Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.29, NO.1, Feb. 2019 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2019.29.1.149

전기통신금융사기 사고에 대한 이상징후 지능화(AI) 탐지 모델 연구

정 의 석,† 임 종 인‡
고려대학교 정보보호대학원

Study on Intelligence (AI) Detection Model about Telecommunication Finance Fraud Accident

Eui-seok Jeong,† Jong-in Lim‡
Graduate School of Information Security, Korea University

요 약

Digital Transformation과 4차 산업혁명 등 변화의 시대에 급변하는 기술 변화에 맞게 전자금융서비스는 안전하게 제공하여야 한다. 그러나 전기통신금융사기(보이스피싱) 사고는 현재진행형 이어서 사고의 지속적 증가, 지능화 및 고도화 현상을 대응하려 법률 제·개정 및 정책 제도 개선등 사고 근절을 위해 다양한 노력을 기울이고 있다. 더불어 금융회사는 이상금융거래탐지 시스템 개선 및 고도화를 통한 전기통신금융사기 사고 방지에 노력하고 있으나, 그 대응 결과는 그리 밝지 않다. 이러한 노력에도 불구하고 전기통신금융사기 사고는 관련 대책에 맞서 변화하며 진화를 거듭하고 있다. 본 연구에서는 보이스피싱에 의한 금융거래 사고발생 방지를 위해 시나리오 기반의 Rule 모델과 인공지능 알고리즘을 통해 모델링 된 지능형 이상금융거래 시스템을 설계하고 금융기관의 전자금융거래 시스템 에 실제 설치·운용해 본 결과를 바탕으로 인공지능형 이상금융거래 탐지시스템의 구현 모델과 분석·탐지 결과를 차단·대응 할 수 있는 고도화 된 대응 모델을 제안하고자 한다.

ABSTRACT

Digital Transformation and the Fourth Industrial Revolution, electronic financial services should be provided safely in accordance with rapidly changing technology changes in the times of change. However, telecommunication finance fraud (voice phishing) accidents are currently ongoing, and various efforts are being made to eradicate accidents such as legal amendment and improvement of policy system in order to cope with continuous increase, intelligence and advancement of accidents. In addition, financial institutions are trying to prevent fraudulent accidents by improving and upgrading the abnormal financial transaction detection system, but the results are not very clear. Despite these efforts, telecommunications and financial fraud incidents have evolved to evolve against countermeasures. In this paper, we propose an intelligent over - the - counter financial transaction system modeled through scenario - based Rule model and artificial intelligence algorithm to prevent financial transaction accidents by voice phishing. We propose an implementation model of artificial intelligence abnormal financial transaction detection system and an optimized countermeasure model that can block and respond to analysis and detection results.

Keywords: : AI, Fraud Detection System

Received(12. 04. 2018), Modified(01. 11. 2019), Accepted(01. 11. 2019)

† 주저자, vtamjes@hanmail.net
‡ 교신저자, jilim@korea.ac.kr(Corresponding author)

I. 서 론

비대면 채널에서의 전자금융 불법이체사고가 2013년 이후부터 급격히 증가하여 이에 전자금융 사고를 예방하기 위하여 금융위원회는 2013년 '금융전산 보안 강화 종합대책'을 통해 이용자 보호 강화 및 이상금융거래 탐지 시스템(Fraud Detection System : FDS) 구축을 권고 하였으며[1],금융감독원에서는 2014년에 금융권 FDS 추진 협의체 출범 및 FDS 고도화 1.0을 통해 2016년까지의 로드맵을 제시하였고[2], 2018년 현재는 모든 금융권에서 FDS 시스템을 자체 구축 운영하여 전자금융 불법이체 사고를 효과적으로 예방하고 있다. 그 결과 전자금융 불법이체사고가 2017년 이후 현저히 감소하고 있으며, 2018년 상반기에는 대다수의 금융회사에서 전자금융 불법이체사고가 발생하지 않고 있다. 이러한 노력으로 사기범들은 어렵고 힘든 고객의 금융정보탈취를 통한 전자금융 불법이체보다 효과적으로 고객의 금융자산을 탈취할 수 있는 다양한 사회공학적 방법 특히, 전기통신금융사기(보이스피싱) 수법을 집중적으로 시도하고 있다. 금융감독원 보고서에 의하면 2017년 보이스피싱 피해액은 2,423억원으로 전년 대비 26.0%(499억원↑) 증가하였으며, 피해건수는 49,948건으로 전년 대비 8.8%(4,027건↑) 증가하였다. 이처럼 건당 피해액이 증가한 원인은 보이스피싱 수법이 '정부기관 사칭형'에서 '대출빙자형'으로 전환된 것으로 보고 있다. 이처럼 사칭형에서 카드론 대출, 그리고 현재 대출빙자형에 이르기까지 보이스피싱의 범죄수법은 서민을 대상으로 한 소위 '햇살론' 등 과 같은 대출금 상환을 빙자한 수법이 계속적으로 증가 추세에 있다. 또한 경찰청·금융위원회·금융감독원·금융회사가 공동으로 보이스피싱 및 대포통장 근절을 위한 대대적인 홍보 및 모니터링 강화에도 불구하고 보이스피싱에 의한 보안사고는 지속적으로 증가하고 있으며 사기범의 전달책으로 이용되는 대포통장 이용 범죄 행위가 근절되지 않고 있다. 사기범들은 각종 전기통신사기(피싱, 파밍, 스미싱 등)방법을 이용하여 고객 금융정보 탈취하여 다양한 사기 수법으로 대포통장을 획득 후 보이스피싱을 통한 고객의 금융자산 탈취를 피해고객 별 상황에 맞는 시나리오로 불법이체사고를 시도하고 있다. 즉, 사고대상 고객별 공격유형과 맞춤형 시나리오를 만드는 등 공격방법의 고도화되어 있는 실정이다. 현재 금융권에서 운영중인 빅데이터 기반 이상징후 탐지 시스템(Fraud Detection System,이하 FDS 시스템)은 비대면 전자금융채널을 이용한 전자금융 불법이체 사고 탐지에 주로 이용되고 있어 보이스피싱 및 대포통장 사고를 예방하는데 많이 미흡하고 전자금융 불법이체 담당 부서와 전기통신금융사기 담당부서 간의 업무협조 및 정보공유 부족 등의 제도적·기술적인 문제점들이 나타나고 있다. 본 연구에서는 전기통신금융사기 사고에 대한 통합형 인공지능 기반의 지능화 이상금융거래탐지 모델을 제시함으로써 개선안을 제시하고자 한다.

II. 관련연구 및 동향

2.1 관련 연구

사회적인 이슈로 자리잡은 전기통신금융사기 사고에 대한 이상거래 탐지에 대한 많은 연구가 이루어지고 있다. 금융회사에서는 블랙리스트기반의 탐지시스템을 기본으로 적용하고 있고 또한 거래행태에 기반한 이상거래 탐지규칙을 적용하고 있어 그에 대한 개선방법에 대한 연구와 이상거래 탐지의 효과성을 향상시키고 지속적인 탐지규칙을 개선하는 방법을 꾸준히 연구하고 있다. 최의순[3]은 국내 금융 환경을 반영한 통계적 기법, 빅데이터 분석기법, 매체 정보 중심 탐지기법을 통해서 운영해 본 결과, 이상거래 탐지에 필요한 데이터 축적이 미미하여 오탐율이 높아 민원발생 빈도가 높은 것을 확인하였다. 최의순의 연구에서는 실제 사고 내역과 FDS를 통한 사고를 비교·분석하여 False Negative와 False Positive를 감소시키기 위해 사용자 프로파일 구성, 탐지 규칙 등을 제시하였다. 박은영[4]은 기존 블랙리스트 패턴기법의 사후 분석기법의 문제점을 도출하고 개선하기 위해 상태전이기법을 적용한 '사전적 사고 예방을 위한 이상금 융거래 탐지시스템'을 제시하였다. 지도학습의 한계를 극복하기 위해 비지도학습인 오토인코더 신경망을 활용한 딥러닝 기술로 이상거래 탐지 모델 적용 방안을 제시하였고, 이 FDS 시스템을 통해 이상거래 탐지율 향상, 신속한 대응, 자동화된 탐지 결과 반영 등을 기대하고 있다고 했으나, 현재 운영 중인 FDS 시스템에 비 지도학습을 시킨 모델을 적용하면 사고 데이터의 부족으로 정탐율이 현저히 낮아지게 된다. 윤해성[5]은 보이스피싱 사기유형에 따른 법률적용여부, 보이스피싱 관련 법률 개정안을 제시하였고, 구체적 피해 유형에 따른

예방과 대응방안, 전기통신금융사기 피해자 지원방안 등 법률적 차원에서 기술하였다. 정고은[6]은 인간기반 사회적공학 방법의 공격을 기술하고 Gragg(2003)가 제안한 심리적 트리거를 적용하여 각 사고 사례들에 대한 해석을 시도하고 제안한 사회공학이면의 심리적트리거와 주요 보이스피싱 사례 사이의 관련성을 기술하였다. 현재까지의 전기통신금융사기 이상금융거래 탐지 선행연구의 경향을 살펴보면, 기술적으로 과거 발생한 사고 정보의 블랙리스트 기반 탐지는 정탐율은 높으나 발생 빈도수가 거의 없고, 패턴 및 행위분석에 의한 탐지는 정 탐율이 낮아 민원 발생율이 높아지게 된다. 그리고 이상금융거래가 발생한 후 이 이상 금융거래 분석·확인·등록까지의 시간이 필요하여 블랙리스트 탐지에 사고 단말정보가 등록되기 전까지 동일한 사고 발생의 개연성이 높으며 신속한 탐지 및 예방이 불가능한 상황이다. 또한 프로파일링 및 Rule기반 탐지는 블랙리스트기반 탐지에 비해 정탐율은 높으나, 정상적인 거래절차를 기준으로 행위를 패턴화하여 이상금융거래를 탐지하기 때문에 오탐율도 높아지고, 사고의 유형이 빠르게 변화하면서 대응시간이 현저히 느려질 수 있다. 특히 대부분의 연구가 대응방안, 실태분석, 법적 및 제도적 장치마련 그리고 보이스피싱에 대한 사회공학적 접근에 초점을 두어왔다. 보이스피싱 발생 사고 시 실질적으로 소비자를 보호하는 효과가 나타나는 연구가 필요하게 된 것이다. 또한 이 연구는 금융기관 입장에서 전기통신금융사기 사고에 대해 이상금융거래에 대한 탐지를 지능화 모델 연구를 시도하는 최초의 연구 논문이다.

