유튜브 썸네일 캐시 기반의 제로-클릭 스테가노그래피 C&C 통신 기법 연구

이형주*, 조영호(교신저자)**
*국방대학교 국방관리대학원 국방과학학부 사이버컴퓨터공학과 석사과정
**국방대학교 국방관리대학원 국방과학학부 사이버컴퓨터공학과 교수
e-mail: younghocho@korea.kr

A Study of YouTube Thumbnail Cache-Based Zero-Click Steganographic C&C Communication Method

Hyungjoo Lee*, Youngho Cho**
*Master's Course, Dept. of Cyber Security and Computer Engineering, Korea National Defense University
**Professor, Dept. of Cyber Security and Computer Engineering, Korea National Defense University

요 약

스테가노그래피(Steganography) 은닉 통신 기술은 디지털 이미지·영상·오디오 등 매체의 통계·지각 특성을 이용해 비밀 정보를 눈에 띄지 않게 은밀히 삽입·전송하는 기술이며, 최근 SNS 플랫폼을 활용한 스테가노그래피 봇넷 구축이 제안되었다. 기존 연구는 주로 업로드·전송 단계의 손실 변환(재압축·스케일링 등)에 견디는 설계에 집중되어 왔지만, 사용자 단말의 브라우저 캐시를 통신 경로로 활용하는 가능성은 충분히 다뤄지지 않았다. 본 논문은 유튜브 채널 재생목록 표지(480×270 JPEG)를 은닉 통신 매체로 하여 스크롤만으로도 수신이 가능한 제로-클릭 봇넷 C&C(Communication & Control, C2) 통신 채널을 제안한다. 제안 기법의 핵심은 저주파 듀얼 캐리어(Dual-Carrier) 및 QIM(Quantization Index Modulation) 기법으로 동일 비트를 두 계수에 이중 기록해 SNR(신호 대 잡음비)을 높이고, 파일럿·Hamming(7,4)·반복 삽입 및 q-스윕·위상 탐색 등의 방식을 조합하여 임계 보정·동기화·강건성을 확보하는 것이다. 초도 실험 결과, 브라우저·플랫폼 변동에도 28 바이트 용량의 봇넷 C2 신호가 사용자(victim)의 별도 클릭없이 반복 복원됨을 입증했다. 이는 유튜브 플랫폼과 브라우저 캐시 계층을 결합한 데이터 은닉·복원 방식이 실질적인 위협 시나리오로 작용할 수 있음을 시사한다.

1. 서 론

스테가노그래피(steganography)는 이미지·영상 등 다양한 디지털 매체에 정보를 은밀히 삽입하는 기술로써, SNS 플랫폼 확산과 함께 악용 가능성 또한 커지고 있다. 특히, 스테가노그래피 봇넷(Stego Botnet)은 SNS 플랫폼에 공유된 평범한 매체를 C&C(Command & Control) 신호의 운반체로 활용하고, 피해자 단말에 상주한 봇 에이전트가 이를 자동 수신·해석하여 지령 등을 복원하는 위협 모델로 주목받고 있다.

기존 스테가노그래피 연구는 주로 업로드·전송·코덱 재압축 등 콘텐츠 유통 경로의 손실 변환에 견디도록 임베딩 기법을 개선하는 데 집중해 왔다[1, 2]. 그러나 실제 사용자 측의 웹 브라우저가 자동 축적하는 클라이언트 캐시를 '전달 채널'로 직접 모델링하거나, 웹 스크롤 행위만으로 신호가 수신되는 위협을 체계적으로 검증한 사례는 충분치 않다.

