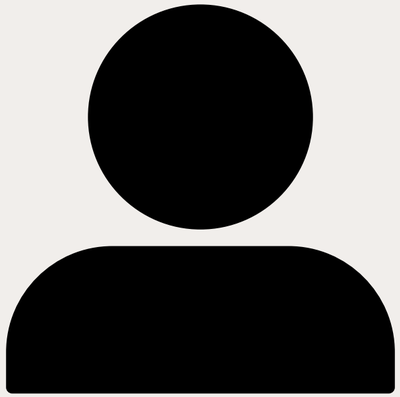
TI-TOTAL

이하영, 김민서, 박승우

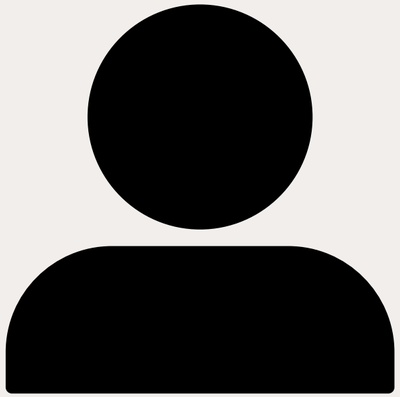
목차

- 팀원 소개
- 주제
- 사용된 도구
- 동작 플로우
- 머신러닝 설계
- 동작 과정
- 기대 효과와 개선 방안

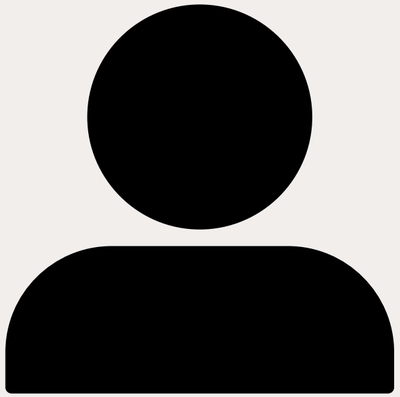
팀원



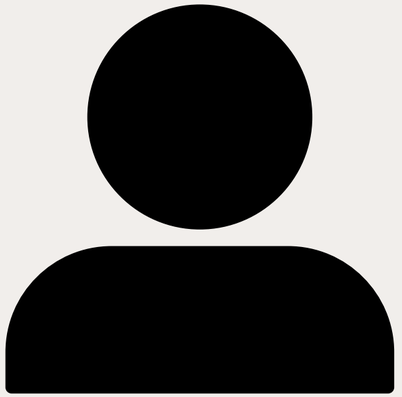
92015386
이하영
ml, backend



92014992
김민서
api 수집, backend



92015128
박승우
DB, backend



92015075
김평안
frontend

주제

TI-Total

TI 통합 검색 엔진

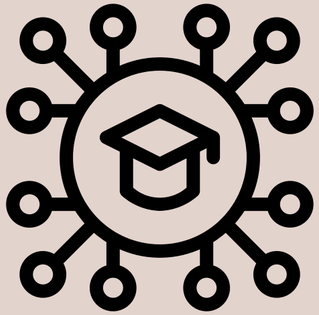
TI(Threat Intelligence) 정보란, 사이버 보안 환경에서
수집된 데이터를 분석하여 구체적인 사이버 위협을
식별하고 대응하는 데 필요한 정보