본 연구에서 이러한 탐지연구들의 단점을 보완하고자 기존의 시나리오 기반의 이상금융거래 탐지방식을 고도화하고 전기통신 금융사기 사고에 대해 인공지능 알고리즘을 적용하여 이상금융거래를 실시간으로 탐지가 되도록 지능화하여 고도화된 최적의 모델을 제안하고자 한다.

2.2 전기통신금융사기 사고동향

2.2.1 전기통신금융사기 유형

전기통신금융사기는 사회공학적 측면에서 컴퓨터기반 공격은 공격자가 공격대상에게 악성코드, 컴퓨터 프로그램 그리고 웹사이트 등을 이용하여 접근하는 경우이며 인간기반 공격은 공격자 스스로 공격 대상에게 직접적으로 혹은 전화 등 통신 장비를 이용하여 접근하는 경우를 의미[6]한다. 피싱, 파밍, 메모리해킹, 스미싱, 큐싱, 이메일을 통한 랜섬웨어 등이 컴퓨터기반 공격으로, 보이스피싱, 대포통장은 인간기반 공격으로 분류할 수 있다. 연구에서는 인간기반 공격 유형으로서 전기통신사기의 가장 대표적인 보이스피싱사기를 중점적으로 살펴보고자 한다. 보이스피싱 사기 수법은 그동안 다양하게 진화해왔다. 초기의 피싱수법은 국세청 등 공공기관을 사칭하여 세금 환급을 빌미로 피해자를 현금지급기(ATM)로 유도하는 것이었으나, 이런 수법이 널리 알려지자 대학 등록금 환급, 경품행사 당첨 등의 다양한 수법이 등장하였고, 피해자에게 신뢰할 수 있는 기관으로 사칭하기 위해 사전에 입수한 개인정보를 활용하게 되고, 휴대폰 문자메세지를 대량으로 발송하고 통화가 연결되면 신용카드 제3자 불법사용을 빙자하는 등으로 그 사기수법과 화법이 계속 진화 하였다[7]. 또한 진화된 유형을 보면 메르스를 빙자한 공공기관 사칭형 수법, 금융당국의 제도 개선 내용을 범죄에 역이용하는 수법, 저금리로 정부지원자금 대출을 받게 해준다며 대출금을 편취하는 수법, 금융회사 직원을 사칭한 수법,대검찰청 공식 홈페이지를 악용한 수법,대출금 상환을 이용한 사기 수법등 새로운 유형들이 등장하고 있다[8].

2.2.2 전기통신 금융사고의 동향

2.2.2.1 사회 전반적 대처 능력 제고에 따른 변화

보이스피싱은 인터넷전화(VoIP)와 현금자동입·출금기(ATM)등 사회적 인프라가 잘 갖추어져야 가능한 선진국형 범죄라고 할 수 있다. 2015년 보이스피싱을 민생침해 5대 금융 악으로 규정한 이후 2016년에는 보이스피싱 피해금액과 발생건수가 줄어들었으나 최근에는 금리인상, 경기 위축에 따른 서민들의 가계대출 수요 증가를 악용, 대출빙자형이 크게 증가하고 있다[9]. 이는 금융감독원이 2018년 상반기에 발표한 자료에서도 확인이 되는데, '17년 피해액은 2,423억원으로 '16년 대비 26.0%(499억원↑) 증가하였다. 대출빙자형 사기는 1,895억원으로 전체 피해액 대비 74.5%를 차지하는데 발신번호 변작, Auto call을 통한 무차별적인 문자메시지 발송과 더불어 금융회사의 실제 영업과 식별이 어려울 정도로 그 수법이 한층 정교화·지능화되고 있다.

반면, 대포통장 발생 건수는 대포통장 근절을 위한 금융권의 노력에 힘입어 '17년에는 45,422건으로 전년 대비 2.6% 감소(1,204건↓)하는 등 '15년 이후 발생 건수는 지속적으로 감소하고 있다[10].

2.2.2.2 사회 취약 계층 표적 접근

보이스피싱지킴이 2018년 9월 자료에 의하면 연령대별 보이스피싱 피해액은 규모의 차이는 있으나 전 연령대에서 피해가 발생하고 있다.

정부기관 등 사칭형의 경우 여성의 피해금액(363억원)은 남성(152억원)의 2.4배로 여성의 피해가 크며, 20~30대 여성이 여전히 이러한 수법에 취약한 이유는 이 연령대의 여성이 결혼자금 등 목돈을 모았을 가능성이 높고, 사회 초년생으로 사기 등과 같은 사회 경험이 적은 상황에서 신분상 불이익 등을 우려하여 사기범에 쉽게 속는 경향이 있어 특히 사칭형의 주요 표적이 되고 있다[6]. 이들의 피해금액은 175억원으로 전체 피해금액(201억원)의 87%를 차지하고, 동년배 남성(26억원)에 비해 7배 가까운 금액이다.

반면, 대출빙자형의 경우는 자녀 결혼 등 목돈 마련 등의 이유로 대출 필요성이 많은 남성 및 40·50대의 피해가 큰데, 40·50대의 피해금액이 가장 큰 비중(67.2%)을 차지한다[11].

III. 이상징후탐지(FDS) 운영현황과 기술적 한계

3.1 이상징후탐지(FDS) 운영현황

이상징후탐지(이하 이상금융거래탐지)시스템은 이용자와 기업의 금전적 손실 또는 정보 유출 등 여러 이상금융거래를 통해 발생될 수 있는 악의적인 행위들을 탐지하고 차단하기 위해 개발된 시스템이다.

Fig. 1.은 실제적으로 전자금융 거래에서 이상금융거래를 탐지하고 처리하는 흐름을 나타낸 것이다. 보는바와 같이 전자금융 각각의 채널에서 인입되는 모든 거래에 대해 이용자단말, 거래로그, 기타 여·수신 거래 정보, 공인인증서 발급정보 등을 실시간으로 수집하여 1단계로 로그인 시에는 블랙리스트 여부를 판단하고, 2단계로 이체 시에는 수집된 정보를 가지고 이상금융거래를 판단하여 각 거래마다 대응 보안 수준을 정의하게 된다. 그 보안 수준에 따라 지급정지, 이체정지, 고객확인 Call 수행 등 대응하게 구성되어 있다. 또한 이체시 이상금융거래거래 탐지는

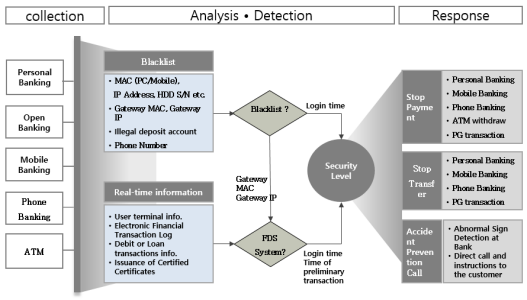

Fig. 1. A Flowchart of financial transaction detection and processing (example)

예비거래[1] 시에 탐지하게 되는데 그 이유는 본거래(이체) 전에 판단하여 불법이체 사고를 사전에 차단하여 최소화하고자 하기 때문이다.

3.1.1 금융기관 이상금융거래 탐지 대응 현황

은행과 증권사는 2014년부터 FDS 구축을 추진하였으며, 2017년말 기준 총 46개금융회사(은행 20개사·증권회사 26개사)가 구축 완료하여 운영 중이다. 사고예방 건수 및 금액은 3,665건, 445.8억원으로 1개사 기준 연평균 79.6건, 9.7억원의 예방 효과를 나타내었다.(건당 12백만원 예방)

금융감독원의 발표자료에 따르면 이상금융거래 시도에 대한 사고 예방률(평균) 또한 95.4%로 2017년 1분기(94.8%) 이후 증가 추세[2]로 FDS 탐지 정확도가 점차 향상되고 있으며, 사고 시도에 대하여 효과적으로 차단 및 예방하고 있으며, 미탐율 또한 평균 2.3% 수준으로 2017년 2분기 이후 감소[3]하고 있다. 그리고 이상금융거래 시도에 대하여 사고를 탐지하였으나 사고가 발생한 비율인 탐지후사고율(평균) 역시 2.3%로 1분기(2.9%) 이후 감소 추세[4]이다[12].

