JDFR Journal of Data Forensics Research
데이터포렌식연구, 3권 1호, 2026년 6월, 55-76면
Journal of Data Forensics Research, Volume 3, Issue 1, June 2026, Pages 55-76.
https://doi.org/10.12972/JDFR.2026.3.1.3
Print ISSN 3058-9460
Online ISSN 3058-9479

원저

한국 피싱 범죄 7단계 킬 체인의 실증적 정립: 판결문 범죄사실 방향성 시퀀스 분석을 중심으로

한상현[1], 김지온[2]

[1]한림대학교 인공지능융합학부 학사 과정, [2]한림대학교 융합과학수사학과 교수

교신저자: 김지온, jion972@hallym.ac.kr

요약

본 연구는 한국 피싱 범죄의 구조를 7단계 킬 체인으로 실증적으로 정립하기 위해, 보이스피싱·스미싱·메신저피싱·투자리딩방·로맨스스캠·팀미션 사기의 총 6개 피싱 유형에 해당하는 형사 판결문 290건에서 「범죄사실」 텍스트를 추출하고, TF-IDF 기반 상향식 수법 사전 구축과 규칙 기반 방향성 시퀀스 추출을 통해 방향성 가중 네트워크를 구성하였다. Hutchins et al.(2011)의 침입 킬 체인(Intrusion Kill Chain)을 피싱 도메인에 맞게 재정의한 7단계(정찰/대상선정, 접촉/유인, 신뢰 구축/사칭, 통제/지시, 자금이전, 세탁/은닉, 흔적제거)를 분석 골격으로 삼아, 34개 정규화 수법 노드와 448개 방향성 엣지로 구성된 킬 체인 네트워크를 구축하였다. 분석 결과, 단계 간 전이는 순방향 45.6%, 동일 단계 유지 20.0%, 역방향 34.4%로 나타났다. 이는 한국형 피싱 킬 체인이 1단계에서 7단계로 엄격하게 진행하는 선형 사슬이 아니라, 전반적인 순방향 진행 경향 위에서 특정 단계를 중심으로 빈번히 순환하는 구조임을 보여준다. 또한 판결문이라는 단일 출처는 자금이전(5단계)·세탁/은닉(6단계)에 서술이 집중되는 반면, 정찰(1단계, 12.1%)과 흔적제거(7단계, 7.4%)는 상대적으로 희소하게 기재된다는 정보 밀도 편향 또한 확인되었다. 본 연구는 기존 사기 수법 네트워크 연구의 무방향 공출현 분석 한계를 방향성 시퀀스 분석으로 극복한 첫 한국어 피싱 범죄 실증 연구로서, 수사 실무의 단계별 차단 전략 수립에 활용 가능한 구조적 토대를 제공한다.

주제어

피싱 범죄, 킬 체인, 판결문 분석, TF-IDF, 방향성 시퀀스 네트워크

Received: June 08, 2026
Revised: June 29, 2026
Accepted: June 29, 2026
Published: June 30, 2026

Original Article

Empirical establishment of a 7-stage kill chain for Korean phishing crime: A directional sequence analysis of criminal facts in court verdicts

Sanghyun Han[1], Jion Kim[2]

[1]Undergraduate Student, Department of Artificial Intelligence Convergence, Hallym University, Republic of Korea
[2]Professor, Department of Forensic Science, Hallym University, Republic of Korea

Corresponding Author: Jion Kim, jion972@hallym.ac.kr

ABSTRACT

This study empirically established a 7-stage kill chain framework for Korean phishing crimes. Drawing on 290 criminal court verdicts covering six phishing types (voice phishing, smishing, messenger phishing, investment-recommendation rooms, romance scams, and team-mission fraud), we extracted the "criminal facts" section from each verdict and constructed a directed weighted network using a term frequency–inverse document frequency (TF-IDF)-based bottom-up method dictionary and rule-based directional sequence extraction. Adapting Hutchins et al.'s (2011) intrusion kill chain to the phishing domain, we defined seven stages: target reconnaissance, contact/luring, trust-building/impersonation, control/direction, fund transfer, laundering/concealment, and trace removal. The resulting network comprised 34 normalized method nodes and 448 directed edges. Stage transitions were 45.6% forward, 20.0% same-stage, and 34.4% backward, demonstrating that the Korean phishing kill chain is not strictly sequential but a structure that, while showing an overall forward progression, frequently cycles through particular stages. Information density analysis revealed that court verdicts were heavily concentrated on fund transfer (Stage 5) and laundering (Stage 6), with reconnaissance (Stage 1: 12.1%) and trace removal (Stage 7: 7.4%) being sparsely documented. This study is the first empirical Korean-language phishing crime analysis to overcome the co-occurrence limitations of prior studies through directional sequence modeling, providing a structural foundation for evidence-based interdiction strategies.

KEYWORDS

Phishing Crime, Kill Chain, Court Verdict Analysis, TF-IDF, Directed Sequence Network

I. 서론

1.1. 연구 배경 및 필요성

사기 범죄는 오늘날 한국에서 가장 빠르게 증가하는 범죄 유형 중 하나다. 2019년부터 2024년까지 6년간 피싱 관련 사기의 누적 피해액은 약 155조 원을 상회하며, 2024년 한 해에만 약 28조 원을 넘어섰다[1]. 같은 기간 사기 범죄가 전체 범죄에서 차지하는 비중은 18.9%에서 26.6%로 지속 상승한 반면, 검거율은 73.9%에서 60.2%로 급락하였다[2]. 이러한 역전 현상은 사기 수법이 단일 수법에서 복합·연쇄 수법으로 빠르게 진화하고 있음을 보여주며, 특히 최근에는 악성앱 설치와 스미싱을 결합한 디지털 채널 기반 수법이 빠르게 확산되고 있다[3]. 이처럼 수법이 고도화되는 가운데, 사후 검거 중심의 현행 대응 전략은 구조적 한계에 도달했음을 시사한다.

이 문제에서 핵심적으로 빠져 있는 것은 피싱 범죄가 어떤 순서로, 어떤 단계를 거쳐 완성되는가에 대한 실증적 분석과 검증이다. 경찰청이 제공하는 '10대 악성사기' 분류는 사기 유형을 목록화하지만, 범행 단계 간의 선후 관계나 차단 우선순위를 제시하지는 않는다[4]. 기존의 지식그래프 기반 사이버 범죄 분류 연구는 키워드 간의 정적인 관계를 정의하는 공출현 분석에 주로 의존하고 있어, 수법 간의 시간적 방향성을 분석에 충분히 활용하지 못하고 있다[5]. 방향성이 없으면 어느 지점에서 개입해야 하는지 알 수 없으므로, 차단 전략은 구조의 이해에서 출발해야 한다.

1.2. 문제 제기 및 연구 목적

본 연구는 이 공백을 메우기 위해 사이버보안 분야에서 정립된 침입 킬 체인(Intrusion Kill Chain; 공격이 표적 식별부터 최종 목적 달성까지 거치는 일련의 단계를 순서대로 연결한 분석 틀, Hutchins et al. 2011[6]) 개념을 피싱 사기 도메인으로 적응시킨 7단계 킬 체인을 정립하고, 형사 판결문 290건에서 추출한 방향성 시퀀스로 그 구조를 실증적으로 검증한다. 분석 대상은 보이스피싱·스미싱·메신저피싱의 전통 피싱 3종과 투자리딩방·로맨스스캠·팀미션 사기의 신종 피싱 3종으로 구성하였다. 이들은 가짜 신원 구축, 비대면 디지털 채널 활용, 단계적 신뢰 누적 후 자금 탈취라는 공통 구조를 가지므로 단일 킬 체인 프레임 안에서 비교 분석하기에 적합하다.

본 연구를 통해 검증할 세 가지 가설은 다음과 같다.

첫째(Q1), 사이버보안에서 정립된 킬 체인이 피싱이라는 정보통신망 이용(cyber-enabled) 범죄에도 적용될 것이다. 침입 킬 체인은 본래 시스템 공격을 분석하기 위한 틀인데, 인간의 신뢰를 공격하는 피싱에도 7단계 구조가 실제로 나타나는지를 290건의 판결문 데이터로 검증한다. 기존 연구가 가정에 머물렀던 단계 구조를 실데이터로 증명하려는 것이다.

둘째(Q2), 판결문이라는 단일 출처는 어느 단계를 잘 담고 어느 단계를 놓치는가? 데이터 출처의 한계를 명시적으로 측정함으로써, 향후 어떤 데이터를 보완해야 하는지에 대한 실천적 함의를 도출한다.

셋째(Q3), 피싱 킬 체인의 단계 전이는 사이버보안의 침입 킬 체인이 전제하는 선형적 진행 구조를 따르는가? 록히드 킬 체인의 핵심 전제는 공격이 단계를 순차적으로 통과하며, 따라서 어느 한 단계만 끊어도 전체가 좌절된다는 단일 차단의 유효성이다. 그러나 이 전제가 인간을 대상으로 하는 피싱 범죄에서도 성립하는지는 실증된 바 없다. 본 연구는 단계 간 전이 방향을 정량적으로 분석하여 피싱 킬 체인이 이러한 선형적 진행을 따르는지 검증하고, 그렇지 않다면

어떠한 구조적 특성을 보이는지를 규명한다.

II. 이론적 배경 및 관련 연구

2.1. 이론적 배경

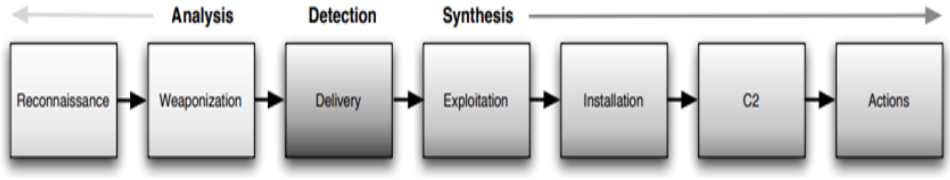
2.1.1. 피싱 범죄의 개념과 진화

피싱(Phishing)은 신뢰할 수 있는 기관이나 인물을 사칭하여 피해자로부터 금전 또는 개인·금융정보를 편취하는 사회공학적 사기 수법을 총칭한다. 초기의 피싱이 위조 웹사이트나 이메일을 통한 정보 탈취에 머물렀다면, 한국의 피싱 범죄는 전화를 매개로 한 보이스피싱, 문자메시지를 활용한 스미싱, 메신저 계정을 도용한 메신저피싱으로 매체를 확장해 왔다. 나아가 최근에는 투자리딩방·로맨스스캠·팀미션 사기와 같이 장기간에 걸쳐 피해자와 관계를 구축한 뒤 자금을 탈취하는 신종 유형이 등장하면서, 단일 수법이 아닌 복합·연쇄 수법으로 진화하는 양상을 보이고 있다[3].