도메인 정보, ip정보, 네트워크 정보, 인증서

사용한 도구

Back, Front

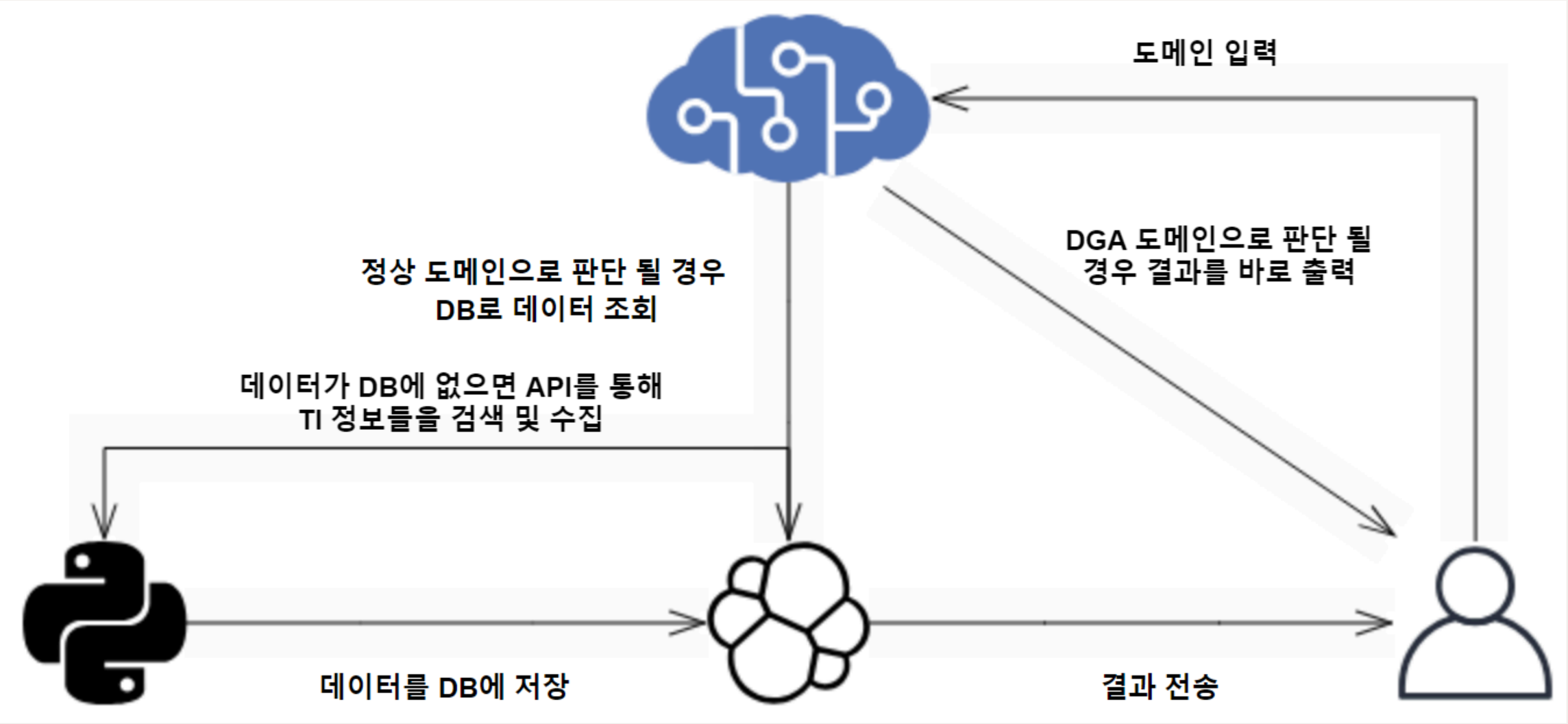
Flask



DB

redis

동작 플로우


도메인 입력
정상 도메인으로 판단 될 경우
DB로 데이터 조회
DGA 도메인으로 판단 될
경우 결과를 바로 출력
데이터가 DB에 없으면 API를 통해
TI 정보들을 검색 및 수집
데이터를 DB에 저장
결과 전송

DGA란?

DGA.Changer, “샌드박스 쯤은 우습게 우회해요”