이러한 감소 추세를 유지하기 위해 금융회사들은 신종 사고유형 탐지를 위해 FDS탐지 룰(Rule) 개선 등의 시스템 고도화를 추진하고 FDS 전담운영인

1) 본거래(이체) 전에 준비 거래로서 정상계좌 여부, 비밀번호 정상여부, 계좌잔액 등 사전에 본거래 가능여부를 확인하는 거래
2) 예방율 : 1분기(94.8%) → 2분기(94.8%) → 3분기(96.6%) → 4분기(97.5%)
3) 미탐율 : 1분기(2.3%) → 2분기(2.9%) → 3분기(1.6%) → 4분기(1.4%)
4) 1분기(2.9%) → 2분기(2.3%) → 3분기(1.8%) → 4분기(1.1%)

력 확보하는 등을 지속적인 노력을 기하고 있으나, 위의 내용들은 대부분 전자금융 불법이체 위주의 내용이며 전기통신금융사기에 대한 탐지 내용은 극히 일부가 반영 것이다.

3.2 이상금융거래 탐지의 기술적 한계

현재는 대부분의 금융회사가 빅데이터 기반 실시간으로 처리 가능한 고도화된 FDS 시스템을 운영중에 있다. 하지만 고도화된 FDS 시스템으로 탐지하는 대상은 전자금융 불법이체 사고에 대하여 국한되어 운영되고 있는 것이 현실이다. 반면 전기통신금융사기 관련 탐지는 금융회사 내 요구불 계좌의 입·출금을 처리하는 상품(계정계)시스템에서 단순 시나리오에 의한 실시간 처리나 배치를 통한 단순 모니터링 수준으로 이상거래 탐지를 수행하기에 오탐과 과탐 비율이 높게 발생한다. 이러하기에 전기통신금융사기 관련 이상거래 탐지 시 실시간으로 지급정지 등의 임시조치를 적용하기 어려운 상황으로 현재 다음과 같은 기술적 한계가 존재한다.

첫째, 현재 금융회사에서 운영 중에 있는 FDS 시스템으로는 전기통신금융사기를 실시간으로 거래정보를 대량으로 수집하여 이상거래를 탐지하기에는 기술적 아키텍쳐 한계와 처리 용량의 부족이 존재한다.

FDS 시스템에서 전기통신금융사기 사고를 처리하기 위해서는 매우 중요한 부분으로 전기통신금융사기 사고에 대한 대응을 위해서는 하나의 금융거래에 대한 대량의 거래정보를 기반으로 판단해야만 실질적 대응 모델을 제시 할 수 있기 때문이다. 또한 전기통신금융사기 이상금융거래 탐지에 필요한 수집정보의 처리를 위해서는 운영 중인 시스템의 약 2~3배이상 성능과 용량을 보유한 시스템을 재 구축해야 필요한 수집정보의 처리가 되며, 원활한 기술적 처리를 가능하게 하기 위한 아키텍쳐를 재 구성해야 전기통신금융사기 사고에 대한 적절한 대응이 가능하게 된다. 둘째, 사고 유형별 고도화 된 맞춤형 시스템이 필요하다. 대부분의 금융회사가 전기통신금융사기 탐지를 위하여 Rule 기반 단순 시나리오를 적용하다보니 지능화되고 지속적으로 변화된 사기범들의 보이스피싱 및 대포통장 사기 수법을 탐지하기가 불가능에 가까우며, 탐지를 하더라도 과탐이 많이 발생하기에 실시간 지급정지 등의 임시조치 보다는 모니터링 위주의 사후 조치를 수행하고 있다. 결국 고도화된 Rule 기반 체계 및 지능화된 사기 수법을 사전에 예방할 수 있는 지능화(AI : Artificial Intelligent) 탐지 체계가 통합된 시스템이 필요하게 되었다. 셋째, 사고 별 탐지 대응 프로세서가 이원화되고 단순화되어 적절하게 대응하지 못하는 문제점이 있다. 현재 금융회사 별 전자금융 불법이체 사고 및 전기통신금융사기 탐지 대응 프로세스가 사고별로 각각 대응하는 이원화체계로 운영되고 있으며, 대응 수준도 시나리오로 단순하게 처리되고 있어 고도화되고 지능화된 사고에 대한 실질적인 대응은 못하고 있는 실정이다. 따라서 어떤 유형의 보안사고 발생 시 에도 다양하고 통합되어 처리할 수 있는 프로세스가 필요하게 된다. 앞에서 설명 기술했듯이 최근의 모든 사고의 유형은 전자금융의 모든 채널과 ATM, 창구, 전자상거래, 가상화폐 등을 가리지 않고 피해자의 상황에 맞게 맞춤형으로 발생하고 불법이체 및 편취를 하고 있다. 이에 대응 기관인 금융회사도 여기에 맞는 대응체계를 조속히 마련해야 한다. 넷째, 국내 금융회사는 전자금융 불법이체 사고 탐지는 24*365로 운영하고 있으나 전기통신금융사기 탐지는 기술적 및 조직적인 대응프로세스 한계로 평일 일과시간(09:00 ~ 18:00)에만 적용하여 운영하고 있다. 일과시간 이후 발생하는 전기통신금융사기가 지속적으로 증가하고 있지만, 대부분의 금융회사가 현실적으로 탐지 대응을 못하고 있는 실정이다.

IV. 인공지능 탐지 모델 연구

4.1 통합 지능화 이상금융거래 탐지시스템 구성

앞에서 전기통신금융사기 사고 대응의 기술적 한계에서도 언급했지만 다양하고 지능적인 보안사고 시도를 예방하기 위해서는 고도화된 Rule 기반 체계와 지능화 탐지 시스템이 필요한 것이다.

이에 통합 지능화 FDS 시스템 구성을 위해서는 첫째, 금융회사 별 상이한 대량의 거래 데이터를 실시간으로 수집할 수 있는 방안이 있어야 한다. 둘째, 탐지 및 분석이 용이하게 실시간으로 수집한 거래 데이터를 필터링하고 정규화 할 수 있는 전처리 기능이 필요하다. 셋째, 이상금융거래를 Hybrid(블랙리스트기반, 시나리오기반, 인공지능기반)기반 신속하게 탐지 및 분석 처리가 가능한 기능이 필요하다. 넷째, 인공지능기반 이상금융거래 탐지 지원을 위하여 사전에 모델링 작업을 위한 인공지능 모델링서버 구성이 필요하다. 다섯째, 이상금융거래 탐지 방식(Inline

방식, External방식)을 금융회사 업무시스템 상황에 맞게 처리 할 수 있는 기능과 탐지 결과를 업무채널에 제공하기 위한 인터페이스 기능이 필요하다. 마지막으로, 상기의 기능을 종합적으로 관리하고 운영할 수 있는 관리포털 기능이 필요하다.

4.1.1 다양한 거래 데이터 수집 방안

전기통신금융사기를 종합적으로 대응하기 위해서는 대량의 거래 데이터 수집이 필요하며, 금융회사별 다양한 업무채널에서 발생하는 거래 데이터를 실시간으로 수집하거나 거래 데이터 특성 상 일일배치 등의 비 실시간으로 수집하는 방법 등 다양한 수집 기능이 가능해야 한다.

첫째, 업무채널에서 거래 데이터를 파일로 제공 가능한 경우는 Agent기반으로 거래 데이터를 실시간 수집이 가능한 방식이다. 둘째, FDS 시스템 에서 제공한 Socket 인터페이스를 통하여 업무채널에서 거래 데이터를 직접 전송하는 방식이다. 셋째, 금융회사 별 표준으로 이용하는 통합 거래 중계채널(MCA, EAI 등)과 연동하여 거래 데이터를 수집하는 방식이다. 이는 중계채널에서 FDS 시스템으로 거래 데이터를 실시간으로 전송 할 수는 인터페이스가 추가 되어야하며, FDS 시스템에서는 중계채널에서 제공하는 표준 거래 인터페이스에 맞게 수신할 수 있는 기능 개발이 필요하다. 넷째, 업무채널에서 거래 데이터가 데이터베이스(DB)에 저장되는 경우는 데이터베이스에 접속하여 거래 데이터를 DB to DB 또는 DB to File 형태로 수집하는 방식이다. 다섯째, 전기통신금융사기의 효율적 탐지를 위해서는 고객정보, 자동화기기 위치정보 등을 비 실시간(일일배치 등)으로 수집하는 방식이 필요하다.

위의 5가지 유형의 데이터를 수집함으로써 이 대량의 데이터를 처리 할수 있는 빅데이타 기반의 플랫폼이 필요하게 된다. Fig. 2.은 FDS 시스템에서 다양한 형태의 거래 데이터를 수집하기 위한 수집흐름도 및 수집방법을 나타내고 있다.

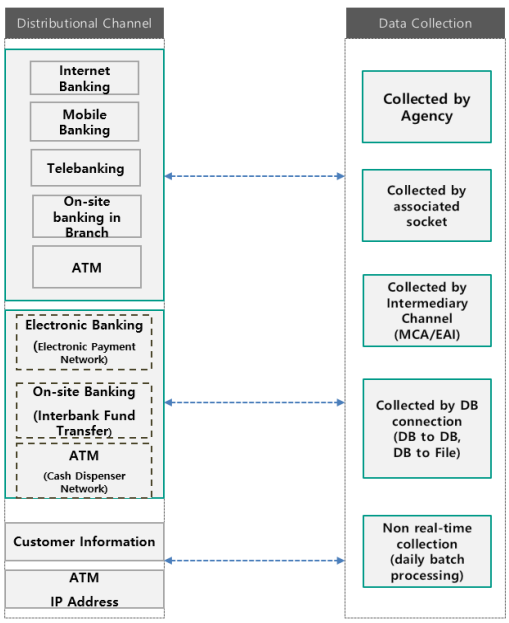

Fig. 2. Various transaction data collection flowcharts and collection methods

4.1.2 빅데이터 기반 실시간 전처리 서버 구성

다양한 방식으로 수집한 업무채널 거래 데이터는 실시간으로 전처리 서버로 전송되며 전처리 서버에서는 탐지 및 분석서버에서 효과적으로 실시간 이상금융거래를 탐지 할 수 있도록 전달하여 주는 역할을 수행하고 이중화로 구성하여 다음과 같은 기능을 처리한다. 첫째, 전문형태의 거래 데이터를 파싱(Parsing) 및 Key-Value 형태로 가공하여 탐지 및 분석서버로 전송한다. 대량의 일일 거래 데이터가 전송되므로 빅데이타 기반으로 구성해야 한다. 둘째, 불필요한 거래 데이터를 제거하거나, 중요 거래 데이터 정보에 대하여 암·복호화 기능을 수행한다.