이러한 진화의 공통된 특징은 범행이 일회적 기만으로 종결되지 않고, 정보 수집부터 접촉, 신뢰 구축, 자금 탈취, 흔적 제거에 이르기까지 다단계로 전개된다는 점이다. 특히 디지털 채널을 기반으로 한 비대면 수법이 보편화되면서, 악성앱 설치를 통한 원격 제어나 대포통장·가상자산을 경유한 자금 세탁 등 기술적 요소가 범행 단계에 결합되는 경향이 뚜렷해지고 있다[3]. 그럼에도 현행 대응은 개별 수법의 적발과 사후 검거에 집중되어 있어, 범행이 단계적으로 어떻게 연결되는지에 대한 구조적 이해는 충분히 축적되지 못한 실정이다.

2.1.2. 침입 킬 체인의 개념과 방어 논리

킬 체인(Kill Chain)은 본래 군사 작전에서 유래한 개념으로, 하나의 공격이 표적 식별부터 최종 타격까지 거치는 일련의 단계를 시간 순서로 연결한 사슬을 가리킨다. 사이버보안 분야에서는 2011년 록히드 마틴 연구진이 지능형 지속 위협(Advanced Persistent Threat, APT) 분석에 이 개념을 도입하면서 본격적으로 활용되기 시작하였다[6]. Hutchins 등(2011)이 제안한 침입 킬 체인은 공격자의 관점에서 APT 침입을 정찰(Reconnaissance), 무기화(Weaponization), 전달(Delivery), 취약점 공격(Exploitation), 설치(Installation), 원격 명령·통제(Command and Control), 목적 달성(Actions on Objectives)의 7단계로 재구성한 분석 프레임이다<Figure 1>[6].



<Figure 1> Hutchins(2011)의 킬 체인(7단계)

이 모형이 방어 전략에 갖는 함의는 두 가지 명제로 압축된다. 하나는 단계 통과 가정으로, 공격자가 최종 목적을 달성하려면 7단계 전부를 순차적으로 통과해야 한다는 것이다. 다른 하나는 단일 차단의 유효성으로, 체인의 어느 한 단계만 끊어도 전체 공격이 좌절된다는 비대칭적 방어 논리다[6]. 이 두 명제는 방어자에게 전체 사슬을 빠짐없이 차단해야 한다는 부담을 덜어

주고, 가장 효율적인 단일 개입 지점을 식별하는 것만으로 충분하다는 전략 방향을 제시한다는 점에서 실천적 가치를 지닌다. 한편 사이버 침입 분석에는 킬 체인 외에도 공격자·피해자·역량·인프라의 관계를 모델링하는 Diamond Model[7]과 같은 대안적 프레임도 존재하나, 이는 단계의 시간적 순차성보다 공격 구성 요소 간의 관계 규명에 초점을 둔다는 점에서 본 연구의 분석 목적과는 차이가 있다.

2.1.3. 킬 체인의 피싱 도메인 적응

Hutchins 등(2011)의 원형 모형은 시스템·소프트웨어를 겨냥한 기술적 침입을 전제로 설계되었으므로, 인간의 신뢰를 공격하는 피싱 사기 도메인에 그대로 이식하기는 어렵다. 본 연구는 '공격 대상이 시스템에서 인간으로 전환될 때 각 단계의 기능이 어떻게 변형되는가'라는 일관된 기준에 따라, 원형의 7단계를 피싱 도메인의 7단계로 다음과 같이 재정의하였다.

원형의 첫 단계인 정찰(Reconnaissance)은 공격 대상을 물색하고 정보를 수집하는 단계로, 피싱의 '정찰/대상선정'과 직접 대응한다. 이어 무기화(Weaponization)는 원형에서 악성코드 등 공격 도구를 제작하는 단계이나, 인간을 공격하는 피싱에서는 독립적인 코드 제작이 존재하지 않고 사회공학 스크립트의 구성으로서 이후의 신뢰구축 단계에 흡수된다. 전달(Delivery)은 공격 수단을 표적에게 도달시키는 단계로, 피싱에서는 전화·문자·메신저 등을 통해 피해자에게 최초로 접근하는 '접촉/유인' 단계에 해당한다. 취약점 공격(Exploitation)은 원형에서 시스템의 기술적 취약점을 공략하는 단계이나, 공격 대상이 시스템이 아닌 인간인 피싱에서는 기관 사칭과 신뢰 형성을 통해 피해자의 인지적 취약점을 공략하는 '신뢰구축/사칭' 단계가 그 등가물이 된다. 설치(Installation)와 원격 명령·통제(Command and Control)는 원형에서 공격자가 피해 시스템에 대한 지속적 장악력을 확보하는 단계인데, 피싱에서는 악성앱 설치·원격제어·단체대화방 지시 등을 통해 피해자의 행동을 장악·통제한다는 점에서 '통제/지시' 단계로 통합된다.

원형의 마지막 단계인 목적 달성(Actions on Objectives)은 두 가지 측면에서 재정의가 필요하다. 사이버 침입에서 공격자의 최종 목적은 데이터 탈취, 시스템 파괴, 금전 요구 등 다양하게 나타나며 원형은 이를 단일 단계로 포괄하지만, 피싱의 종점 대부분은 피해자로부터 편취한 자금이다. 이 자금은 탈취에 그치지 않고 대포통장·가상자산 등을 경유하는 세탁·은닉 과정을 거쳐 비로소 범행이 완성되며, 자금의 이전과 세탁이 시간적으로 구분되는 별개의 행위로 판결문에 뚜렷이 관찰된다. 이에 본 연구는 원형의 단일 목적 달성 단계를 '자금이전'과 '세탁/은닉'의 두 단계로 분할하였다. 나아가 사이버 침입에서는 흔적 제거가 별도 단계로 명시되지 않지만, 피싱에서는 연락 차단·계정 폐기·명의 폐기 등의 행위가 사건 진행에서 명확히 관찰되고 재범 방지 관점에서도 중요한 의미를 가지므로, 원형에 없는 '흔적제거'를 7단계로 신설하였다.

이상의 대응 관계는 단순한 명칭의 치환이 아니라 각 단계의 기능적 등가성에 근거하여 도출된 것으로, 그 결과를 원형 모형과 정리하면 <Table 1>과 같다. 이 7단계 골격이 실제 판결문 데이터와 정합하는지는 IV장에서 실증적으로 검증한다.

<Table 1> Hutchins(2011)의 킬 체인과 본 연구의 7단계 피싱 킬 체인 간의 대응 관계

록히드 마틴 원형 (2011)	본 연구 피싱 7단계	매핑 관계
1. 정찰(Reconnaissance)	1. 정찰/대상선정	대상 물색·정보 수집 기능 동일
2. 무기화(Weaponization)	(3단계에 흡수)	사회공학 스크립트로 통합
3. 전달(Delivery)	2. 접촉/유인	공격 수단을 표적에 도달시키는 기능 → 피해자 최초 접촉
4. 취약점 공략 (Exploitation)	3. 신뢰구축/사칭	취약점 공략 대상이 시스템 → 인간 인지로 전환
5. 설치(Installation)	4. 통제/지시	시스템 장악 → 피해자 행동 장악으로 기능 전환
6. 원격 명령·통제 (Command and Control)	(4단계에 통합)	지속적 통제 기능이 통제/지시 단계에 흡수
7. 목적 달성 (Actions on Objectives)	5. 자금이전 + 6. 세탁/은닉	자금 흐름이 이전·세탁의 별개 행위로 관찰되어 분할
(없음)	7. 흔적제거	연락 차단·계정 폐기 등 피싱 특이 행위로 신설

2.2. 선행연구 분석

본 연구가 분석 골격으로 차용한 킬 체인 개념의 직접적 기원은 Hutchins 등(2011)의 침입 킬 체인 연구이다[6]. 이 연구는 APT 공격을 정찰부터 목적 달성까지 7단계로 구조화하고, 방어자가 어느 한 단계만 차단해도 공격 전체가 좌절된다는 정보 기반 방어(intelligence-driven defense) 논리를 제시하였다. 이후 이 프레임은 사이버 침입 분석의 표준 모형으로 자리 잡았으나, 그 적용은 대체로 시스템·네트워크에 대한 기술적 공격에 한정되어 왔다. 인간의 신뢰를 공격하는 피싱·사회공학 범죄에 킬 체인을 적용한 시도는 제한적이며, 특히 형사 판결문과 같은 실제 사건 데이터로 그 단계 구조를 실증한 연구는 국내에서 보고된 바 없다.

국내에서 피싱 범죄의 단계 구조를 가장 체계적으로 분석한 선행 연구는 김은정(2022)의 대면편취형 보이스피싱 6단계 범행 과정 모형이다[8]. 이 연구는 가해자 시점에서 범행 경과를 명시적으로 단계화했다는 점에서 의의가 있으나, 세 가지 한계를 지닌다. 첫째, 분석 대상이 대면편취형 보이스피싱 1종에 국한되어 있어 비대면 자금이전 기반의 신종 피싱에 적용하기 어렵다. 둘째, 가해자의 범행 과정에 집중한 나머지 자금 세탁·은닉 및 흔적 제거 단계가 분석 범위 밖에 남겨졌다. 셋째, 악성앱 설치·원격 제어·가상자산 환전 등 디지털 채널 특이 수법이 단계 구조에 충분히 반영되지 못하였다[8]. 본 연구의 7단계 킬 체인은 이 세 한계를 직접적으로 보완하는 방향으로 설계되었다.

