JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 31 No. 11, pp. 327-341, November 2026
https://doi.org/10.9708/jksci.2026.31.11.327

Analysis of Patent Trends and Technological Structures in Malicious Behavior Detection for Encrypted Traffic

Byungkyu Kim*, Min-Woo Park*, Beom-Jong You**, Taewoong Kwon***, Jun Lee****

*Principal Researcher, AI Digital Security Research Team, KISTI, Daejeon, Korea
**Research Fellow, AI Digital Security Research Team, KISTI, Daejeon, Korea
***Team Head, AI Digital Security Research Team, KISTI, Daejeon, Korea
****Director, Science & Technology Security Research Center, KISTI, Daejeon, Korea

[Abstract]

The encryption of network traffic limits direct payload inspection, increasing the importance of detecting malicious behavior using traffic information observable without decryption. This study examined filing and applicant trends in 392 patent publications corresponding to applications filed between 2006 and 2025. Using classification schemes for data types and detection technologies, it analyzed data use, temporal changes across five technology groups and 26 subtypes, and direct-use relationships linking traffic data units and representations to features. It also examined the co-occurrence structures of patent classifications, detection technologies, and features, as well as co-occurrence relationships between detection technologies and features. Of the publications, 70.7% corresponded to applications filed in 2021-2025, while non-decryption approaches accounted for 92.9%. The relative emphasis shifted from rule/statistics-based approaches and classical machine learning toward deep-learning-based representation learning and hybrid model/feature fusion, with the latter two groups accounting for 80.3% of publications with 2024-2025 filing dates. Earlier technology groups persisted, indicating a reconfiguration in which multiple detection approaches coexisted. Detection approaches also differed in how observable traffic information was organized and used for decision making. In later periods, feature types changed little, while links among existing features increased and their relationships with detection technologies became more diverse. Of the distinct co-occurrence links between detection technologies and features observed in 2022-2025, 69.7% were absent from the cumulative 2006-2021 network. By integrating long-term patent trends and relationships among technological elements, this study provides an analytical basis for comparing the structures of malicious-behavior detection technologies for encrypted traffic and exploring future R&D directions.

▸ **Key words**: Encrypted traffic, Malicious behavior detection, Patent analysis, Artificial intelligence, Co-occurrence networks

• First Author: Byungkyu Kim, Corresponding Author: Byungkyu Kim
*Byungkyu Kim (bk.kim@kisti.re.kr), AI Digital Security Research Team, KISTI
*Min-Woo Park (pminwoo@kisti.re.kr), AI Digital Security Research Team, KISTI
**Beom-Jong You (you@kisti.re.kr), AI Digital Security Research Team, KISTI
***Taewoong Kwon (taewoong.kwon@kisti.re.kr), AI Digital Security Research Team, KISTI
****Jun Lee (jun.lee@kisti.re.kr), Science & Technology Security Research Center, KISTI
• Received: 2026. 06. 25, Revised: 2026. 10. 02, Accepted: 2026. 11. 02.
• This article is an extension of the conference paper: "Analysis of Patents on Malicious Encrypted Traffic Detection" (the 73rd conference of the Korean Society of Computer Information)

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

[요 약]

암호화 통신은 페이로드의 직접 관찰을 제한하므로, 복호화 없이 관측 가능한 트래픽 정보를 이용한 악성행위 탐지의 중요성이 커지고 있다. 본 연구는 2006년부터 2025년까지 출원된 특허공보 392건을 대상으로 출원 및 출원주체 동향을 살펴보고, 데이터 유형과 탐지기술의 분류체계를 적용하여 데이터 활용, 5대 기술군과 26개 세부유형의 시기별 변화, 트래픽 데이터 단위 및 표현과 특징정보의 직접 사용관계를 분석하였다. 또한 특허분류, 탐지기술 및 특징정보 그리고 탐지기술과 특징정보 사이의 동시출현 구조를 분석하였다. 전체 공보의 70.7%가 2021년부터 2025년까지 출원되었으며, 비복호화 접근은 92.9%를 차지하였다. 기술군의 상대적 중심은 규칙·통계 기반 접근과 전통적인 머신러닝에서 딥러닝 기반 표현학습형과 복합모델·특징융합형으로 이동했으며, 두 기술군은 최근 2년간 출원 공보의 80.3%를 차지하였다. 기존 기술군도 최근까지 지속되어, 기술 변화는 특정 방식의 단순한 대체보다 여러 탐지방식이 공존하며 기술구성이 재편되는 양상을 보였다. 또한 탐지방식에 따라 관측 가능한 트래픽 정보를 구성하고 판정에 활용하는 방식에도 차이가 나타났다. 동시출현 네트워크에서는 후반 시기에 특징정보의 종류가 크게 달라지지 않은 반면 기존 특징정보 사이의 연결은 증가하였고, 탐지기술과 특징정보의 관계도 더욱 다양해졌다. 특히 탐지기술과 특징정보 간의 동시출현 네트워크에서 2022년부터 2025년까지 관찰된 고유 연결의 69.7%는 2006년부터 2021년까지의 누적망에 없던 관계였다. 본 연구는 장기 특허 동향과 기술요소 간 관계를 통합적으로 분석하였으며, 탐지기술의 구조적 비교와 향후 연구개발 방향 탐색을 위한 기반을 마련하였다.

▸ **주제어**: 암호 트래픽, 악성행위 탐지, 특허분석, 인공지능, 동시출현 네트워크

I. Introduction

전송계층보안 (Transport Layer Security; TLS) 기반 HTTPS와 QUIC 기반 HTTP/3 등 암호화 통신의 확산은 통신의 기밀성과 무결성을 강화하는 한편, 네트워크 중간 지점에서 애플리케이션 페이로드를 직접 관찰하기 어렵게 만들었다. 공격자 역시 이러한 암호화 채널을 악성코드 통신, 명령 및 제어 (command and control; C2), 데이터 유출, 은닉 터널링 등에 활용한다. 따라서 평문 페이로드 분석이나 콘텐츠 기반 시그니처 검사에 의존하는 기존 탐지방식만으로는 암호 트래픽의 악성행위를 식별하는 데 한계가 있다. 이러한 한계에 대응하여 복호화 없이 관측 가능한 트래픽 정보를 활용하는 악성행위 탐지기술이 발전해 왔다 [1], [2].

암호화 통신에서 악성 트래픽을 탐지하는 기술은 관측 정보를 구성하고 판정에 활용하는 방식의 변화와 인공지능 (artificial intelligence; AI)의 발전에 따라 고도화되어 왔다. 딥러닝 (deep learning; DL)은 머신러닝 (machine learning; ML)의 한 분야이지만, 본 연구에서는 기술적 특성과 변화 양상을 비교하기 위해 전통적인 머신러닝 기법 (이하 고전 ML)과 DL을 구분하였다.

고전 ML은 플로우와 세션 통계, 프로토콜 문맥 등 명시적으로 구성한 특징정보를 분류에 활용해 왔으며 [3], DL은 입력정보로부터 특징표현을 학습하는 방향으로 발전하였다 [4]. 최근에는 사전학습을 활용한 표현학습 접근이 확대되고 있으며 [5], Transformer 기반 시퀀스 모델, 그래프 신경망 (Graph Neural Network; GNN), 멀티모달 융합 등 다양한 AI 기반 탐지기술도 제시되고 있다 [4], [6]. 이러한 변화는 탐지기술의 발전을 특정 모델의 교체만으로 설명하기보다 관측정보의 구성과 특징표현, 판정기술의 결합방식을 함께 살펴볼 필요가 있음을 보여 준다. 데이터셋의 규모와 편향, 서로 다른 환경에 대한 일반화, 실시간 처리 및 실제 운영환경 적용은 여전히 중요한 과제로 남아 있다 [7].

해당 분야의 기존 학술연구에서는 이러한 문제를 포함하여 탐지방법의 설계와 성능평가를 폭넓게 다루어 왔다. 한편 특허공보에는 발명이 해결하려는 기술적 과제와 청구항 및 상세설명에 기재된 기술구성, 출원 시기와 출원주체 등이 기록되어 있어 다수의 공보를 종합하면 기술구조와 출원 동향의 장기적 변화를 분석할 수 있다 [8]. 따라서 학술연구와 특허분석은 암호 트래픽 악성행위 탐지기술의 발전과 구조를 이해하는 데 상호보완적인 관점을 제공한다.

특허문헌은 공개공보와 등록공보 등을 포괄하며, 본 연

www.kci.go.kr

구에서는 공보번호로 식별한 최종 선정 공보를 계량분석의 기본 단위로 설정하였다. 출원 동향은 각 공보의 출원일을 기준으로 집계하였으며, 동일한 발명이 여러 국가에서 공개된 경우에도 각국의 공보를 분석대상에 포함하였다. 또한 연도별 출원 추세를 살펴보기 위해 개별 특허공보 수와 동일 특허패밀리에 속한 공보를 1건으로 집계한 결과를 함께 제시하였다.

본 연구의 목적은 암호 트래픽 악성행위 탐지 관련 특허공보의 장기 출원 동향과 기술구조를 분석하는 데 있다. 이를 위해 2006년부터 2025년까지 출원된 관련 특허공보 392건을 선별하고, 기술요소를 구조화한 분석 데이터셋을 구축하여 네 가지 측면에서 분석하였다. 첫째, 출원 동향과 출원주체의 구성을 분석하였다. 둘째, 데이터 출처와 확보방식을 파악하고 트래픽 데이터 단위 및 표현과 표준 피처의 직접 사용관계를 분석하였다. 셋째, 5대 기술군과 26개 세부유형의 시기별 변화를 살펴보았으며, 마지막으로 국제특허분류 (International Patent Classification; IPC)와 협력특허분류 (Cooperative Patent Classification; CPC), 탐지기술과 특징정보의 동시출현 구조를 분석하였다.

II. Related Work

1. Studies on Malicious Behavior Detection in Encrypted Traffic

암호 트래픽 악성행위 탐지는 암호화된 콘텐츠를 복호화하거나 암호 기능을 이용해 검사하는 능동형 접근과, 복호화 없이 관측 가능한 트래픽 특징을 분석하는 수동형 접근으로 구분할 수 있다 [2], [9]. 복호화 기반 검사는 평문 콘텐츠를 직접 분석할 수 있지만 키와 인증서 관리, 개인정보 보호, 처리 부하, 배치 위치 등의 운영상 제약을 수반한다 [2], [9]. 반면 비복호화 방식은 암호화된 통신에서도 관측 가능한 핸드셰이크 관련 정보 등을 이용하여 악성 여부를 판정한다 [10]. 이때 통신 규모와 시간 특성, 프로토콜 정보, 행동 패턴 등이 주요 분석정보로 활용된다.