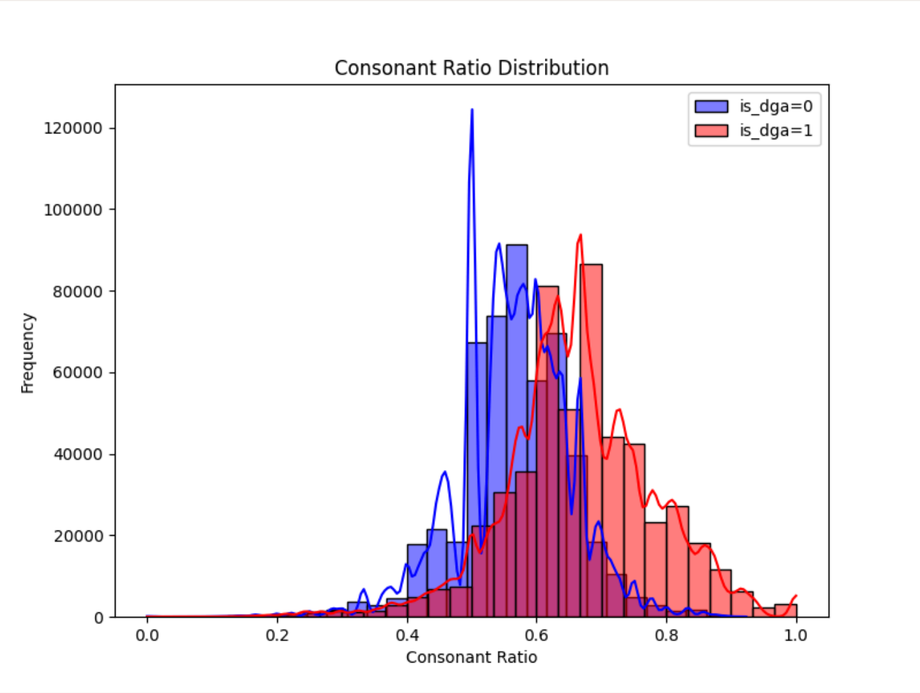
입력 : 2015-08-11 19:00

발견이 어려움은 물론 발견해도 가짜일 확률 높아
누구든 마음만 먹으면 얼마든지 구매하여 사용 가능

[보안뉴스 주소형] 접속 도메인 명을 생성하는 DGA(Domain Generation Algorithm)라는 기술이 공격으로 악용되고 있다. 지속적으로 도메인 명을 변경해줌으로서 보안전문가들을 따돌리는 것이다. 이 같은 공격은 ‘DGA.Changer’라고 불리고 있으며 이스라엘 사이버보안 회사인 시큘러트(Seculert)에 의해 드러났다.

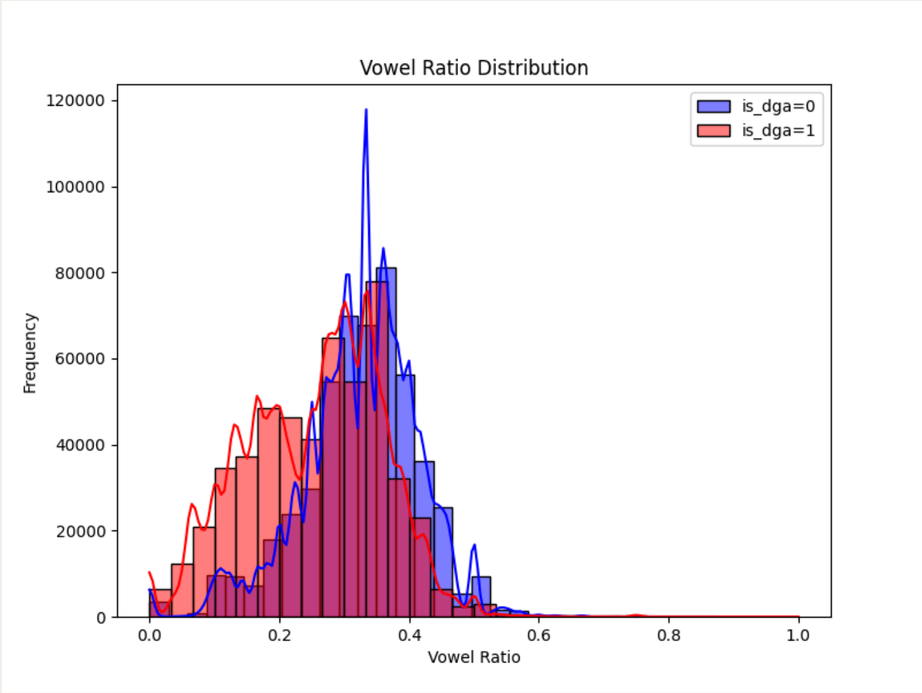
DGA (Domain generation algorithm)은 다수의 도메인 주소를 생성하기 위해 설계된 알고리즘.
특정 seed 값에 따라 도메인을 생성한다.