방법론의 측면에서, 사기 수법 네트워크 분석 분야는 판결문 또는 신고 데이터에서 수법 키워드를 추출하고 공출현(co-occurrence) 관계를 무방향 네트워크로 시각화하는 접근이 주를 이루어 왔다[5,9]. 이 방법은 어떤 수법들이 함께 등장하는지는 보여주지만, 어느 수법이 먼저 발생하고 어느 수법이 뒤따르는지를 알려주지 못한다. 사기를 예방하거나 수사 자원을 특정 단계에 집중 투입하려면 수법 간의 선후 관계가 핵심 정보임에도, 기존 접근은 이 정보를 분석에 활용하지 못한 셈이다. 이러한 방법론적 공백이 본 연구가 무방향 공출현 분석 대신 방향성 시퀀스 분석을 채택한 직접적인 이유다.

한편 법률 텍스트 마이닝 분야의 기존 연구는 주로 판결문의 법적 쟁점 추출이나 개체 간 지식 그래프(Knowledge Graph) 구축에 집중되어 왔다[5]. 그러나 사이버 범죄 형사 판결문을 기반으로 범행 준비부터 실행, 자금 세탁에 이르는 구체적 수법의 시계열적 전개 과정을, 수사 단

서 확보를 목적으로 한 방향성 그래프로 구조화하여 분석한 연구는 아직 충분히 축적되지 않았다. 종합하면, 기존 연구들은 ① 단계 모형을 제시하더라도 단일 유형에 국한되거나, ② 네트워크 분석을 수행하더라도 방향성을 결여하거나, ③ 텍스트 마이닝을 적용하더라도 범행의 시계열적 전개를 다루지 못한다는 한계를 지닌다. 본 연구는 이 세 공백을 동시에 겨냥하여, Hutchins 등(2011)의 킬 체인을 피싱 도메인으로 적응시키되 개념적 차용에 그치지 않고 290건의 판결문 데이터로 단계 구조의 타당성을 실증한다는 점에서 차별화된다.

III. 연구 방법론

3.1. 연구 가설 및 분석 설계

본 연구는 서론에서 제시한 세 연구 가설(Q1, Q2, Q3)에 답하기 위해, 형사 판결문에서 추출한 방향성 시퀀스를 분석 단위로 하는 텍스트 마이닝 기반 네트워크 분석을 설계하였다. 분석의 전체 흐름은 ① 판결문 수집 및 「범죄사실」 텍스트 추출, ② TF-IDF 기반 상향식 수법 사전 구축, ③ 문맥 규칙을 적용한 방향성 시퀀스 추출, ④ 방향성 가중 네트워크 구축 및 사회연결망분석(SNA), ⑤ 수동 검수를 통한 매칭 결과 검증의 5단계로 구성된다.

연구 설계 측면에서 본 연구는 연역과 귀납을 결합한 가설-연역적(hypothetico-deductive) 접근을 취한다. 즉 7단계 골격은 Hutchins 등(2011)의 프레임으로부터 이론적으로 도출하되 이를 확정된 전제가 아닌 검증 대상 가설로 설정하고(연역적 요소), 골격을 채우는 수법 사전은 판결문 코퍼스 자체에서 TF-IDF로 후보를 추출하는 상향식 절차로 구축하였다(귀납적 요소). 단계 구조 자체를 데이터로부터 완전 귀납적으로 도출하는 대안도 가능하나 그 경우 기존 침입 킬 체인 및 단일 차단 방어 논리와의 비교 가능성을 잃게 되므로, 본 연구는 검증 가능한 골격을 먼저 세우고 데이터가 이를 기각하거나 수정할 수 있도록 설계하였으며 그 판정은 IV장에서 이루어진다. 아울러 수법 추출에 임베딩 기반 언어모델을 도입하는 대신 결정적 규칙 기반 파이프라인을 우선 적용하였다. 290건 규모의 도메인 특화 코퍼스에서 임베딩 접근은 분류 근거를 문장 수준에서 역추적하기 어려워 수사 실무 활용의 전제인 설명 가능성을 충족하기 어렵기 때문이며, 언어모델과의 결합은 본 연구로 구조화된 시퀀스 데이터를 토대로 5.6의 후속 연구 과제로 설정한다.

이러한 설계는 두 가지 방법론적 원칙을 따른다. 첫째는 재현 가능성이다. 본 연구는 사전과 규칙이 명시적으로 공개되어 동일 입력에 대해 항상 동일한 출력이 보장되는 결정적(deterministic) 파이프라인을 채택하였으며, 이는 블랙박스적 분류 모델과 달리 후속 연구자가 결과를 검증·재현할 수 있도록 하기 위함이다. 둘째는 추적 가능성이다. 추출된 모든 수법 노드에는 매칭 근거가 되는 원문 키워드와 문장이 함께 기록되어, 각 분석 결과가 판결문의 어느 서술에서 비롯되었는지 역추적할 수 있도록 하였다. 이하 3.2부터 3.6까지 각 단계를 순서대로 기술한다.

3.2. 분석 자료 구성

3.2.1. 자료의 선정과 적합성

본 연구는 형사 판결문을 1차 자료로 사용하였다. 판결문은 법원의 사실인정 절차를 거친 자료로서 범죄 사실의 신뢰성이 확보되어 있으며, 「범죄사실」 항목이 통상 시간 순서로 서술된다는 점에서 방향성 시퀀스 추출의 합리적 근사 기반이 된다. 다만 판결문은 단일 출처로서 특정 범행 단계가 다른 단계보다 상세히 기재될 수 있다는 자료적 특성을 지니며, 본 연구는 이를 분

석의 전제로 인식하였다. 이 정보 밀도의 편향과 그 해석상 유의점은 V장에서 논의한다.

3.2.2. 자료 수집 및 표본 구성

자료 수집은 국내 판결문 열람 서비스인 엘박스(L-Box)를 통해 수행하였다. 검색 키워드로는 '보이스피싱(전기통신금융사기)', '스미싱', '메신저피싱', '투자리딩방', '로맨스스캠', '팀미션 사기'의 6가지를 사용하였으며, 수집 범위는 2023년 1월부터 2026년 5월까지의 확정 판결로 한정하였다.유형별 검색 결과를 최신순으로 정렬한 뒤, 모집단 규모에 비례한 등간격으로 판결문을 추출하는 체계적 표집(systematic sampling) 방식으로 수행하였다. 그 결과 보이스피싱(약 25,000건), 메신저피싱(3,991건), 투자리딩방(895건), 로맨스스캠(449건), 스미싱(136건)의 5개 유형에서 각 50건을 추출하였으며, 가용 판결문이 44건에 불과한 팀미션 사기는 전수를 수집하였다. 이후 동일 사건이 복수 유형 폴더에 중복 수집된 건을 사건 단위로 통합하여, 최종 290건의 고유 판결문을 분석 대상으로 확정하였다. 다만 원자료인 판결문 전문은 개인정보 및 저작권상의 제약으로 그대로 공개하기 어려우므로, 재현을 위해서는 본문에 명시한 수집 출처(엘박스)·검색 키워드·수집 기간·선정 기준을 따라 동등한 표본을 재구성하는 방식으로 갈음한다. 분석의 이론적 골격과 분류 기준을 보완하기 위한 자료로는 Hutchins 등(2011) 원문[6], 김은정(2022) 연구[8], 경찰청 '10대 악성사기' 분류표[4], 대검찰청 「2024 범죄분석」[2], 한국형사법무정책연구원(2023)의 보이스피싱 연구보고서[10]를 참조하였다.

3.3. 텍스트 추출 및 정제

290건의 PDF 원문에서 「범죄사실」 헤더 이하부터 「증거의 요지」 또는 「법령의 적용」 이전까지의 구간만을 추출하였다(extract_judgments.py, PyMuPDF 활용). 이 구간은 범행 경과가 서술되는 핵심 텍스트 블록으로, 당사자 진술이나 법적 판단 부분을 배제할 수 있다는 장점을 지닌다.

추출된 텍스트에는 공백·줄바꿈 정리, 특수문자 정규화, 금액 표기 통일, 동일 수법의 이형태 표준화(예: '카톡'·'카카오톡' → '카카오톡'), 개인 식별 정보 마스킹 등의 전처리를 일괄 적용하였다. 특히 판결문 상단에 위치하는 [기초사실]·[전제사실]·[공모관계]·[범죄전력] 블록은 피고인 개인의 범행이 아니라 범죄 조직 전반을 개관하는 서술로서 오탐의 주요 발생 원천이 되므로, 이를 사전에 제거하고 [구체적 범죄사실]·[범죄사실] 블록만을 분석에 사용하였다. 전처리 결과, 290건 전건에서 고유 사건번호가 확인되어 자료의 중복 없음을 검증하였다.

3.4. 수법 사전 구축: TF-IDF 상향식 접근과 반복 정제

3.4.1. 상향식 사전 구축의 근거

수법 사전의 구축 방식은 본 연구 방법론의 차별점이다. 기존 연구들이 경찰청 분류표 등 사전 정의된 목록에서 출발하는 하향식(top-down) 접근을 택한 것과 달리, 본 연구는 290건의 판결문 전체에 TF-IDF(scikit-learn TfidfVectorizer, 1-gram·2-gram, 형태소 분석은 kiwipiepy)를 적용하여 수법 후보 키워드를 데이터로부터 자동 추출하는 상향식(bottom-up) 방식을 채택하였다. 이는 분석자가 미리 상정한 범주에 데이터를 끼워 맞추는 대신, 판결문 텍스트에서 도출된 표현으로부터 수법 노드를 구성함으로써 분류의 순환논증을 피하기 위함이다. TF-IDF를 채택한 것은 모든 판결문에 공통으로 나타나는 일반어를 억제하고 특정 수법·유형을 변별하는 표현을 후보로 부각하기 위해서이며, 다만 '송금', '계좌'처럼 거의 모든 사건에

등장하는 보편 수법어는 역문서빈도(IDF)가 낮아 TF-IDF 점수만으로는 누락될 수 있다. 이를 보완하기 위해 후보 키워드 선정은 TF-IDF 상위 점수에만 의존하지 않고 원시 출현빈도 상위 어와 연구자 검토를 병행하여 보편 수법어가 사전에서 누락되지 않도록 하였다.