비복호화 악성행위 탐지에서 고전 ML 기반 접근은 패킷 길이와 도착 간격, 플로우 및 세션 통계, TLS 핸드셰이크 정보 등에서 특징을 추출하고, 이를 Random Forest (RF)와 Support Vector Machine (SVM) 등의 분류 모델에 적용하여 정상 및 악성 트래픽을 판별해 왔다 [3], [11]-[12]. 관련 서베이 연구들은 암호 트래픽 분석을 데이터 수집과 전처리, 특징 추출과 선택, 특징 표현, 분류 모델 적용 및 성능평가로 이어지는 과정으로 정리하고 있다 [13]-[15].

DL 기반 연구에서는 합성곱 신경망 (Convolutional Neural Network; CNN)이 암호 트래픽의 국소 패턴을, Long Short-Term Memory (LSTM)를 포함한 순환신경망 계열 모델이 시간적 의존성을 학습하는 데 활용되어 왔다 [4]. 사전학습 기반 접근은 라벨 데이터에 대한 의존도를 줄이고 서로 다른 트래픽 환경에 대한 적응 가능성을 높이기 위한 표현학습 방법으로 연구되고 있다 [5]. 최근에는 Transformer와 그래프 신경망 (Graph Neural Network; GNN) 등 다양한 DL 모델이 활용되고 있으며 [6], [13], 패킷 시퀀스, 그래프, 이미지 등 여러 형태의 데이터 표현을 이용한 암호 트래픽 분석 연구가 활발하다 [13], [15]. 또한 공격 목적, 관측 행동, 탐지모델 사이의 관계를 구조적으로 분석하려는 연구도 제시되었다 [6]. 따라서 탐지기술의 변화를 이해하려면 모델 구조뿐 아니라 관측정보의 구성 단위와 표현 형태, 이를 판정에 활용하는 방식까지 함께 살펴볼 필요가 있다.

암호 트래픽 탐지 및 분류 분야에서는 데이터셋과 특징정보, 탐지모델 및 성능평가를 종합적으로 검토한 문헌고찰과 비교연구가 수행되어 왔으며 [1]-[3], [7], [9], [13]-[15], 저자와 인용, 데이터셋 및 분석방법을 중심으로 한 학술연구 동향 분석도 이루어졌다 [16]. 기존 연구에서는 데이터셋의 품질과 대표성 부족, 연구 간 성능 비교의 어려움이 지적되었으며 [3], [7], 클래스 불균형과 데이터 드리프트, 서로 다른 환경에 대한 일반화 문제도 주요 과제로 다루어졌다 [13]. 실제 적용 측면에서는 연산 및 자원 효율과 실시간 처리의 제약 [1], [7], [13], 오탐에 따른 보안관제 부담 [2] 등이 주요 과제로 제시되었다. 이들 연구는 암호 트래픽 탐지에 활용되는 데이터와 특징정보, 탐지모델의 발전 양상뿐 아니라 성능평가와 실제 적용상의 과제를 이해하는 데 중요한 기초를 제공한다.

2. Patent-based Technology Structure Analysis

특허 기반 기술분석은 출원 건수와 출원주체의 분포를 파악한 데서 나아가 기술영역 간 관계와 장기적인 구조변화를 분석하는 데 활용되어 왔다. 특허 분류정보의 동시출현을 이용하여 기술지식 영역 간 상호작용과 시기별 변화 양상을 분석한 연구가 있으며 [17], 특허 데이터 기반 유망기술 연구방법을 고찰한 문헌에서는 통계 및 네트워크 분석에 더해 텍스트 마이닝과 AI 기반 기법의 활용이 확대되는 흐름을 제시하였다 [18]. 산업별 사례로는 제약산업에서 AI 관련 특허의 기술 활용 분야와 주요 출원 주체의 동향을 분석한 연구 [19], 열에너지 저장기술의 장기 특허동향과 적용 분야를 살펴본 연구 [20] 등이 있다.

www.kci.go.kr

특허분류체계는 이러한 기술분석에서 기술영역을 구분하고 비교하는 주요 기준을 제공한다. IPC는 특허문헌을 기술분야별로 분류하여 검색에 활용하는 국제 분류체계이며 [21], CPC는 IPC를 기반으로 보다 세분된 기술분류를 제공한다 [22]. 두 체계는 분류정보가 부여된 특허문헌의 범위에서도 차이가 존재한다 [23]. 이러한 분류정보는 기술영역별 출원 현황뿐 아니라 기술영역 간 연결구조와 기업 간 경쟁구조를 분석하는 데에도 활용되고 있다. 연구 사례로는 스마트 소방 분야에서 토픽 모델링과 CPC 동시출현 네트워크를 결합하여 기술 간 연관구조를 분석한 연구 [24]와 디스플레이 산업에서 IPC별 출원활동과 특허 품질지표를 시기별로 비교하여 기업 간 경쟁구조의 변화를 분석한 연구 [25]가 있다.

또한 특허 제목과 초록의 의미정보를 네트워크의 구조적 특성과 결합하여 기술기회를 탐색한 연구 [26]와 IPC 동시출현, 특허 네트워크 및 군집분석을 활용하여 기술영역 간 융합구조를 분석한 연구 [27]도 제시되었다. 동시출현 네트워크는 동일 특허공보에서 함께 나타난 분류코드나 기술요소 사이의 관계를 표현하며, 두 요소가 함께 관찰된 공보 수를 간선 가중치로 활용하여 연결구조를 분석할 수 있다.

암호 트래픽 탐지 분야의 학술연구는 관측정보와 특징표현, 탐지모델의 설계와 성능평가, 실제 적용상의 과제를 다루어 왔으며, 특허 기반 연구는 출원정보와 특허분류, 기술요소 간 관계를 활용하여 기술의 장기적 변화와 구조를 분석해 왔다. 이를 토대로 본 연구는 암호 트래픽 악성행위 탐지 관련 특허공보에 기재된 기술요소를 분류하고, 출원 시기별 기술구성의 변화와 청구항 및 상세설명에 근거한 직접 사용관계, 기술요소 간 동시출현 구조를 분석한다. 구체적인 분석방법은 다음 장에서 제시한다.

III. Materials and Methods

1. Patent Collection and Applicant Analysis

암호 트래픽 악성행위 탐지 관련 특허공보는 Google Patents를 중심으로 검색하여 수집하였다. 출원일은 2025년 12월 31일까지로 제한하였으며, 선별된 공보의 출원연도 범위에 따라 분석기간을 2006-2025년으로 설정하였다. 제목 (Title), 초록 (Abstract), 청구항 (Claims)을 대상으로 한 기본 검색식은 TAC=(("encrypted traffic" OR "encrypted network traffic" OR "encrypted flow" OR "malicious encrypted traffic" OR "encrypted malicious traffic" OR "malicious encrypted flow" OR "encrypted malicious flow") AND (malicious OR malware OR anomaly OR abnormal OR attack OR threat) AND (detection OR identification OR classification OR analysis))이다.

기본검색으로 핵심 후보를 확보한 뒤, 인용 및 피인용 관계와 특허패밀리 정보를 활용하고 보완검색을 수행하여 후보를 확대하였다. 보완검색에서는 탐지와 이상행위 관련 유사어를 추가하고, TLS, HTTPS, QUIC 등의 프로토콜명을 트래픽, 위협, 탐지 관련 용어와 조합하였다. 2026년 10월 19일 다운로드한 검색결과 목록 783건을 최종 분석대상 392건과 대조한 결과, 공보번호 일치 217건, 동일 출원의 공개공보와 등록공보 관계 65건, 특허패밀리 관계 44건으로 총 326건이 대응되었다. 최종 분석대상 중 나머지 66건은 인용·피인용 검토(22건)와 보완검색(44건)을 통해 확보한 공보이다.

후보 공보는 제목, 초록, 청구항을 우선 검토하고, 필요한 경우 상세설명을 추가로 확인하였다. 암호화된 네트워크 트래픽을 대상으로 악성행위, 이상행위, 공격 또는 위협을 탐지하거나 판정하는 기술을 주요 내용으로 다룬 공보를 선정하였다. 단순한 서비스 식별, 일반적인 암호화나 복호화 또는 네트워크 관리만을 다룬 공보는 제외하였다. 포함 여부가 불명확한 경우에는 청구항과 상세설명을 재검토하여 판단하였다. 동일 출원의 공개공보와 등록공보가 모두 확인된 경우에는 등록공보를 대표공보로 사용하였다. 세계지식재산기구 (World Intellectual Property Organization; WIPO) 및 서로 다른 관할의 특허청에 공개된 동일 발명은 개별 공보로 유지하되 패밀리로 연결하였다.

최종 분석대상은 특허공보 392건이며, 동일 특허패밀리에 속한 공보를 하나로 집계하면 324개 특허패밀리이다. IPC는 386개 공보, CPC는 354개 공보에서 확인되어 해당 정보가 없는 공보는 각 분류 네트워크에서만 제외하였다. IPC 정보는 별도로 한국 지식재산처의 KIPRIS (지식재산정보 검색 서비스)를 통해 수집하였다. 특허공보에 기재된 출원기관의 명칭은 정식 명칭, 약칭 및 표기 차이를 대조하여 정제하고, 동일 기관의 중복 명칭을 통합하여 221개의 정제 출원인 항목으로 정리하였다. 출원주체는 기업, 대학, 공공연구기관, 정부기관 및 개인으로 구분하였으며, 기업은 보안전문기업, 통신·네트워크기업, 일반 ICT·플랫폼기업, 산업·에너지기업으로 세분화하였다. 공동출원 공보는 참여 주체와 해당 국가에 각각 반영하였다. 본 연구의 특허 수집과 정제, 데이터·기술 분류, 통계와 네트워크 분석의

www.kci.go.kr

전체 연구 수행과정을 정리하면 Figure 1과 같다.

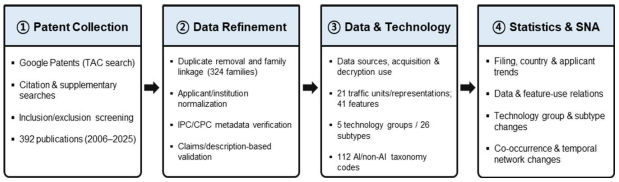


Fig. 1. Research Workflow and Analysis Framework

2. Data-Type and Technology Classification

본 논문에서는 특허에 나타난 데이터 활용과 탐지기술의 구성을 비교하기 위해 데이터 유형과 탐지기술을 분류할 수 있는 분석체계를 구성하였다. 먼저, 데이터 유형은 탐지에 활용되는 정보의 출처와 확보방식, 데이터 단위 및 표현, 그리고 판정에 사용된 특징정보를 구분하여 분석하였다. 이에 따라 데이터 출처 (Dataset source), 데이터 확보방식 (Acquisition method), 트래픽 데이터 단위 및 표현 (Traffic data units and representations), 표준 피처 (Standardized features)의 네 가지 분석 항목을 정의하였다.