4.1.3 Hybrid 기반 실시간 탐지 및 분석 서버 구성

전처리 서버로부터 전달받은 Key-Value 형태의 대량 거래 데이터를 실시간으로 이상금융거래를 탐지하기 위해서는 In-Memory 기반에서 처리 가능할 수 있도록 Fig. 3. 과 같이 구성한다.

Hybrid 기반 탐지 및 분석 서버 내 S/W 상세 기능은 다음과 같다.

Logstash[13]는 이벤트와 로그 메시지를 수집, 처리 및 전달하는 도구로 거래 데이터의 필요 항목만 추출한다. Kafka[14]는 메시지 큐로써 모든 로그 및 거래 정보를 저장 관리하며, Redis[15]는 Remote Dictionary Server의 약자로서, "키-값" 구조의 비정형 데이터를 저장하고 관리하기 위한 오픈 소스 기반의 비관계형 데이터베이스 관리 시스템(DBMS)으로 객체 및 개인화 정보를 저장하는 역할

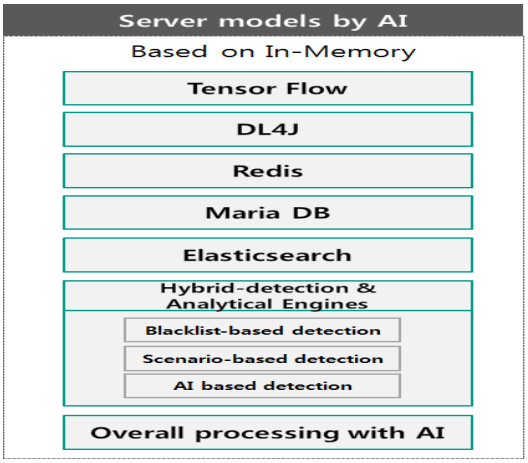

Fig. 3. Hybrid-based detection and analysis server configuration

을 한다. Maria DB는 정형화된 데이터를 관리하는 오픈 소스 기반의 데이터베이스 관리 시스템(DBMS)으로 사용자 정보/정책/포털에 대한 정보를 저장하고, Elasticsearch[16]는 아파치 루신(Apache Lucene)을 기반으로 만든 분산 검색엔진으로 모든 거래 데이터를 저장 및 검색기능을 제공하며, 탐지 및 분석엔진은 이상금융거래를 Hybrid 기반(블랙리스트기반, 시나리오기반, 인공지능기반)으로 종합적으로 분석 및 탐지하고, 인공지능 전처리의 역활은 선정된 인공지능 모델에서 이용할 수 있게 거래 데이터의 특성을 추출하는 기능을 제공한다.

실질적인 대응 모델을 위해서 전자금융 불법이체 사고와 전기통신금융사기 사고를 동시에 실시간으로 In-Memory에서 분석 및 탐지를 해야만 하는데 그러기 위해서는 충분한 시스템 메모리 용량(512G 이상)이 필요하다. 이유는 이상금융거래 탐지 정확도를 위하여 적어도 6개월에서 1년 이내의 다양한 형태의 개인화 프로파일링 작업이 요구되며, 프로파일링 된 대량의 데이터가 메모리 DB에서 운영되어야 만 실시간 처리가 가능하게 된다. 또한 인공지능기반 이상금융거래 탐지도 In-Memory내에서 운영되어야 하기 때문이다.

4.1.4 인공지능 모델링 서버 구성

탐지 및 분석 서버 내 인공지능 전처리 모듈로 처리된 데이터는 모델링 서버로 전송되고 전송된 데이터는 Redis 키 value에 저장되며, 그 즉시 전송된 데이터를 In-Memory 상에서 처리하고, 대용량의 거래 데이터를 학습할 수 있도록 구성한다. Tensor Flow[5][17]의 라이브러리와 함께 다양한 머신러닝 라이브러리를 사용하기 위해 DL4J[6]를 사용하여 크로스플랫폼[18] 사용 시 자바환경에서의 동일한 결과 값을 가지기 위한 목적으로 사용한다. 본 연구에서는 머신러닝 라이브러리의 자유로운 사용을 위한 사용자 인터페이스를 개발하여 Tensor Flow와 DL4J에 대한 각 종 알고리즘 처리를 위한 파라미터에 대한 조정을 쉽게 하도록 구성 한다. 또한 라이브러리 사용자 인터페이스는 다양한 파라미터의 조정을 통해 다양한 모델을 적용할 수 있는 기능을 제공하며, 모델과 모델간의 처리결과를 확인 할 수 있는 기능을 제공한다.

4.1.5 이상금융거래 탐지 방식 및 탐지 결과 전송 인터페이스 구성

금융보안연구원에서 발표한 FDS 기술 가이드에서는 이상금융거래 탐지 방식을 Inline 방식과 External 방식으로 구분하고 있다[19]. Inline 방식은 거래원장이 생성되기 이전에 이상금융거래를 판단하는 방식으로 지시된 금융거래정보가 원장시스템으로 이관되기 까지 매 거래 단계 별(이용자 인증, 예비거래, 본거래 등)로 이상금융거래 유무를 판단하는 것이다. External 방식은 거래원장을 생성하고 난 이후에 이상금융거래를 판단하는 방식으로 지시된 모든 금융거래에 대해 거래원장(DB)을 완성하고, 지시된 금융거래를 최종 승인하기 전에 이상금융거래 유무를 탐지 하는 것이다. 이상금융거래 탐지를 Inline 방식으로 적용하는 것이 보안사고를 사전에 신속하게 탐지할 수 있으나, FDS 시스템 장애 시에는 연계된 업무시스템의 거래에 영향을 주는 문제점을 가지고 있다. 그러하기에 본 연구에서는 Fig. 4.와 같이 전자금융 불법 이체 사고 탐지를 위한 연계 업무시스템에는 Inline 방식을 적용하고, 전기통신금융사기 탐지를 위한 연계 업무시스템에는 개선된 External 방식을 적용하는 방안을 제시한다. 그 이유는 전자금융 불법 이체 사고 탐지를 위한 연계 전

5) 구글이 2011년에 개발을 2015년에 오픈 소스로 공개한 기계학습 라이브러리

6) Deeplearning4j 의 약자이며, 자바와 자바 가상머신용으로 작성된 딥 러닝 라이브러리이며, 딥 러닝 알고리즘을 광범위하게 지원하는 컴퓨팅 프레임워크

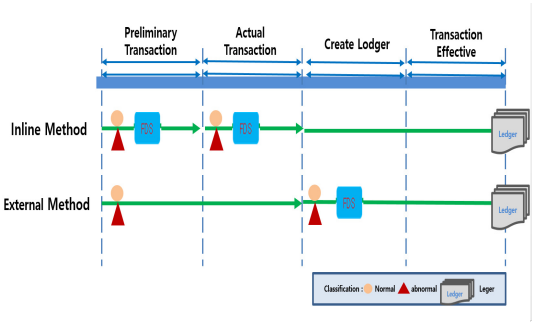

Fig. 4. Application of abnormal financial transaction detection by type of security incident

자금융채널(인터넷뱅킹, 모바일뱅킹, 기업뱅킹 등)에서는 이체 예비거래 시 Inline 방식을 적용하여 이상금융거래 탐지 시에는 전자금융채널에서 추가인증(ARS 인증 등)을 적용 할 수 있는 장점이 있다. 하지만, 전기통신금융사기 탐지를 위해서는 모든 요구불(수신) 계좌에서 출금거래 발생 시 이상금융거래 유무를 판단하기에 출금거래를 요청하는 모든 업무채널을 Inline 방식으로 구성하기에는 대량의 거래를 처리하는 은행 상품(계정)시스템 서비스의 안정성에 대한 Risk가 존재한다. 그러하기에, 예비거래(조회) 시 데이터를 실시간으로 수집하여 탐지할 수 있는 External 방식이 적합하다고 할 수 있다. 또한 지능화 FDS 시스템에는 전자금융 불법이체 사고와 전기통신금융사기에 대하여 실시간으로 탐지한 내역을 보안수준 별로 구분하여 업무채널로 전송하기 위한 인터페이스 기능이 필요하다.

4.1.6 효율적이고 종합적인 관리포털 기능 구성

전기통신금융사기의 이상금융거래 사고탐지내역을 효율적이고 통합 관리하기 위해서는 블랙리스트와 화이트리스트 관리, 사고에 대한 이상금융거래 탐지 내역 관리, 보안정책 관리, 통계관리 등 기능이 필요하다.