3.4.2. 반복 정제 절차

사전 구축은 일회적 선별이 아니라, 목표 정밀도에 도달할 때까지 검증과 수정을 반복하는 정제 절차로 설계하였다. 이는 Hutchins 등(2011)의 지표 생애주기(Indicator Life Cycle: Revealed→Mature→Utilized) 개념을 사전 구축에 적용한 것으로, 다음 네 단계를 반복한다. (1) TF-IDF 후보 키워드를 7단계 phase에 배정하고 동의어·이형태를 단일 정규화 노드로 통합한다(예: '검찰사칭'·'검사 사칭'·'검찰청 직원이라며' → '기관사칭_검찰'). (2) 구축된 사전으로 전량 매칭한 뒤, 층화무작위 표본을 검수하여 오탐(false positive) 패턴을 식별한다. (3) 오탐의 원인(조직 개관 서술, 법조문 인용, 수사 과정 서술에서의 오매칭 등)을 제외 패턴·필수 패턴 등의 문맥 규칙으로 사전에 반영한다. (4) 갱신된 사전으로 재매칭하여 정밀도를 재측정한다.

이 루프를 정밀도가 목표치(75% 이상)에 수렴할 때까지 반복한 결과, 사전은 v1에서 v3까지 진화하였으며 매칭 정밀도는 v1의 45.8%에서 v3의 79.3%로 향상되었다(버전별 추이는 <Table 5>). 최종 사전 v3는 34개 정규화 노드, 7단계 phase 라벨, 키워드·동의어 목록, 그리고 2단계 문맥 규칙(제외 패턴·필수 패턴)으로 구성된다. 본 연구의 재현 가능성을 담보하기 위해, 최종 사전(v3)의 정규화 노드·키워드·동의어 목록과 2단계 문맥 규칙(제외·필수 패턴), 그리고 분석 스크립트(extract_judgments.py, method_context_rules.py 등)는 부록을 통해 제공한다.

3.5. 시퀀스 추출 및 방향성 그래프 구축

3.5.1. 판결문 특유의 오탐 원인

단순 키워드 매칭만으로는 판결문 특유의 세 가지 오탐 원인을 통제할 수 없다. 첫째, 판결문에는 피고인의 행위뿐 아니라 공범·조직의 행위, 범죄단체 구조 개관이 혼재하여 단어만으로는 행위 주체를 구별할 수 없다. 둘째, 판결문은 한 문장이 수백 자에 달하는 만연체로 작성되는 경우가 많아(실측 최대 858자), 서로 다른 단계의 행위가 한 문장에 묶여 무차별 매칭될 위험이 있다. 셋째, 한 판결문에 여러 피해자에 대한 범행이 순차 서술되는 사건(전체의 약 22%, 63건)에서는 피해자 간 시퀀스가 뒤섞일 경우 거짓 역방향 전이가 발생한다.

3.5.2. 세 가지 보정 기법

이 세 문제에 대응하기 위해 다음 세 가지 보정을 적용하였다. (i) 행위 주체 태깅: 각 절의 주어를 기준으로 행위 주체를 '피고인/공범/배경/불명'의 4가지로 분류하여, 범죄단체 구조를 서술하는 '배경' 절은 시퀀스에서 제외하고 피고인·공범·불명 주체의 행위만을 분석 단위로 삼았다. (ii) 절 단위 재분할: 문장 분할 후에도 150자를 초과하는 장문은 한국어 연결어미('-하여', '-하고', '-한 후', '-되자', '-면서' 등)와 쉼표를 기준으로 절 단위로 재분할하였다. (iii) 피해자·병합 사건 블록 분할: '피해자 X에 대한 범행' 헤더를 경계로 각 피해자의 범행을 독립된 시퀀스 단위로 분할하였다.

3.5.3. 2단계 문맥 필터링

추가적인 오탐 제어를 위해 2단계 문맥 필터링 알고리즘(method_context_rules.py)을 구현하였다. 1차로 제외 규칙(Exclude Patterns)을 통해 조직 개관 서술이나 수사 과정 서술이 포함된 절의 매칭을 취소하고, 2차로 필수 규칙(Require Patterns)을 통해 실제 범행 맥락을 구성하는 동사·명사(예: 기관 '사칭/행세', 자산의 '송금/전송')가 병존할 때에만 최종 수법 노드로 인정하였다.

3.5.4. 방향성 가중 그래프 구축

시퀀스 추출은 사건(블록) 내 최초 등장 노드만을 기록하는 방식을 채택하였다. 추출된 각 노드에는 매칭된 원문 키워드(matched_kw), 근거 문장(matched_sentence), 행위 주체(actor)가 함께 기록되어 후속 검증의 추적 가능성을 확보하였다. 사건별 문장 인덱스 순 정렬을 통해 인접 노드 쌍(A→B)을 방향성 엣지로 생성하고, 동일한 쌍이 반복될 때 가중치를 누적하여 Python NetworkX의 DiGraph 객체로 방향성 가중 그래프를 구축하였다.

구축된 그래프는 GraphML 형식으로 변환하여 사회연결망분석(SNA) 전용 소프트웨어인 NetMiner 4에서 분석하였다. NetMiner를 통해 네트워크의 기본 위상 지표(노드 수, 엣지 수, 밀도, 평균 연결정도, 연결 컴포넌트 수)와 노드별 중심성(연결정도 중심성, 매개 중심성)을 산출하였으며, 단계(phase)를 노드 속성으로 지정하여 단계별 색상과 빈도 비례 크기를 적용한 네트워크 다이어그램을 시각화하였다. 산출된 지표와 시각화 결과는 IV장에서 제시한다.

3.6. 매칭 결과 검증

자동 추출 결과의 타당성을 점검하기 위해 층화무작위 표집으로 표본을 선정하고 수동 검수를 수행하였다. 각 매칭 노드에 대해 근거 문장을 대조하여 타당(Y), 부분 타당(P), 오탐(N)으로 판정하였다. 검수는 연구자 본인이 사전에 명문화한 판정 기준에 따라 수행하였으며, 판정의 안정성을 점검하기 위해 동일 표본의 일부를 시차를 두고 재검수하여 판정이 유지되는지를 확인하였다. 검수 기준은 개별 키워드의 100% 정확성이 아니라, 추출된 노드와 엣지 데이터가 전체 피싱 네트워크의 위상(topology)과 전이 흐름을 왜곡하지 않는 수준인가에 두었다. 검수에서 식별된 오탐 패턴은 3.4.2에서 기술한 반복 정제 루프를 통해 사전에 반영하고 290건을 재매칭하였으며, 그 결과는 4.3에서 상술한다.

IV. 분석 및 평가

4.1. 데이터 및 시퀀스 추출 개요

3단계 보정(행위 주체 태깅·절 단위 재분할·피해자 블록 분할)을 적용한 결과, 290건의 판결문으로부터 다중 피해자 블록 분할 및 병합 사건 보정을 거쳐 총 488개의 독립 시퀀스가 추출되었다. 이로부터 34개 정규화 수법 노드와 448개 방향성 엣지로 구성된 킬 체인 네트워크가 구축되었다. 구축된 네트워크의 기본 위상 지표는 <Figure 2>, <Table 2>와 같다. 네트워크 밀도는 0.399로, 34개 수법 노드가 형성 가능한 전체 연결의 약 40%가 실제로 관측되었음을 의미한다. 평균 연결정도는 13.18이며, 연결 컴포넌트가 1개라는 점은 분석 대상 수법 전체가 고립된 부분 없이 하나의 연결된 구조로 통합되어 있음을 보여준다. 본 절의 이하 분석은 이 네트워크를 대상으로 하며, 서론에서 제시한 세 연구 가설(Q1, Q2, Q3)의 순서에 따라 결과를 제시한다.

NETWORK PROPERTIES

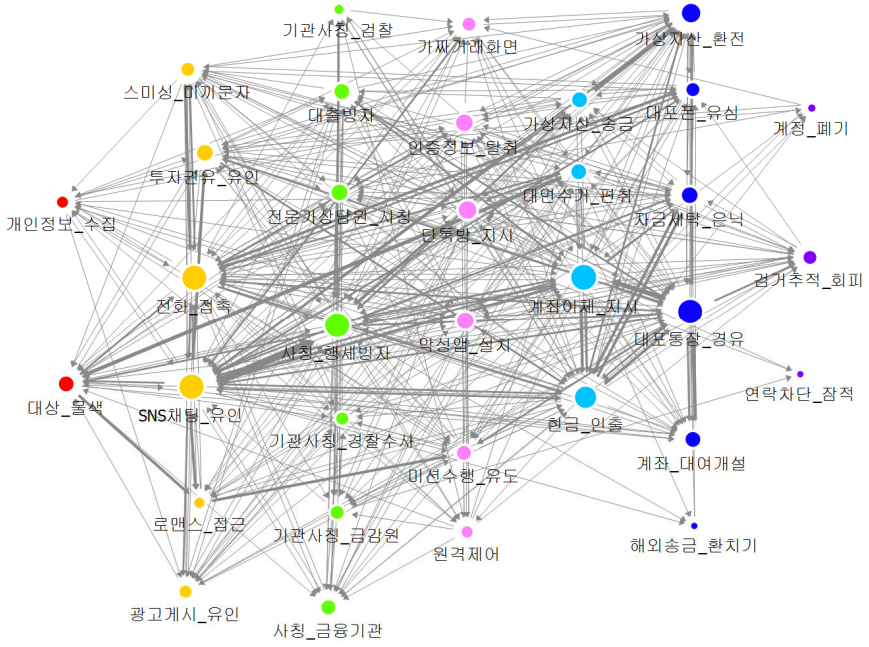
	# of Links : O(m)	Density : O(m)	Average Degree : O(m)	# of Components(Weak) : O(m)
edges	448	0.399	13.176	1

<Figure 2> 네트워크 기본 위상 지표 - NetMiner 4

<Table 2> 네트워크 기본 위상 지표

지표	값
노드 수	34
엣지 수	448
네트워크 밀도	0.399
평균 연결정도	13.18
연결 컴포넌트 수	1

이 네트워크의 전체 구조를 시각적으로 확인하기 위해, 34개 수법 노드와 448개 방향성 엣지를 NetMiner 4로 시각화하였다<Figure 3>. 각 노드는 소속 단계(phase)에 따라 색상을 달리하고 연결정도 중심성에 비례하여 크기를 조정하였으며, 엣지의 방향은 수법 간 전이 방향을 나타낸다. <Figure 3>은 다수의 수법이 서로 조밀하게 연결되어 하나의 통합된 범행 구조를 형성하고 있음을 보여주며, 특정 노드(전화_접촉, SNS채팅_유인, 계좌이체_지시 등)가 크게 표현되어 이들이 네트워크의 중심에 위치함을 시각적으로 드러낸다.



