라도 연결 목적지가 해당 호스트의 승인 범위에 포함되는지는 별도로 확인해야 한다. 본 논문은 이 외부 연결을 관측하고 조직의 승인 정책과 대조하는 문제를 다룬다.

조직의 대응에는 무엇을 먼저 처리할 것인가와 어떤 근거로 관측하고 판정할 것인가라는 두 문제가 있다. 선행연구[2]가 위험 우선순위 모델과 탐지-통제 시차를 다루었다면, 본 논문은 그 우선순위를 다시 산정하지 않고 연결의 발생 주체와 승인 상태를 확인할 근거 및 판정 절차를 구체화한다. 가중합에서 한 요인이 다른 요인의 부재를 상쇄하는 문제는 승인 판정 단계의 설계 과제로 다루고, 관측 근거를 탐지가능성(D) 갱신에 연결하는 방안은 후속 검증 대상으로 둔다.

망분리 개선과 전자금융감독 문헌[3][4], 금융 AI 보안 가이드[5]를 배경으로, 본 연구는 SSL/TLS 복호화를 적용하지 않는 금융 개발 구간을 대상으로 한다. 이 조건에서는 게이트웨이 기록의 가시적 목적지 메타데이터가 연결 대상을 알려 주더라도 요청 내용과 연결 생성 프로세스를 설명하지 못하고, 단말 기록은 프로세스와 계보를 제공하지만 조직의 호스트별 목적지 승인 근거를 갖고 있지 않다. 본 연구는 서로 다른 보안제품의 기록을 별도로 결합해야 하는 이러한 환경을 대상으로 하며, 금융권에서 복호화가 적용되지 않는 구간의 예는 3.1절에서, 금융권 조건이 설계에 반영된 방식은 3.5절에서 정리한다.

본 논문에서 관측 계층은 OSI 계층이 아니라 판정 정보를 제공하는 관측 지점이며, 책임계층은 해당 항목의 판정 기준값을 제공하는 주체이다. 단말 E는 연결이 호스트 또는 하위 프로세스에서 만들어졌는지 식별하고 이미지 등록 상태를 확인하며, 단말 외부의 S는 게이트웨이 목적지 정보를 호스트별 승인 목록과 대조하고 정책 버전 및 유효 시점을 남기고, 패킷 N은 필요할 때 조사를 보조한다. 실행을 위한 이미지가 등록되어 있다는 사실만으로 그 프로그램이 만드는 모든 외부 연결까지 승인되었다는 뜻은 아니므로, 본 구조는 목적지가 승인 목록에 없는 연결을 승인 범위 밖으로 판정하고 다른 정상 단서나 패킷 단서가 목적지의 승인 근거를 대신하지 않도록 조건을 구분한다.

기존 통합 XDR은 벤더 플랫폼 안에서 프로세스와 목적지를 상관하지만, 해당 목적지가 그 호스트에 승인되었는지를 판단할 조직 고유의 기준과 그 버전, 판정의 감사 책임까지 제공하지는 않는다(2.4절). CASB는 발견된 앱 단위로 승인 상태를 관리한다. 이에 비해 본 연구는 승인의 단위를 앱이 아닌 호스트별 외부 목적지 범위로 두고 그 정책의 관리 주체와 적용 버전을 명시하며, 서로 다른 제품의 기록을 동일 연결에 대응시킬 때의 결합 조건과 결손 처리를 정의하고, 점수의 합이 아니라 승인 조건이 우선하는 판정표로 결과를 정한다는 점에서 구별된다.

본 연구의 기여는 다음과 같다. 첫째, 관리 대상 호스트의 외부 연결에 필요한 정보 항목을 구분하고, 연결 식별과 실행 이미지 등록 확인, 목적지 승인 대조에 대한 책임과 기준 정보의 출처를 명세한다. 둘째, 동일한 연결에 대응시킨 사실을 승인 조건 우선의 명시적 조건 행렬로 결합하여 판정 결과와 실제 트래픽에 대한 조치를 구분한다. 셋째, 신규 항목과 관측원을 수용하는 5단계 확장 절차를 기록 결합, 중복 점검, 오탐 및 운영 검증 계획과 연결한다.

적용 예시의 추적 기록은 실제 로그 필드 구조에 따라 구성한 가상 기록이며, 실증 평가는 후속 연구로 둔다. II장은 관련 연구, III장은 구조적 공백 분석, IV장은 제안 구조와 판정 규칙, V장은 적용 예시, 검증 계획 및 논의, VI장은 결론이다.

II. Background and Related Work

본 장에서는 도구 연동 프로토콜의 구조와 전송 특성, AI 관련 위협 환경의 변화, 도구 연동 보안 연구와 암호화 트래픽 및 다계층 상관 연구를 검토한다.

2.1 Structure and Transport of Tool-Integration Protocols

도구 연동 프로토콜은 호스트와 도구 서버의 메시지 및 통신 절차를 정의한다. 대표 사례인 MCP(Model Context Protocol)는 JSON-RPC 2.0을 사용하며 로컬 표준 입출력인 stdio와 원격 전송인 Streamable HTTP를 제공한다[1]. MCP의 호스트는 클라이언트를 관리하는 AI 애플리케이션으로 본 논문에서는 개발도구 실행 프로세스에 대응한다[1]. 원격 TLS 연결은 복호화 없이는 HTTP 경로나 호출 내용을 확인할 수 없고, 동일 호스트가 도구 연동과 일반 API를 함께 제공하면 목적지 정보만으로 두 행위를 구분할 수 없다.

www.kci.go.kr

본 논문은 MCP 사용 여부를 직접 식별하는 방법이 아니라, MCP를 포함한 AI 도구 연동 환경에서 관리 대상 호스트의 외부 연결이 승인 범위에 포함되는지를 판정하는 구조를 다룬다. 실제 도구 사용 여부를 확인하려면 런타임 기록이나 설정 감사 정보가 추가로 필요하며, 전송과 세션의 식별 및 초기화 조건이나 연결 재사용 방식이 바뀌면 관측 전제를 재검토한다.

2.2 Evolution of the Threat Landscape

사용자가 미승인 도구를 선택하는 Shadow IT 문제는 AI 에이전트의 탐색과 실행 문제로 확장되며, OWASP의 에이전트 AI 위협 분류[6]와 MITRE ATLAS[7]가 이를 검토할 배경을 제공한다. 위협의 변화가 통제 갱신보다 빠르면 선행 연구[2]의 탐지-통제 시차가 발생한다. 따라서 고정 목록을 갱신하는 것뿐 아니라 실행 런타임처럼 새로운 정보를 제공하는 관측원을 편입하고, 항목별 책임, 기록 결합, 동작 검증을 함께 갱신해야 한다.