Table 1. Data-Type Classification

Dimension	Categories	Classification
Dataset source	In-house/Non-public, Public, Hybrid	3 types / Single-label
Acquisition method	Operational/observational, Controlled environment, Public resource, External knowledge, Institution-provided data	5 upper categories / 15 methods / Multi-label
Traffic data units and representations	Network observation, Temporal structure, Transformed representation, Record/entity, Identifier, Learned/numeric representation, Event	7 upper categories / 21 items / Multi-label
Standardized features	Protocol/Identification, Statistical/scale, Temporal/behavior, Structural/transformed, Payload/content, Baseline/state, Learned/relation, Decision signal, External knowledge/fingerprint	9 upper categories / 41 items / Multi-label

데이터 유형 분류체계의 전체 구성은 Table 1과 같다. 데이터 출처는 분석에 활용된 데이터의 전체적인 원천을 기준으로 자체 데이터, 공개 데이터, 자체 데이터와 공개 데이터를 함께 사용한 혼합형의 세 유형으로 구분하고 공보별로 하나의 유형을 부여하였다. 데이터 확보방식은 실제 데이터를 어떤 경로와 환경에서 수집, 생성, 제공받거나 결합하였는지를 나타내며, 운영·관측 수집, 통제환경 생성, 공개 자원 활용, 외부 지식 결합, 기관 자료 활용의 5개 상위범주와 15개 세부방식으로 구성하였다. 하나의 공보에서 둘 이상의 데이터 확보방식이 사용된 경우에는 이를 모두 반영하고, 공개 자원을 활용한 경우에는 자료의 유형과 자료명도 확인하였다.

트래픽 데이터 단위 및 표현은 탐지에 활용되는 정보를 분석 단위와 표현 형태에 따라 구분한 것이다. 패킷, 플로우, 세션 등은 트래픽을 관측하거나 집계하는 단위이며, 시계열과 시퀀스는 시간적·순차적 구조를, 행렬, 그래프, 이미지 등은 관측정보를 구조화하거나 변환한 표현을 나타낸다. 레코드, 분석 객체, 식별자, 이벤트, 벡터 등을 포함하여 21개 항목을 7개 상위범주로 정리하였다. 하나의 공보에서 여러 단위나 표현이 활용된 경우에는 각각 반영하였다.

표준 피처는 이러한 트래픽 데이터 단위 및 표현에서 악성행위 탐지와 판정에 활용되는 특징정보를 의미한다. 특허마다 서로 다른 명칭으로 기술된 동일하거나 유사한 특징정보를 원문의 기술적 의미와 탐지과정에서의 역할을 기준으로 통합하여 41개 표준 피처로 정리하였다. 이를 프로토콜·식별, 통계·규모, 시간·행동, 구조·변환, 페이로드·콘텐츠, 기준·상태, 학습표현·관계 등의 9개 상위범주로 구분하였으며, 하나의 공보에서 여러 피처가 활용된 경우에는 해당 피처를 모두 반영하였다.

트래픽 데이터 단위 및 표현과 표준 피처는 서로 구분하여 분류하되, 실제 발명구성에서 두 요소의 직접 사용관계도 별도로 분석하였다. 즉, 특허 원문에서 특정 데이터 단위 또는 표현과 표준 피처 사이의 직접 사용관계가 확인되면 해당 관계를 추출하여 분석자료에 반영하였다. 예를 들어 플로우에서 통계특징을 산출하거나 패킷의 길이정보를 탐지에 활용하는 경우가 이에 해당한다. 동일 공보에서 같은 결합관계는 한 번만 반영하였으며, 특허공보에 사용된 다양한 명칭은 사전에 정의한 21개 트래픽 데이터 단위 및 표현과 41개 표준 피처를 기준으로 정규화하였다. 그 결과 342개의 고유 직접관계 유형과 공보별 관계 출현 3,439회를 확인하였다. 이러한 직접 사용관계는 동일 공보에서 두 항목이 단순히 함께 나타나는 동시출현과 구분된다.

탐지기술은 데이터 유형과 별도의 분석 수준에서 분류하였다. AI 기술분류체계 (AI Technology Taxonomy)는 특허공보에 나타난 학습 및 판정기술의 구체적인 사용 양상을 분석하기 위해 구성한 연구용 분류체계이다. AI 기반 기술을 중심으로 하되 기술구성의 비교를 위해 악성행위 판정에 실질적으로 사용된 규칙, 통계, 임계값, 시그니처 및 정책 기반 기술도 함께 포함하였다. 하나의 공보에서

www.kci.go.kr

여러 기술요소가 사용된 경우에는 해당 요소를 모두 분류에 반영하였으며, 기술명의 단순 출현 여부가 아니라 탐지 및 판정과정에서 해당 기술이 수행하는 기능을 기준으로 판단하였다.

5대 기술군과 26개 세부유형은 탐지기술의 주된 판정방식을 기준으로 각 공보를 분류한 계층형 분류체계로, 각 기술군의 구현 차이를 세부유형으로 구분하고 공보마다 하나의 기술군과 이에 속하는 하나의 세부유형을 부여하였다. 복호화가 악성행위 판정에 필수적으로 사용된 경우는 복호화 기반 검사형, 복호화 없이 규칙, 통계량, 임계값, 지문 또는 시그니처가 주된 판정을 수행하는 경우에는 규칙·통계 기반 비복호화형으로 분류하였다. 명시적으로 구성한 피처를 RF, SVM, XGBoost 등의 전통적 ML에 적용하여 주된 판정을 수행하면 고전 ML 기반 분류형, 신경망의 여러 구성요소가 통합된 구조 안에서 특징표현을 학습하고 이를 주된 판정에 이용하는 경우에는 DL 기반 표현학습형으로 분류하였다. 반면 구별되는 모델이나 분석단계의 결과, 서로 다른 데이터 관점의 특징표현 또는 분산학습모델의 정보를 결합하는 것이 핵심인 경우에는 복합모델·특징융합형으로 구분하였다.

기술군의 경계가 겹치는 경우에는 청구항과 상세설명을 대조하여 복호화의 필수성, 주된 판정기술의 역할 및 모델·단계 간 결과 결합 여부를 확인하였다. 복호화된 정보가 최종 판정에 필수인 경우에는 복호화 기반 검사형으로 분류하였으며, 단일 신경망 내부의 공동학습은 DL 기반 표현학습형, 독립 모델·분석단계의 결과 결합은 복합모델·특징융합형으로 구분하였다. 26개 세부유형은 각 기술군 내부의 구현 차이를 비교하기 위해 구성하였으며 이를 정리하면 Table 2와 같다.

복호화 사용 여부는 기술군과 별도로 비복호화, 복호화, 혼합의 세 유형으로 분류하였다. 기술군은 공보의 주된 최종 판정구조를 나타내는 반면, 복호화 사용 여부는 탐지 및 판정과정에서 복호화된 정보가 활용되는지를 나타낸다.

기술구조의 시기별 변화는 초창기 (2006-2017), 성장전반 (2018-2019), 성장후반 (2020-2021), 최근전반 (2022-2023), 최근후반 (2024-2025)의 5개 출원기간으로 구분하여 분석하였다. 5대 기술군은 기간별 공보 수와 구성비를 비교하였으며, 26개 세부유형은 각 기간의 유형별 공보 수와 순위, 관찰된 세부유형 수 및 상위 두 유형의 합계 비중을 통해 변화 양상을 분석하였다.

Table 2. Technology Groups and Technology Subtypes

Technology Group	No. of Subtypes	Technology subtypes (publications)
DL-based Representation Learning	8	Composite DL Backbone / Internal Fusion (42) AE / Generative / Self-Supervised (23) CNN Spatial / Byte Representation (21) General Neural Network / Embedding (15) GNN Graph Representation (15) RNN Temporal Representation (12) Transformer Long-Context Representation (8) SSM / Mamba (1)
Hybrid Model/Feature Fusion	5	Classical ML-DL Hybrid (35) Multi-View Feature Fusion (23) Classical ML Ensemble / Multi-Stage (20) DL-to-DL Hybrid (6) Federated Learning / Model Aggregation (3)
Classical ML-based Classification	5	Handshake / Context Feature Classification (42) Clustering / Profiling (15) Statistical-Feature Classification (14) Multi-Feature Fusion Classification (8) TLS Structure Inference (4)
Rule/Statistics-based Non-decryption	4	Protocol / Entropy Analysis (22) Baseline / Threshold Detection (17) Certificate / Domain Fingerprinting (11) Threat-Intel Linked Policy (10)
Decryption-based Inspection	4	Decryption + ML/DPI Combination (10) Post-Decryption Reinspection (7) Selective Decryption (7) Proxy / MITM Decryption (1)

3. Co-occurrence Network Analysis

본 절에서는 IPC, CPC, AI 기술분류체계 및 표준 피처의 동시출현 관계와 시기별 변화를 분석하기 위한 네트워크 구성과 측정방법을 제시한다. 동일 특허공보에서 두 요소가 함께 나타난 경우 간선을 생성하고 동일 노드쌍은 공보당 한 번만 반영하였으며, 간선 가중치는 두 요소가 함께 나타난 공보 수로 정의하였다. 이를 바탕으로 IPC, CPC, AI 기술분류체계, 표준 피처의 4개 단일모드 네트워크와 AI 기술분류체계-표준 피처 이분 네트워크 (bipartite network)를 구성하였다. IPC는 분류정보가 확인된 386개 공보, CPC는 354개 공보를 사용하였다.

네트워크 특성은 노드 수, 간선 수, 밀도, 총연결강도 (total link strength)와 주요 동시출현 관계를 이용하여 측정하였다. 노드 수는 네트워크에 나타난 기술요소의 수, 간선 수는 동시출현한 고유 노드쌍의 수, 밀도는 가능한 연결 중 실제 연결이 나타난 비율을 의미하며, 총연결강도는 한 노드에 연결된 간선 가중치의 합으로 산출하였다. 네트워크 통계는 NetworkX 3.6.1로 산출하고, 4장에서 제시하는 Figure 8-10의 전체기간 네트워크 맵은 VOSviewer로 시

www.kci.go.kr

각화하였다 [28]. 시기별 변화는 노드 및 간선 수와 함께 2022-2025년 고유 간선 중 2006-2021년 누적망에 없던 관계의 비율로 측정하였다. 동시출현은 동일 공보에 두 요소가 함께 나타난 관계로 정의하여 원문에서 확인한 직접 사용관계와 구분하였으며, 네트워크 지표는 각 네트워크 내부의 구조와 시기별 변화를 기술하는 데 사용하였다.

IV. Results

1. Filing Trends by Country and Institution

출원연도와 특허패밀리, 국가별 참여, 기관유형과 주요 출원주체, 공동출원 구조에 대한 분석결과를 살펴보면 다음과 같다. Figure 2는 연도별 특허공보 수와 동일 특허패밀리에 속한 공보를 1건으로 집계한 결과를 함께 보여 준다.