결국 전자금융 불법이체 사고와 전기통신금융사기를 통합하여 탐지 할 수 있는 지능화 FDS 시스템을 구성하기 위해서는 Fig. 5. 와 같이 다양한 업무채널과 유연하게 연결가능하며, 거래로그를 실시간으로 수집하고 이상금융거래를 Hybrid 기반(블랙리스트, 시나리오, 인공지능)으로 실시간 탐지 할 수 있는 체계가 필요하다.

4.2 인공지능기반 이상금융거래(FDS) 탐지 모델 연구

인공지능 지도학습 알고리즘을 통한 탐지 모델 연구를 위하여 A 은행에서 2017년도에 발생한 보안사고 데이터와 시나리오기반으로 탐지한 거래 데이터 및 정상 거래 데이터를 이용하였으며, 본 연구를 위하여 2018년 2월부터 8월까지 인공지능기반 FDS 탐지 모델을 위한 인프라 구축 및 고도화된 모델 개발 연구를 진행 하였다. 전기통신금융사기 인공지능 탐지 모델 연구를 위하여 Deep Learning 기반 지도학습 방식의 합성곱 신경망(Convolutional Neural Network, CNN) 알고리즘을 선택 하였다. 이유는 CNN 알고리즘의 경우에는 컴퓨터 비젼 분야에서 잘 적용되었을 뿐만 아니라, 각각의 성공적인 적용 사례에 대한 문서화 또한 잘 되어 있다[20]. 더욱 최근에는 CNN 알고리즘이 자동음성인

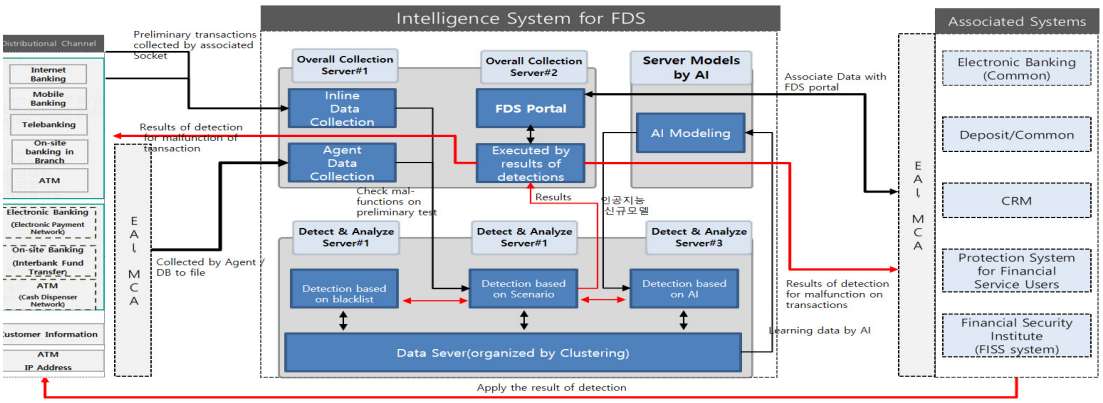

Fig. 5. Intelligent FDS system configuration diagram

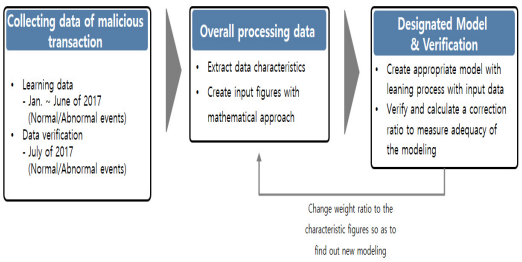

Fig. 6. Development of Artificial Intelligence Model for Supervised Learning

식(Automatic Speech Recognition, ASR)을 위한 음향 모델 분야에 적용되었으며, 기존의 모델들보다 더욱 성공적으로 적용되었다는 평가를 받고 있다[21]. 또한 CNN은 인간의 신경망을 유사하게 구현한 인공 신경망의 일종으로 다른 신경망과 달리 입력 데이터의 필터링을 수행하며, 이미지 인식을 위한 특징을 자동으로 학습하고 이미지 형태 변이를 효과적으로 학습하는 알고리즘이다. CNN 알고리즘을 이용한 최적의 탐지 모델 개발을 위하여 Fig. 6. 과 같은 단계적인 과정을 수행하였다.

4.2.1 사고 데이터 수집 단계

인공지능 모델 개발을 위하여 A은행의 2017년 1월부터 12월까지 거래된 요구불(수신)계좌의 모든 실 거래내역을 수집하였으며, 지도학습을 위하여 2017년 1월부터 6월까지 거래된 계좌 중 정상계좌 최대 5000개, 사고계좌 최대 1000개를 확보하였다. 또한 학습된 모델의 정탐율을 검증 확인하기 위하여 2017년 7월에 거래된 정상계좌 최대 200개, 사고계좌 최대 100개를 확보하여 이용 하였다.

4.2.2 데이터 전처리 단계

데이터 전처리 단계에서는 의미 있는 특성(feature) 값을 추출하며, 추출된 특성 값을 컴퓨터가 이해 할 수 있는 수학적 수치 값으로 변환하는 과정을 수행한다. 이 단계는 지도학습 모델 개발과정 중에서 가장 중요한 단계로, 그 중 의미 있는 특성 값을 추출하는 것이 중요하다.

결국 데이터 전처리를 위한 의미 있는 특성 값은 한 번에 추출하기는 어렵고 모델 생성 및 검증을 통한 많은 반복과정을 통하여 의미 있는 특성 값 추출이 가능하다. 약 2개월간의 모델 개발을 통하여

Table 1. Collection data information for development of artificial intelligence supervised learning model

classification		Explanation
Total collected data		Actual transaction history of debit accounts traded between January-December 2017.01
Learning data	normal	Normal transactions traded between January-June 2017 (Up to 5,000 normal accounts)
	accident	Transactions from the accident account that occurred from January to June 2017 before the reporting date(accidents account up to 1000 units)
Verification data	normal	Normal transactions traded in July 2017 (up to 200 regular accounts)
	accident	Transaction history from the accident account that occurred in July 2017 before the reporting date(up to 100 accidents)

Fig. 7.과 같은 의미 있는 총 특성 값 21개를 추출하였으며, 모델 및 생성 검증 단계에서 다양한 방법으로 학습을 진행하기 위하여 대분류 4개(feature1, feature2, feature3, feature4) 및 소분류 6개(feature1, feature2_1, feature2_2, feature3_1, feature3_2, feature4)로 구분 적용하였다. 반복적인 작업을 통해 모델링 및 검증 단계 결과에 따라 데이터 전처리 단계에서 최종적으로 사용할 특성 값이 확정된다.

'Alteration/Exception'은 사기의심계좌(대포통장)로 사용 가능한 사전 징후를 탐지하기 위하여 이용하며, 특성 값 11개는 거래 TR 발생 시 2일(48시간) 이내 해당 TR에 해당하는 계좌정보에 변경 기록이 있으면 각 column 값에 '1' 로 tagging 하며, 변경 기록이 없으면 '0' 으로 tagging 한다.

'Riskiness Remark'는 사전 블랙리스트[7]로 등록된 사고 위험 개연성이 있는 개인 또는 업체와 거래가 있는지 여부를 판단하기 위한 것으로, 특성 값

7) 사전에 기 구축 된 시스템을 통한 모니터링 결과 불법 이체 징후 업체 및 개인 리스트

Categories				Title of column	Details of column	Remarks
Frequency	Basic	feature1	feature1	LN_DPS_TRSC_KIND_CD	Transaction code of Loan/Deposits	- Measure Frequencies by means of listing transactions in order of occurrence recently
				CHNL_TYP_CD	Channel code	
				OTSD_BIZ_CD	External Reporting Code	
				THR_ACCT_NO	Their Account Number	
				RMRK	Remark	
				SUMM_PSBK_RMRK	Summary of passbook remarks	
				TRT_BR_NO	Transaction Branch Number	
tagging	Alteration/ Exception	feature2	feature2_1	T_1032	ATM withdrawal	- for the duration of window.size Change and exception behavior tagging - default window size = 172800 - (2 days)
				T_1155	Change password	
				T_1156	Without Passbook, Without Card, Change password on ATM	
				T_1547	Accounts registered to monitoring list	
				T_1619	Lon-term dormancy or suspended account	
				T_3055	Passbook reissued with brought forward	
				T_3057	Change seal /signature, Reissue	
			feature2_2	T_1019	Daily withdrawal limit of ATM	
				T_1605	Push alarming of fund transfers	
				T_3051	Reissue passbook and sea due to lost	
				T_5077	SMS of fund transfer	
	Riskiness Remark	feature3	feature3_1	T_RMRK_1	Remark of riskiness	Risk Factors: Loaner, Cradit, Savings Bank, Capital, Limited Company, Unemployment Benefit Employment Department, Korea Scholarship Foundation
			feature3_1	T_RMRK_2	Remark of riskiness among all transactions	
	Balance	feature4	feature4	T_BAL	Calculating current and previous balance	1 - (|Small Amount | / |Large Amount)

Fig. 7. Meaningful feature value classification for data preprocessing

2개는 거래 TR 발생 시 사전 등록된 블랙리스트와 연관성이 있으면 각 column 값에 '1' 로 tagging 하며, 연관성이 없는 경우에 '0' 으로 tagging 한다.

'Balance'은 거래가 거의 없는 계좌에서 사기의심 계좌(대포통장)로 이용 시 갑자기 입·출금이 발생하는 변동폭을 확인하기 위한 특성 값이다.