2.3 Tool-Integration Security Research

도구 연동 보안 연구는 위협 분석과 평가 방법을 구체화해 왔다. 공급망과 도구 오염의 위협 지형이 체계화되고[8] 공격 및 구현 평가 벤치마크[9]가 제시되었으며, 분석 범위는 주변 구성요소의 공격 벡터[10]와 명세 취약점[11]으로 넓어졌다. 자동 레드팀[12]과 다단계 방어[13]는 평가와 방어의 방향을 제시한다. 이들 연구가 도구와 실행 환경 내부의 통제를 다룬다면, 본 논문은 조직의 보안 기록으로 외부 연결의 발생 주체와 승인 여부를 확인하는 문제를 다룬다.

2.4 Encrypted Traffic Analysis and Multi-Layer Correlation

암호화 트래픽을 분류하는 기존의 연구[14]와 다계층 탐지[15][16], EDR[17], 제로 트러스트[18]는 서로 다른 특징과 출처를 결합한다. 통합 XDR은 벤더 플랫폼 안에서 연결 생성 프로세스와 목적지 필드를 함께 제공한다[19]. 다만 그 목적지의 승인 여부를 판단할 기준과 버전은 조직이 별도로 정해야 한다. CASB는 발견된 앱의 승인 상태를 관리하고 연동된 통제 수단으로 사용을 제한하며[20][21], 앱별 방화벽은 실행 경로나 PolicyAppId 태그로 통제한다[22]. 본 연구는 이기종 기록의 상관 기능 자체를 기여로 주장하지 않는다. 이는 상관 자체가 SIEM이나 통합 XDR로도 구현될 수 있기 때문이다. 본 연구가 명세하는 것은 상관된 기록을 판정 근거로 쓰기 위한 결합 조건(식별 우선, 결합 보류, 결손 보고)과, 그 판정에 적용할 승인 기준 및 버전이다(Table 1).

Table 1(a). Related approaches compared with this design

Approach/Condition	Available basis / reflection in the design	What this design addresses / section
Alert correlation / SIEM [16]	Alerts and multi-source event relationships	Item-level responsibility and evidence linkage for the same connection
EDR / integrated XDR [17][19]	Initiating process and remote endpoint within one vendor platform	Host-level approval basis, policy version and audit responsibility
CASB app governance [20][21]	App-level sanctioned/unsanctioned tags and integrated enforcement	Approval at host-level external-destination scope rather than app level
Host-based per-application firewall [22]	App-scoped local enforcement; central GPO/CSP management	Decision audit with external policy version, separated from enforcement
Proposed heterogeneous-layer design	E host attribution; S host-specific approval; optional N	Priority-ordered non-compensatory matrix; auditable five-step extension

Table 1(b). Financial-sector conditions (C1-C4 in Section 3.5) reflected in the design

Approach/Condition	Available basis / reflection in the design	What this design addresses / section
C1: Decryption absent or restricted in some segments (certificate-pinning exceptions; development networks newly opened under network-separation reform [3][4])	Decision without payload: host attribution plus destination approval; general-function entries to prevent false positives; FQDN-based approval units; packet features as annotations only	3.1, 3.2
C2: Layer-specific security products from different vendors operated in parallel	Cross-product join by 5-tuple and UTC time; explicit gap and non-estimation reporting	4.2
C3: The approving authority sits outside the endpoint and must justify approval to auditors	S checks destination approval (non-compensatory: no other cue can substitute)	4.1, 4.3
C4: Supervisory inspection requires reproducible evidence; blanket blocking of development traffic is operationally difficult	Policy version, validity time and raw records retained per decision; detection-only operation; D6 is not a normality guarantee	4.3, 5.3

www.kci.go.kr

III. Structural Gap in Observation and Decision

본 장에서는 외부 연결의 발생 주체와 목적지 승인 정보가 서로 다른 관측 지점에 분산될 때의 공백을 분석하고, 적용 범위, 위협 유형, 금융권 조건이 설계에 반영된 방식을 정리한다.

3.1 Problem Statement and Scope

2장의 연구들은 프로토콜 내부의 취약점을 분석하거나 여러 보안 기록을 결합하는 방법을 다루어 왔다. 금융회사의 보안 담당자가 실제로 답해야 하는 질문은 "관리 대상 개발도구가 만든 이 외부 연결은 조직이 허용한 범위 안에 있는가?"이다. 복호화를 적용하지 않은 게이트웨이는 목적지는 알지만 연결을 만든 프로그램을 알지 못하고, 단말 기록은 프로그램과 계보를 보여 주지만 그 목적지가 허용된 곳인지 판단할 조직의 기준은 갖고 있지 않다. 따라서 두 기록이 같은 연결을 가리키는지 확인한 뒤 그 호스트에 적용되는 승인 기준과 대조해야 판단이 가능하다.

대상은 조직이 관리 대상으로 지정한 AI 개발도구 호스트와, 호스트가 실행한 로컬 도구 서버를 포함한 하위 프로세스가 만드는 외부 연결이다. 분석 환경은 SSL/TLS 복호화를 적용하지 않으며 단말 관측 기록과 경계 게이트웨이 기록을 확보할 수 있다고 전제한다. 경계 게이트웨이를 거치지 않는 내부 연결은 기본 범위에서 제외하고, 이를 포함하는 확장 구성은 4.4절에서 다룬다.

금융권에서 복호화가 적용되지 않는 구간은 다음과 같은 경우에 나타난다. 첫째, 인증서 고정(pinning)이나 자체 인증서 저장소를 사용하는 개발도구, 패키지 관리자, API 클라이언트는 중간 복호화 시 동작이 실패하므로 검사 예외로 등록되는 경우가 있다. 둘째, 망분리 규제 개선[3][4]에 따라 외부 연결이 새로 허용된 연구개발망 구간에 복호화 장비가 아직 배치되지 않았거나 처리 용량이 부족한 경우다. 셋째, 개인정보나 인증정보의 노출 우려 또는 내부 정책에 따라 특정 범주의 트래픽을 복호화 대상에서 제외하는 경우다. 여기에 암호화된 Client Hello[23]나 DNS over HTTPS[24]가 쓰이면 복호화 미적용 구간에서 관측할 수 있는 목적지 이름 정보가 더 줄어든다.