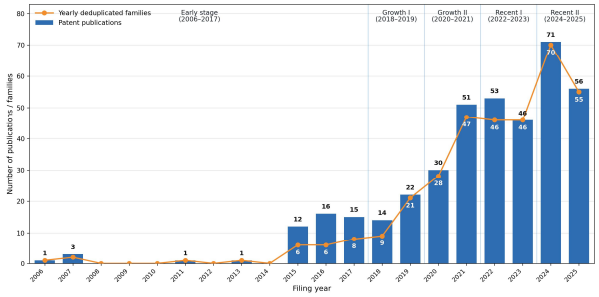


Fig. 2. Annual Patent Filings and Deduplicated Patent Families

출원 공보는 초창기 49건, 성장전반 36건, 성장후반 81건, 최근전반 99건, 최근후반 127건으로 기간 후반에 뚜렷하게 증가하였다. 특히 2021년부터 2025년까지의 출원 공보가 277건 (70.7%)이었고, 2024년이 71건으로 가장 많았다. 패밀리 중복을 제거한 연도별 흐름도 전반적으로 유사한 증가 추세와 최근의 출원 집중 경향을 보였다. 다만 2025년 출원 건수는 특허의 공개 시차로 인해 최종 출원 규모와 차이가 있을 수 있다.

Figure 3은 국가별 참여와 시기별 변화, 기관유형 및 세부유형별 참여 공보와 참여기관 수를 보여 준다. 국가별 참여 공보는 중국 283건, 미국 53건, 영국 20건, 대한민국 16건, 이탈리아 10건, 이스라엘 8건, 싱가포르와 남아프리카 공화국 각 1건이었다. 기관유형별로는 기업이 230건 (58.7%)으로 가장 많았으며, 대학 141건 (36.0%), 공공연구기관 38건 (9.7%), 정부기관 6건 (1.5%), 개인 2건 (0.5%) 순이었다. 참여기관 수도 기업 133개로 대학 64개보다 많았다. 기업 세부유형별 참여 공보는 보안전문기업 124건, 통신·네트워크기업 67건, 산업·에너지기업 26건, 일반 ICT·플랫폼기업 23건과 같았다. 세부유형별 합계는 240건이나, 서로 다른 기업 유형이 함께 참여한 공보를 중복 제거하면 기업 전체의 참여 공보는 230건이다. 이는 암호 트래픽 악성행위 탐지기술의 연구개발이 대학뿐 아니라 보안·통신·ICT와 산업·에너지 분야의 다양한 기업에서도 활발하게 진행되고 있음을 보여 준다.

국가별 참여 양상도 시기에 따라 달라졌다. 초창기에는 미국 26건과 영국 17건이 중국 4건보다 많았으나, 중국은 성장전반 17건, 성장후반 64건, 최근전반 83건, 최근후반 115건으로 증가하였다. 중국 참여 공보의 92.6%가 2020년 이후에 나타났고, 영국의 20건은 모두 2019년까지였다. 대한민국도 16건 중 12건이 2022년 이후에 나타났다.

개별 출원기관 중 기업에서는 Cisco Technology Inc가 23건으로 가장 많았으며, British Telecommunications PLC 20건, Beijing Guancheng Technology Co Ltd 18건 순이었다. 대학에서는 Beijing University of Posts and Telecommunications가 8건으로 가장 많았고, 공공연구기관에서는 Institute of Information Engineering, Chinese Academy of Sciences가 6건의 공보에 참여하였다. 전체 392개 공보 중 공동출원은 51건 (13.0%)이었고 공보당 평균 출원기관 수는 1.19개로, 단독출원이 우세하였다.

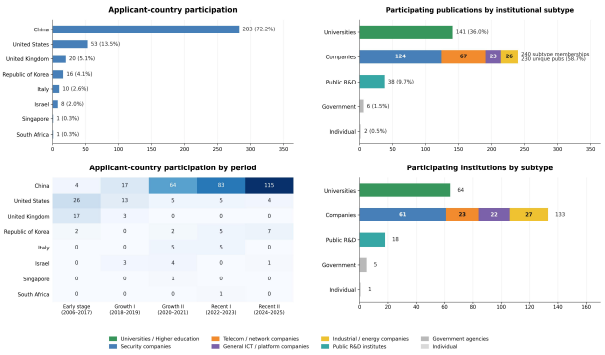


Fig. 3. Applicant Country and Institutional Subtype Structure

Table 3은 국가와 기관 세부유형에 따른 기술군 구성의 차이를 보여 준다. 중국 참여 공보에서는 DL형과 복합형의 비중이 높았고, 미국에서는 고전 ML형과 규칙·통계형, 영국에서는 규칙·통계형, 대한민국에서는 DL형의 비중이 상대적으로 높았다. 기관유형별로 대학과 공공연구기관에서는 DL형과 복합형의 합계가 각각 78.7%와 76.3%였으며, 보안전문기업과 통신·네트워크기업에서는 고전 ML형과 규칙·통계형의 합계가 각각 53.2%와 50.7%였다. 산업·에너지기업 26건에서도 DL형이 46.2%로 가장 높은 비중을 차지하였다. 다만 이러한 차이는 분석대상 특허공보에서 관찰된 것이므로 국가나 기관유형별 실제 기술 역량을 비교하기 위해서는 추가 연구가 필요하다.

Table 3. Technology Group Composition by Applicant Country and Institutional Subtype

Applicant country / institution subtype	Participating publications, n	DL	Hybrid	Classical ML	Rule/stat.	Decryption
China	283	119 (42.0%)	77 (27.2%)	54 (19.1%)	22 (7.8%)	11 (3.9%)
United States	53	1 (1.9%)	9 (17.0%)	22 (41.5%)	15 (28.3%)	6 (11.3%)
United Kingdom	20	7 (35.0%)	0 (0.0%)	0 (0.0%)	13 (65.0%)	0 (0.0%)
Republic of Korea	16	10 (62.5%)	1 (6.2%)	1 (6.2%)	2 (12.5%)	2 (12.5%)
Italy	10	0 (0.0%)	0 (0.0%)	6 (60.0%)	0 (0.0%)	4 (40.0%)
Israel	8	0 (0.0%)	0 (0.0%)	0 (0.0%)	7 (87.5%)	1 (12.5%)
Singapore	1	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (100.0%)
South Africa	1	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (100.0%)	0 (0.0%)
University	141	74 (52.5%)	37 (26.2%)	20 (14.2%)	7 (5.0%)	3 (2.1%)
Security Company	124	24 (19.4%)	19 (15.3%)	36 (29.0%)	30 (24.2%)	15 (12.1%)
Telecom/ Networking Company	67	18 (26.9%)	12 (17.9%)	14 (20.9%)	20 (29.9%)	3 (4.5%)
General ICT/Platform Company	23	9 (39.1%)	8 (34.8%)	3 (13.0%)	2 (8.7%)	1 (4.3%)
Industrial/Energy Company	26	12 (46.2%)	5 (19.2%)	6 (23.1%)	1 (3.8%)	2 (7.7%)
Public Research Institute	38	20 (52.6%)	9 (23.7%)	6 (15.8%)	2 (5.3%)	1 (2.6%)
Government Agency	6	3 (50.0%)	2 (33.3%)	1 (16.7%)	0 (0.0%)	0 (0.0%)
Individual	2	0 (0.0%)	1 (50.0%)	0 (0.0%)	0 (0.0%)	1 (50.0%)

2. Data Acquisition and Feature Representation

본 절에서는 데이터 출처와 확보방식, 트래픽 데이터 단위 및 표현, 표준 피처의 분포와 이들 간의 관계에 대한 분석 결과를 제시한다. 데이터 출처별로는 자체 데이터만 활용한 공보가 325건 (82.9%)으로 가장 많았으며, 공개 자원만 활용한 공보는 36건 (9.2%), 두 유형을 함께 활용한 공보는 31건 (7.9%)이었다. 데이터 확보방식의 상위범주에서는 운영·관측 수집이 261건(66.6%)으로 가장 많았고, 공개 자원 활용 67건 (17.1%), 통제환경 생성 63건 (16.1%), 외부 지식 결합 24건 (6.1%), 기관 자료 활용 2건 (0.5%) 순이었다. 다만 자체 데이터만 활용한 공보 중 62건 (전체의 15.8%)은 구체적인 확보방식을 확인할 수 없어 미확인으로 처리하였다. 하나의 공보에서 여러 확보방식이 확인된 경우에는 각각 반영하였으며, 비율은 전체 392건을 기준으로 산출하였다.

공개 자원을 활용한 67건 중 47건 (전체의 12.0%)에서 공개 벤치마크 데이터셋이 확인되었다. 주요 데이터셋은 CTU-13, CICIDS 2017, USTC-TFC 2016, ISCX VPN-nonVPN, CIRA-CIC-DoHBrw-2020 등이었으며, 공개 자원에는 봇넷·침입 탐지, 트래픽 분류, VPN 및 DNS over HTTPS (DoH) 분석용 데이터셋과 도메인·위협 정보 참조목록 등이 포함되었다.

트래픽 데이터 단위 및 표현에서는 네트워크 관측 단위가 382건 (97.4%)으로 가장 많았으며, 시간구조 227건 (57.9%), 학습·수치표현 179건 (45.7%), 기록·개체 136건 (34.7%), 변환표현 121건 (30.9%), 식별자 110건 (28.1%) 순이었다. 세부적으로는 플로우 257건, 패킷 201건, 세션 148건이 주요 관측 단위였으며, 벡터 179건, 시퀀스 103건, 시계열 100건, 행렬 69건, 그래프 43건도 확인되었다.

표준 피처의 상위범주에서는 프로토콜·식별 296건 (75.5%), 통계·규모 273건 (69.6%), 시간·행동 268건 (68.4%)이 많았으며, 구조·변환 142건, 페이로드·콘텐츠 117건, 학습 표현·관계 116건으로 확인되었다. 세부 피처에서는 통계특징 236건, 패킷길이 167건, 핸드셰이크정보 135건, 헤더필드와 시간간격 각각 134건, 시퀀스특징 132건이 상위 항목이었다.