본 논문은 유튜브 채널의 '재생목록(Playlist)' 화면을 스크롤하는 것만으로 자동 생성·축적되는 480×270 JPEG 썸네일 캐시를 은닉 매체로 활용하는 제로-클릭(Zero-Click) 스테가노그래피 C&C 통신 기법을 제안한다. 구축된 제안 채널은 사용자의 추가 행위(클릭 또는 다운로드 등) 없이도 은닉된 메시지가 주기적으로 유입되는 자동전달(autodelivery) 특성을 갖는다. 상기 채널은 은닉 통신(지령 전달 등)과 C&C 통신에 모두 적용될 수 있으나, 본 연구에서는 C&C 통신 채널로 한정지어 활용한다. 제안 기법으로는 유튜브 웹 브라우저 캐시의 관측 특성에 정합되는 임베딩·복원 파이프라인을 제시한다. 해당 기법은 브라우저나 플랫폼 변환(리사이즈, 재압축, 메타데이터 스트립 등)에도 은닉 데이터가 생존함과 동시에 스크롤만으로 쌓인 캐시로부터 최대 28 바이트의 소용량 C&C 신호를 자동적으로 복원하도록 설계되었다.

또한, 본 논문은 최종적으로 유튜브 재생목록 구간의 JPEG 캐시 경로를 위협 채널로 모델링한 브라우저 캐시 기반의 C&C 통신이라는 실전형 위협 시나리오를 제시한다. 이를 입증하기 위해 실제 상용 브라우저 환경에서 은닉된 메시지가 자동적으로 전달 및 복원됨을 실험으로 보인다. 이를 통해 사용자가 유튜브 영상 목록을 스크롤하는 것만으로도 작동하는 C&C 채널이 사이버 보안 분야에 있어 새로운 위협이 될 수 있음을 초도 실험을 통해 나타낸다.

2. 배경지식 및 관련 연구

2.1 스테가노그래피 봇넷

스테가노그래피 봇넷은 C&C 메시지를 멀티미디어 콘텐츠(이미지·영상 등) 파일에 은닉하여 전달하는 형태로, 탐지 회피·은닉성 측면에서 강점을 가진다. 공격자(Bot master)가 SNS처럼 공개된 플랫폼에 평범해 보이는 스테고 매체를 피해자 단말(victim device)에 전송하고 봇 악성에이전트(SW)가 해당 스테고 매체로부터 명령을 추출하여 복원한다[3, 4, 5].

성공적인 봇넷 C&C 통신을 위해서는 C2 메시지의 연속적인 전송이 핵심이다. 비콘(beacon)·토큰 교환 등 신호가 주기적으로 흐르지 못하면 제어가 불안정해지고, 클릭·다운로드 같은 중간 상호작용이 개입될 경우 통신 지연 및 실패 확률이 증가할 수 있다. 따라서 본 논문은 자동 로드·저장되는 웹 브라우저 캐시 계층을 C2 전달 채널로 사용하여 연속적인 신호 교환이 가능한 스테가노그래피 C&C 통신을 제시한다.

2.2 유튜브 썸네일 캐시 생성과정 및 알고리즘

[그림 1] 유튜브 영상 업로드시 썸네일 지정 화면

캐시란 동일(또는 유사)한 요청의 지연을 줄이고 네트워크 사용을 절감하기 위해, 클라이언트·중간 경유지·서버 등에 일정기간 동안 응답을 저장해 재사용하는 메커니즘이다. 유튜브에서는 영상 업로드 시 지정된 포맷(JPEG/PNG/GIF)으로 대표 썸네일을 등록할 수 있으며(그림 1), 유튜브 웹은 화면 스크롤 시 화면에 진입한 카드의 썸네일을 요청하는 패턴을 보인다. 이때 내려받은 응답은 브라우저별 HTTP 캐시 정책에 따라 사용자의 메모리·디스크에 저장된다[6].

최신 브라우저(Chrome v140, Firefox v142)상에서 유튜브 홈·검색·영상 시청 등 대다수의 영역의 썸네일 캐시 포맷은 AVIF·WebP 형태이다. 해당 포맷은 컨테이너 완결성을 전제로 동작하므로, 응답 스트림이 혼입·분절되어 기록되는 경우에는 은닉한 데이터가 훼손될 가능성이 있다.