피쳐선정



Consonant Ratio

자음의 비율

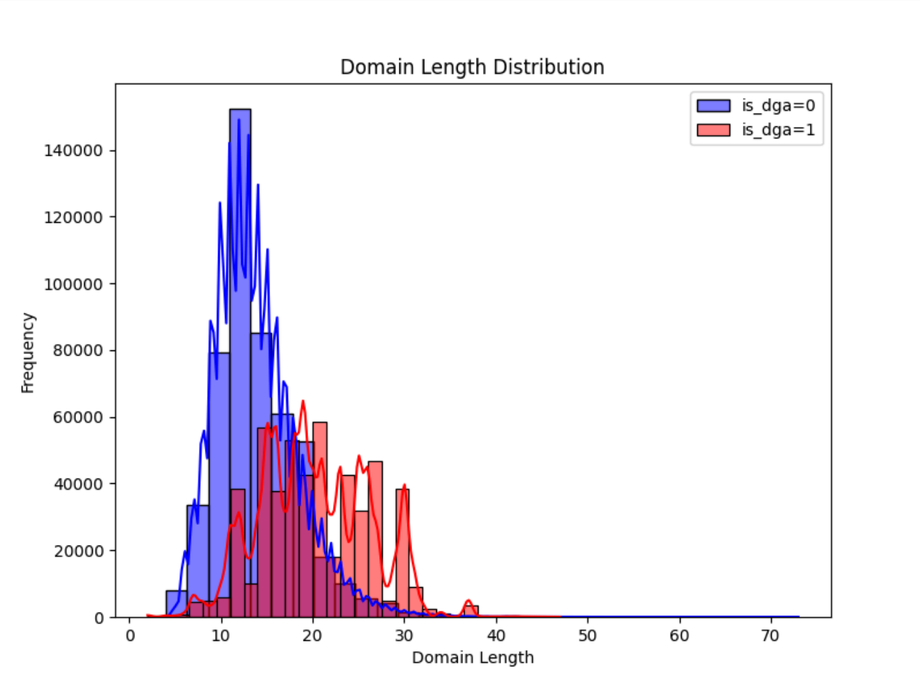


Vowel Ratio

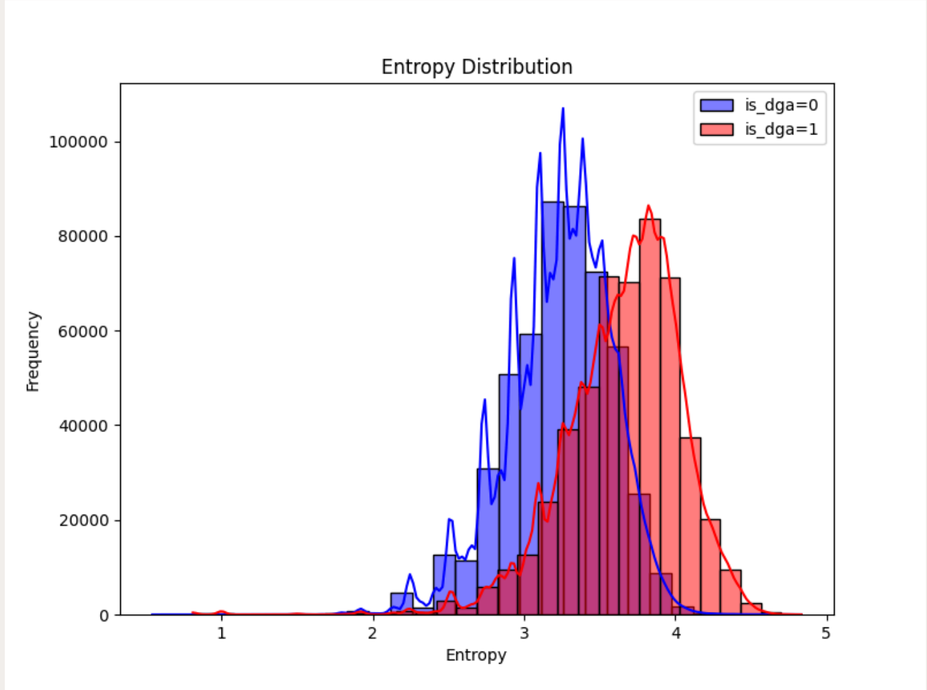
모음의 비율

정상적 도메인은 언어의 자음과 모음의 배치를 따르지만, DGA 도메인은 무작위로 생성되기 때문에 비율이 비정상적이다.