판단의 기준은 조직이 미리 준비하는 호스트별 외부 연결 승인 목록이다. 각 항목에는 호스트 정책 키, 관측 가능한 FQDN 또는 IP, 용도, 승인 상태, 버전과 유효 시점이 들어간다. 호스트 정책 키는 호스트 유형이나 이름처럼 오래 유지되는 관리 식별자로 정하고, 실행할 때마다 바뀌는 프로세스 번호는 쓰지 않는다. 용도에는 도구 연동 목적지뿐 아니라 LLM API, 업데이트, 텔레메트리, 패키지 수신처럼 평소 사용하는 일반 기능도 함께 적어, 정상적인 업데이트나 패키지 다운로드를 미승인 연결로 잘못 판단하는 일을 줄인다. 목록이 완전하다는 것은 인터넷의 모든 목적지를 미리 수집했다는 뜻이 아니라, 해당 호스트의 외부 연결에 어떤 승인 항목과 정책 범위를 적용할지 조직이 정해 두었다는 뜻이다. 목록은 차단 정책에도 사후 모니터링에도 쓸 수 있으므로 네트워크 전체에 허용 목록 방식의 방화벽을 적용할 필요는 없다. 실행 이미지 등록 여부는 별도의 이미지 등록 목록으로 관리한다.

승인 항목의 목적지 식별자는 FQDN 또는 IP로 기재되지만, CNAME 체이닝, CDN의 가변 IP, DNS 응답 시점과 연결 생성 시점의 차이 때문에 관측된 식별자가 승인 항목과 같은 대상을 가리킨다고 단정하기 어려운 경우가 있다. 이때 본 구조는 이를 승인으로 간주하지 않고, 이름 해석 결과를 연결 시점과 대응시킬 수 없으면 UNKNOWN으로 두어 판정을 보류한다. 승인 단위는 원칙적으로 FQDN으로 두되, 고정 IP로만 제공되는 목적지는 IP 항목으로 승인하며, 관측 시점의 이름 해석 결과와 그 출처를 함께 보존하는 운영을 전제한다.

S는 E가 식별한 호스트의 정책 키와 연결에서 식별된 목적지를 연결 시점에 유효한 버전의 목록에 대조하여 목적지 상태를 네 가지로 구분한다. 적용 정책이 명시적으로 금지한 목적지는 비인가, 도구 연동 또는 일반 기능 항목에 포함된 목적지는 승인이다. 게이트웨이가 연결을 통과시켰지만 확정된 정책 범위의 어떤 승인 항목에도 없는 목적지로의 연결은 '승인 범위 밖 목적지 연결'이라 부르며, 조직 정책에 따라 통제하거나 모니터링한다. 게이트웨이의 트래픽 허용과 목적지 승인은 서로 다른 사실이다. 목적지를 식별할 수 없거나 호스트와 목적지의 대응 또는 적용 정책이 없거나 불완전하면 미분류(UNKNOWN)로 둔다. 다만 처음 관측된 목적지라는 이유만으로 미분류로 두지 않으며, 정책 범위가 확정되어 있으면 처음 보는 목적지도 승인 범위 밖으로 판단한다.

3.2 Observation Limits without Decryption

승인 여부를 판단하려면 각 관측 지점에서 확보할 수 있는 정보와 확인하기 어려운 정보를 구분해야 한다. 단말 E는 연결을 만든 프로세스와 호스트 계보를 제공하고[25][26], 단말 외부의 S는 게이트웨이 기록을 해당 호스트의 승인 목록과 대조하며[27], 필요하면 패킷 N의 관측 정보로 조사를 보완한다. Table 2는 복호화를 적용하지 않

www.kci.go.kr

Table 2. Layer-wise observability and decision responsibility without decryption

Item	Endpoint(E)	Gateway(S)	Packet(N)	Note
Destination hostname	Partial	Conditional	Conditional	Depends on DNS/SNI visibility; distinguish encrypted DNS [24], hidden SNI [23] and direct IP connections
Destination authorization	Unavailable in base configuration	Responsible layer	Unavailable	S joins the attributed host policy key and observable FQDN/IP to a versioned host-specific approval list
Session volume and duration	Partial	Responsible layer	Observable	Directional counts and duration; supplementary investigation context
Packet inter-arrival times; TLS fingerprint	Unavailable in base logs	Partial (TLS fingerprint only)	Responsible layer	Actual packet timestamps; packet intervals are not tool-call intervals. A visible handshake is required for the fingerprint; both are investigation aids [28].
Connection-creating process	Responsible layer	Unavailable	Unavailable	Host or child attribution from E records; distinct from image-registration signal

는 기본 구성에서 각 정보의 관측 가능성과 판정 책임을 정리한 것이다.

3.3 Structural Gap in Local-Process Transport

원격 연결 관측만으로는 개발도구가 로컬에서 실행하는 도구 서버의 동작을 충분히 설명하기 어렵다. stdio 통신 자체는 네트워크에 나타나지 않으므로 본 구조의 범위 밖에 두지만, 로컬 서버가 외부 API를 호출하거나 침해되어 데이터를 외부로 보내는 연결은 관측할 수 있다. 로컬 서버의 실행 이미지가 미등록이면 호스트 계보에서 단말 신호(σE, 부록 A)가 성립하고, 등록된 이미지라면 σE=0이지만 연결을 만든 프로세스와 호스트 계보는 확인되므로 목적지가 승인 범위 밖이면 4.3절의 판정 규칙에 따라 승인 범위 밖 연결로 판정된다. 즉 로컬 서버의 외부 연결은 이미지 등록 여부와 관계없이 목적지 승인 대조의 대상이다.