[그림 2] 유튜브 채널(YTN) 재생목록 구성 화면

[그림 3] 재생목록에서 생성된 JPEG 캐시(좌. Firefox 우. Chrome)

반면, 유튜브 채널 내 '재생목록' 배너(그림 2)에서는 두 브라우저 모두 JPEG(480×270) 썸네일 캐시를 생성한다(그림 3). JPEG 포맷은 다른 형태와 달리 SOI 시그니처(0xFF D8)로 시작하며, 캐시 스트림이 훼손되어도 SOI부터 핵심 헤더가 남아 있으면 부분적으로 복원이 가능하다. 본 연구는 이러한 재생목록에서의 JPEG 캐시 경로를 통신 채널로 활용한다.

2.3 유튜브 플랫폼에서의 트랜스코딩(Transcoding)

유튜브는 업로더가 지정한 단일 썸네일을 그대로 배포하지 않는다. 그 대신 서비스 품질과 전송 효율을 동시에 충족하기 위해, 사전 정의된 규격에 따라 서로 다른 해상도·비트레이트·포맷의 렌디션을 일괄적으로 생성한다. 이 단계에서 리사이즈(보간), 색공간 변환(RGB↔YCbCr), 양자화(손실 및 재압축), 메타데이터 스트립(ICC/EXIF 제거) 등이 발생한다. 결과적으로 이러한 변환 과정은 삽입 신호가 놓인 공간을 변형하고 노이즈를 추가하여, 은닉된 데이터를 훼손할 수 있다.

2.4 기존 연구 및 제한사항

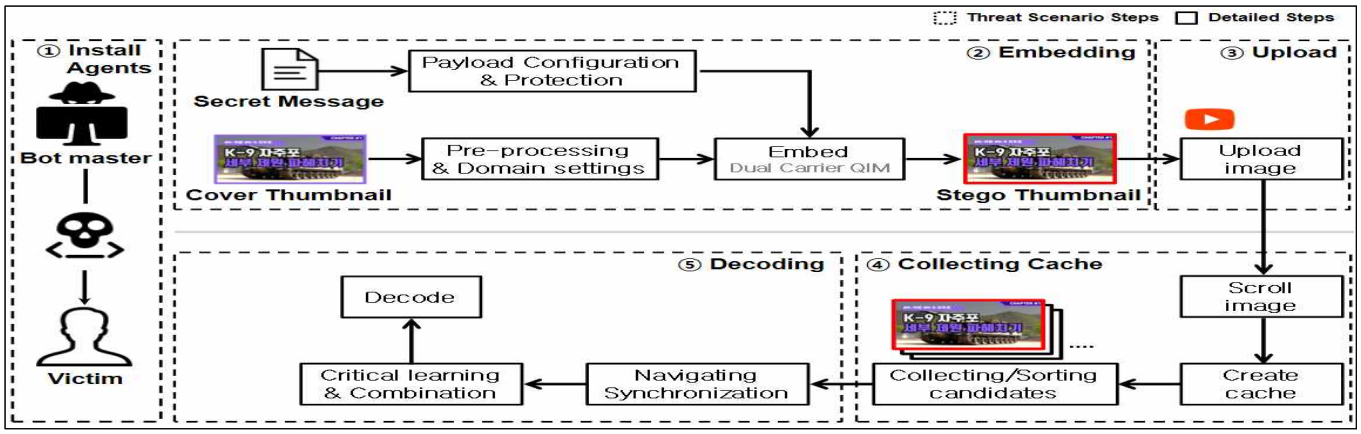
소셜미디어·플랫폼 환경에 강건한(robust) 스테가노그래피 연구는 업로드·전송·트랜스코딩 과정의 손실 변환에 견디도록 발전해 왔다. 대표적으로 PCA(Principal Component Analysis) 기반의 비디오 은닉·복원 방식[1], 플랫폼의 프리뷰·썸네일 스트림에 적응형 임베딩을 결합한 접근[2] 등이 보고되었다.