피쳐선정



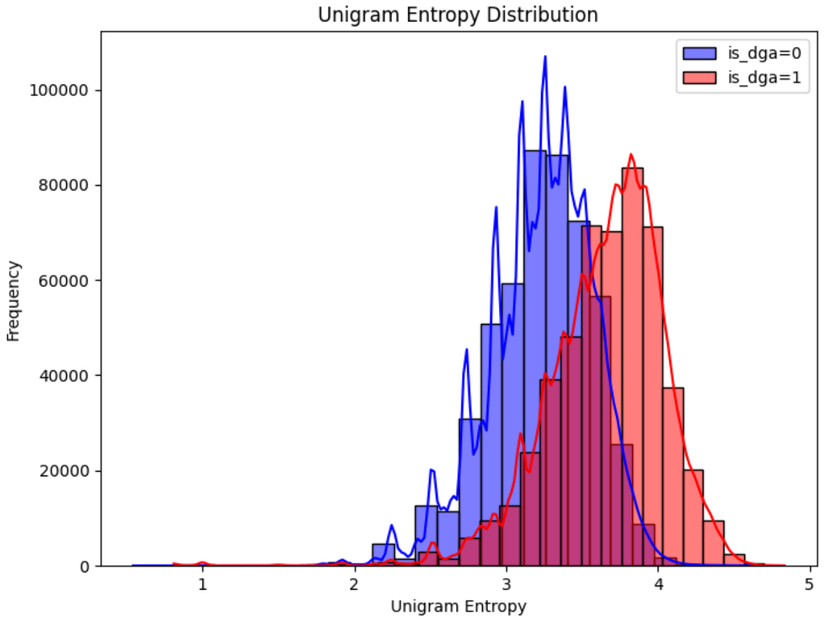
Domain Length
도메인의 길이

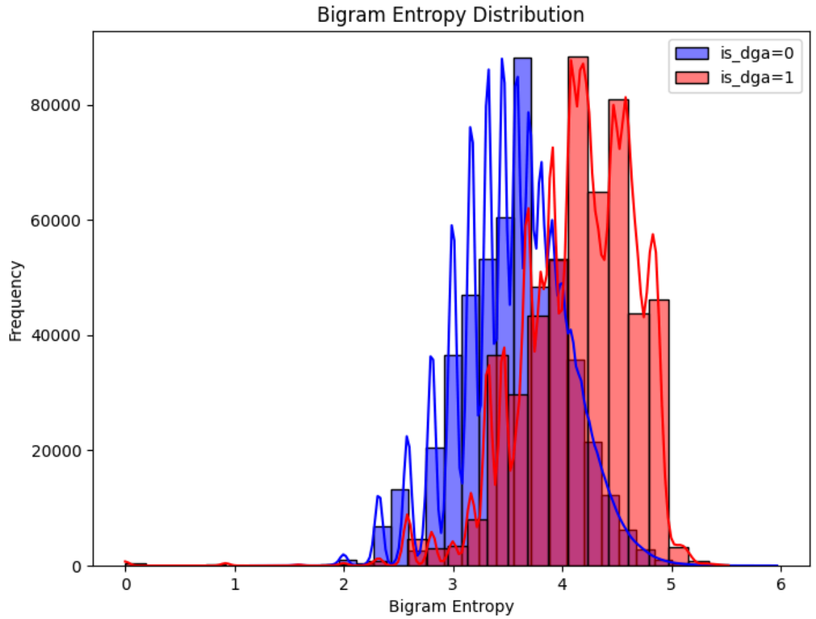


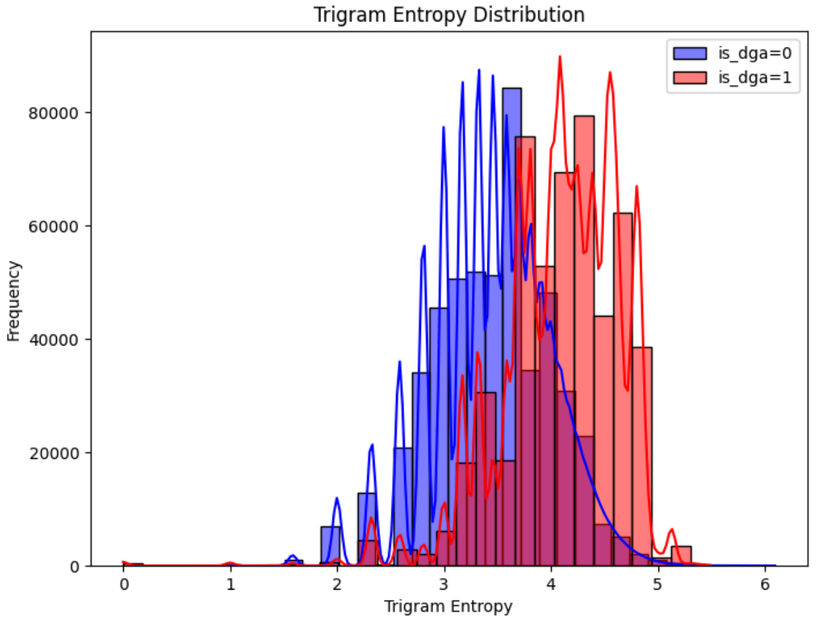
average Entropy
평균 엔트로피

DGA로 생성된 도메인은
고정된 길이이거나
비정상적으로 길 수 있고,
높은 엔트로피 값은
도메인이 무작위적으로
생성되었음을 나태내고,
이러한 점은 DGA의 특징이다

피쳐선정







Unigram Entropy
개별 단어의 빈도 기반 엔트로피

Bigram Entropy
2단어의 빈도 기반 엔트로피

Trigram Entropy
3단어의 빈도 기반 엔트로피

DGA 도메인은 무작위성이 높아, 위와 같이 차이가 크게 나타나기 때문에
DGA 도메인과 정상 도메인을 구분하는 피쳐로 적절하다.

