대한민국 국가 도메인(.kr) 사칭 피싱 도메인 분석*

송예지, 박제만
경희대학교 소프트웨어융합학과
skaehdlf0318@khu.ac.kr, jeman@khu.ac.kr

Analysis of Phishing Domains Impersonating South Korea's ccTLD(.kr)

Yaeji Song, Jeman Park
The Department of Software Convergence, KyungHee University

요 약

본 연구는 대한민국 국가 도메인(.kr)의 신뢰도를 악용하는 사칭형 피싱 도메인을 식별하고 체계적으로 분류하였다. 먼저 Subfinder를 활용해 10개의 사칭 의심 도메인 후보를 선정하였고, httpx 기반으로 약 2,000여 개의 도메인 Dataset을 구축하였다. 이후 각 도메인에 대해 .kr, .com, 공식 2단계 도메인과의 유사 도메인 존재 여부를 검증하고, 리다이렉션 체인, 상태 코드, 서버 메타데이터를 분석하여 피싱 행위 가능성을 평가하였다. 이를 바탕으로 의심 도메인을 사칭형·정상형·오류형·악용형으로 분류하였다. 그 결과, 사칭형이 전체의 54.5%로 가장 높은 비율을 차지하였고, 악용형(10.2%) 분석으로 국가 도메인의 신뢰도 악용 실태가 확인되었다. 단순 문자열 비교 기반의 탐지로는 ccTLD 악용형 피싱을 효과적으로 식별하기 어렵다는 점을 실증적으로 제시하며, 시각적 유사도, 시계열 분석과 같은 위협 인텔리전스의 필요성을 강조하고 있다.

1. 서 론

인터넷의 핵심 주소 체계인 DNS는 온라인 서비스의 신뢰성과 접근성을 보장하는 기반 인프라로 기능한다. 특히 대한민국의 ccTLD(TLD List)인 .kr[1]은 공공기관, 금융기관, 언론사 등 주요 국가 기반 서비스가 사용하는 공식 도메인으로서 국민에게 높은 신뢰를 제공한다. 그러나 최근 몇 년간 이러한 신뢰 체계를 악용하여 .kr 도메인을 사칭하는 피싱 도메인이 급격히 증가하고 있다. 공격자는 n-e.kr, p-e.kr처럼 합법적인 .kr 도메인처럼 보이지만 실제로는 공격자가 등록한 일반 2단계 도메인을 활용함으로써 도메인에 대한 사용자의 신뢰를 교묘히 악용하고 있다.

이러한 .kr 사칭형 피싱 도메인은 단순한 스팸 수준을 넘어, 정부기관 사칭형, 금융·포털 브랜드 사칭형, pseudo-ccTLD 형으로 분화되어 운영되고 있다.

이들은 합법적인 서비스와 동일하거나 시각적으로 유사한 웹페이지로 사용자의 인증 정보를 탈취하거나, 악성코드를 유포시킨다. 특히 .kr은 국내 이용자들에게 "국가가 관리하는 안전한 도메인"이라는 인식이 강하기 때문에, 사용자는 도메인에 '.kr'이 포함되어 있다는 이유만으로 웹사이트를 신뢰하는 경향이 있다. 이로 인해 피싱 성공률이 gTLD(.com, .net)보다 훨씬 높게 보고되고 있으며, 공격자는 이를 노린 정교한 사회공학적 공격을 지속적으로 전개하고 있다.

기존의 피싱 탐지 연구는 주로 웹페이지 시각적 요소같은 콘텐츠 기반 분석 또는 일반 도메인 속성 기반(길이, TTL, WHOIS 정보 등)에 집중되어 있다. 하지만 .kr 사칭형 피싱의 경우, 공격자가 실제 대한민국 ccTLD를 사용하는 것처럼 보이기 때문에 기존 기법의 탐지 정확도가 크게 저하된다. 또한 공격자들은 n-e.kr, p-e.kr과 같은 짧은 도메인을 이용하여, "국가기관용 2단계 도메인(re.kr, or.kr)"과 혼동되도록 설계함으로써 기계학습 기반 필터나 블랙리스트 기반 필터링을 회피하는 사례도 지속적으로 보고되고 있다.[2]