의미있는 특성 값에 대해 설명하면, 기본 특성 값

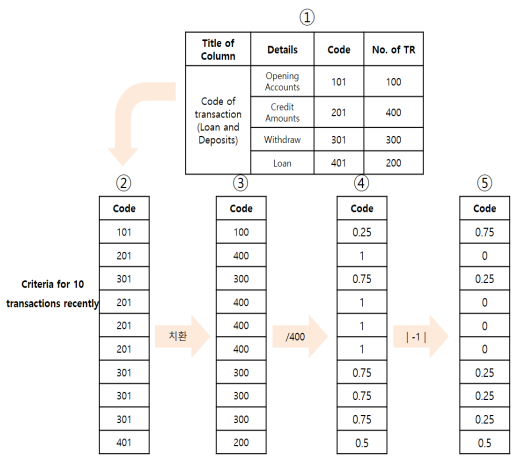

Fig. 8. Calculate frequency by default property value column(example)

7개는 필수 학습 데이터로 거래 트랜잭션(TR) 발생 시 각 특성 값 column 별 빈도수를 계산한다. 각 column 별 빈도수는 다음과 같은 과정을 거쳐 계산한다. 〔Fig. 8.〕

① 기존 거래(2017년 1월 ~ 6개월)에서 계좌당 각 column에 해당하는 값들의 분류코드 별 발생한 TR 총 개수를 산출한다.

② 계좌당 각 column에 해당 하는 값들의 분류코드 별 발생한 최근 TR을 발생 순서별로 수집한다. 최근 TR 수집 갯수는 데이터 전처리 속도 및 정탐율 모두 충족할 수 있는 수로 결정한다. 본 연구에서는 수많은 반복적인 작업을 거쳐 최근 TR 30개로 최종 결정하여 적용 하였다. TR수의 결정은 시스템 용량에 따른 속도와 정탐율을 충족하는 수준으로 결정한다.

③ 최근 수집된 TR에 해당하는 분류코드 값을 ①에서 기 수집된 TR 총 개수로 치환한다.

④ 치환된 분류코드 TR 값 중 가장 큰 수로 각각의 TR 값을 나누기 한다.

⑤ ④ 결과 값에 숫자 1을 빼고 절대 값을 적용한다. 최종 숫자를 이미지화하고 그 이미지를 알고리즘에 학습시킨다.

4.2.3 모델 생성 및 검증 단계

모델 생성 및 검증 단계에서는 Fig. 9. 와 같이 학습조건 5개, 검증조건 2개, batch number 공통 기준을 100개, learning number 공통 기준을 2000번, 특성(feature)은 8개, 직전(최근) TR 30개 기준으로 정의하였다.

학습조건과 feature 값을 다양한 형태로 구성하여 CNN 알고리즘을 통하여 학습시킨 지능화 모델을 생성한 후 생성된 모델에 검증 데이터로 정상계좌를 정상계좌로 탐지하는 비율과 사고계좌를 사고계좌로 탐지하는 비율 결과를 가지고 최적의 지능화 모델을 선정한다. 예를 들어, 학습조건1 과 feature1을 가지고 지능화 모델을 생성하는 순서는 다음과 같다 〔Fig.10.〕

① 학습용 정상계좌 600개와 사고계좌 1000개를 혼합한 후 랜덤하게 100개 계좌 추출
② 추출된 100개 계좌에 대해 각 계좌 별 feature1의 특성 값 7개 조건에 맞게 직전 TR 30개를 추출 (계좌 별 210개의 특성 생성)
③ 추출된 100개 계좌에 대해 각 계좌 별 생성된 210개의 특성을 수학적 수치 값으로 변환 후 이미지 값 생성
④ 생성된 이미지 값을 CNN 알고리즘에 학습 후 정상거래 패턴과 사고거래 패턴으로 분류
⑤ ①~④ 과정을 2000번 반복 수행하여 최종 지능화 모델 생성

Learning Condition : Jan to June 2017 | Verified condition: June of 2017 | Batch number | Learning number | Feature condition | Right before/recently(TR

Categories	Normal Accounts	Malicious Accounts
Learning Condition 1	600	1,000
Learning Condition 2	2,000	1,000
Learning Condition 3	3,000	1,000
Learning Condition 4	1,000	1,000
Learning Condition 5	5,000	1,000

Categories	Normal Accounts	Malicious Accounts
Verified condition 1	100	100
Verified condition 2	200	100

batch _num	Learning _num	feature		Right before TR
100	2,000	feature1		30
		feature2	feature2_1	
			feature2_2	
		feature3	feature3_1	
			feature3_2	
		feature4		

Fig. 9. Learning and Verification Conditions for Optimal Model Creation

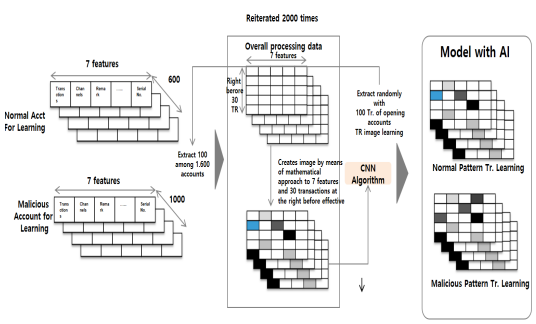

Fig.10. Intelligent Model Creation Flowchart

미지 값 생성
④ 생성된 이미지 값을 CNN 알고리즘에 학습 후 정상거래 패턴과 사고거래 패턴으로 분류
⑤ ①~④ 과정을 2000번 반복 수행하여 최종 지능화 모델 생성

생성된 지능화 모델이 고도화 모델인지를 선정하기 위하여 검증용 정상계좌와 사고계좌를 데이터 전처리를 통한 후 지능화 모델에 적용하여 정상계좌를 정상계좌로, 사고계좌를 사고계좌로 탐지하는 비율을 검증한다.

예를 들어, 검증조건2를 생성된 지능화 모델에 적용하여 검증하는 순서는 다음과 같다 〔Fig. 11.〕

① 검증용 정상계좌와 사고계좌 중 1개 계좌 추출
② 추출된 계좌에 대한 feature1의 특성 값 7개 조건에 맞게 직전 TR 30개 추출 (계좌 당 210개의 특성 생성)
③ 생성된 210개의 특성을 수학적 수치 값으로 변환 후 이미지 값 생성
④ 생성된 이미지 값을 지능화 모델에 적용하여 정상계좌인지 사고계좌인 판별
⑤ ①~④ 과정을 300번 수행하여 정상계좌(200개)를 정상계좌로 탐지하는 비율과 사고계좌(100개)를 사고계좌로 탐지하는 비율을 각각 산정

본 연구에서는 Fig. 12.와 같은 조건으로 총 61개의 지능화 모델을 생성 후 검증을 수행하였다.

총 61개의 생성된 지능화 모델 중 Fig. 13. 와 같이 검증데이터를 통하여 정상계좌를 정상계좌로 탐지한 정탐율 90% 이상과 사고계좌를 사고계좌로 탐지한 정탐율 90% 이상인 모델 6개를 선정하였다.

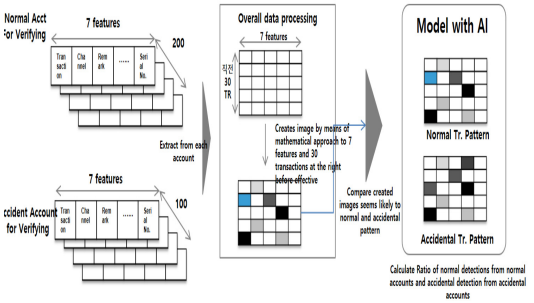

Fig. 11. The generated intelligent model verification flow chart

NO	Learning condition		Verified condition		Batch colmn (Unit: count)	No. of Learning	feature	Right before TR (Unit: times)	Correction ratio	
	Normal (Unit: number)	Malicious (Unit: number)	Normal (Unit: number)	Malicious (Unit: number)					Normal (Unit: %)	Malicious (Unit: %)
model_1	600	1000	100	100	100	2,000	feature1	30	82%	86%
model_2	600	1000	100	100	100	2,000	feature1 + feature2	30	74%	92%
model_3	2000	1000	100	100	100	2,000	feature1 + feature2	30	98%	74%
model_4	600	1000	200	100	100	2,000	feature1 + feature2 + feature4	30	89%	93%
model_5	600	1000	100	100	100	2,000	feature1 + feature2 + feature3	30	92%	86%
model_6	2000	1000	100	100	100	2,000	feature1 + feature2_1 + feature3	30	96%	72%
model_7	600	1000	200	100	100	2,000	feature1 + feature2_2 + feature3	30	91%	94%
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
model_61	600	1000	100	100	100	2,000	feature1 + feature2 + feature2_1 + feature4	30	92%	88%

Fig. 12. Intelligent model generation and verification history(sample)

NO	Learning condition		Verified condition		Batch colmn (Unit: count)	No. of Learning	feature	Right before TR (Unit: times)	Correction ratio	
	Normal (Unit: number)	Malicious (Unit: number)	Normal (Unit: number)	Malicious (Unit: number)					Normal (Unit: %)	Malicious (Unit: %)
model_7	600	1000	200	100	100	2,000	feature1 + feature2_2 + feature3 + feature4	30	91%	94%
model_16	600	1000	200	100	100	2,000	feature1 + feature2_1 + feature4	30	90%	92%
model_19	600	1000	200	100	100	2,000	feature1 + feature2_1 + feature3 + feature4	30	90%	93%
model_28	600	1000	200	100	100	2,000	feature1 + feature2_2 + feature4	30	90%	93%
model_32	600	1000	200	100	100	2,000	feature1 + feature2_2 + feature3 + feature4	30	92%	95%
model_33	1000	1000	200	100	100	2,000	feature1 + feature2_2 + feature3 + feature4	30	96%	90%

Fig. 13. Intelligent models with over 90% of normal account and accident account specifications

지도학습 지능화 모델 연구 결과에 의하여 정상계좌를 정상계좌로 탐지한 최고 정탐율(96%) 조건은 model_33 이며, 사고계좌를 사고계좌로 탐지한 최고 정탐율(95%)은 model_32 조건임을 알 수 있다. 하지만, 실제 A은행의 하루 평균 입·출금 거래가 일어나는 계좌 수는 약 2,088,000 개로 model_32 적용하여 정상 계좌임을 탐지하면 8%인 약 166,970개를 과탐으로 오인하여 탐지 할 수 있다. 이런 결과를 실제 금융회사 운영환경에 지도학습 인공지능 모델만을 적용하면 금융회사는 과탐 거래의 민원을 처리하느라 업무가 마비 될 것이고 국민들 또한 금융혼란에 빠지게 되어 전기통신금융사기를 탐지하는 모델로는 불가능하다는 것을 알 수 있었다.