3.4 Threat Model

금융권 개발환경의 위협을 연결의 발생 주체와 목적지 승인 상태에 따라 구분하면 Table 3과 같다. 본 구조의 주된 적용 대상은 관리 대상 호스트의 외부 연결에 해당하는 유형 2~4이다. 유형 2와 3은 연결의 발생 주체와 목적지를 확인하고 호스트별 승인 목록과 대조하여 승인 범위 밖 연결을 판정하며, 유형 4는 반복되는 판정 결과로 조사 우선순위를 높인다. 각 유형은 선행 연구[2]의 후보 위협과 대체로 대응한다(유형 1은 T1, 유형 2는 T2, 유형 3은 T3과 T5′, 유형 4는 T4). 유형 1은 기존 프록시 및 DLP 통제와, 유형 5는 서비스 감사 기록과 함께 다룬다. 조직이 업무 시스템을 도구 서버로 노출하여 AI 에이전트가 호출하는 반대 방향의 경우는 유출이 아니라 호출자의 권한과 위임 범위의 문제(T4 영역)이며, 도구 이름, 호출 인자, 호출자 신원은 MCP 서버의 감사 로그를 권위 출처로 삼는다.

3.5 Financial-Sector Conditions Reflected in the Design

제안 구조의 적용 대상을 금융권으로 둔 것은 기술 자체가 금융권에만 적용되기 때문이 아니라, 설계 선택이 금융권의 운영 조건에서 도출되었기 때문이다. 이 구조가 필요한 환경은 다음 네 조건(C1~C4)을 동시에 만족한다. 첫째

Table 3. Threat types classified by observation characteristics

Type	Scenario	Available evidence	Decision implications
1	Source code entered directly into a web-based generative AI service (no tool integration)	Destination decision only; endpoint signal not applicable	Outside the main scope; destination approval alone does not assess submitted content
2	Connection to an unapproved remote tool server for development	Host attribution and destination comparison; σE depends on image registration; packet clues optional	D3 when the attributed destination is outside the host approval scope
3	Exfiltration via a compromised local tool server	Registered image: σE=0, yet attribution and destination comparison apply	D3 for a destination outside the host approval scope; content sent to approved destinations needs additional evidence
4	Autonomous AI probing and attempting tool-integration paths	D3 or D5 repeated for the same host policy key and destination	Investigation priority rises; the underlying decision is unchanged
5	Tool integration inside an approved SaaS (e.g., Microsoft 365 Copilot)	External connection records only; service-internal actions not visible	Service audit records required

www.kci.go.kr

(C1), 복호화가 적용되지 않거나 제한되는 구간이 존재한다. 둘째(C2), 관측 지점이 서로 다른 제품에 분산되어 있다. 셋째(C3), 목적지 승인을 판단할 권한이 단말 외부에 있다. 넷째(C4), 판정 근거의 재현과 감사가 요구된다. 금융권은 규제와 감독 요구가 이 조건을 강화하는 대표 사례이며, 같은 조건을 가진 공공, 의료, 제조 구간에도 적용할 수 있다. Table 1(b)는 이 조건들이 설계에 반영된 방식과, 이와 함께 고려한 운영상 제약을 정리한 것이다. 일반 기업의 개발망에 적용하는 경우 정책 버전 관리와 판정 재현성 요건은 조직의 감사 수준에 맞추어 조정할 수 있다.

IV. Proposed Observation and Decision Structure

본 장에서는 분산된 정보를 동일한 연결에 대응시키고 해당 호스트의 승인 범위와 대조하여 판정하는 구조를 제안한다.

4.1 Who Is Responsible for Each Item

앞 장에서 확인한 정보 분산을 보완하려면 서로 다른 기록을 모으는 것뿐 아니라 각 정보의 판정 기준을 어느 관측 지점에서 제공할지 정해야 한다. 본 구조는 동일한 정보를 중복 반영하거나 연결의 발생 주체를 잘못 판단하지 않도록 다음 다섯 원칙을 적용한다.

- 책임 분리: 항목마다 기준값을 제공할 책임계층 하나를 지정하고 중복 관측은 보조 대조로 둔다.
- 식별 우선: 프로세스 생성 기록과 연결 기록으로 호스트 또는 하위 프로세스를 확인한다. 등록 여부는 별도이며 IP 및 시각 근접성만으로 판단하지 않는다.
- 시간축 결합: 출발지 포트를 포함한 5-튜플과 UTC 시각으로 대응시키고 동기화 상태에 따른 허용 오차를 둔다.
- 결손 보고: 결과 미매칭, 복수의 후보, 미산출 이유를 원시 근거와 함께 보존한다.
- 확장 개방성: 신규 항목과 관측원을 추가할 때 책임, 결합, 운영 영향을 재검증한다.

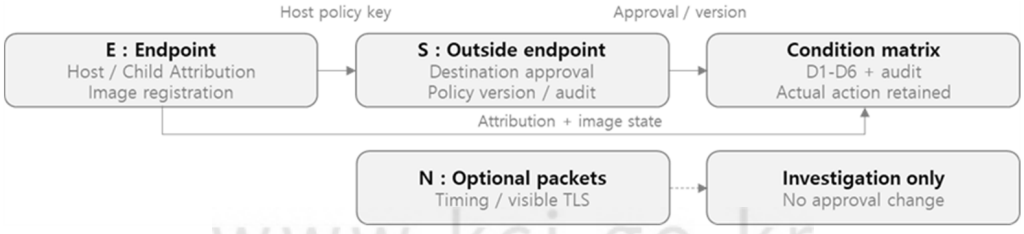
E는 호스트 식별과 이미지 등록을, S는 호스트별 목적지 승인 대조를 책임진다. 원시 기록과 승인 목록, 적용 정책 및 신호 산출 기준의 버전을 남겨 동일 연결의 판정을 재현할 수 있도록 한다. Fig. 1은 단말 E의 식별 정보를 S의 목적지 승인 대조에 연결하여 판정하고, N의 관측 정보는 후속 조사를 보조하는 흐름을 보여 준다.

4.2 How Records Are Combined

정보별 책임을 배정한 다음에는 각 관측 지점의 기록이 동일한 연결에 관한 것인지 확인해야 한다. 다른 연결의 프로세스 정보와 승인 정보를 결합하면 잘못된 판정으로 이어지기 때문이다. 먼저 단말의 연결 기록을 프로세스 생성 기록과 대조하여 연결을 만든 프로세스를 확인하고, 5-튜플과 기록 시각으로 같은 연결의 S 세션을 찾는다. 기록 시각의 차이는 ±Δ까지 허용하고, NAT, 컨테이너, 프록시 환경에서 대응 후보가 여러 개이거나 변환 전후의 연결을 확인할 기록이 없으면 결합을 보류하며, 차단된 연결 시도는 완료된 세션과 구분하여 기록한다. 각 계층의 신호는 Table 4와 같이 정의하고 기호는 부록 A에 정리하였다. 신호를 산출하지 못한 경우에는 그 이유를 기록하여 신호가 성립하지 않은 경우와 구분한다.