<Figure 3> 실측 네트워크 그림 - NetMiner 4

4.2. 주요 연구 가설 검증

4.2.1. 7단계 킬 체인 구조의 데이터 정합성 (Q1)

Q1은 사이버보안 분야에서 정립된 침입 킬 체인이 인간의 신뢰를 공격하는 피싱이라는 정보

통신망 이용(cyber-enabled) 범죄에도 구조적으로 적용되는가를 묻는다. 이에 답하기 위해, TF-IDF로 추출하고 선별된 34개 수법 노드가 사전 정의한 7단계 골격에 어떻게 배분되는지를 검토하였다.

분석 결과, 34개 노드는 7단계 전체에 빠짐없이 배분되었으며, 그 구성은 <Table 3>과 같다. 단계별 노드 수는 정찰/대상선정 2개, 접촉/유인 6개, 신뢰구축/사칭 7개, 통제/지시 6개, 자금이전 4개, 세탁/은닉 6개, 흔적제거 3개로 나타났다. 신뢰구축/사칭 단계의 노드가 7개로 가장 많은 것은, 한국 피싱이 검찰·금융감독원·은행·전문가 등 사칭 대상을 가장 다양하게 분화시켜 왔음을 반영한다.

<Table 3> 최종 7단계 킬 체인 분류표 (v3)

단계	단계명	대표 수법 노드	노드 수
1	정찰/대상선정	대상_물색, 개인정보_수집	2
2	접촉/유인	SNS채팅_유인, 전화_접촉, 스미싱_미끼문자, 투자권유_유인, 로맨스_접근 등	6
3	신뢰구축/사칭	사칭_행세빙자, 사칭_금융기관, 기관사칭_경찰수사, 대출빙자, 전문가상담원_사칭 등	7
4	통제/지시	단톡방_지시, 인증정보_탈취, 미션수행_유도, 악성앱_설치, 가짜거래화면 등	6
5	자금이전	계좌이체_지시, 현금_인출, 가상자산_송금, 대면수거_편취	4
6	세탁/은닉	대포통장_경유, 가상자산_환전, 자금세탁_은닉, 대포폰_유심 등	6
7	흔적제거	검거추적_회피, 계정_폐기, 연락차단_잠적	3

주목할 점은, 상향식으로 추출된 34개 수법 노드가 어느 하나도 7단계 골격 밖으로 남거나 별도 단계의 신설을 요구하지 않고 기존 골격 안에 모두 수용되었다는 것이다. 다만 7단계라는 골격 자체는 분석에 앞서 설정된 라벨이므로, 노드가 7단계에 빠짐없이 배분되었다는 사실만으로 단계 구조의 실재가 증명되는 것은 아니다. 따라서 본 절의 결과는 골격의 실증적 증명이라기보다, 침입 분석을 위해 고안된 킬 체인 프레임이 피싱 사기 도메인의 실제 수법 어휘를 무리 없이 담아낼 만큼 적합함을 확인한 것으로 읽는 것이 정확하다. 7단계 구조가 피싱 범죄의 실제 작동을 어떻게 반영하는지에 대한 본격적 실증은 단계별 정보 밀도(4.2.2)와 단계 간 전이 구조(4.2.3) 분석에서 이루어진다.

이 결과는 범죄 분석상 실천적 함의를 갖는다. 그간 보이스피싱·투자리딩방·로맨스스캠 등은 수법이 이질적이라는 이유로 개별 분석되어 왔으나, 본 연구는 이들이 단일한 7단계 골격으로 통합 기술될 수 있음을 보였다. 이는 유형이 다른 사건이라도 동일한 단계 틀로 비교·분류할 수 있는 공통 분석 언어를 제공하며, 수사 실무에서 신종 피싱이 등장하더라도 기존 단계 구조 안에서 위치를 식별하여 분석할 수 있게 된다.

4.2.2. 단계별 정보 밀도 분포 (Q2)

Q2는 판결문이라는 단일 사법 출처가 7단계에 걸쳐 어떤 정보 밀도 분포를 보이는가를 묻는다. 각 단계가 488개 시퀀스 중 몇 건에서 등장했는지를 측정하여 단계별 등장 비율을 산출하였으며, 그 결과는 <Table 4>와 같다.

<Table 4> 단계별 정보 밀도/등장 비율표

단계	단계명	시퀀스 등장 비율	밀도 분류
1	정찰/대상선정	12.1%	희소
2	접촉/유인	55.3%	풍부
3	신뢰구축/사칭	35.1%	중간
4	통제/지시	35.1%	중간
5	자금이전	48.2%	풍부
6	세탁/은닉	41.6%	풍부
7	흔적제거	7.4%	희소

분석 결과, 자금이전(5단계, 48.2%)·세탁/은닉(6단계, 41.6%)·접촉/유인(2단계, 55.3%)이 높은 빈도로 관측된 반면, 정찰(1단계, 12.1%)과 흔적제거(7단계, 7.4%)는 현저히 낮은 빈도를 보였다. 이는 자금이전·세탁 단계를 중심으로 한 범행의 실행부가 판결문에 가장 풍부하게 포착됨을 보여준다. 한편 정찰·흔적제거 단계의 낮은 빈도가 갖는 해석상의 유의점에 대해서는 V장에서 별도로 논의한다.

이 결과는 두 가지 함의를 갖는다. 첫째, 방법론적으로 특정 단계의 낮은 출현 빈도를 그 단계의 실제 범행상 비중이 낮다는 결론으로 곧바로 해석해서는 안 된다. 출처의 기재 관행과 실제 발생 빈도가 혼재되어 있어 본 연구의 데이터만으로는 양자를 구별하기 어렵기 때문이다. 둘째, 1단계(정찰)와 7단계(흔적제거)의 실태를 정확히 규명하려면 판결문만으로는 한계가 있으며, 피해자 신고 데이터, 통신사 로그, 금융 이상거래 탐지 데이터 등 복수 출처의 결합이 요구된다.

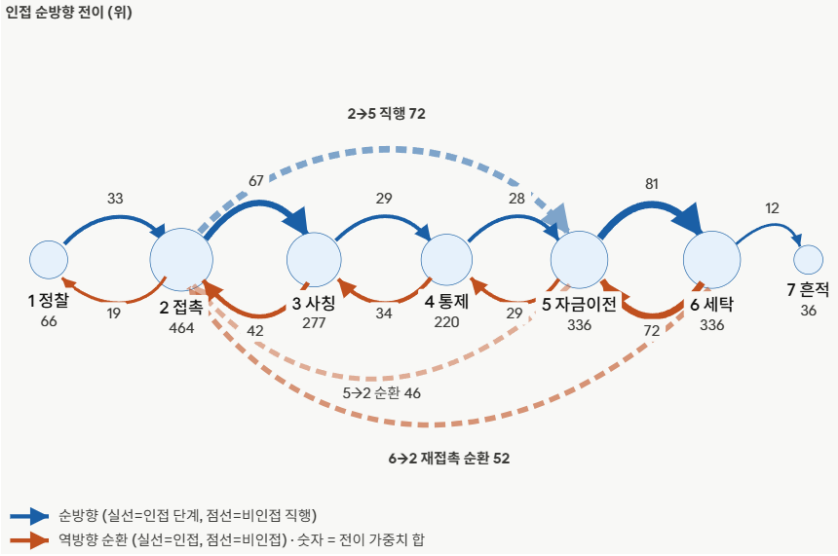
4.2.3. 단계 간 전이 방향과 순환 구조 (Q3)

Q3은 피싱 킬 체인의 단계 전이가 사이버보안 침입 킬 체인의 선형적 진행 구조를 따르는지를 묻는다. 이에 답하기 위해 488개 시퀀스에서 인접 수법 노드 쌍의 phase 전이를 집계하였다(배경 actor 제외).

집계된 전이 1,317건 기준 순방향(phase 증가) 45.6%(601건), 동일 단계 유지 20.0%(263건), 역방향(phase 감소) 34.4%(453건)로 나타났다. 역방향 비율이 34.4%에 달한다는 사실은 처음에는 데이터 품질 문제로 오해될 수 있다. 그러나 역방향이 집중되는 단계 쌍을 살펴보면 단순한 추출 오류가 아님이 드러난다. 빈도 상위 역방향 전이 쌍은 6→5(72건), 6→2(52건), 5→2(46건), 3→2(42건) 순이었는데, 이 패턴은 피싱 범죄의 실제 작동 방식을 반영한다. 투자리딩방 사기를 예로 들면, 피해자가 처음 입금(5단계)한 뒤 수익 화면을 보고(4단계) 신뢰를 더 쌓고(3단계) 다시 접촉 채널로 돌아가(2단계) 추가 입금을 유도받는(5단계) 순환이 반복된다. 로맨스스캠 역시 신뢰 구축과 감정 조작이 자금 요청과 번갈아 등장하며, 요청 실패 후 신뢰 단계로 복귀하는 패턴이 빈번하다. 접촉·사칭·통제 행위가 칼로 자르듯 분절되지 않고 중첩·순환하는 것은 피싱 도메인의 실질적 특성이지 분석 방법의 결함이 아닌 것으로 해석된다.

이러한 비선형적 순환 구조가 사이버보안의 침입 킬 체인과 달리 나타나는 까닭은 두 도메인의 공격 대상이 본질적으로 다르기 때문으로 해석된다. 시스템을 공격하는 침입에서는 일단 통과한 단계(예: 취약점 공격)를 되돌릴 이유가 없어 단계가 단방향으로 진행된다. 반면 인간의 신뢰를 공격하는 피싱에서는 피해자의 의심·저항·신뢰 동요가 수시로 발생하며, 공격자는 이를 무마하기 위해 앞선 신뢰구축·접촉 단계로 순환한다. 즉 공격 대상이 상태가 고정된 시스템에서 심리가 변동하는 인간으로 바뀌면서, 단방향 사슬이 순환적 구조로 변형된 것이다.