스테가노그래피 은닉 통신 영역에서는 메신저(카카오톡, 텔레그램)에서 오디오 파일을 매체로 사용해 C&C 은닉 통신을 구현한 연구가 있다[3, 4]. 또한, 플랫폼을 확장시켜 메타버스(로블록스) 환경에서 은닉 통신을 구현한 연구도 존재한다[5].

그러나 기존 선행연구들은 공통적으로 사용자 단말에 생성되는 캐시를 통신 채널로 직접 모델링하지 않으며, 사용자의 직접적인 클릭, 다운로드 등의 행위를 전제로 한다. 본 논문은 이러한 제한사항에 주목하여, 웹브라우저 캐시를 은닉 통신의 전달 경로로 활용하는 새로운 기법을 제안함으로써 데이터 은닉·복원 과정의 자동화 및 강건성을 동시에 향상하고자 한다.

3. 제안 기법

3.1 위협 시나리오 및 동작 절차



[그림 4] 위협 시나리오 흐름도

그림 4는 본 논문에서 제안하는 C&C 통신 기법을 활용한 위협 시나리오를 설명한다. 기본적인 흐름은 다음과 같다.

① 악성 에이전트 설치 단계(가정): 공격 개시 전, 공격자의 스피어피싱 등을 통한 해킹으로 피해자(Victim) PC에 악성 에이전트가 설치되었다고 가정한다.

② 데이터 은닉 단계: 공격자는 제안 기법을 통해 지령이나 명령어 등과 같은 데이터를 페이로드로 구성 및 보호한 후, 썸네일 원본 이미지에 삽입하여 스테고 썸네일을 생성한다.

③ 썸네일 업로드 단계: 공격자 관점에서 피해자를 채널로 유도하는 현실적 경로로 타겟(예: 방산업계 종사자 등)이 자주 방문할 만한 주제의 유튜브 채널(예: K-방산 소개 채널)을 구축하는 방식을 택한다. 공격자는 생성한 스테고 썸네일을 자신의 채널 재생목록 영역에 노출한다.

④ 캐시 수집 단계: 피해자가 해당 채널의 재생목록 화면을 스크롤하면 썸네일 이미지가 자동 요청·렌더링되며, 브라우저 캐시 디렉터리에 해당 파일들이 자동으로 저장된다. 봇 에이전트는 캐시에서 썸네일 규격(JPEG, 480×270)이 일치하는 후보 썸네일들을 자동 수집·정렬해 복원 대상 풀을 구축한다.

⑤ 데이터 복원 단계: 에이전트는 후보 썸네일에 대해 복원 알고리즘으로 오프셋·양자화 강도 등 복원 파라미터를 탐색 후, 비트열을 판정한다. 이후 오류정정 및 무결성 검증을 거쳐 페이로드를 파싱하고 본문을 해독해 데이터를 복원한다.

3.2 설계 내용

3.2.1 임베딩(삽입) 기법

① 페이로드 구성 및 보호 단계: 은닉 데이터는 MAGIC(4B)|LEN(4B)|BODY(본문)|CRC32(4B) 구조로 패키징한다. 채널 부호화로 Hamming(7,4)을 적용해 단일 비트 오류를 정정하고, 반복 삽입(repeat 3)으로 현장 노이즈에 대한 신뢰도를 보강한다(반복 수는 조정 가능하다).

② 전처리 및 도메인 설정 단계: 커버 썸네일을 480×270으로 정규화 후, JPEG 8×8 부호화에 맞춰 Y(밝기)채널을 격자 정렬한다. 은닉 가능 용량은 이미지당 1,980bit가 된다. 은닉 도메인은 Y채널의 8×8 DCT(Discrete Cosine Transform)이며, 지각 왜곡이 작고 재압축 영향이 적은 저주파 AC 계수 쌍 (2,1) /(1,2)을 캐리어로 선택한다. 임계·양자화 강도 q는 강건성/왜곡 균형을 고려해 설정한다. 파일럿(예: 0101… 패턴 등)은 복원 단계에서 설명할 임계 학습과 동기화 복구의 기준이 된다.