이에 따라 본 연구는 대한민국 국가 도메인(.kr)을 사칭하는 피싱 URL의 구조적·행태적 특징을 체계적으로 분석하고, 이를 기반으로 한 도메인 신뢰도 검증 및 위장 도메인 판별 체계를 제안한다. 연구는 KISA[3], AhnLab[2], VirusTotal, PhishTank 등의 공개 인텔리전스 데이터를 수집하여 .kr 도메인군 중 비정상 도메인을 선별하고 httpx[4], subfinder[5], urlscan.io 도구를 활용해서 실제 피싱 페이지의 특성을 통계적으로 분석한다. 또한 .kr 사칭 피싱 도메인이 실제로 어떤 국가기관·브랜드를 모방하고 있는지, 어떤 네트워크 인프라와 결합되어 유포되는지를 분류함으로써, 향후 국가 단위의 피싱 대응 체계의 토대를 마련하고자 한다.[3]

결국 본 연구는 ".kr 도메인에 대한 신뢰"가 공격 벡터로 전환되는 현상을 실증적으로 규명하고, 이를 기술적·정책적으로 차단할 수 있는 방안을 제시하는 것을 목표로 한다. 이를 통해 향후 .kr 도메인 기반의 사이버 위협에 대한 인식 제고와 함께, 국가 차원의 도메인 보안 정책 수립에 기여하고자 한다.

* "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2025년도 SW중심대학사업의 결과로 수행되었음"(2023-0-00042)

2. 본 론

2.1 사칭 의심 도메인 후보 선정

먼저 TLD 리스트[1], AhnLab[2], KISA(한국인터넷진흥원)[3], 지자체 보안 권고문을 참고하여 대한민국 .kr 사칭 도메인 후보를 선정했다. 실제 대한민국 국가기관용 2단계 도메인으로 사용되는 도메인과 혼동 가능한 도메인, .kr과 유사한 도메인, 실제 피싱 사례에서 .kr을 포함하고 있던 도메인을 후보로 선정했다. 이후에는 r-e.kr과 같은 형태를 서브도메인 혹은 2단계 도메인이라고 지칭한다.

표 1: 사칭 도메인 후보 유형

후보 유형	사칭도메인 후보	합법 도메인
국가기관용 2단계 도메인	r-e.kr	re.kr
	o-r.kr	or.kr
	m-s.kr	ms.kr
	g-o.kr	go.kr
	n-e.kr	ne.kr
	p-e.kr	pe.kr
	c-o.kr	co.kr
.kr 유사 도메인	kr.com	.kr
실제 피싱 도메인	kor.kr	.kr
	kro.kr	.kr

2.2 사칭 의심 도메인 Dataset 구축

사칭 의심 도메인을 포함한 서브도메인과 해당 도메인의 http 정보를 오픈소스 도구 ProjectDiscovery httpx[4], subfinder[5]로 수집했다. subfinder는 한 번이라도 관측되었던 서브도메인 리스트를 보여주는 도구이고, httpx는 특정 도메인의 http 정보를 보여주어 도메인 통계 분석 및 피싱 패턴 탐색에 적합하다.

표 2: 사칭 의심 도메인 정보 수집 명령어

```
subfinder -d n-e.kr -silent \
  | sort -u \
  | httpx -silent -json -title -status-code -favicon
-follow-redirects -timeout 10 \
  > n-e_kr.json
```

1,974개의 사칭 의심 서브도메인을 각 사칭 도메인 후보 유형별로 수집하였으며, 수집 결과를 존(zone) 단위 JSON 파일로 구성하여 체계적으로 관리하였다. JSON 파일에는 각 서브도메인에 대한 주요 필드가 포함되어 있으며, port 번호, 최종 랜딩 페이지 주소, 현재 도메인 상태, 수집 시각, 리다이렉트 홉 개수, 웹사이트의 favicon 해시 값 등의 정보가 포함된다.

2.3 유사 도메인 존재 여부 분석

수집된 dataset에서 url을 기반으로 사칭 의심 도메인과 유사하지만 TLD만 다른 도메인의 존재 여부를 분석했다. 예를 들어 *https://homefarm[.]p-e[.]kr* 이 사칭 도메인 후보일 때 *http://homefarm[.]pe[.]kr*이나 *http://homefarm[.]kr*, *http://homefarm[.]com* 중 하나라도 접속이 가능하거나 이전에 접속 가능했던 히스토리가 존재할 때 유사 도메인이 존재한다고 판별했다.