트래픽 데이터 단위 및 표현과 표준 피처의 직접 사용관계에서는 플로우-통계특징이 157건으로 가장 많았으며, 패킷-패킷길이 109건, 패킷-헤더필드와 벡터-임베딩표현이 각각 83건으로 뒤를 이었다. 세션-핸드셰이크정보와 식별문자열-도메인정보는 각각 69건, 시퀀스-시퀀스특징은 65건, 시계열-시간간격과 플로우-패킷수는 각각 60건으로 나타났다. 이 밖에도 플로우-핸드셰이크정보 57건, 벡터-통계특징과 행렬-행렬변환 각각 53건, 상태레코드-모델점수 41건, 그래프-그래프노드엣지특징 37건 등이 확인되었다. 이러한 결과는 플로우의 통계 및 프로토콜 정보, 패킷의 길이와 헤더 정보가 두드러지는 가운데, 트래픽의 시간 간격과 순서 정보, 벡터와 행렬의 학습 및 변환 표현, 그래프의 연결관계뿐 아니라 트래픽 상태를 평가하는 기준과 탐지모델의 분석 결과 등도 활용되고 있음을 보여 준다.

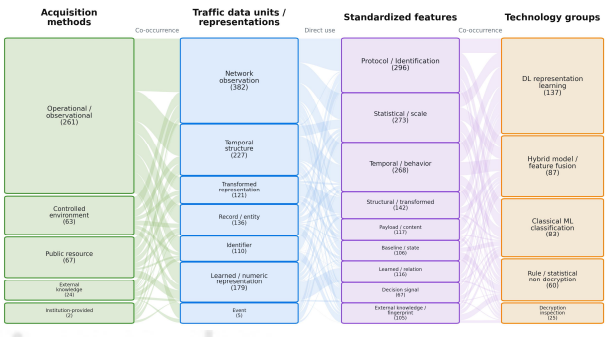


Fig. 4. Links from Data Acquisition Methods to Technology Groups

www.kci.go.kr

Figure 4는 데이터 확보방식, 트래픽 데이터 단위 및 표현, 표준 피처의 상위범주와 5대 기술군 사이의 관계를 보여 준다. 양쪽 연결구간은 동일 공보에서 두 범주가 함께 나타난 동시출현을, 가운데 연결구간은 데이터 단위 및 표현과 표준 피처의 직접 사용관계를 나타낸다. 연결선의 굵기는 두 범주의 관계가 확인된 공보 수에 비례한다. 확보방식이 미확인인 62건은 첫 번째 연결구간에서만 제외하고 나머지 분석에는 포함하였다.

데이터 확보방식-트래픽 데이터 단위·표현 동시출현에서는 운영·관측 수집-네트워크 관측 단위가 256건으로 가장 많았다. 데이터 단위·표현-표준 피처의 상위범주별 직접 사용관계에서는 네트워크 관측 단위와 프로토콜·식별 279건, 통계·규모 224건, 시간·행동 212건이 주요 연결로 나타났으며, 시간구조-시간·행동 176건과 변환표현-구조·변환 111건이 뒤를 이었다. 동일 공보 내 같은 상위범주 조합은 한 번만 집계하였다.

표준 피처-기술군 동시출현에서 DL형은 시간·행동 106건 (77.4%), 구조·변환 84건 (61.3%), 프로토콜·식별 81건 (59.1%), 통계·규모 80건 (58.4%) 순이었으며, 학습표현·관계도 73건 확인되었다. 복합형은 시간·행동과 통계·규모가 각각 72건 (82.8%), 프로토콜·식별이 71건 (81.6%)이었으며, 판정신호도 35건 확인되었다. 고전 ML형은 프로토콜·식별 69건과 통계·규모 63건, 규칙·통계형은 프로토콜·식별 53건, 통계·규모 41건, 기준·상태 36건이 주요 범주였다. 복호화 기반 검사형에서는 프로토콜·식별 22건과 페이로드·콘텐츠 18건이 확인되었다.

기술군 분류기준과 함께 보면 규칙·통계형은 기준이나 규칙과의 비교, 고전 ML형은 명시적으로 구성한 피처를 이용한 분류, DL형은 시퀀스·행렬·그래프 등 다양한 표현에서의 특징 학습, 복합형은 독립된 모델이나 분석단계의 결과 결합, 복호화 기반 검사형은 복호화된 트래픽 내용의 검사가 중심이었다. DL형과 복합형에서도 프로토콜·통계·시간·행동 정보가 폭넓게 확인되었으며, 기술군별 차이는 이러한 정보를 표현하고 탐지 판정에 활용하는 방식에서도 나타났다.

3. Technology Groups and Subtype Changes

복호화 사용 여부와 5대 기술군, 26개 세부유형의 시기별 변화 추이 분석 결과는 다음과 같다. 5대 기술군은 DL 기반 표현학습형 137건 (34.9%), 복합모델·특징융합형 87건 (22.2%), 고전 ML 기반 분류형 83건 (21.2%), 규칙·통계 기반 비복호화형 60건 (15.3%), 복호화 기반 검사형 25건 (6.4%)과 같았다. 복호화 사용 여부는 비복호화 364건 (92.9%), 복호화 13건, 혼합 15건으로 분석되었다. 기술군은 공보의 주된 최종 판정구조를, 복호화 사용 여부는 처리과정에서 복호화된 정보가 사용되었는지를 기준으로 분류하였다.

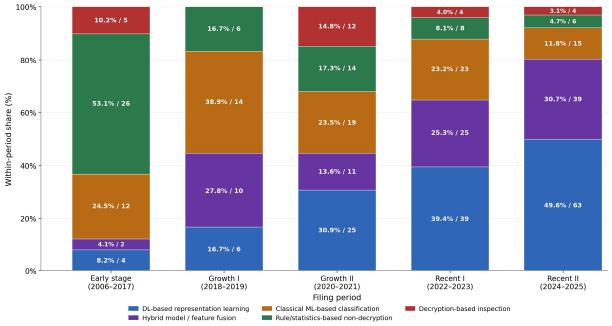


Fig. 5. Technology Group Shares by Filing Period

Figure 5는 시기별 기술군의 구성비 변화를 보여 준다. 초창기에는 규칙·통계형이 49건 중 26건 (53.1%)으로 가장 높았고, 성장전반에는 고전 ML형이 36건 중 14건 (38.9%)으로 가장 높은 비중을 차지하였다. 성장후반에는 DL형이 25건으로 고전 ML형 19건보다 많았다. 최근전반에는 DL형 39건과 복합형 25건이 합계 64건 (64.6%), 최근후반에는 DL형 63건과 복합형 39건이 합계 102건 (80.3%)을 차지하였다. 규칙·통계형과 고전 ML형도 최근까지 관찰되어, 여러 기술방식이 공존하는 가운데 DL형과 복합형의 비중이 커지는 양상을 보였다.

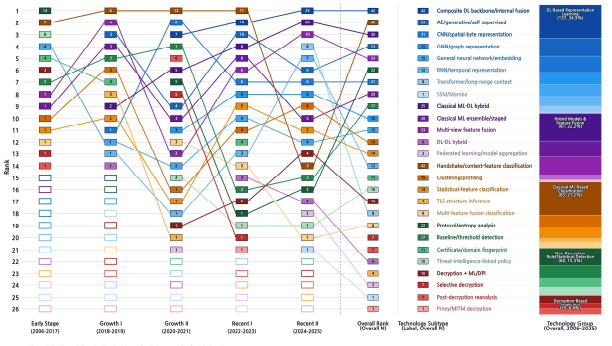


Fig. 6. Technology Subtype Counts and Ranks by Filing Period

Figure 6은 이러한 변화를 26개 세부유형 수준에서 보여 준다. 관찰된 세부유형은 초창기 14개, 성장전반 14개, 성장후반 21개, 최근전반 21개, 최근후반 21개였다. 초창기에는 프로토콜·엔트로피 분석형 14건과 핸드셰이크·문맥 특징 분류형 9건이 많았고, 성장전반에는 핸드셰이크·문맥 특징 분류형 6건과 다중관점 특징융합형 4건이 상위였다. 성장후반에는 핸드셰이크·문맥특징 분류형 13건, CNN·공간/바이트 표현형 8건, 복합 DL 백본·내부융합형 7건이

www.kci.go.kr

주요 유형이었다. 최근전반에는 핸드셰이크·문맥특징 분류형 11건, AE·생성·자기지도형과 복합 DL 백본·내부융합형이 각각 10건으로 많았다. 최근후반에는 복합 DL 백본·내부융합형 25건과 고전 ML-DL 결합형 20건이 상위 두 유형으로, 합계 45건 (35.4%)을 차지하였다. 최근후반에는 Transformer 기반 장거리 문맥형 8건, RNN·시계열 표현형 7건, GNN·그래프 표현형 6건이 확인되어 입력과 표현, 모델의 결합방식이 한층 다양해졌다.

26개 세부유형 중 23개는 2021년까지 이미 확인되었으며, 최근에는 일부 학습·결합형의 출원이 증가하면서 세부유형별 분포가 달라졌다. 구체적으로 GNN 기반 그래프 표현형은 전체 15건 중 12건, 복합 DL 백본·내부융합형은 42건 중 35건, 고전 ML-DL 결합형은 35건 중 28건이 2022년 이후 출원되었다. Transformer 기반 장거리 문맥형도 전체 8건이 모두 이 시기에 관찰되었다.

4. Co-occurrence Network Analysis Results

마지막으로 IPC, CPC, AI 기술분류체계와 표준 피처의 동시출현 네트워크 분석결과를 바탕으로 주요 연결구조와 시기별 변화를 비교하였다. 전체 기간 네트워크 규모는 IPC 177노드/1,054간선, CPC 220/2,113, AI 기술분류체계 112/793, 표준 피처 41/721, AI 기술분류체계-표준 피처 153/2,098이었다. Table 4는 각 네트워크의 규모와 주요 연결요소를 비교한 결과이다.

IPC와 CPC에서는 네트워크 보안과 탐지, 기록, 대응 관련 분류가 높은 연결강도를 보였다. AI 기술분류체계에서는 특징표현 및 표현학습이 가장 높은 총연결강도를 보였고, CNN과 통계 및 임계값 판정이 뒤를 이었다. 표준 피처 네트워크에서는 통계특징, 패킷길이, 시간간격이 주요 연결요소로 나타났으며, AI 기술분류체계-표준 피처 이분망에서는 통계 및 임계값 판정과 특징표현·표현학습이 통계특징, 패킷길이, 시퀀스특징 등과 폭넓게 함께 나타났다. 즉 네트워크 보안이라는 응용영역 위에서 통계적 판단, 표현학습, 프로토콜·시간·길이 정보가 서로 다른 분석 수준에서 반복적으로 함께 나타나는 연결구조가 확인되었다.

2022-2025년 고유 간선 중 2006-2021년 누적망에 없던 관계의 비율은 표준 피처 25.8%, AI 기술분류체계 86.4%, AI 기술분류체계-표준 피처 69.7%였다. 이는 최근 시기에 이전에는 관찰되지 않았던 기술요소 간 동시출현 관계가 다수 나타났음을 보여 준다. 다만 신규 간선 비율은 최근 출원 공보의 집중과 네트워크별 구성 차이를 고려하여 해석할 필요가 있다.