결국, 전기통신금융사기 지도학습 지능화 모델 연구 결과로 보면

지능화 모델만으로는 한계가 존재하기에 시나리오와 결합한 Hybrid 기반으로 적용하여 운영하는 것이 적절하다고 판단하였다. 전기통신금융사기를 Hybrid 기반으로 탐지하기 위하여 연구한 지능화 모델 중 model_32를 최종 선택하였다. 그 이유는 지능화 모델로 전기통신금융사기 즉, 사고계좌 탐지 비율을 높이고 시나리오 적용을 통하여 정상계좌 과탐 비율을 최소화 할 수 있다고 판단하기 때문에 모델 중 가장 사고계좌 정탐율이 높은 model_32를 선택하게 된 것이다.

V. 고도화 된 지능화(AI) 탐지 모델 제안

5.1 Hybrid기반 지능화 탐지 모델 제안

전기통신금융사기 지도학습 지능화 탐지 모델 연구 결과에서 알 수 있듯이 실제 금융환경 적용 시에 너무 많은 과탐(오탐)이 발생한다는 것이다. 그러하기에 Fig. 14. 과 같이 과탐을 최소화하기 위하여 우선 1차로 시나리오기반으로 이상금융거래를 탐지 후 2차로 인공지능 모델을 통하여 최종 이상금융거래를 탐지하는 Hybrid 방안을 제안한다. 이유는 1차 시나리오기반 탐지를 통하여 2차 인공지능 모델이 탐지 할 수 있는 거래 데이터를 최소화하여 과탐을 줄이고 이상금융거래 정탐율을 높일 수 있기 때문이다. 단, 1차 시나리오기반 탐지에서 실제 사고데이터(이상금융거래) 95% 이상을 포함하여 탐지하는 전제조건이 있어야한다.

제안한 Hybrid 기반 지능화 탐지 모델의 실효성을 검증하기 위하여 A은행에서 2017년 7월 상품(계정)시스템에서 시나리오기반으로 탐지한 이상금융거래 3,000건과 7월에 실제 발생한 전기통신금융사기 60건 중 실제 시나리오로 예방한 13건을 가지고 다음과 같은 조건으로 검증을 수행하였다〔Fig. 15.〕 검증 수행결과 1차 시나리오 기반으로 탐지한 것 보다 2차 인공지능 모델을 통하여 Hybrid 기반으로 탐지한 결과 탐지수가 3,000건에서 460건으로 과탐이 약 85% 감소하였음을 알 수 있으며, 실제사고 예방 탐지율도 21%에서 83%로 약 4배 증가하는

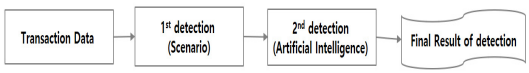

Fig. 14. Hybrid-based Intelligent Telecom finance fraud Detection Model Flowchart

Categories	Number of detection	Number of accidents	Prevention	Detection Ratio
1st detection (Scenario)	3,000	60	13	**21%**

3,060

Categories	Number of detection	Number of accidents	Prevention	Detection Ratio
2nd detection (Artificial Intelligence)	460	60	50	**83%**

Fig. 15. Hybrid based Intelligent Detection Model Verification Result

효과가 있음을 알 수 있다.

결과적으로 전기통신금융사기의 효과적인 예방을 위해서는 시나리오기반과 인공지능기반이 결합한 Hybrid 기반 지능화 탐지 모델이 실제 금융회사 환경에서 적용 가능한 모델임을 확인할 수 있다.

5.2 고도화된 통합 대응 프로세스 제안

전기통신금융사기 사고를 신속하고 정확하게 탐지하는 것도 중요하지만, 탐지 후 보안사고 유형 별 효과적으로 대응할 수 있는 프로세스도 중요하다. 효과적인 대응 프로세스를 통하여 이상금융 거래 오탐 시 발생하는 소비자의 민원을 최소화하고 실제 이상금융거래 탐지 시 임시 지급정지 등의 다양한 대응 프로세스로 고객의 피해를 예방할 수 있다.

본 연구에서는 지능화 FDS 시스템에서 전자금융 불법이체 사고와 전기통신금융사기를 통합 대응 가능한 고도화 된 통합 대응프로세스를 Fig. 16. 와 같이 제안한다.

대응 프로세스에 대해 간단히 설명하면,

'Security Management(보안관제)'대응은 불특정 다수가 대량으로 부정접속 발생 시 해당 접속 IP 주소를 24*365 운영중인 관제 조직과 연동하여 해당 IP주소를 네트워크 보안시스템에서 차단하는 프로세스 이다.

'Stop Fund Transfer(이체정지)' 대응은 FDS 시스템에서 이상금융거래로 탐지한 대부분 거래는 고객번호 및 이용자 아이디 기반으로 추가인증(ARS 인증)하는 프로세스를 거쳐, 실패한 경우는 전자금융 채널에서의 출금을 제한하게 된다.

'Delaying Fund Transfer(지연이체)' 대응은 보이스피싱 전용 대응프로세스로 실시간 자금 이체의 경우 보이스피싱 이상금융거래로 탐지되면 1차 추가인증(ARS 인증 등)을 적용하며, 추가인증 후 본거래 성공 시 1시간동안 지연이체가 적용되는 프로세스이다. '지연이체' 로 탐지된 경우는 보안사고 담당부서에서 반드시 1시간이내 Outbound Call(가칭 : CareCall)을 통하여 실제 사고 여부를 확인하게 된다.

'Stop Payment(지급정지)' 는 이상금융거래가 확실하다고 판단된 거래에 대하여 고객번호 기준으로 모든 채널에 대하여 출금을 제한하는 프로세스 이다.

'Partially Stop Payment(일부지급정지)' 는 사기의심계좌에 특정 금액 입금 시 탐지된 경우 특정 입금된 금액에 대해서만 임시 출금을 차단하는 프로세스이다.

'Stop ATM Payment(ATM지급정지)' 는 자·타행 ATM에서 출금 시도 시 전기통신금융사기로 탐지된 경우로 ATM에서 임시 출금을 차단하는 프로세스이다.

'Suspend Accounts(창구출금불가)' 는 자행 영업점에서 출금 시도 시 전기통신금융사기로 탐지된 경우로 임시 영업점에서 출금을 차단하는 프로세스이다.

'Stop Payment through on-site banking in branch(전부지급정지)' 는 실제 전기통신금융사기로 판단한 경우 해당 사기의심계좌에 대하여 모든 채널에서의 출금을 차단하는 프로세스이다.

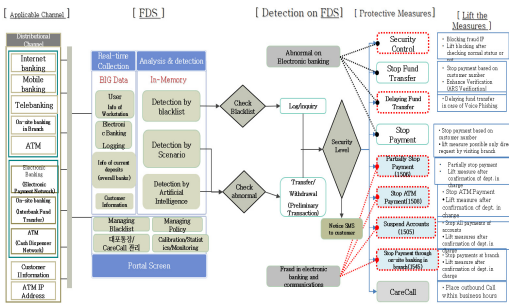

Fig. 16. Optimized Integrated Response Process

VI. 지능화(AI) 탐지 모델 실증적 검증

본 논문에서 제안한 고도화된 통합 지능화 FDS 시스템과 Hybrid 기반 지능화 탐지 모델 방안의 효과성을 검증하기 위해서 A은행 실제 운영환경에 적용(2108.10.01. ~ 11.16)하여 전기통신금융사기 사고현황 및 예방 추이를 연구하였다. A은행의 하루 평균 요구불(수신) 계좌의 입·출금 거래는 약 5,925,000건 이며, 하루 평균 입·출금 거래가 발생하는 계좌는 2,088,000개 이다. A은행에서 기존 전기통신금융사기 탐지 운영 방식(Before)과 실증 검증을 위하여 본 연구에서 제안한 Hybrid 기반 지능화 탐지 모델 적용 후 방식(After)의 차이점을 비교하였는데, 탐지방식은 'Inline 방식'에서 'External'으로 , 탐지대상은 '전체 고객의 70% 수준'에서 '전체고객'으로, 탐지방법은 '단순시나리오 탐지(임시지급정지 2개, 모니터링1개)'에서 'Hybrid 기반 지능

화 탐지(복합시나리오 + 지능화 모델)'로 변경하여 적용하였다.