③ 임베딩 단계(Dual Carrier QIM): 은닉 데이터는 선택된 블록의 (2,1)/(1,2) 두 계수(듀얼 캐리어 방식)에 QIM 기법으로 동시에 임베딩된다. 최종 스테고 썸네일은 JPEG 포맷으로 인코딩해 업로드에 사용한다.

3.2.2 추출(복원) 기법

① 탐색 및 동기화 단계: 후보 썸네일을 Y 채널 8×8 격자에 정렬한 뒤, 업로드·재압축 과정에서 발생 가능한 0..7 범위의 블록 시작 위치 오프셋(dx,dy)과 양자화 강도 q 후보를 전수 탐색한다. 각 조합에 대해 듀얼 캐리어 관측치를 수집하고, 매직 헤더의 존재 여부를 탐지해 동기화한다.

② 임계 학습 및 결합 단계: 파일럿 구간에서 관측 분포를 추정해 임계값을 자동 학습(보정)하고, 듀얼 캐리어 점수를 가중 결합해 블록별 0/1 판정을 수행한다. 반복 삽입된 표본은 신뢰도 순 정렬 후 트림 다수결로 결합해 오류를 감소시킨다.

③ 복원 단계: 비트열에 Hamming 복호를 적용하고, 페이로드 구조의 검증을 통과한 데이터만 채택되어 본문이 복원된다. 복원 실패 시 다음 후보(오프셋·q)에 대해 복원이 이루어진다. 복원된 데이터는 에이전트의 정책에 따라 처리된다.

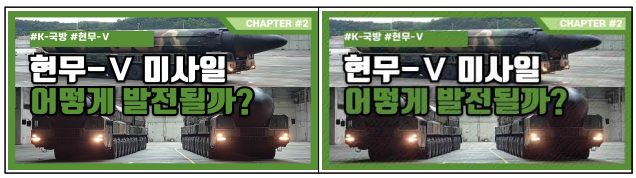
4. 초도 실험 결과

4.1 실험 목적 및 방법

초도 실험의 목적은 유튜브 채널 재생목록(480×270) JPEG 썸네일 캐시만으로 C2 명령이 클릭 없이 전달·복원되는지를 입증하는 것이다. 구체적인 목적은 다음과 같다. 첫째, 스크롤만으로 브라우저 캐시 자동 생성 여부와 스테고 썸네일 추출가능성을 확인한다. 둘째, 변환·재압축·블록 위상 변동에도 듀얼 캐리어 QIM

과 파일럿·반복·Hamming·CRC 체인이 정상 동작하는지 점검한다. 마지막으로 설계 파라미터(q≈56, repeat =3, pilot=96)에서 은닉 본문 상한(약 28B)을 검증한다.

[그림 5] 실험에 사용된 이미지(4장)



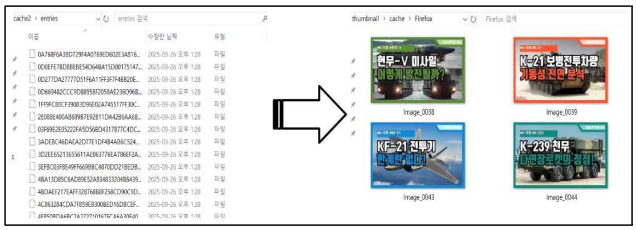
[그림 6] 원본 이미지와 스테고 이미지의 육안 비교(좌. 원본 우. 스테고)

임베딩 데이터셋은 그림 5처럼 직접 제작한 썸네일 이미지 4장으로 구성했다. 제안 기법을 활용하여 각각의 이미지에 서로 다른 데이터가 은닉된 스테고 이미지를 생성했으며, 그림 6에서 보이는 것처럼 육안 구별이 어려움을 알 수 있다.

[그림 7] 공격자 유튜브 채널 내 재생목록 썸네일 예시

공격자 채널을 실제로 구현하기 위해 실험용 유튜브 채널 내에 총 4개의 재생목록을 신설하였으며, 각 재생목록별로 스테고 이미지를 썸네일로 지정한 영상을 대표 이미지로 지정하였다(그림 7).