단계 간 전이의 흐름을 요약하면 <Figure 4>와 같다. 인접 단계 간 전이를 기본 골격으로, 자금이전에서 세탁으로의 순방향 전이(5→6: 81)와 세탁에서 자금이전·접촉으로 회귀하는 역방향 전이(6→5: 72, 6→2: 52)가 가장 굵게 나타나, 자금이전·세탁·접촉 단계를 중심으로 한 순환 구조가 시각적으로 드러난다.



<Figure 4> 피싱 킬 체인 단계 간 전이 흐름

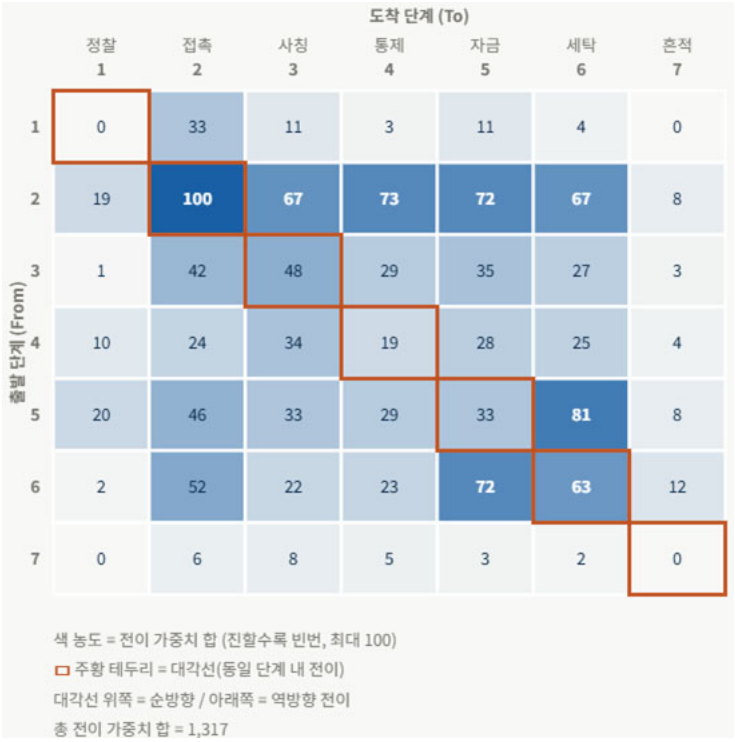
<Figure 5>의 전이 행렬은 단계 간 전이의 전체 분포를 보여준다.

첫째, 접촉(2단계)에서 나가는 전이의 가중치 합이 406으로 가장 높으며, 사칭·통제·자금이전·세탁 등 후속 단계 전반으로 고르게 전이된다(2→3: 67, 2→4: 73, 2→5: 72, 2→6: 67). 이는 접촉 단계가 다수 범행 경로가 분기하는 출발 허브임을 시사한다.

둘째, 자금이전(5)과 세탁(6) 단계 간 전이가 순방향(5→6: 81)과 역방향(6→5: 72) 모두 가장 빈번하여, 이 구간이 킬 체인에서 가장 활발한 순환 고리를 형성한다. 나아가 세탁 단계에서 접촉 단계로 회귀하는 전이(6→2: 52)가 두드러지는데, 이는 자금 세탁 이후 새로운 피해자 접촉으로 범행이 재순환하는 구조를 보여준다.

셋째, 정찰(1단계, 나가는 전이 합 62)과 흔적제거(7단계, 24)는 전이 빈도가 낮아 행렬의 양 끝에서 옅게 나타난다. 다만 이 낮은 빈도는 해당 단계의 실제 희소성과 판결문의 기재 관행이 혼재된 결과일 수 있다는 점을 감안해야 할 것이다.

종합하면, 전이는 순방향(45.6%)이 주를 이루나 동일 단계(20.0%)와 역방향(34.4%)이 상당한 비중을 차지하며, 특히 인접·비인접을 막론한 역방향 순환이 자금이전·세탁·접촉 단계를 중심으로 집중된다. 이는 한국 피싱 킬 체인이 엄격한 선형 사슬이 아니라, 전반적인 순방향 진행 경향 위에서 특정 단계를 중심으로 빈번히 역행·순환하는 구조임을 행렬 수준에서 재확인시켜 준다.



<Figure 5> 피싱 킬 체인 단계 간 전이 행렬

이러한 비선형성은 킬 체인 프레임 자체의 유효성을 약화시키지 않는다. 록히드 마틴의 침입 킬 체인을 실제 APT 사례에 적용한 연구들에서도 단계가 항상 엄격한 순서로 진행되지 않음은 이미 알려져 있다. 공격자는 정찰과 무기화를 동시에 수행하거나 침투 실패 후 정찰로 복귀하며, 여러 경로를 병렬로 시도한다. 그럼에도 킬 체인이 핵심 분석 프레임으로 활용되는 것은, 그 목적이 "공격이 항상 이 순서로 일어남을 예측하는 것"이 아니라 "공격을 구성하는 단계를 식별하고 각 단계에서 차단 지점을 찾는 것"에 있기 때문이다. 따라서 본 연구의 결과는, 한국형 피싱 킬 체인이 1단계에서 7단계로 엄격히 진행하는 선형 사슬이 아니라, 전반적인 순방향 경향과 더불어 접촉·사칭·통제 단계를 중심으로 한 순환이 공존하는 구조임을 보여준다.

이러한 순환 구조는 차단 전략에 직접적 함의를 갖는다. 록히드 킬 체인의 핵심 전제인 '단일 단계 차단으로 전체 사슬이 좌절된다'는 명제는, 단계가 순환하는 피싱 도메인에서는 제한적으로만 성립한다. 단일 시점 차단만으로는 범행이 접촉 단계로 복귀하여 재개될 수 있기 때문이다. 따라서 차단 전략은 단일 지점이 아니라 가장 빈번히 경유되는 순환 고리(특히 6↔2, 3↔2)를 반복적으로 겨냥해야 효율적이며, 이는 후속 연구의 차단 우선순위(BPS) 분석 및 매개 중심성 분석으로 이어질 명제다.

4.3. 매칭 검증 결과

자동 추출의 신뢰성은 3.4.2.에서 기술한 반복 정제 루프를 통해 단계적으로 검증·향상되었다. 각 라운드에서 전체 매칭 결과 중 오탐 가능성이 높은 매칭(조직 개관·법조문·수사 과정 서술에서 비롯된 것)을 우선 추출하여 검수하고, 식별된 오탐 패턴을 사전의 제외·필수 패턴에 반영한 뒤 290건을 재매칭하여 정밀도를 재측정하였다.

그 결과 매칭 정밀도는 사전 v1의 45.8%에서 출발하여, 조직 개관·결론문 오탐을 제거하고

15건 표본 검수(배경 actor 제외)를 반영한 v2에서 60.7%, 행위 주체 게이트와 문맥 규칙(exclude/require)을 강화하고 의심 매칭 140건을 O/X 검수한 v3에서 79.3%로 향상되어 사전에 설정한 목표 정밀도를 충족하였다(버전별 추이는 <Table 5>). v1·v2는 동일 15건(142 매칭행)에서 정밀도를 Y/(Y+N) 기준으로 비교하였고, v3는 전수 매칭 중 오탐 의심만 선별한 표본에서 O/(O+X)로 산출하였다. 따라서 v1·v2와 v3의 수치는 검수 표본과 산출 기준이 서로 달라 동일 척도의 절대 비교가 아니며, 정제가 진행될수록 오탐이 줄어드는 방향성을 보여주는 지표로 해석되어야 한다. 특히 v3의 표본은 오탐 가능성이 높은 매칭을 우선 선별한 것이므로, 이 표본에서 측정된 정밀도는 전수 매칭에 대한 실제 정밀도의 보수적 하한에 해당한다. 즉 v3의 79.3%는 전체 매칭의 정밀도를 과대평가한 값이 아니라, 오탐이 농축된 구간에서도 그 수준이 확보되었음을 의미한다. 한편, phase 라벨 정확도는 타당 판정된 노드 전원에서 100%를 유지하였다.

다만 v3에서의 79.3%는 오탐 우선 검수 표본에 대한 수치이며, 규칙 기반 접근만으로는 복합 문장 처리에서 추가적인 정밀도 상승 여지가 제한된다. 잔존 오탐의 대부분은 키워드 자체의 오인식이 아니라, '피고인은 공범 B와 함께…'처럼 한 절에 복수 주체가 중첩되는 복합 문장을 규칙으로 완전히 분리하지 못하는 데서 비롯된다. 이러한 문맥 의미 수준의 분리는 규칙 기반의 본질적 한계이며, 의미 기반 언어모델과의 결합을 통한 보완은 후속 연구 과제다(V장 참조). 한편 본 연구의 분석 목적이 개별 매칭의 완벽한 재현이 아니라 전체 네트워크의 위상과 단계 간 전이 흐름이라는 거시 구조의 규명에 있음을 고려할 때, phase 라벨 정확도 100%와 거시 지표(순방향 45.6%·역방향 34.4%·동일 phase 20.0%, 단계별 밀도, 최종 분석 기준 290건·488시퀀스)의 안정적 산출은 추출된 네트워크가 전체 구조를 왜곡하지 않는 수준임을 뒷받침한다.