구현 예시로 Sysmon을 사용하는 경우, 프로세스 생성 기록(Event 1)의 실행 파일, 명령행, 부모 프로세스 정보와 프로세스 식별자를 이용해 호스트와 하위 프로세스의 생성 관계를 확인한다. 이후 연결 기록(Event 3)의 프로세스 식별자를 생성 기록과 대조하여 연결을 발생시킨 프로세스를 확인하고, 5-튜플과 발생 시각을 이용해 게이트웨이의 대응 기록을 찾는다[25].

다른 단말 기록을 사용하는 경우에는 동일한 식별과 결합에 필요한 정보를 제공하는지 확인한다[26]. 게이트웨이 기록에서는 적용 규칙과 허용 및 차단 결과를 확인하고[27], 관측 가능한 목적지 정보를 해당 호스트의 외부 연결 승인 목록과 대조한다. 패킷 계층의 조사 단서는 실제 패킷 관측 시각을 이용해 산출한다.



www.kci.go.kr

Fig. 1. Endpoint attribution and host-specific approval decisions with optional packet annotations

Table 4. General layer signals and their meanings

Signal / Layer	General condition and meaning
σE / E	1 when an unregistered child executable image creates an external connection in the managed host lineage; 0 when registered. Preserve image-list version and attribution separately.
σT / S	1 when the identifiable destination is unapproved or prohibited under the applicable host policy; 0 when approved; UNKNOWN when identity, host mapping or policy scope is unavailable.
σδ, σF / N	Optional investigation context: packet-interval criterion $\delta \geq \tau\delta$ under declared baseline, estimator and minimum-sample conditions, and visible TLS-fingerprint or reputation criterion with visibility and source/version. Record non-estimation separately.

4.3 Explicit Decisions and Investigation Annotations

동일한 연결의 기록을 결합하면 연결을 만든 프로세스의 등록 상태와 해당 호스트에 대한 목적지 승인 상태를 함께 확인할 수 있다. Table 5는 비인가 정책을 먼저 적용한 뒤 발생 주체 식별 여부, 실행 이미지의 등록 상태와 목적지 승인 상태에 따라 판정 결과를 구분한다. 이 규칙에서는 실행 이미지가 등록되어 있거나 다른 정상 단서가 있더라도 목적지의 승인 근거를 대신하지 못한다. 이러한 비보상적 판정은 선행 연구[2]에서 다룬 가중합의 상쇄 문제에 대한 대응이다. 임계논리[29]와 분류기 결합[30]이 여러 입력을 종합하는 방법에 관한 배경을 제공한다면, 본 구조는 점수의 합보다 승인 조건의 적용 순서를 명확히 하는 데 초점을 둔다.

D3에서 σE도 성립하면 승인 범위 밖 목적지 연결을 주 판정으로 두고 미등록 이미지 사실도 보존한다. 비인가 정책은 차단 또는 경보 전용일 수 있으며 실제 allow/drop 과 성공 여부를 따로 기록하고, 다른 트래픽 규칙의 drop 은 승인 정책의 비인가를 뜻하지 않는다. 실행 이미지의 등록 상태를 확인하지 못한 경우에는 σE를 0으로 간주하지 않으며, 목적지에 따른 D1 또는 D3 판정은 기록하되 이미지 등록 상태가 필요한 나머지 판정은 확인될 때까지 보류한다.

같은 (호스트 정책 키, 목적지, 판정결과)의 D3 또는 D5가 W 안에 m회 나타나면 조사를 우선한다(예시 W=24시간, m=3). 반복은 근거 수를 늘리지 않으며 D5를 D3로 바꾸거나 자동 차단하지 않는다. 정상 반복과 승인 갱신 대기는 제외하고, 첫 관측 신규성은 조회 기간을 명시한 보조 단서로만 둔다.

4.4 How the Structure Is Extended

새로운 위협이나 관측원이 등장하면 판정에 필요한 정보도 달라진다. 자율형 AI가 실행을 반복하는 경우에는 연결의 반복 주기와 동일 실행 이미지의 재사용이 조사 단서가 되고, 런타임 감사 기록이 확보되면 기존 기록에서 확인하지 못한 실행 정보를 추가할 수 있다. 본 프레임워크는 이러한 변화를 다섯 단계로 수용한다. ① 판정이나 조사에 필요한 정보 항목을 도출하고, ② 각 정보를 제공할 책임계층을 배정하며, ③ 사용할 원시 기록과 신호의 성립 조건을 정하고 정보 부족으로 신호를 산출하지 못하는 경우의 처리 방법을 정의한다. ④ 새 기록을 기존 연결에 정확하게 대응시킬 수 있는지, 같은 정보를 중복 반영하지 않는지, 관측원들이 동일한 권한 아래에서 함께 변조될 가능성과 오탐은 없는지 검토하고, ⑤ 추가된 정보에 기존 판정 규칙을 적용할 수 있는지와 수집과 승인 관리에 필요한 운영 부담을 확인한다.

이 절차는 기존 계층에 항목을 추가할 때도 적용한다. 내부 목적지까지 범위를 넓히는 경우에는 기록과 승인 기준을 확보할 수 있는 내부 정책 지점에 S의 책임을 배정하되, TLS를 처리하는 지점이 있다는 사실만으로 감사 기록이 확보되었다고 판단하지 않는다. 프로토콜이나 전송 방

Table 5. Priority-ordered conditions, decisions, and actions

Condition(Priority order)	Result	Action / retained facts
Applicable host policy prohibits destination; any attribution or σE	D1: Prohibited	Apply block or alert-only policy first; retain actual action and attempt/session status
Not prohibited; host attribution unavailable	D2: Unattributed	Retain destination assessment; do not assign the connection to a host instance
Attributed; destination outside applicable approval scope; σE any	D3: Outside approval scope	Policy review/control; retain any simultaneous unregistered-image signal
Attributed; σE=1; destination approved or UNKNOWN	D4: Unregistered image	Review executable registration; preserve destination state
Attributed; destination UNKNOWN with σE=0, or σE not estimated with destination not outside scope	D5: Assessment pending	Obtain missing host-destination-policy context
Attributed; σE=0; destination approved	D6: No defined violation	Retain decision and policy version; not a normality guarantee

www.kci.go.kr

식이 바뀌면 패킷 특징의 유효성을 다시 검토하여 N의 신호를 조정하거나 사용을 중지하고, 확보하지 못한 정보는 별도로 기록한다.