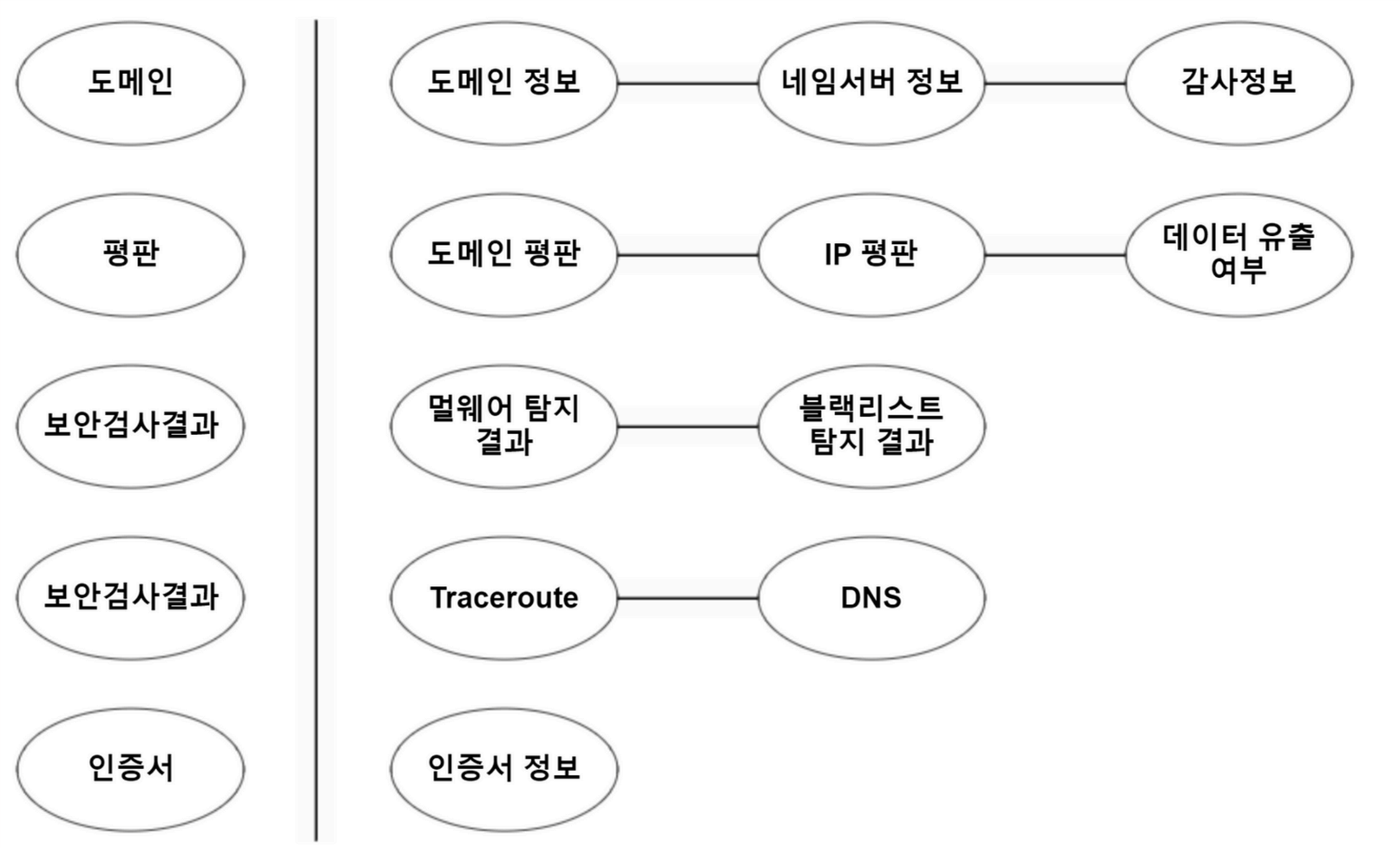
탐지 결과

```
Accuracy: 0.9094496257429028
Precision: 0.9275016441458239
Recall: 0.888640600741833
F1 Score: 0.9076553563608859
```