〔표 2〕는 본 연구에서 제안한 Hybrid 기반 지능화 탐지 모델 적용 전과 적용 후의 하루 평균 전기통신금융사기 탐지 수, 사고예방 수, 실제 사고 수, 예방율, 예방금액을 나타낸다.

제안한 Hybrid 기반 지능화 탐지 모델 적용을 통하여 적용 전 상품(계정)시스템을 통한 일일 평균 과탐비율은 약 79% 감소, 사고예방은 약 3배 증가, 실제사고는 약 4.2배 감소, 예방율은 약 3.6배 증가한 효과를 알 수 있다. 또한 오탐율은 약 12% 감소하였으며, 특히 적용 전 대비 일일평균 4배 이상 증가한 약 9천 7백원 정도의 금액을 예방함을 알 수 있다. 하지만, A은행 실제 운영환경에서의 검증과정을 통하여 '모델 생성 및 검증' 단계에서 발견하지 못한 중요한 문제점을 확인 할 수 있었다. Hybrid기반 지능화 탐지 모델의 탐지과정 중 1단계 인 시나리오 탐지에서 95% 이상의 실제사고를 포함한 이상금융거래를 탐지하더라도 2단계 지능화 모델을 통한 탐지 시 실제사고 예방율이 85% 이상 넘지 않는 것에 의문점을 가졌고, '모델 생성 및 검증' 단계에서 2가지 경우를 확인하지 않고 모델링 작업을 한 문제점

Table 2. Comparison results before and after application of intelligent detection model〔By: Daily average〕

classification	Before	After
Term	2018.09.01. ~ 09.29	2018.10.01. ~ 11.16
Detectable count	534	114
accident Preventive count	5	15
real count of Accidents	17	4
Prevention rate	22%	**79%**
Prevention Amount (KRW 1,000)	24,800	**97,000**
False Rate	99%	**87%**

을 발견하였다. 첫 번째 문제점은 학습한 사고데이터에 거짓 사고계좌가 일부 포함되어 있는 경우이며, 두 번째는 학습한 사고데이터의 시점이 최초 사고가 발생한 시점이 아니라 일부 사고계좌에 대해서는 2~3번째 발생한 사고시점을 기준으로 학습이 된 경우이다. 결국 사고데이터에 쓰레기(Garbage) 데이터가 포함되어 학습이 되었으며, 학습된 지능화 모델로 85% 이상 사고 예방을 하는데 한계가 존재했다는 사실을 확인 할 수 있었다. 또한 Before 대비 오탐율은 감소하였지만 목표치인 60% ~ 70% 범위에 들어오기 위해서는 1단계인 시나리오 탐지에서 고객 및 계좌 별 개인화 프로파일링을 이용한 정교한 Rule 설정이 필요함을 알 수 있었다. 앞으로 위에서 언급한 '모델 생성 및 검증' 단계에서의 2가지의 문제점을 제거하고 1단계 시나리오의 정교한 탐지 Rule을 동시에 적용 한다면 90%이상의 예방율을 달성 할 것이라 판단한다.

VII. 결 론

앞에서 기술 한 바와 같이 전기통신금융사기 사고의 지속적 증가, 그리고 지능화 및 고도화 현상을 고려하면, 보이스피싱에 대한 인공지능 기반의 대응 연구는 좀더 일찍 그리고 실질적으로 이루어졌어야 했다고 보여진다. 본 연구에서는 현행 전기통신금융사기 사고에 대한 이상금융거래 탐지의 기술적 한계 대하여 살펴보고, 이를 극복하기 위한 효과적인 이상금융거래 탐지 지능화 모델 구축에 있어서 시나리오 기반과 인공지능 기반의 결합 된 Hybrid 형태의 최적의 지능화 모델을 제안하였다.

연구 결과의 의의로는 첫 번째로, 기존의 전자적 매체를 이용한 전자금융 불법이체 사고에 대한 이상금융거래 탐지를 대응 체계에서 벗어나 그동안 대응이 어렵고 힘들고 효과도 미미한 사회공학적 방법의 전기통신금융사기 사고에 대해서도 기술적 한계를 해결하고 기존의 Rule방식을 고도화하고 새로운 개념의 인공지능 기반탐지 모델을 시스템화하여 지능화된 대응이 가능하게 했으며, 그 적용 결과도 매우 뛰어난 결과를 도출한 연구라는 점이다. 두 번째로는 최적의 이상금융거래 탐지 모델을 위해 하이브리드 엔진을 설계하고 시나리오 기반의 룰엔진과 인공지능 데이터분석 머신러닝 알고리즘을 적용한 지능화 FDS 모델을 통합한 Hybrid 형의 새로운 이상금융거래 탐지 모델을 제시했다는 것이다. 세 번째로는

국내의 전자금융거래 사고 중 가능 높은 금융사고율을 보이고 있는 보이스피싱을 통한 금융사기 사고에 대해 기존의 해당 사고의 탐지, 차단율을 약 400% 정도의 상승한 정탐율과 탐지 차단율을 구현하여 기대이상의 효과를 얻으므로서 본 연구의 유효성을 확보하였다는 점이다. 또한 본 연구의 결과가 넓게는 국가의 숙원 과제인 전지통신금융사기로 발생하는 피해를 대폭 방지하고 피해금액 감소는 물론 국가 금융기반인 국내 금융회사에 대한 신뢰성을 제고 할 수 있을 것이라 기대한다.

향후 앞에서 제시한 전기통신금융사기 사고 탐지 모델에 더불어 전자금융 불법이체 사고에 대한 인공지능 기반의 이상금융거래탐지 모델의 연구도 진행되어야 한다고 본다. 또한 지능화 모델링 작업도 수동적 작업이 아닌 실시간의 자동화 작업 모델 연구도 지속되어야 할 것이다.

References

[1] FSC(Financial Services Commission), "Financial Security Comprehensive Plan", , Vow. 2013.

[2] FSS(Financial Supervisory Service), "FDS upgrade oadmap for financial industry", Dec. 2014

[3] Choi E.S., Lee K.H, "A Study on Improvement of Effectiveness Using Anomaly Analysis rule modification in Electronic Finance Trading", *Journal of the Korea Institute of Information Security & Cryptology*, 25(3), pp. 615-625 (11 pages), Jun. 2015.

[4] Park E.Y., Yoon J.W., "A Study of Accident Prevention Effect through Anomaly Analysis in E-Banking", *The Journal of Society for e-Business Studies*, Vol.19, No.4, pp.119-134, November 2014

[5] Yoon H.S., Kim Y.G., "A Study on Specific Prevention Methods for Voice Phishing Damage Types", *The Supreme Prosecutors' Office report*, 2017.

[6] Jung K.E., Kim Y.R., Min Y.K., "Application of Psychological Triggers to Voice Phishing : Focusing on Social Engineering", *Social Science Research,* Volume 28, Issue 4, 181-194 (14 pages), Oct. 2017

[7] Yoon H.S., Kwak D.K., "Study on Prevention and Countermeasure of Voice Phishing", *Criminal Policy Institute Research Series*, pp. 9-118., Dec. 2009

[8] Cho H.D., "Voice Phishing Occurrence and Counterplan", *The Korea Contents Society*, 12(7), pp. 176-182, Dec. 2012

[9] Shin S.C., "An Analysis on the Activities of Taiwanese Voice Phishing Crime Organizations in Korea", *Asian Research*, 21(3), pp. 151-191, Aug. 2018

[10] Financial Supervisory Service, National Police Agency, "Voice phishing guard Announcement statistics", Feb. 2018

[11] Financial Supervisory Service, National Police Agency, "Voice phishing guard Announcement statistics", Sep. 2018

[12] Financial Supervisory Service, "Status and supervision of abnormal financial transaction detection system of banks and securities companies in 2017", Jun. 2018

[13] Wikipedia, https://wikitech.wikimedia.org/wiki/Logstash

[14] Apache, https://kafka.apache.org/

[15] Wikipedia, https://wikitech.wikimedia.org/wiki/Redis

[16] Wikipedia, https://ko.wikipedia.org/

[17] Namuwiki, https://namu.wiki/w/텐서플로우/

[18] Namuwiki, https://namu.wiki/w/크로스플랫폼

[19] Financial Supervisory Service, "FDS technical guide", Aug. 2014

〔20〕 LeCun,y., "Gradient-based learning applied to document recognition", *Preceedings of the IEEE*, 86(11), pp. 2278-2324, Nov. 1998

〔21〕 T. Sainath et al., "Convolutional neural networks for LVCSR," ICASSP, 2013.

〈저자소개〉

정 의 석(Eui-seok Jeong) 정회원
1989년 2월: 광운대학교 전자계산기공학과 졸업
2004년 2월: 고려대학교 정보보호대학원 석사
現 KEB하나은행 정보보호부 부장
〈관심분야〉 전자금융보안, FDS, 네트워 보안, 개인정보보호, 클라이언트, 융합기술보안 등

임 종 인 (Jong-In Lim) 종신회원
1980년 2월: 고려대학교 수학과 졸업
1982년 2월: 고려대학교 수학과 석사
1986년 2월: 고려대학교 수학과 박사
現 고려대학교정보보호대학원교수 고려대학교사이버국방학과교수 개인정보보호위원회 위원, 대검찰청 디지털수사자문위원회 위원장, 국방부 정보화책임관자문위원, 한국저작권위원회 위원 등
〈관심분야〉 사이버국방, 정보법학, 디지털포렌식, 개인정보보호, 융합기술보안 등