V. Application Examples, Validation Plan, and Discussion

본 장에서는 적용 예시로 판정 과정과 확장 절차를 설명하고, 실제 환경에서 확인할 검증 항목과 운영상 활용 및 한계를 논의한다.

5.1 Application Examples

본 절에서는 호스트가 직접 연결하는 경우(유형 2)와 로컬 하위 프로세스가 연결하는 경우(유형 3)의 두 사례를 통해, 실행 이미지가 등록되어 있더라도 목적지가 호스트의 승인 범위 밖이면 판정할 수 있음을 보인다. Table 6의 추적 기록은 측정 결과가 아니라 판정 절차를 설명하기 위한 가상 기록이다. 각 항목은 Sysmon Event 1과 3[25]과 게이트웨이 트래픽 로그[27]의 필드 구조를 따랐으며, 판정에 사용한 필드를 표에 함께 적었다. 주소와 도메인은 문서 및 예시용 예약 대역(203.0.113.0/24, 198.51.100.0/24)과 .example 도메인을 사용하였다. 두 사례 모두 목적지가 해당 호스트의 승인 목록에 없도록 설정하였으므로, 트래픽이 허용되고 실행 이미지가 등록되어 있어도 D3로 판정된다. 기록 결합의 시간 오차 ±2초는 절차 설명을 위한 예시 값이다.

등록 인터프리터가 실행하는 스크립트와 패키지 정보를 단말 E에 추가하는 경우에도 4.4절의 다섯 단계를 그대로 적용한다. 실행 이미지의 등록 대조만으로는 인터프리터가 실행한 내용을 구분할 수 없으므로, 이 확장에서는 실행 이미지와 스크립트 또는 패키지의 등록 위반을 구분하여 기록하고, ProcessGuid와 5-튜플로 실행 맥락과 연결을 결합한 뒤 확장 전후의 동일 기록을 재생하여 오탐과 승인 관리 부담을 비교한다.

5.2 Validation Plan

실제 운영에 적용하려면 기록을 정확히 결합할 수 있는지, 승인 범위 밖 연결을 올바르게 판정하는지, 정상 개발 활동에 불필요한 경보가 발생하지 않는지를 확인해야 한다. 평가는 연결별 정답 정보와 해당 시점의 호스트별 승인 목록을 기준으로 수행하며, 표본 구성, 측정 단위, 적용 정책 버전과 허용 가능한 오류 수준을 평가 전에 정한다 (Table 7). 사례의 충실도는 필드 가용성, 기록 완전성, 분포 현실성의 세 축으로 확인하며, 본 논문은 실제 로그 필드 대응으로 첫째 축을 제시하고 나머지 두 축은 이 계획에 따라 측정한다.

본 절의 검증 계획은 제안 구조를 실제 환경에서 평가하기 위한 기준이며, 현재는 측정 결과를 제시한 단계가 아니다. N의 지표 정의와 산출 조건은 부록 A에 정리하였고,

Table 6. Illustrative attribution and decision traces (virtual records, not measured)

Item	Type 2: Host connection outside destination approval scope	Type 3: External connection in a hypothesized local-server compromise	Source fields used
E attribution	Registered AI development-tool host connects directly; 10.0.5.21:51544 → 203.0.113.25:443	Registered child; host ancestry confirmed; 10.0.5.21:52210 → 198.51.100.40:443	[Sysmon] · Event 1: ProcessGuid, Parent ProcessGuid, Image; · Event 3: ProcessGuid, SourceIp /Port, DestinationIp/Port, UtcTime
Record correlation	E and S matched by 5-tuple and ±2 s; N added only if available	Host lineage and S session matched by 5-tuple and ±2 s; optional N flow linked	· Gateway traffic log: source/destination address and port, NAT source port, session start time
S authorization comparison	Complete applicable host policy; tools.unapproved.example absent from integration and general-function approvals; actual action: allow; $\sigma T=1$	Complete host policy; transfer.unapproved.example outside all approval scopes; actual action: allow; $\sigma T=1$	· Gateway: rule name, action; destination name from SNI/DNS records; approval list: host policy key, entry, version, validity
E registration comparison	Registered host: $\sigma E=0$; attribution established	Registered child image: $\sigma E=0$; ancestry established; compromise not established	· Event 1: Image, Hashes vs. image-registration list version
N evidence	Optional hypothetical $\delta \geq \tau\delta$; investigation annotation only	Optional $\delta \geq \tau\delta$; compromise and content unverified	· Packet timestamps (optional)
Decision and investigation	D3 ($\sigma E=0$, $\sigma T=1$): policy review; N does not change approval; no protocol-use finding	D3 ($\sigma E=0$, $\sigma T=1$): destination policy review; N only annotates investigation	· Retained with policy and image-list versions

www.kci.go.kr

Table 7. Validation targets, boundary cases, and measurement items

Validation target	Method and boundary cases	Measurement items
Attribution and correlation	Trace raw records → normalization → E attribution → S comparison → decision; concurrent flows, NAT, containers, proxies and clock errors	Correlation success over all connections; mismatch rate among joined records; multiple candidates, unmatched counts and gap handling
Decision correctness and reproducibility	Replay identical records and approval snapshots; compromised registered images, incomplete policy/mapping, hidden SNI, sparse packets; normal development, approved general functions, normal tool integration and threat scenarios under the same host policy	Scope-correct decisions; reproducibility; non-estimation rates; utility and limits of N annotations
False positives at normal-operation boundaries	Approved LLM APIs, updates, telemetry and package destinations; pending approvals, version changes and repeated new normal destinations	Erroneous decisions, actions and priority increases; UNKNOWN rate from name-resolution mismatch; sensitivity to W, m and novelty; approval-maintenance workload and delay
Extension and trust domains	Replay before/after the extension in Section 4.4; E and runtime audit records sharing permissions; omissions and tampering; expert review of policy scope and matrix	ProcessGuid correlation stability, duplicates, version-update false positives; evidence under joint failures; validity of responsibility assignment and extension steps

식별 및 결합 성공률은 전체 연결을 기준으로, 오결합률은 결합된 기록을 기준으로 측정한다. 확보 및 결손 근거와 성립률은 선행 연구[2]의 탐지가능성(D) 갱신 후보이며, 위험 점수와의 연계는 별도 검증 대상이다.