Table 4. Structural Characteristics of Co-occurrence Networks

Network	Nodes / edges	Density	Leading nodes (total link strength)	Major co-occurrences (weight)
IPC	177 / 1,054	0.068	H04L9/40 (639) G06N3/08 (267) G06N3/0464 (218)	G06N3/0464 - H04L9/40 (41) G06N3/08-H04L9/40 (41) G06F18/214-H04L9/40 (26)
CPC	220 / 2,113	0.088	H04L63/1416 (1,018) H04L63/1425 (994) H04L63/1441 (455)	H04L63/1416 - H04L63/1425 (119) H04L63/1416-H04L63/1441 (46) H04L63/0428-H04L63/1416 (44)
AI technology taxonomy	112 / 793	0.128	Feature Representation, Representation Learning & Embeddings (187) ConvolutionalNeuralNetworks(CNN) (177) Statistics/Threshold/Scoring (154)	Protocol/HandshakeAnalysis - Statistics/Threshold/Scoring (25) Mitigation/PolicyControl-Statistics/Threshold/Scoring (24) Mitigation/PolicyControl-Protocol/HandshakeAnalysis (23)
Standardized features	41 / 721	0.879	Statistical Features (1,441) Packet Length (1,177) Inter-arrival Time (957)	Statistical Features - Packet Length (129) Inter-arrivalTime-PacketLength (107) Inter-arrivalTime-StatisticalFeatures (105)
AI technology taxonomy-standardized features	153 / 2,098	0.457	**Feature:** Statistical Features (725); Packet Length (552); Inter-arrival Time (451) **AI:**Statistics/Threshold/Scoring(562); Mitigation/Policy Control(495);RepresentationLearning (484)	Statistics/Threshold/Scoring - Statistical Features (57) Statistics/Threshold/Scoring-BaselineFeatures (52) Mitigation/PolicyControl-HeaderFields (45)

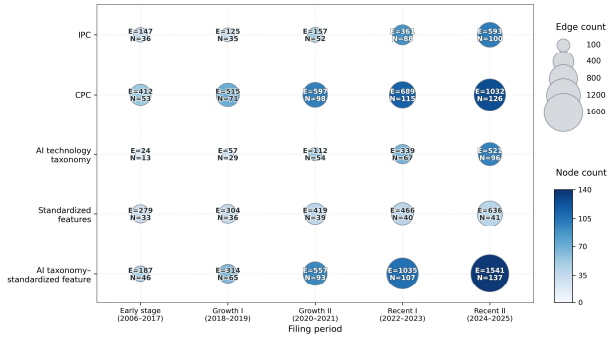


Fig. 7. Co-occurrence Network Size by Filing Period

Figure 7은 5개 출원 시기별 네트워크의 노드 수 (N)와 간선 수 (E)를 비교한 결과이다. 최근후반 (2024-2025)에는 IPC 100노드/593간선, CPC 126/1,032, AI 기술분류체계 96/521, 표준 피처 41/636, AI 기술분류체계-표준 피처 137/1,541로 나타났다. 전반적으로 최근 구간으로

www.kci.go.kr

갈수록 네트워크에서 관찰되는 기술요소와 동시출현 관계가 확대되는 양상을 보였다.

네트워크별 변화에는 차이가 있었다. AI 기술분류체계는 초창기 13노드/24간선에서 최근후반 96/521로 확대되었다. 반면 표준 피처는 성장후반에 이미 39개가 관찰되어 최근후반 41개와 큰 차이가 없었지만, 간선 수는 419개에서 636개로 증가하였다. 이는 새로운 피처의 추가보다 기존 피처 간 동시출현 조합의 확대가 두드러졌음을 보여 준다. AI 기술분류체계-표준 피처 네트워크도 초창기 46노드/187간선에서 최근후반 137/1,541로 확대되어 기술항목과 피처 간 연결 범위가 넓어졌다.

아래 Figure 8, 9, 10은 전체기간 IPC, CPC, AI 기술분류체계의 동시출현 관계를 시각화하여 네트워크 지도로 제시한 것이다. 노드 크기는 총연결강도, 간선은 동일 공보에서의 동시출현 관계를 나타내며, 색상은 연결구조에 따른 군집을, 노드 간 상대적 거리는 분류항목 사이의 연관성을 시각적으로 나타낸다.

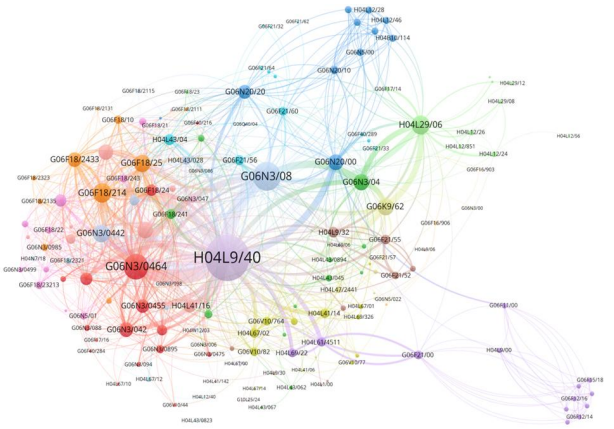


Fig. 8. IPC Code Co-occurrence Network

Figure 8은 분석대상 특허공보에서 IPC 기술분류가 동일 공보에 동시출현하는 구조를 보여 준다. 전체기간에서 H04L9/40 (보안 통신/네트워크 보안 관련)의 총연결강도는 639로 가장 높았고, G06N3/08 (신경망 학습 방법) 267, G06N3/0464 (합성곱 신경망) 218이 뒤를 이었다. 이는 분석대상 특허공보가 네트워크 보안 관련 기술영역을 중심으로 학습 및 분류 관련 기술영역과 반복적으로 결합되어 있음을 보여 준다. 이러한 IPC 수준의 연결구조는 암호 트래픽 악성행위 탐지가 네트워크 보안과 인공지능 기반 분석기술이 교차하는 기술구조로 형성되어 있음을 확인할 수 있다.

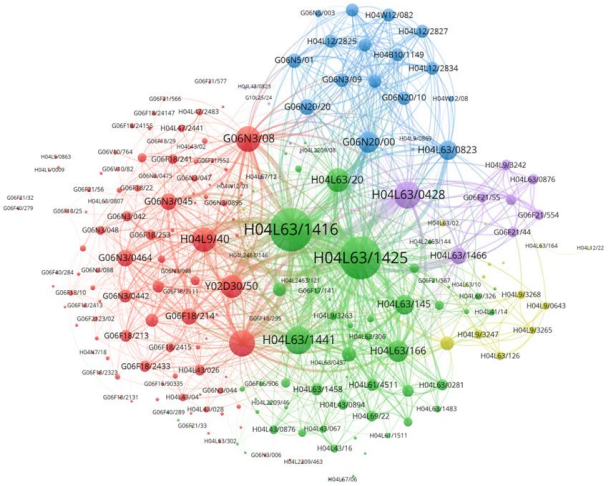


Fig. 9. CPC Code Co-occurrence Network

Figure 9는 IPC보다 세분화된 CPC 기술분류의 동시출현하는 구조를 보여 준다. H04L63/1416 (악성 트래픽 시그니처 탐지)의 총연결강도는 1,018로 가장 높았고, H04L63/1425 (이상 탐지 및 트래픽 로깅) 994, H04L63/1441 (악성 트래픽에 대한 대응책) 455였다. 이러한 구조는 분석대상 특허공보에서 탐지, 기록 및 대응 관련 보안 기능 관련 분류가 동일 공보에서 반복적으로 함께 출현하고 있음을 보여 준다. 따라서 CPC 네트워크는 암호 트래픽 악성행위 탐지 관련 특허가 탐지뿐 아니라 기록·모니터링·대응 등 연관 보안기능과도 밀접하게 연결되어 있음을 보여 준다.

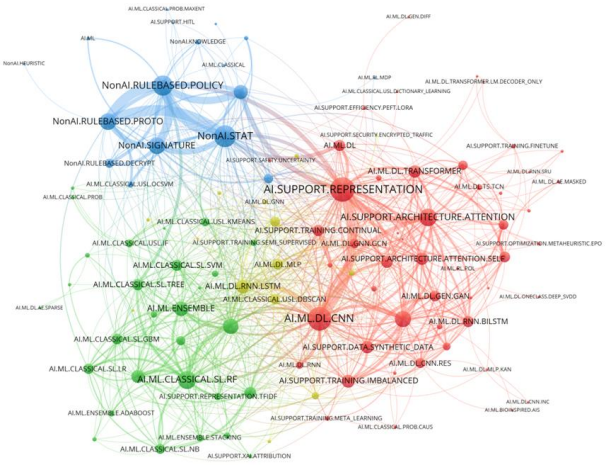


Fig. 10. AI Taxonomy Code Co-occurrence Network

Figure 10은 AI 기술분류체계의 기술항목이 동일 특허공보에서 형성하는 동시출현 구조를 보여 준다. 총연결강도는 Feature representation / representation learning (특징표현 및 표현학습)이 187로 가장 높았고, CNN (합성곱

www.kci.go.kr

신경망) 177, Statistical / threshold scoring (통계 및 임계값 판정) 154, Policy / rule-based control (정책·규칙 기반 제어) 130, RF (Random Forest) 106, Attention (어텐션) 103, Protocol / handshake analysis (프로토콜·핸드셰이크 분석) 94, Rule / signature matching (규칙·시그니처 매칭) 93 순이었다. 대표적인 동시출현 관계는 프로토콜·핸드셰이크 분석-통계·임계값 판정 25건, 정책·규칙 기반 제어-통계·임계값 판정 24건, 규칙·시그니처 매칭-통계·임계값 판정 14건과 같았다.

이러한 결과는 규칙·통계 기반 판정, 고전 ML, DL 기반 표현학습 등 서로 다른 기술계열의 요소가 동일 특허공보에서 함께 나타나며 다양한 연결관계를 형성하고 있음을 보여 준다. 특히 특징표현·표현학습과 CNN 등 DL 관련 항목이 높은 총연결강도를 보이는 가운데, 통계·임계값 판정과 정책·규칙 기반 기술도 주요 연결요소로 나타나, 다양한 탐지방식이 공존하는 복합적인 기술구조를 보여 준다.