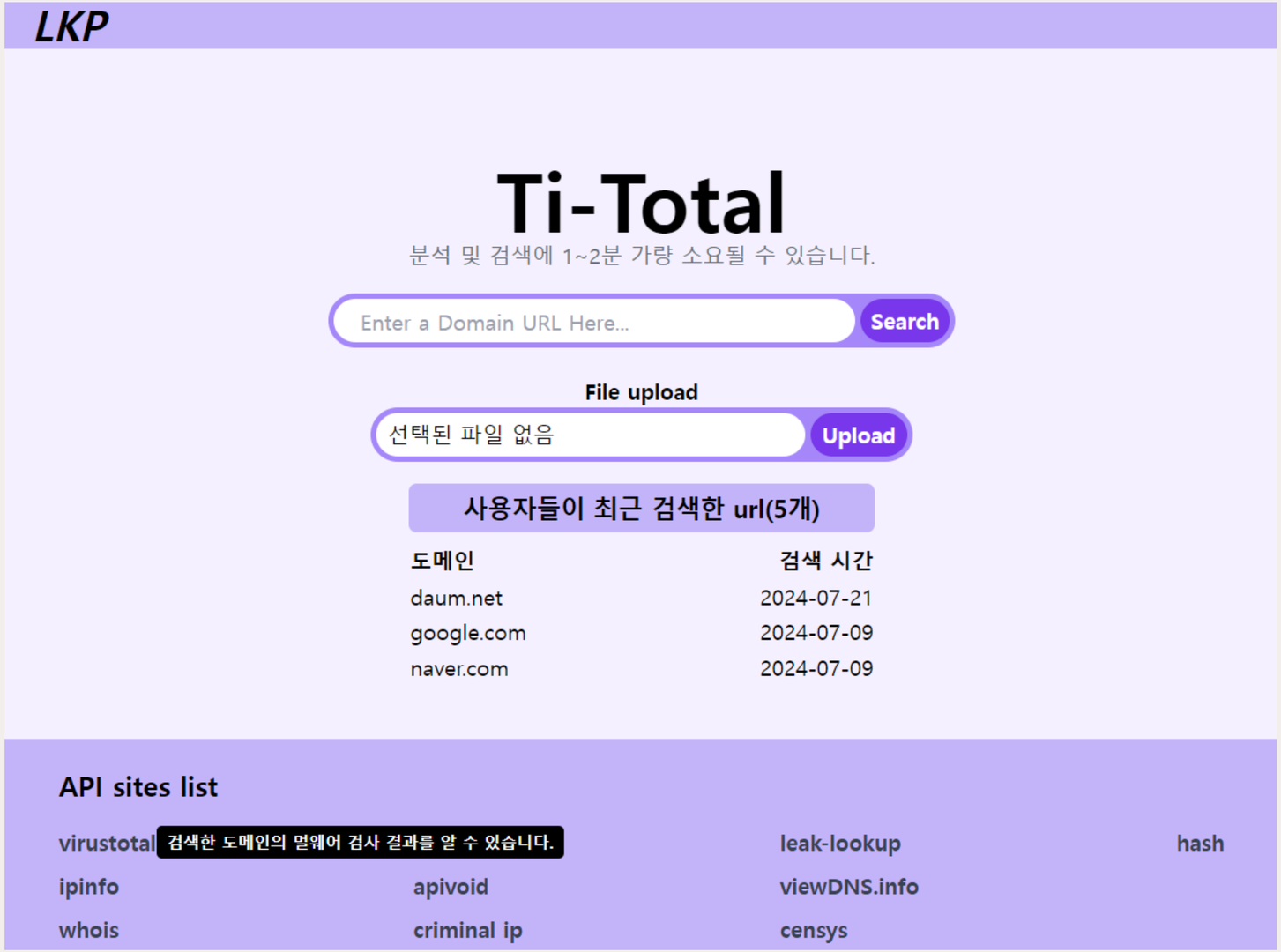
실제 도메인 데이터를 약 3만개 가량
입력 하였을 때, 정확도 91%로 탐지를 하였다.
또한, F1-Score는 약 90.8이 나온 것을 확인 할 수 있다.

Domain	Result	DGA	not DGA
mortiscon	Not DGA	1.928562	98.07144
cvyh1po6	DGA	99.97965	0.02035
plasticbag	Not DGA	7.52724	92.47276
mzltrack.c	DGA	67.46748	32.53252
miss-slim.r	Not DGA	0.322912	99.67709
txumyqrut	DGA	96.4333	3.566701
myhosting	DGA	77.1249	22.8751
ixekrihagi	Not DGA	3.995031	96.00497
rjyuosmhf	DGA	66.20192	33.79808
djqrmautt	DGA	80.10161	19.89839
tqbmiuyw	DGA	82.0957	17.9043
brotherne	Not DGA	12.07622	87.92378
download	Not DGA	0.689877	99.31012
lrushteeoi	Not DGA	43.72518	56.27482