<Table 5> 수법사전 버전별 매칭 정밀도

버전	정밀도(%)	검수 표본	주요 조치
V1	45.8	15건·142행, Y/(Y+N)	34노드 초안, 키워드 매칭
V2	60.7	15건·142행, Y/(Y+N), 배경 제외	조직·결론·수사 오탐 제거
V3	79.3	Tier2 의심 140건, O/(O+X)	주체 게이트·문맥 규칙

V. 논의

5.1. 주요 발견 요약

본 연구는 형사 판결문 290건에서 추출한 방향성 시퀀스를 통해 한국 피싱 범죄의 7단계 킬 체인 구조를 실증하고, 그 구조적 성격을 세 연구 가설에 따라 규명하였다. 주요 발견은 다음과 같이 요약된다. 첫째, Hutchins 등(2011)의 침입 킬 체인을 피싱 도메인에 맞게 재정의한 7단계 골격은 상향식 추출 노드를 별도 단계의 신설 없이 수용하였다(Q1). 둘째, 판결문이라는 단일 출처는 자금이전·세탁 단계에 서술이 집중되는 반면 정찰·흔적제거 단계는 희소하게 포착되는 정보 밀도의 편향을 보였다(Q2). 셋째, 단계 간 전이는 순방향 45.6%, 동일 단계 20.0%, 역방향 34.4%로 나타나, 피싱 킬 체인이 엄격한 선형 사슬이 아니라 특정 단계를 중심으로 순환하는 구조임이 확인되었다(Q3). 이하에서는 이들 발견이 피싱 범죄 연구와 수사 실무에 갖는 함의를 이론·데이터·차단 전략의 세 측면에서 논의한다.

5.2. 이론적 함의: 피싱 킬 체인의 순환성

이론적 측면에서 본 연구는 피싱 범죄의 단계 구조가 순서가 고정된 선형 과정으로 개념화될 수 없음을 실증한다. 범행은 접촉(2단계)에서 사칭(3단계)을 거쳐 통제(4단계)로 진행하되, 신뢰 회복이 필요할 때 사칭 단계로 복귀하거나 새로운 피해자를 접촉하는 단계로 회귀하는 반복 순환이 사건 내에서 자연스럽게 발생한다. 이는 자금이전 단계를 일회적으로 차단하더라도 범행이 접촉·사칭 단계로 복귀하여 재개될 수 있음을 의미한다.

이러한 비선형성이 킬 체인 프레임 자체의 유효성을 부정하는 것은 아니다. 록히드 마틴의 침입 킬 체인을 실제 APT 사례에 적용한 연구에서도 공격 단계가 항상 엄격한 순서로 진행되지는 않음이 알려져 있으며, 그럼에도 킬 체인이 핵심 분석 프레임으로 기능하는 것은 그 목적이 단계의 순서를 예측하는 데 있는 것이 아니라 공격을 구성하는 단계를 식별하고 각 단계의 차단 지점을 찾는 데 있기 때문이다. 다만 록히드 킬 체인의 핵심 전제인 '단일 단계 차단으로 전체 사슬이 좌절된다'는 명제는, 단계가 순환하는 피싱 도메인에서는 제한적으로만 성립한다. 따라서 향후 피싱 시나리오를 모형화할 때 이 순환적 특성이 전제로 반영되어야 하며, 차단 전략 또한 단일 시점이 아니라 반복되는 순환 고리 자체를 겨냥하는 방향으로 재설정될 필요가 있다.

5.3. 데이터 활용 측면의 함의: 정보 밀도 편향과 복수 출처의 필요성

판결문의 정보 밀도 편향은 데이터 활용 측면에서 중요한 함의를 갖는다. 정찰(1단계)과 흔적 제거(7단계)가 데이터에 낮은 빈도로 포착된다는 것은, 수사 실무에서 이 두 단계의 실태를 판결문만으로 파악하기 어렵다는 것을 의미한다. 특히 1단계에서 이루어지는 개인정보 수집·명단 매집 행위는 이후 모든 단계의 전제가 됨에도, 이 부분이 사법 기록에 충분히 남지 않는다면 사전 차단 전략 수립에 필요한 정보가 구조적으로 부족해진다.

다만 이 낮은 빈도가 해당 행위의 실제 희소성에서 비롯된 것인지, 판결문의 기재 관행에서 비롯된 것인지는 본 연구의 데이터만으로 단정하기 어렵다(4.2.2 참조). 어느 쪽이든, 1단계와 7단계의 실태를 정확히 규명하기 위해서는 판결문이라는 단일 출처를 넘어 피해자 신고 데이터, 통신사 로그, 금융 이상거래 탐지 데이터 등 복수 출처를 결합할 필요가 있다. 이는 본 연구의 분석 경험에서 도출된 실천적 제언이자, 후속 연구의 데이터 설계 방향이기도 하다.

5.4. 차단 전략 설계의 함의

본 연구의 킬 체인 네트워크는 차단 전략 설계를 위한 거시적 구조 지도로 직접 활용될 수 있다. 순환 구조는 차단 지점 선정의 기준 자체를 바꾼다. 선형 사슬이라면 임의의 한 단계를 차단하는 것으로 충분하지만, 순환 구조에서는 가장 빈번히 경유되는 고리를 끊는 것이 효율적이기 때문이다.

역방향 전이 빈도 상위 쌍인 6→5(72건), 6→2(52건), 3→2(42건) 등의 패턴은 자금 세탁(6단계)과 접촉(2단계) 노드가 다수 사건의 경로에서 반복적으로 경유되는 허브 역할을 수행함을 시사한다. 이는 이 노드들에 탐지·차단 자원을 우선 집중하는 것이 전체 피싱 연쇄를 끊는 데 더 효율적일 수 있다는 가설로 이어진다. 다만 본 연구는 이러한 허브의 차단 효과를 직접 검증하지는 않았으므로, 해당 가설은 후속 연구에서 매개 중심성(betweenness centrality) 분석과 가상 차단 시뮬레이션을 통해 검증되어야 할 명제로 남는다.

5.5. 연구의 한계

본 연구는 다음과 같은 한계를 지닌다.

첫째, 표본의 대표성이다. 엘박스에서 수집한 290건은 한국 전체 피싱 판결문의 일부이므로, 표본 편향 가능성을 완전히 배제할 수 없다.

둘째, 시간 순서 근사의 가정적 위험이다. 본 연구는 판결문의 문장 인덱스를 범행의 시간 순서에 대한 근사로 사용하였으나, 판결문의 서술 순서가 실제 범행 순서와 항상 일치하지는 않을 수 있다.

셋째, 규칙 기반 매칭의 문맥 한계다. 본 연구가 채택한 TF-IDF·규칙 기반 매칭은 단어의 표면적 출현에 의존하므로, '검사를 사칭'과 '검사 결과'처럼 동일 표현이 상이한 맥락에서 쓰일 때 이를 완전히 구별하지 못하는 본질적 한계가 있다.

넷째, 단일 출처에 따른 정보 밀도의 편향 가능성이다. 본 연구의 자료인 판결문에서 정찰(1단계)과 흔적제거(7단계)는 상대적으로 낮은 빈도로 관측되었다. 이 낮은 빈도가 해당 행위의 실제 희소성에서 비롯된 것인지 판결문의 서술 관행에서 비롯된 것인지는 본 연구의 데이터만으로 단정하기 어렵다. 다만 조직적 전기통신금융사기 사건의 판결문은 양형 및 가담 정도 판단을 위해 범행 전 단계의 조직 구성 및 모의 과정과 범행 후의 증거 인멸 및 도주 정황을 비교적 상세히 기술하는 경향이 있어, 이 편향이 분석 결과의 거시 구조를 크게 왜곡하지 않을 가능성도 있다. 그럼에도 두 단계의 실태를 정확히 규명하기 위해서는 피해자 신고 데이터, 통신사 로그 등 복수 출처의 결합이 요구된다.

다섯째, 평가 방식의 한계다. 본 연구는 매칭의 정밀도(precision)에 초점을 두었으며, 누락된 수법을 포착하는 재현율(recall)은 직접 측정하지 않았다. 이는 분석 목적이 개별 매칭의 완전성보다 네트워크 거시 구조의 규명에 있기 때문이다. 아울러 수동 검수가 단일 검수자에 의해 수행되어 검수자 간 신뢰도를 산출하지 못한 점 역시 한계로 남는다.

다만 규칙 기반 매칭의 문맥 한계(셋째)는 방법론적 선택의 이면이기도 하다. 규칙 기반 접근은 문맥 이해 능력에서 언어모델에 미치지 못하지만, 동일 입력에 대해 항상 동일한 출력을 보장하는 완전 결정적(deterministic) 파이프라인이라는 점에서 재현 가능성의 분명한 강점을 지닌다. 즉 후속 연구자가 사전과 규칙을 수정·확장하며 결과를 재현하고 검증하는 것이 원칙적으로 가능하다. 본 연구가 정확성의 일부를 양보하면서도 규칙 기반을 채택한 이유가 여기에 있으며, 이 재현 가능한 시퀀스 데이터를 토대로 문맥 이해력을 보완하는 방향이 다음의 후속 연구 과제로 이어진다.

5.6. 향후 연구 제안

본 연구의 한계와 발견은 다음과 같은 후속 연구로 확장될 수 있다.

첫째, 규칙 기반과 의미 기반 모델의 하이브리드 접근이다. 본 연구로 1차 정제·구조화한 시퀀스 데이터를 학습 기반으로 삼아, 한국어 법률·범죄 도메인에 적응된 사전학습 언어모델(예: KoELECTRA, KLUE-BERT 계열)로 수법 노드의 의미론적 분류와 문맥 검증을 수행할 수 있다. 이때 규칙 기반 파이프라인의 재현성과 언어모델의 문맥 이해력을 결합함으로써, 정확성과 검증 가능성을 동시에 확보하는 방향이 바람직하다.

둘째, 차단 우선순위의 정량적 검증이다. 5.4에서 제시한 허브 노드 차단 가설을 검증하기 위해, 매개 중심성 분석과 가상 차단 시뮬레이션을 통해 차단 우선순위 점수(BPS)를 산출하고 그 효과를 정량적으로 평가할 수 있다.

셋째, 피싱 유형별 서브그래프 비교다. 본 연구는 6개 피싱 유형을 단일 킬 체인 프레임으로 통합 분석하였으나, 유형별로 서브그래프를 분리하여 보이스피싱·투자리딩방·로맨스스캠 등이 동일한 7단계 골격 안에서 어떻게 상이한 전이 패턴을 보이는지를 비교 분석할 수 있다.