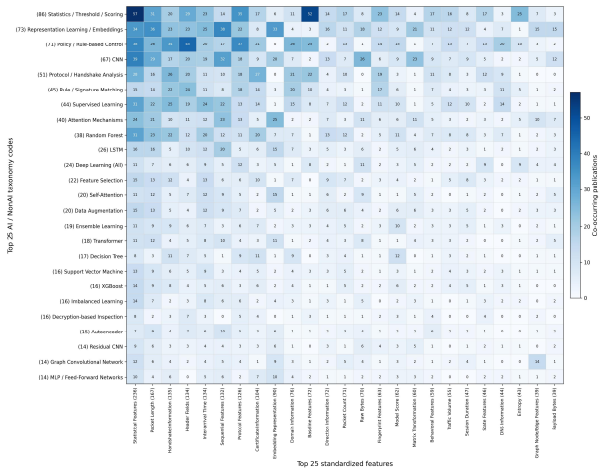


Fig. 11. AI Technology Taxonomy–Standardized Feature Co-occurrence Heatmap

Figure 11은 전체 392개 특허공보에서 확인된 AI 기술분류체계 항목 112개와 표준 피처 41개 가운데 공보 출현 빈도가 높은 항목을 각 25개씩 선정하여 동일 공보 내 동시출현 관계를 25 × 25 행렬로 나타낸 것이다. 행과 열의 괄호 안 수치는 각 항목이 나타난 전체 공보 수이고, 셀의 수치와 색상 농도는 두 항목이 같은 공보에서 함께 나타난 횟수를 뜻한다.

동시출현 구조에서는 여러 탐지방식이 공유하는 기초 관측정보와 기술계열별 특징적인 피처 구성이 함께 확인되었다. 통계 및 임계값 판정은 통계특징 57건과 기준선특징 52건에서 높은 동시출현을 보였다. RF는 통계특징 31건, 패킷길이 23건, 핸드셰이크정보 22건, 인증서정보 20건과 함께 나타났다. 규칙 및 시그니처 계열은 헤더필드 24건, 핸드셰이크정보 22건, 도메인정보 20건, 인증서정보 14건과 반복적으로 동시출현하였다. 한편 특징표현 및 표현학습은 시퀀스특징 38건, 패킷길이 36건, 통계특징 34건, 임베딩표현 33건과 함께 나타났고, CNN은 통계특징 39건, 시퀀스특징 32건, 패킷길이 29건, 원시바이트 26건, 임베딩표현 20건과 동시출현하였다. 통계, 길이, 프로토콜 정보가 여러 기술에서 공통적으로 관찰되는 가운데, 판정기술에 따라 함께 나타나는 정보 구성에는 차이가 있었다.

LSTM, self-attention, Transformer 등의 순차 및 문맥형 기술은 시간간격, 시퀀스특징, 행렬변환, 임베딩표현 등의 표준 피처와 반복적으로 동시출현하였다. 특히 그래프 합성곱 신경망 (Graph Convolutional Network; GCN)은 Graph Node/Edge Features와 14개 공보에서, Graph Structural Features와 9개 공보에서 동시출현하여 그래프 기반 구조정보와의 결합이 두드러졌다.

이는 암호 트래픽 악성행위 탐지기술에서 공통의 트래픽 관측정보를 활용하면서도 특징정보를 구성하고 이를 분석 및 판정기술과 연계하는 방식이 다양함을 시사한다.

V. Discussion

분석대상 특허공보의 70.7%가 2021-2025년에 출원되었으며, 중국 소재 출원기관의 참여도 크게 증가하였다. 대학과 공공연구기관에서는 DL형과 복합형, 보안전문기업과 통신·네트워크기업에서는 고전 ML형과 규칙·통계형의 비중이 상대적으로 높았다. 기술군의 상대적 중심은 규칙·통계형과 고전 ML형에서 DL 기반 표현학습형과 복합모델·특징융합형으로 이동하였으며, 두 기술군은 최근후반 (2024-2025) 공보의 80.3%를 차지하였다. 비복호화 접근은 전체의 92.9%였다.

이러한 비중 변화에도 규칙·통계형과 고전 ML형은 최근까지 관찰되었으며, 26개 세부유형 중 23개가 2021년까지 확인되었다. 따라서 최근의 기술 변화는 기존 탐지방식의 단순한 대체라기보다 여러 방식이 공존하는 가운데 표현학습과 모델·특징 융합을 활용하는 유형의 비중이 커지면서 기술구성이 재편되는 양상으로 해석할 수 있다. 이는 표현학습, Transformer와 GNN, 사전학습 및 융합기술을 다룬 선행연구들과 맥락을 같이한다. 본 연구는 출원 시기와 출원주체별 기술구성의 차이, 발명 구성요소 간 관계를 분석하여 이러한 변화를 구체화함으로써 기존 연구를 보완하였다.

데이터 활용 측면에서는 21개 트래픽 데이터 단위 및 표현과 41개 표준 피처 사이에서 342개 유형의 직접 사용관계

www.kci.go.kr

가 확인되었으며, 공보별 관계 출현 횟수는 총 3,439회였다. 규칙·통계형과 고전 ML형에서 주요하게 나타난 프로토콜 관련 정보와 통계정보는 DL형과 복합형에서도 폭넓게 관찰되었다. 다만 이러한 정보를 표현하고 판정에 활용하는 방식에는 차이가 있었다. 기술군의 분류기준을 함께 살펴보면 규칙·통계형은 관측정보에 규칙이나 통계적 기준을 적용하여 판정하고, 고전 ML형은 명시적으로 구성한 피처를 분류에 활용한다. DL형은 시퀀스, 행렬, 그래프 등 다양한 형태로 표현된 데이터에서 특징을 학습하며, 복합형은 구별되는 모델이나 분석단계의 결과, 서로 다른 관점의 특징표현 또는 분산 학습모델의 정보를 결합한다. 이처럼 탐지기술의 구조적 차이를 이해하려면 관측정보의 종류뿐 아니라 그 표현형태와 판정방식도 함께 살펴볼 필요가 있다.

동시출현 네트워크에서는 동일 공보에 함께 나타나는 기술요소 간 연결구조를 확인하였다. IPC에서는 네트워크 보안과 학습 및 분류 기술이, CPC에서는 탐지, 기록, 모니터링, 대응 관련 분류가 주요 연결을 형성하였다. AI 기술분류체계에서는 특징표현·표현학습과 CNN 등 DL 관련 항목뿐 아니라 통계·임계값 판정과 정책·규칙 기반 기술도 주요 연결요소로 나타났다. AI 기술분류체계와 표준 피처 간 동시출현 네트워크에서는 통계특징, 패킷길이, 프로토콜 관련 정보가 여러 탐지기술에서 공통적으로 관찰되었으며, 피처와의 동시출현 양상은 탐지기술에 따라 차이를 보였다. 2022-2025년 고유 간선 중 2006-2021년 누적망에 없던 관계의 비율은 표준 피처 25.8%, AI 기술분류체계 86.4%, AI 기술분류체계-표준 피처 69.7%였다. 성장후반과 최근후반을 비교하면 표준 피처의 종류가 크게 늘지 않았는데도 피처 간 간선 수는 증가하였다. 이는 기존 관측정보를 활용하는 피처와 탐지기술의 조합이 다양해졌음을 시사한다.

다만 신규 간선의 비율을 해석할 때에는 시기별 특허공보 수에 따라 동시출현의 관측 기회가 달라질 수 있음을 고려해야 한다. 국가 및 기관유형별 기술군 구성 역시 출원 시기와 참여 규모, 공동출원에 따른 중복 참여의 영향을 받을 수 있다. 따라서 본 연구에서 확인한 기술구성과 출원 동향의 변화만으로 그 원인이나 국가 및 기관유형 간 기술적 우열을 직접 판단하기는 어렵다.

향후 연구개발에서는 특허에 제안된 기술구성이 다양한 실제 환경에서도 유효한지 검증할 필요가 있다. 전체 392건 중 자체 데이터만 활용한 공보는 325건 (82.9%)이었으며, 데이터 확보방식에서는 운영·관측 수집이 261건 (66.6%)으로 가장 많았다. 예를 들어, 등록특허 CN118337533B [29]는 실제 네트워크에서 수집한 정상·악성 트래픽과 모의공격 환경에서 확보한 악성 트래픽을 공개 데이터셋과 함께 활용한다. 이 특허에서 제안한 방법은 양방향 세션의 초기 30개 패킷에서 유효부하 크기, 시간간격, 전송 방향 등의 특징을 추출하고, 이를 LSTM과 Multi-Head Self-Attention으로 분석하여 악성 여부를 판정한다. 이러한 사례와 데이터 편향 및 일반화 문제를 다룬 선행연구를 고려하면, 후속 연구에서는 동일한 패킷 특징을 이용한 모델 구성별 성능을 비교하고 데이터, 네트워크 및 프로토콜 환경 변화에 따른 일반화와 강건성을 검증할 필요가 있다. 아울러 제한된 자원에서의 연산 효율, 실시간 처리 성능, 오탐, 설명가능성 등 운영 적용성도 평가해야 한다. 이는 최근 비중이 커진 DL형과 복합형 기술의 실제 활용 가능성을 판단하는 데 중요한 과제다.

본 연구의 데이터 유형 및 탐지기술 분류체계는 트래픽 데이터의 분석 단위와 표현 형태를 더욱 명확히 구분하고 기술 발전과 적용 범위의 변화를 반영하여 지속적으로 보완하는 한편, 분석대상 요소를 정확하고 일관되게 식별 및 분류할 수 있도록 기준과 절차를 정교화할 필요가 있다.

VI. Conclusion

본 연구는 지난 20년간 출원된 암호 트래픽 악성행위 탐지 관련 특허공보 392건을 선별하고 기술요소를 구조화한 분석 데이터셋을 구축하였다. 이를 바탕으로 출원 및 출원주체 동향, 데이터 활용, 5대 기술군과 26개 세부유형의 변화, 특허 원문에서 확인한 직접 사용관계와 주요 기술요소의 동시출현 구조를 분석하였다. 분석 결과, 비복호화 접근이 높은 비중을 차지하였으며, 기술군의 상대적 중심은 규칙·통계형과 고전 ML형에서 DL 기반 표현학습형과 복합모델·특징융합형으로 이동하였다. 그러나 기존 기술군도 최근까지 관찰되어, 기술 변화는 특정 방식의 단순한 대체보다 여러 탐지방식이 공존하는 가운데 기술구성이 재편되는 양상으로 나타났다. 또한 탐지기술에 따라 관측정보를 구성하고 판정에 활용하는 방식에 차이가 있었으며, 최근에는 기술요소 간 동시출현 관계도 다양해졌다.

본 연구의 의의는 출원 동향과 시기별 기술구성의 변화를 분석하고, 발명 구성요소 간 직접 사용관계와 동일 공보 내 동시출현 관계를 구분하여 살펴보았다는 데 있다. 이를 통해 기존 학술연구를 보완하는 특허 기반의 기술동향 및 기술구조 분석 결과를 제시하였다. 본 연구의 분석체계와 결과는 향후 탐지기술의 변화와 기술요소 간 연계 양상을 파악하고 후속 연구개발 방향을 구체화하는 데 활용될 수 있다.

www.kci.go.kr

ACKNOWLEDGEMENT

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) funded by Korean Government [Ministry of Science and Information and Communications Technology (MSIT)] through the Development of Security Monitoring Technology-Based Network Behavior Against Encrypted Cyber Threats in ICT Convergence Environment under Grant RS-2023-00235509.