표 3: TLD만 다른 도메인 존재 여부

	존재함	존재하지 않음
개수	1,156(58%)	818(42%)

분석 결과, 1,974개 중 1,054개가 TLD가 다른 도메인이 존재했다. 구체적인 예시로, .kr.com 도메인을 사용하는 http://cannaboxcontainer[.]kr[.]com과 유사한 http://cannaboxcontainer[.]com 웹페이지가 발견되었다. 악성 사이트를 알려주는 VirusTotal과 수동 분석으로 이중 확인한 결과 .com으로 끝나는 웹페이지는 정상적인 사이트임이 확인되었다. 이처럼 합법 사이트와 혼동할 수 있는 유사 url이 절반을 넘는 58%의 수치를 기록했다.

2.4 리다이렉션 분석

유사 도메인이 존재하지 않고 TLD만 다른 880개의 도메인을 심층 분석했다. 기존 사칭 의심 도메인의 url과 최종 접속 페이지인 final_url이 다른 도메인은 ccTLD(.kr)의 신뢰를 악용할 가능성이 높은 도메인일 것이라는 가설을 세우고 이에 맞는 118개의 url을 추출했다. 그중이 url들의 리다이렉션 홉 수 분포를 보니 다음과 같았다.

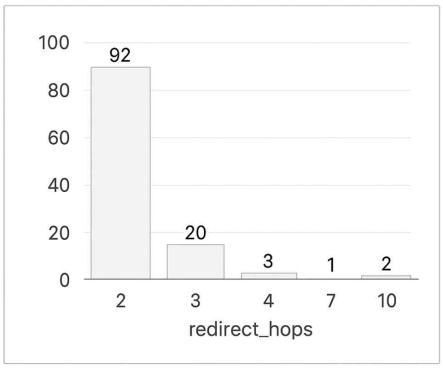

그림 1 리다이렉션 홉 수 분포

리다이렉션 홉 수는 최소 2부터 10까지로 확인되었으며, 약 78%는 2의 값을 가진다. 880개의 약 13%인 118개의 도메인은 신뢰도 높은 .kr 도메인을 사용하는 것처럼 url을 생성하고 실제로는 다른 url로 리다이렉션되는 현상이 발생하고 있음을 알 수 있다. 리다이렉션되는 도메인

대부분 final_url을 확인해보면 사칭 의심 도메인이 아닌 전혀 다른 .xyz, .net, .com과 같은 TLD를 사용하는 웹페이지로 이동하는 것을 볼 수 있었으며, 과반수 이상이 /login, /auth 하위페이지가 포함되어 있는 경로로 이동하는 특징을 보였다. 리다이렉션 홉 수가 증가할수록 payload 포함된 url이 등장하거나 /download, .app이 포함된 경로로 리다이렉션되는 경우가 자주 관찰되었다.

2.5 의심 도메인 유형 분류(사칭형/정상형/오류형/악용형)

데이터셋의 서브도메인을 Dataset 분석(2.3)의 의거하여, 사칭형·정상형·오류형·악용형으로 분류한다. ***사칭형***은 유사 도메인이 존재하며, 이 도메인이 VirusTotal에서 정상 사이트로 분류되고 실제 피싱 사례에 포함되어있지 않을 때 '사칭형'이라고 명시한다. ***정상형***은 유사 도메인이 존재하지 않으며 리다이렉션 홉 수가 2 이하이고, 악성 기능이 없을 때 '정상형'이라고 판별한다. ***오류형***은 도메인 만료 혹은 접속 불가 상태 등의 이유로 악성 여부를 판단할 수 없는 도메인을 '오류형'이라고 판단한다. 마지막으로, ***악용형***은 정상 도메인을 사칭하지 않으며 리다이렉션 홉 수가 3 이상이거나 VirusTotal 에서 악성 사이트로 분류되거나 악성 기능이 발견되면 이를 '악용형'이라고 칭한다. 분류 유형에 대한 명확한 정의는 다음과 같다.