4.2 실험 결과 및 분석



[그림 8] 생성된 브라우저 캐시 및 추출된 JPEG 이미지

실험 결과는 다음과 같다. 첫째, 피해자가 공격자의 유튜브 채널 재생목록 화면을 스크롤만 해도 다양한 캐시 파일이 브라우저별 디렉터리 내에 자동으로 생성되었다. 또한, 수집된 캐시 파일로부터 재생목록 표지에 노출된 스테고 썸네일 이미지를 선별하여 성공적으로 추출됨을 보였다(그림 8).

[그림 9] 복원 알고리즘 결과(은닉 텍스트 정상 복원 완료)

둘째, 캐시에서 추출한 썸네일에 대해 디코더가 q·오프셋 전수 탐색과 파일럿 임계 추정을 거쳐 듀얼 점수를 정상 분리했고, 트림 다수결 및 Hamming·CRC32 체인을 이상 없이 통과했으며, 모든 이미지에서 은닉 본문(C2 지령 등 4개 본문)을 성공적으로 복원했다(그림 9). 이로써 브라우저 캐시로 소용량 C2 신호를 송·수신할 수 있음을 실증적으로 보여준다.

기본 파라미터(q≈56, repeat=3, pilot=96) 설정에서 28B 용량의 텍스트 데이터를 은닉할 경우, 썸네일 이미지 1장당 수용 가능한 1,980bit 슬롯 중 약 1,974bit를 사용하게 되며, 실용적인 본문 상한이 28B임을 나타낸다. 이는 비록 소용량이나, 단순 지령(명령)·세션 키·토큰·플래그 등 지휘·제어에 필요한 신호를 전달하기에는 충분한 수준이다.

5. 결론 및 향후 연구계획

본 연구는 유튜브 채널 재생목록 화면을 스크롤하는 것만으로 생성·축적되는 썸네일 캐시를 C&C 통신 경로로 활용하여, 소용량 C2 신호(최대 28B)를 은닉하고 자동 복원할 수 있음을 입증했다. 구체적으로 듀얼 캐리어·QIM, 파일럿 임계 추정, 8×8 경계 위상 탐색, Hamming·CRC32를 결합한 알고리즘을 통해 Chrome, Firefox 브라우저에서의 실험을 성공시켰으며, 캐시 계층이 보안 위협 표면이 될 수 있음을 실증적으로 보였다.

향후 연구 방향성은 다음과 같다. 첫째, 기존 파라미터를 정밀 조정하고 다중 썸네일 체이닝·반복 노출을 적용해 용량 고도화를 추진할 것이다. 둘째, 텍스트 중심에서 이미지·비디오·오디오 등 은닉 데이터 모달리티를 다양화할 계획이다. 셋째, AVIF·WebP 등의 포맷을 대상으로 알고리즘을 일반화해 범용적인 은닉 기법을 확립할 예정이다.

참고문헌

[1] P. Fan et al., “Robust video steganography for social media sharing based on PCA,” EURASIP J. Inf. Security, 2022.

[2] Y. Wang, “Hiding Data within Thumbnail Videos: An Adaptive Downsampling-Resilient Video Steganography Method,” IEEE Access, 2024.

[3] J. Jeon and Y. Cho, “Construction and Performance Analysis of Image Steganography-Based Botnet in KakaoTalk Openchat,” Computers, 2019.

[4] J. Jeon and Y. Cho, “텔레그램 메신저 기반의 오디오 스테가 노그래피 봇넷 구축,” Journal of Internet Computing and Services (JICS), 2022.

[5] D. Yun and Y. Cho, “로블록스 메타버스 환경에서의 스테가 노그래피 기반 은닉통신 기법,” Journal of the Korea Institute of Information Security & Cryptology (JKIISC), 2023.

[6] M. El-Tayeb et al., “Live-Streamed Video Reconstruction for Web Browser Forensics,” Ingénierie des Systèmes d’Information(ISI), 2022.