5.3 Discussion

호스트별 승인 범위와 정보 책임을 명세하면 서로 다른 제품의 기록을 같은 관리 대상에 대조하여 승인 예외와 범위 밖 연결을 감사할 수 있다. 탐지 전용으로 운영하면 개발 연결을 일괄 차단하지 않고 승인 범위 밖 연결을 검토할 수 있으며, 일반 기능 목적지와 예외의 버전을 함께 관리하면 정상 업데이트를 도구 미승인으로 처리하는 오류를 점검할 수 있다. 원시 기록과 적용 정책 버전을 보존하면 어떤 연결을 어떤 정책으로 판정했는지 재구성할 수 있어 감사와 점검, 감독과 인증 대응을 위한 내부 증적 정리에 활용할 수 있고, 통합 XDR을 운영하는 환경에서도 승인 목록과 원시 기록의 대응을 검증하는 기준으로 사용할 수 있다.

본 연구의 범위와 한계는 다음과 같다. 첫째, 본 구조가 판정하는 것은 외부 연결의 승인 범위이며, 승인된 목적지로 전송되는 내용이나 서비스 내부 행위의 안전성은 확인하지 않는다. 둘째, 판정의 신뢰성은 단말 기록의 정확성, 연결 기록의 대응 가능성, 승인 목록의 적시 갱신에 좌우되므로 운영 시 이 조건을 함께 관리해야 한다. 셋째, 5.1절의 사례는 실제 로그 필드 구조에 따른 가상 기록으로 판정 절차를 설명한 것이며, 판정 정확도, 결합 성공률, 정상 활동의 오탐과 승인 목록의 관리 부담은 아직 측정되지 않았다. 이는 Table 7의 계획에 따라 후속 연구에서 검증한다.

VI. Conclusion

금융권에서 AI 개발도구의 활용이 확대되면서 개발 단말의 승인 범위 밖 외부 연결도 통제 대상이 되었다. 그러나 SSL/TLS 복호화 없이 서로 다른 보안제품의 기록을 쓰는 환경에서는 연결을 만든 프로세스와 목적지 승인 정보가 각각 관리되어 어느 한 기록만으로 적정성을 판단하기 어렵다. 본 논문은 이 분산을 보완하기 위해 단말의 프로세스 식별 정보와 단말 외부의 호스트별 목적지 승인 대조를 동일 연결 단위로 결합하는 관측 및 판정 구조를 제안하였다.

제안 구조의 주요 기여는 등록된 AI 개발도구나 그 하위 프로그램이 만든 연결이라도 목적지가 해당 호스트의 승인 범위 밖이면 이를 식별하여 검토 대상으로 삼는 데 있다. 게이트웨이가 통신을 허용했다는 사실이나 실행 프로그램이 등록되어 있다는 사실만으로 연결을 승인된 것으로 취급하지 않고, 도구 연동과 일반 기능에 허용된 목적지를 함께 대조하여 판단한다. 벤더 플랫폼 안에서 상관하는 통합 XDR이나 앱 단위로 승인을 관리하는 CASB와 달리, 호스트별 승인 범위를 판정 기준으로 삼고, 결합 조건과 정책 버전을 명세한다는 점에서 구별된다.

또한 연결의 발생 주체, 목적지 승인 상태와 적용 정책 버전을 함께 기록하여 판정 근거를 재구성할 수 있도록 하였으며, 이는 개발 업무를 일괄 차단하지 않는 탐지 모드에서 승인 범위 밖 연결을 조사하고 그 근거를 감사와 정책 검토에 활용하는 데 의미가 있다. 새로운 위협이나 관측 정보가 등장하면 5단계 확장 절차에 따라 정보 항목과 책임, 결합 조건, 운영 영향을 재검토하도록 하였다.

본 연구는 설계 구조를 제시하고 실제 로그 필드 구조에

www.kci.go.kr

따른 가상 기록으로 판정 절차를 설명한 단계이며, 제시한 사례는 실측 결과가 아니다. 실제 환경에서의 탐지 성능과 운영 효과는 후속 검증이 필요하다. 향후에는 5.2절의 검증 계획에 따라 승인 범위 밖 연결에 대한 판정 정확도, 프로세스 식별과 기록 결합의 성공률, 정상 개발 활동에서의 오탐과 승인 목록의 관리 부담을 평가할 예정이다.

REFERENCES

[1] Model Context Protocol, "Model Context Protocol Specification, revision 2026-07-28 (incl. SEP-2567 and SEP-2575)," https://modelcontextprotocol.io/specification (accessed Sep. 30, 2026).

[2] S. M. Jin and D. J. Hong, "A repeatedly applicable risk prioritization model for increasingly advanced AI threats in the financial sector," Journal of The Korea Society of Computer and Information, vol. 31, no. 9, pp. 107-116, Sep. 2026. DOI: 10.9708/jksci.2026.31.09.107

[3] Financial Services Commission (Korea), "Roadmap for improving network separation regulations in the financial sector," FSC Press Release, Aug. 2024.

[4] Financial Services Commission (Korea), "Regulation on Electronic Financial Supervision," Article 15, FSC Notice, 2025.

[5] Financial Security Institute (Korea), "Financial AI Security Guidebook" (in Korean), FSI Guide, 2025. https://www.fsec.or.kr/bbs/222 (accessed Sep. 30, 2026).

[6] OWASP GenAI Security Project, "Agentic AI - Threats and Mitigations," Version 1.0, Feb. 2025. https://genai.owasp.org/

[7] MITRE, "ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems," https://atlas.mitre.org/ (accessed Sep. 30, 2026).

[8] X. Hou, Y. Zhao, S. Wang, and H. Wang, "Model Context Protocol (MCP): Landscape, security threats, and future research directions," ACM Transactions on Software Engineering and Methodology, 2026 (presented at FSE 2026 Journal-First). arXiv:2503.23278.

[9] D. Zhang, Z. Li, X. Luo, X. Liu, P. Li, and W. Xu, "MCP Security Bench (MSB): Benchmarking attacks against Model Context Protocol in LLM agents," in Proc. International Conference on Learning Representations (ICLR), 2026.

[10] H. Song, Y. Shen, W. Luo, L. Guo, T. Chen, J. Wang, B. Li, X. Zhang, and J. Chen, "Beyond the protocol: Unveiling attack vectors in the Model Context Protocol (MCP) ecosystem," arXiv preprint arXiv:2506.02040, 2025.