- ***사칭형***: 정상 도메인과 혼동하도록 조작한 도메인
- ***정상형***: 악용 가능성이 없는 정상 도메인
- ***오류형***: 접속이 불가능하거나 에러 등의 이유로 악성 여부를 판단할 수 없는 도메인
- ***악용형***: 정상 도메인을 사칭하지 않았지만, ccTLD(.kr)의 신뢰도를 이용해 악성 행위를 시도하는 도메인

표 4: 의심 도메인 유형별 분포

유형	사칭형	정상형	오류형	악용형
개수	1,075	303	395	201

수집·분류한 1,974개의 의심 도메인을 유형별로 집계한 결과, 사칭형이 1,075건(약 54.5%)으로 가장 큰 비중을 차지하였다. 정상형은 303건(약 15.3%), 오류형은 395건(약 20.0%), 악용형은 201건(약 10.2%)으로 나타났다. 이러한 분포는 .kr 계열의 의심 도메인 생태계가 크게 실존하는 정상 도메인의 형태적 유사성을 이용해 사회공학적 혼란을 노리는 '사칭형'과 기술적 문제·임시 호스팅·악의적 전파 등으로 판별되는 '악용형' 두가지 축으로 구성되어 있음을 시사한다.

악용형(약 10.2%)은 광고 리다이렉션, 피싱 전 단계로 사용되는 랜딩 페이지 등의 다양한 유형이 혼재되어 있다. 특히 리다이렉션 홉이 길어지는 경우는 탐지 회피나 트래픽 분산을 통한 분석 회피를 노리는 전형적인 피싱 전략으로 해석된다. 오류형의 비율(약 20%)이 비교적 높은 이유는 다각적이다. 도메인 만료, DNS 불가·해상 실패, HTTP 5xx 계열의 서버 오류 등이 주된 원인으로 확인되었다. 웹페이지 확인이 불가해 악성 여부를 판단할 수 없지만, 도메인 등록정보를 살펴봤을 때 사칭·피싱 목적으로 준비 중인 예비 인프라거나 이미 악용된 후 폐기되었다고 판단된다면 잠재적 사칭형, 악용형으로도 분류 가능하다.

3. 결론 및 향후 연구 방향

본 연구에서는 대한민국 국가 도메인(.kr) 계열을 대상으로 실존하는 정상 도메인을 사칭하거나 위장하는 피싱형 도메인을 대규모로 수집·분석하였다. 10종의 의심 도메인 후보군에 대해 서브도메인 기반의 실측 데이터를 수집하였으며, 각 URL의 HTTP 응답 특성, 리다이렉션 구조, 콘텐츠 메타데이터, 그리고 VirusTotal·피싱 데이터베이스와의 교차검증을 통해 유형을 분류하였다.

분석 결과, 전체 1,974개 도메인 중 사칭형이 1,075개(54.5%)로 가장 높은 비중을 차지하였다. 이는 ccTLD(.kr)의 신뢰도를 악용하여 정상 기관이나 서비스의 서브도메인을 위조하는 사례가 광범위하게 발생하고 있음을 보여준다. 또한 오류형(20.0%)과 악용형(10.2%)의 비율로 보아 등록 후 방치되거나 악성 리디렉션 인프라로 활용되는 경우가 많음을 확인하였다.

이러한 결과는 단순 문자열 기반의 도메인 탐지로는 사칭형 위협을 완전하게 식별하기 어렵다는 점을 시사한다. 향후 연구에서는 정적·동적 행위을 결합해서 분석할 것이다. 또한, 오류형 도메인의 재등록 및 재활성화 여부를 시계열 기반 데이터를 수집해 주기적으로 모니터링하여 잠재적 공격 인프라를 조기에 탐지하고자 한다. 이러한 탐지 파이프라인을 오픈소스 인텔리전스 도구로 공개해 국가 차원의 도메인 보안 정책 수립에 도움이 되고자 한다.

참고문헌

[1] TLD-LIST, https://tld-list.com/tld/kr

[2] Ahnlab, "국내 유명 금융 앱 사칭한 피싱공격", Dec, 12. 2022. https://asec.ahnlab.com/ko/44225/

[3] KISA, "인터넷 주소 관리" https://www.kisa.or.kr/1020703

[4] httpx, https://github.com/projectdiscovery/httpx

[5] subfinder, https://github.com/projectdiscovery/subfinder