넷째, 복수 출처 데이터의 결합이다. 5.3에서 지적한 정보 밀도 편향을 보완하기 위해, 판결문에 피해자 신고 데이터·통신사 로그·금융 이상거래 데이터를 결합하여 정찰·흔적제거 단계의 실태를 보강하는 연구가 요구된다. 예를 들어, 더치트와 같은 피해 신고 데이터에는 판결에 이르지 않은 미수 사건이나 범행 초기의 접촉·정찰 정황이 판결문보다 풍부하게 남아 있다. 따라서 이러한 출처를 결합하면 본 연구에서 낮게 관측된 정찰(1단계)과 흔적제거(7단계)의 등장 비율이 높아지고, 단계 간 전이 분포도 달라질 수 있다. 본 연구가 제시한 단계별 밀도와 전이 비율은 어디까지나 판결문이 포착한 범위 안에서의 결과이기 때문이다. 다만 수치가 달라지더라도, 상향식으로 추출된 수법 노드들이 7단계 골격 안에 수용된다는 구조적 결론까지 흔들릴 가능성은 낮다고 판단된다.

Ⅵ. 결론

6.1. 연구 요약 및 시사점

본 연구는 형사 판결문 290건에서 추출한 방향성 시퀀스를 통해 한국 피싱 범죄의 7단계 킬 체인 구조를 실증적으로 정립하였다. 핵심 발견 세 가지를 요약하면 다음과 같다.

첫째, Hutchins 등(2011)의 침입 킬 체인을 피싱 도메인에 맞게 재정의한 7단계 분류표는 실제 판결문 데이터를 타당하게 담아낸다. 상향식으로 추출된 34개 수법 노드가 별도 단계의 신설이나 통합 없이 7단계 골격 안에 모두 수용되었다.

둘째, 한국형 피싱 킬 체인은 순방향 진행 경향과 순환이 공존하는 구조를 가진다. 순방향 전이 45.6%, 역방향 34.4%라는 수치는 피싱이 단방향 사슬이 아니라 접촉·사칭·통제 단계가 상호 순환하는 구조임을 보여준다. 이는 단일 단계 차단으로 전체 범행이 좌절된다는 킬 체인의 고전적 전제가 피싱에서는 제한적으로만 성립함을 의미하며, 차단 전략이 단일 지점이 아니라 순환 고리를 겨냥해야 함을 시사한다.

셋째, 판결문이라는 단일 출처는 자금이전·세탁 관련 후반 단계(5~6단계)에 정보가 집중되는 반면, 정찰(1단계, 12.1%)과 흔적제거(7단계, 7.4%)는 희소하게 포착된다. 이는 데이터 수집 설계의 실천적 함의를 제기하며, 두 단계의 실태 규명을 위한 복수 출처 결합의 필요성을 환기한다.

이러한 발견은 다음과 같은 시사점을 갖는다. 첫째, 본 연구는 그간 수법이 이질적이라는 이유로 개별적으로 분석되어 온 보이스피싱·투자리딩방·로맨스스캠 등의 다양한 피싱 유형이 단일한 7단계 골격으로 통합 기술될 수 있음을 보였다. 이는 유형이 다른 사건이라도 동일한 단계 틀로 비교·분류할 수 있는 공통 분석 언어를 제공하며, 수사 실무에서 신종 피싱이 등장하더라도 기존 단계 구조 안에서 그 위치를 식별할 수 있게 한다. 둘째, 피싱 범죄가 순환적 구조를 가진다는 본 연구의 발견은 차단 전략의 패러다임 전환을 요청한다. 단일 시점의 차단만으로는 범행이 접촉 단계로 복귀하여 재개될 수 있으므로, 가장 빈번히 경유되는 순환 고리에 탐지·차단 자원을 우선 집중하는 전략이 보다 효율적일 수 있다. 셋째, 정찰·흔적제거 단계가 사법 기록에 희소하게 남는다는 사실은, 사전 예방과 재범 방지에 필요한 정보가 판결문만으로는 구조적으로 확보되기 어려움을 드러낸다. 이는 수사·정책 차원에서 복수의 데이터 출처를 결합한 통합적 분석 체계의 구축이 필요함을 시사한다.

6.2. 연구의 의의 및 향후 과제

본 연구의 의의는 세 층위에서 평가될 수 있다. 학술적으로는 무방향 공출현 분석에 머물던 기존 사기 수법 연구를 방향성 시퀀스 분석으로 확장하였다는 점에서 방법론적 진전을 이룬다. 응용적으로는 수사 실무에서 차단 우선순위를 설계하기 위한 구조적 기반을 제공한다. 데이터 인프라의 측면에서는 290건의 판결문에서 정제·구조화한 피싱 시퀀스 데이터를 산출함으로써, 향후 한국어 사기 탐지 인공지능 연구를 위한 학습 자원으로 활용될 토대를 마련하였다.

이러한 성과를 바탕으로, 본 연구의 킬 체인 네트워크는 후속 연구의 출발점이 될 수 있다. 앞서 논의에서 제시한 바와 같이, 규칙 기반 파이프라인과 의미 기반 언어모델을 결합한 하이브리드 분석, 차단 우선순위 점수(BPS) 산출과 가상 차단 시뮬레이션을 통한 허브 노드 차단 효과의 정량적 검증, 그리고 6개 피싱 유형별 서브그래프 비교 분석이 후속 과제로 이어진다. 본 연구가 정립한 7단계 킬 체인이 이러한 후속 연구의 공통 분석 골격으로 기능함으로써, 한국 피싱 범죄에 대한 구조적 이해와 실효적 차단 전략 수립에 기여할 수 있기를 기대한다.

감사의 글

본 연구에 사용된 원천 데이터는 엘박스(L-Box)로부터 제공받았습니다.

참고문헌 (References)

[1] Sisa Journal. 2025. 3. Damages from fraud crimes including voice phishing and investment scams exceed 155 trillion won over six years [보이스피싱·투자사기 등 사기 범죄 피해액 6년간 155조 원 돌파]. Sisa Journal.

[2] Supreme Prosecutors' Office of Korea. 2025. Analytical report on crime 2025 [2025 대검찰청 범죄분석]. Crime Statistics Division, Supreme Prosecutors' Office of Korea.

[3] Korea Internet & Security Agency (KISA). 2024. 2024 cybersecurity threat trend report: analysis of smishing and malicious applications [2024 사이버보안 위협 동향 보고서: 스미싱·악성앱 분석]. KISA, Naju, Korea.

[4] Korean National Police Agency, Cyber Investigation Bureau. 2024. First meeting of the dedicated task force on multi-victim fraud held (Economic Crime Investigation) [다중피해사기 대응 전담반 첫 회의 개최(경제범죄수사)]. Korean National Police Agency, Seoul, Korea.

[5] Jung H, Lee D, Kim S, et al. 2025. Cybercrime classification using keyword-based knowledge graphs and graph neural networks [키워드 기반 지식그래프와 GNN을 활용한 사이버 범죄 분류]. Journal of Data Forensics Research, 2(2), 175-195. https://doi.org/10.12972/JDFR.2025.2.2.4

[6] Hutchins EM, Cloppert MJ, Amin RM. 2011. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Proceedings of the 6th International Conference on Information Warfare and Security, Washington DC, pp. 113-125.

[7] Caltagirone S, Pendergast A, Betz C. 2013. The diamond model of intrusion analysis. Center for Cyber Intelligence Analysis and Threat Research, Hanover, MD, USA. Technical Report ADA586960.

[8] Kim EJ. 2022. A crime process analysis of the voice phishing crimes: Based on crime script analysis of the face-to-face defraudation method [대면편취형 보이스피싱 범죄의 범행과정 분석: 범죄스크립트 분석을 중심으로]. Criminal Investigation Studies, 8(2), 31-54. http://doi.org/10.46225/CIS.2022.12.8.2.31

[9] Kim H, Lim H. 2022. A named entity recognition model in criminal investigation domain using pretrained language model [사전학습 언어모델을 활용한 범죄수사 도메인 개체명 인식]. Journal of the Korea Convergence Society, 13(2), 13-20. http://doi.org/10.15207/JKCS.2022.13.02.013

[10] Korean Institute of Criminology and Justice. 2023. A study on the response plan for each stage of voice phishing crime [보이스피싱 범행단계별 대응방안 연구]. Korean Institute of Criminology and Justice, Seoul, Korea.

부록(Appendix)

Appendix 1. 원시 데이터 및 분석 파이프라인 저장소

본 연구의 결정적 파이프라인에 대한 재현 가능성과 학술적 투명성을 확보하기 위해, 판결문 텍스트 추출, 문맥 필터링 규칙, 그리고 상향식 수법 사전에 사용된 모든 소스 코드와 데이터셋을 외부 공개 저장소에 보존한다. 후속 연구자는 아래의 깃허브(GitHub) 저장소를 통해 본 연구의 데이터 전처리 및 시퀀스 네트워크 구축 과정을 그대로 재현하거나 확장할 수 있다.

- 오픈소스 저장소 URL:
https://github.com/Eddy-Han/korean-phishing-killchain/tree/main

Appendix 2. 저장소 구성 파일 목록 및 핵심 기능

1. extract_judgments.py: PyMuPDF 라이브러리를 활용하여 형사 판결문 PDF 원문으로부터 「범죄사실」 헤더 이하의 핵심 텍스트 블록만을 정밀하게 분리 및 추출하는 전처리 파이썬 스크립트이다. 판결문 상단의 조직 개관 서술 등 오탐 원천 블록을 사전 제거하는 로직이 포함되어 있다.

2. method_context_rules.py: 추출된 문장을 대상으로 행위 주체 태깅(피고인·공범 분류), 장문 절 단위 분할(연결어미 기준), 피해자별 블록 분할의 3단계 보정을 수행한 후, 제외 규칙(Exclude Patterns) 및 필수 규칙(Require Patterns)의 2단계 문맥 필터링을 적용하여 오탐을 제어하는 알고리즘 스크립트이다.

3. 수법사전_v3.csv: TF-IDF 상향식 접근과 반복 정제 루프를 통해 최종 확정된 34개 정규화 노드, 7단계 Phase 라벨, 키워드 및 동의어 이형태 목록을 담고 있는 핵심 수법 사전 데이터셋이다. 본 연구의 최종 매칭 정밀도(79.3%) 산출의 근거가 된다.