REFERENCES

[1] A. Sharma and A. H. Lashkari, "A survey on encrypted network traffic: A comprehensive survey of identification/classification techniques, challenges, and future directions," Computer Networks, Vol. 257, Art. no. 110984, 2025. DOI: 10.1016/j.comnet.2024.110984.

[2] C. Oh, J. Ha, and H. Roh, "A Survey on TLS-Encrypted Malware Network Traffic Analysis Applicable to Security Operations Centers," Applied Sciences, Vol. 12, No. 1, Art. no. 155, 2022. DOI: 10.3390/app12010155.

[3] Z. Wang, K.-W. Fok, and V. L. L. Thing, "Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study," Computers & Security, Vol. 113, Art. no. 102542, 2022. DOI: 10.1016/j.cose.2021.102542.

[4] B. Gong, J. Liu, X. Tang, M. Yu, H. Gong, and M. Huang, "Intelligent Analysis Technologies for Encrypted Traffic: Current Status, Advances, and Challenges," Journal of Electronics & Information Technology, Vol. 48, No. 3, pp. 1180-1197, 2026. DOI: 10.11999/JEIT250416.

[5] W. Dong, J. Yu, X. Lin, G. Gou, and G. Xiong, "Deep learning and pre-training technology for encrypted traffic classification: A comprehensive review," Neurocomputing, Vol. 617, Art. no. 128444, 2025. DOI: 10.1016/j.neucom.2024.128444.

[6] Y. Kim, C. Lee, and K. Sohn, "A Survey of AI-Based Encrypted Traffic Detection: Multi-Level Taxonomy and Structural Analysis of Intent-Behavior-Model Coupling," Computer Modeling in Engineering & Sciences, Vol. 148, No. 1, Art. no. 5, 2026. DOI: 10.32604/cmes.2026.083669.

[7] I. H. Ji, J. H. Lee, M. J. Kang, W. J. Park, S. H. Jeon, and J. T. Seo, "Artificial Intelligence-Based Anomaly Detection Technology over Encrypted Traffic: A Systematic Literature Review," Sensors, Vol. 24, No. 3, Art. no. 898, 2024. DOI: 10.3390/s24030898.

[8] T. van Rijn and J. K. Timmis, "Patent landscape analysis-Contributing to the identification of technology trends and informing research and innovation funding policy," Microbial Biotechnology, Vol. 16, No. 4, pp. 683-696, 2023. DOI: 10.1111/1751-7915.14201.

[9] J. Hou, H. Lu, F.-A. Liu, X.-W. Wang, and Z.-H. Tian, "Detection and Countermeasure of Encrypted Malicious Traffic: A Survey," Journal of Software, Vol. 35, No. 1, pp. 333-355, 2024. DOI: 10.13328/j.cnki.jos.006891.

[10] K. Keshkeh, A. Jantan, K. Alieyan, and U. M. Gana, "A Review on TLS Encryption Malware Detection: TLS Features, Machine Learning Usage, and Future Directions," in Advances in Cyber Security, Communications in Computer and Information Science, Vol. 1487, pp. 213-229, Springer, 2021. DOI: 10.1007/978-981-16-8059-5_13.

[11] I. A. Alwhbi, C. C. Zou, and R. N. Alharbi, "Encrypted Network Traffic Analysis and Classification Utilizing Machine Learning," Sensors, Vol. 24, No. 11, Art. no. 3509, 2024. DOI: 10.3390/s24113509.

[12] A. P. Singh and M. Singh, "A Comparative Review of Malware Analysis and Detection in HTTPs Traffic," International Journal of Computing and Digital Systems, Vol. 10, No. 1, pp. 111-123, 2021. DOI: 10.12785/ijcds/100111.

[13] S. Han, H. Zhang, M. Lu, S. Guo, B. Tian, and J. Wang, "A comprehensive survey on encrypted network traffic classification," Computer Networks, Vol. 287, Art. no. 112524, 2026. DOI: 10.1016/j.comnet.2026.112524.

[14] Y. Fu, T. Liu, K. Wang, and Y. Yu, "Survey of research on encrypted traffic classification based on machine learning," Journal on Communications, Vol. 46, No. 1, pp. 167-191, 2025. DOI: 10.11959/j.issn.1000-436x.2025006.

[15] Y. Choi, J. Yoo, K. Koo, and D. Moon, "Trends of Encrypted Network Traffic Analysis Technologies for Network Anomaly Detection," Electronics and Telecommunications Trends, Vol. 38, No. 5, pp. 71-80, 2023. DOI: 10.22648/ETRI.2023.J.380507.

[16] B. Kim, M. Park, B.-J. You, and J. Lee, "Analysis of Academic Research on Encrypted Malicious Traffic Detection," Proceedings of the Korea Society of Computer Information Conference, Vol. 33, No. 1, pp. 495-497, Cheonan, Republic of Korea, January 2025.

[17] A. Block and C. H. Song, "Exploring the characteristics of technological knowledge interaction dynamics in the field of solid-state batteries: A patent-based approach," Journal of Cleaner Production, Vol. 353, Art. no. 131689, 2022. DOI: 10.1016/j.jclepro.2022.131689.

[18] H.-S. Noh, D.-H. Noh, and T. Kim, "Patent-Based Trend Analysis of Research on Promising Technologies: A PATC Framework Approach," Journal of Intellectual Property, Vol. 20, No. 4, pp. 165-199, 2025. DOI: 10.34122/jip.2025.20.4.165.

[19] S. Sakaoka and S. Kano, "Quantitative insights on artificial intelligence in the pharmaceutical industry: A patent-basis analysis of technological trends and key players," World Patent Information, Vol. 82, Art. no. 102381, 2025. DOI: 10.1016/j.wpi.2025.102381.

[20] Z. Abbas, L. Yong, Y. Li, and R. Wang, "Patent-based trend analysis for advanced thermal energy storage technologies and their applications," International Journal of Energy Research, Vol. 44, No. 7, pp. 5093-5116, 2020. DOI: 10.1002/er.5148.

www.kci.go.kr

[21] W. G. Vijvers, "The international patent classification as a search tool," World Patent Information, Vol. 12, No. 1, pp. 26-30, 1990. DOI: 10.1016/0172-2190(90)90285-S.

[22] T. Montecchi, D. Russo, and Y. Liu, "Searching in Cooperative Patent Classification: Comparison between keyword and concept-based search," Advanced Engineering Informatics, Vol. 27, No. 3, pp. 335-345, 2013. DOI: 10.1016/j.aei.2013.02.002.

[23] B. Degroote and P. Held, "Analysis of the patent documentation coverage of the CPC in comparison with the IPC with a focus on Asian documentation," World Patent Information, Vol. 54, pp. S78-S84, 2018. DOI: 10.1016/j.wpi.2017.10.001.

[24] D. H. Kim and J. H. Jeon, "Analysis of Patent Convergence Structure and Network Characteristics in Smart Firefighting Technologies: Focusing on Topic Modeling and CPC Co-Occurrence Network," Journal of Industrial Innovation, Vol. 42, No. 1, pp. 1-22, 2026. DOI: 10.22793/indinn.2026.42.1.001.

[25] S. Cho, D. Kim, and K. Cho, "A Comparative Study of the Technological Innovation Competitive Structure in the Global Display Industry: An IPC-Based Patent Portfolio Analysis," Journal of Industrial Innovation, Vol. 42, No. 1, pp. 217-248, 2026. DOI: 10.22793/indinn.2026.42.1.016.

[26] J. Sun, Y. Zhai, X. Lin, D. Zhang, R. Miao, and M. Li, "Patent-based technology opportunity identification via semantic-structural coupling entropy: A retrospective validation study of unmanned surface vehicle technologies," World Patent Information, Vol. 85, Art. no. 102456, 2026. DOI: 10.1016/j.wpi.2026.102456.

[27] J. Han and C.-B. Kim, "Exploring the convergent structure of Korea's space and port technologies using IPC co-occurrence, patent network, and cluster analyses," Asian Journal of Technology Innovation, pp. 1-32, 2026. DOI: 10.1080/19761597.2026.2667488.

[28] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," Scientometrics, Vol. 84, No. 2, pp. 523-538, 2010. DOI: 10.1007/s11192-009-0146-3.

[29] China Academy of Information and Communications Technology (CAICT), "A method and system for detecting malicious encrypted traffic based on data packet representation learning," Chinese Patent CN118337533B, Sep. 20, 2024.

Authors



Byungkyu Kim received his B.S. and M.S. in Computer Science from Chungnam National University in 2001 and 2003, respectively, and his Ph.D. in Computer Science & Engineering from the same university in 2018. Dr. Kim joined the Korea Institute of Science and Technology Information (KISTI) in 2002. He is currently a Principal Researcher in the AI Digital Security Research Team at KISTI. His research interests include scientometrics, disaster and safety management, and cybersecurity.



Min-Woo Park received the B.S. and M.S. degrees from the Division of Computer Convergence, Chungnam National University, South Korea, in 1992 and 2004, respectively. Since 1996, he has been with the Korea Institute of Science and Technology Information (KISTI) in South Korea. He is currently a Principal Researcher in the AI Digital Security Research Team. His research interests include system architecture, information security, the Internet of Things, smart city, and artificial intelligence.



Beom-Jong You received a B.S. in Electronic Engineering from Sogang University, Korea, in 1984, and M.S. and Ph.D. degrees in Library & Information Science from Chungnam National University, Korea, in 2000 and 2005, respectively. Dr. You joined the System Engineering Research Institute in Korea in 1987. He is currently a Research Fellow in the AI Digital Security Research Team at KISTI. His research interests include Data Science, disaster & safety platforms, and cybersecurity.



Taewoong Kwon received the B.S. degree in computer science and engineering from Soongsil University, Seoul, South Korea, in 2012, and the M.S. degree in information security from Korea University, Seoul, in 2014. He has been a Senior Researcher at the Korea Institute of Science and Technology Information (KISTI) in Daejeon, South Korea, since 2014. He currently serves as the Head of the AI Digital Security Research Team. His research interests include information security, network security, and data visualization.



Jun Lee received his B.S. in Computer Science from Korea Aerospace University in February 2010, his master's degree in Information Engineering in February 2012, and his doctorate in Information Engineering in August 2017, all from Korea Aerospace University. Dr. Lee was a Research Assistant with the Artificial Intelligence Cloud Research Team, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan, from April 2017 to November 2017. From December 2017 to December 2019, he was a Postdoctoral Researcher with the Artificial Intelligence Research Center (AIRC), AIST. He is currently the Director of the Science and Technology Security Research Center, Korea Institute of Science and Technology Information (KISTI). His research interests include knowledge discovery, natural language processing, data mining, and cybersecurity.

www.kci.go.kr