[11] N. Maloyan and D. Namiot, "Breaking the protocol: Security analysis of the Model Context Protocol specification and prompt injection vulnerabilities in tool-integrated LLM agents," arXiv preprint arXiv:2601.17549, 2026.

[12] P. He, C. Li, B. Zhao, T. Du, and S. Ji, "Automatic red teaming LLM-based agents with Model Context Protocol tools," IEEE Transactions on Information Forensics and Security, 2026 (early access). arXiv:2509.21011.

[13] W. Xing, Z. Qi, Y. Qin, Y. Li, C. Chang, J. Yu, C. Lin, Z. Xie, and M. Han, "MCP-Guard: A multi-stage defense-in-depth framework for securing Model Context Protocol in agentic AI," in Findings of the Association for Computational Linguistics: ACL 2026, 2026.

[14] S. Rezaei and X. Liu, "Deep learning for encrypted traffic classification: An overview," IEEE Communications Magazine, vol. 57, no. 5, pp. 76-81, 2019.

[15] R. Ramachandran, S. Neelakantan, and A. S. Bidyarthy, "Behavior model for detecting data exfiltration in network environment," in Proc. 2011 IEEE 5th International Conference on Internet Multimedia Systems Architecture and Application (IMSAA), Bangalore, India, Dec. 2011, pp. 1-5.

[16] J. Rahmani, S. D. Çakmakçı, K. O. Detken, and A. Sikora, "A SIEM-based framework for multi-layer data collection and anomaly detection in OT-networks," in Proc. 2025 IEEE 13th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Gliwice, Poland, Sep. 2025, pp. 261-266.

[17] W. U. Hassan, A. Bates, and D. Marino, "Tactical provenance analysis for endpoint detection and response systems," in Proc. 2020 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, May 2020, pp. 1172-1189.

[18] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, Aug. 2020.

[19] Microsoft, "DeviceNetworkEvents table in the advanced hunting schema," https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-devicenetworkevents-table (accessed Oct. 5, 2026).

[20] Microsoft, "Govern discovered apps," https://learn.microsoft.com/en-us/defender-cloud-apps/governance-discovery (accessed Oct. 5, 2026).

[21] Microsoft, "Govern discovered apps using Microsoft Defender for Endpoint," https://learn.microsoft.com/en-us/defender-cloud-apps/mde-govern (accessed Oct. 5, 2026).

[22] Microsoft, "Windows Firewall rules," https://learn.microsoft.com/en-us/windows/security/operating-system-security/network-security/windows-firewall/rules (accessed Oct. 5, 2026).

[23] E. Rescorla, K. Oku, N. Sullivan, and C. A. Wood, "TLS Encrypted Client Hello," IETF RFC 9849, Mar. 2026.

[24] P. Hoffman and P. McManus, "DNS Queries over HTTPS (DoH)," IETF RFC 8484, 2018.

[25] Microsoft, "Sysmon," Sysinternals Documentation, https://learn.microsoft.com/sysinternals/downloads/sysmon (accessed Sep. 30, 2026).

[26] Microsoft, "Event 5156(S): The Windows Filtering Platform has permitted a connection," Windows Security Auditing Reference, https://learn.microsoft.com/en-us/windows/security/threat-protecti

www.kci.go.kr

on/auditing/event-5156 (accessed Oct. 24, 2026).
[27] Palo Alto Networks, "Traffic Log Fields," PAN-OS 11.1 Administrator's Guide, https://docs.paloaltonetworks.com/pan-os/11-1/pan-os-admin/monitoring/use-syslog-for-monitoring/syslog-field-descriptions/traffic-log-fields (accessed Oct. 24, 2026).
[28] J. Althouse, J. Atkinson, and J. Atkins, "TLS fingerprinting with JA3 and JA3S," Salesforce Engineering Blog, 2017.
[29] S. Muroga, Threshold Logic and Its Applications, New York: Wiley-Interscience, 1971.
[30] L. I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, 2nd ed., Hoboken, NJ: Wiley, 2014.

Appendix A. 기호정리 (Notation)

본문과 표에서 사용하는 계층, 신호, 판정 및 파라미터의 기호를 표로 정리하였다.

기호	정의 (관련 절)
E / S / N	연결 식별(E), 호스트별 승인 대조(S), 패킷 조사 보조(N)의 관측 계층 (3.1, 3.2절)
σE, σT	미등록 실행 이미지의 외부 연결 신호(σE)와 목적지 승인 대조 신호(σT) (Table 4)
σδ, σF	패킷 도달 간격 임계 신호와 TLS 지문 및 평판 보조 신호 (Table 4)
D1~D6	판정 결과: 비인가, 발생 주체 미확인, 승인 범위 밖, 미등록 이미지, 판정 보류, 위반 없음(Table 5)
δ, τδ	δ는 변동계수(CV)에 변환 함수를 적용한 도달 간격 지표, τδ는 δ의 판정 임계값 (4.2, 5.2절)
Δ, W, m	시각 허용 오차, 반복 판정 창, 조사 우선순위 상향 반복 횟수 (4.2, 4.3절)
호스트 정책키, ProcessGuid	호스트별 승인 목록의 관리 식별자와 프로세스 생성 기록의 고유 식별자 (3.1, 4.2절)
UNKNOWN	목적지 식별, 호스트/목적지 대응 또는 정책 범위가 확정되지 않은 상태(미산출과 구분) (3.1, 4.3절)

Authors



Seungman Jin received the B.S. and M.S. degrees in Computer Engineering from Jeonbuk National University and Yonsei University, Korea, in 2002 and 2006, respectively, and completed the Ph.D. coursework in Computer Engineering from Jeonbuk National University, Korea, in 2022. Mr. Jin performed information security duties at the Korea Internet & Security Agency (KISA) in 2002, Samsung Electronics in 2006, and FSEC in 2007. He joined Jeonbuk Bank in 2011, where he is currently serving as the Information Security Officer.



Deukjo Hong received the B.S. and M.S. degrees in Mathematics from Korea University in 1999 and 2002, respectively, and the Ph.D. degree in Information Security from Korea University in 2006. Mr. Hong is currently a Professor in the Department of Computer Science & Artificial Intelligence at Jeonbuk National University. His research interests include cryptography and network security.

www.kci.go.kr