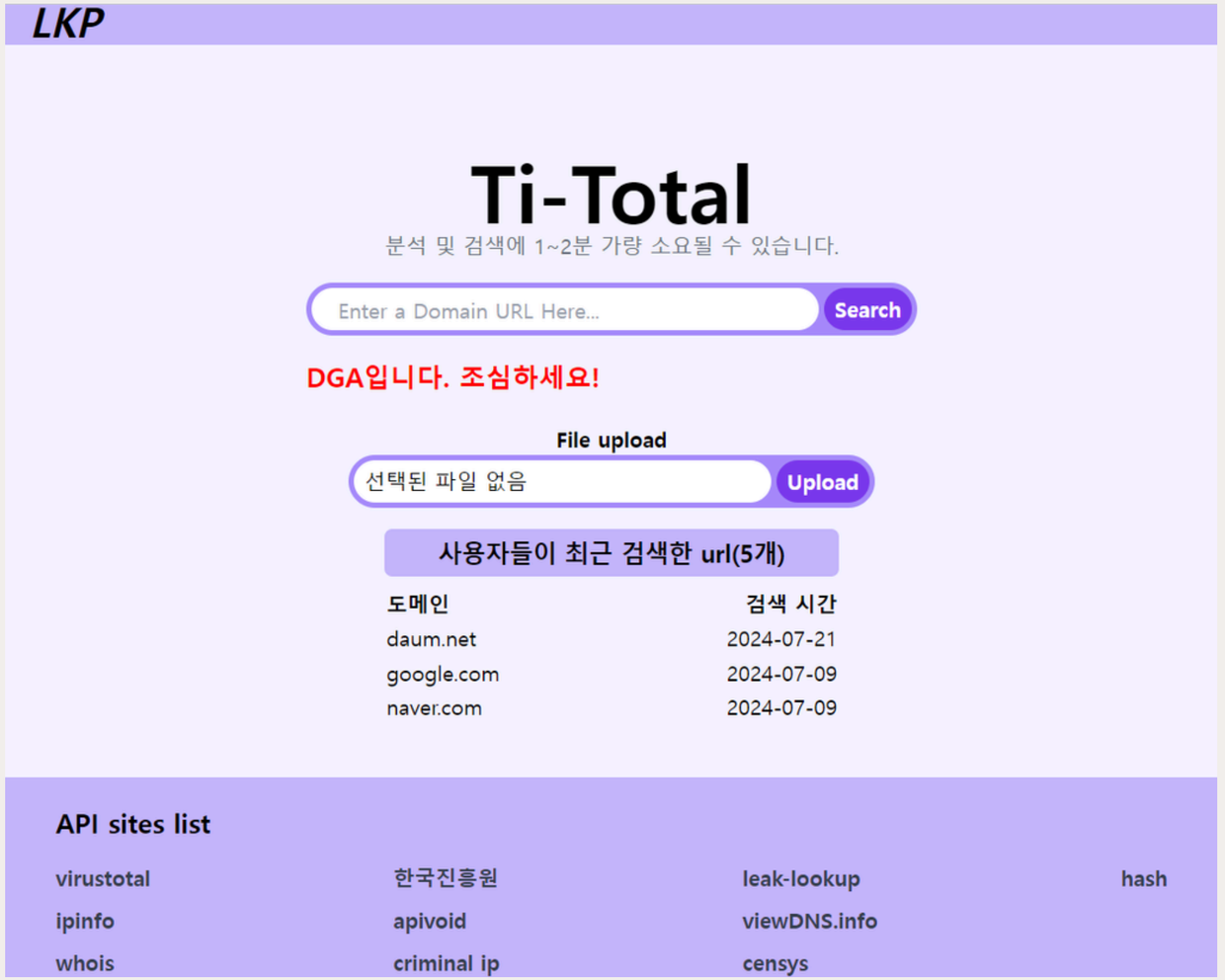
분류 카테고리


도메인
도메인 정보
네임서버 정보
감사정보
평판
도메인 평판
IP 평판
데이터 유출 여부
보안검사결과
멀웨어 탐지 결과
블랙리스트 탐지 결과
보안검사결과
Traceroute
DNS
인증서
인증서 정보

동작 과정(입력창)


LKP
Ti-Total
분석 및 검색에 1~2분 가량 소요될 수 있습니다.
Enter a Domain URL Here...
Search
File upload
선택된 파일 없음
Upload
사용자들이 최근 검색한 url(5개)
도메인
검색 시간
daum.net
2024-07-21
google.com
2024-07-09
naver.com
2024-07-09
API sites list
virustotal
검색한 도메인의 멀웨어 검사 결과를 알 수 있습니다.
leak-lookup
hash
ipinfo
apivoid
viewDNS.info
whois
criminal ip
censys

동작 과정(입력창)


LKP
Ti-Total
분석 및 검색에 1~2분 가량 소요될 수 있습니다.
Enter a Domain URL Here...
Search
DGA입니다. 조심하세요!
File upload
선택된 파일 없음
Upload
사용자들이 최근 검색한 url(5개)
도메인
검색 시간
daum.net
2024-07-21
google.com
2024-07-09
naver.com
2024-07-09
API sites list
virustotal
한국진흥원
leak-lookup
hash
ipinfo
apivoid
viewDNS.info
whois
criminal ip
censys

동작 과정(도메인 결과창)

도메인

- 도메인 정보
- 네임서버 정보
- 감사정보

도메인 및 IP평판

- 도메인 평판
- IP 평판
- 데이터 유출 여부

보안 검사 결과

- 멀웨어 탐지 결과
- 블랙리스트 탐지 결과

네트워크

- Traceroute
- DNS

인증서 정보

☰ *LKP*

입력한 도메인: naver.com

데이터 저장 날짜: 2024-07-09 14:20:48 (12일 전)

도메인 정보

- 도시: Seoul
- 국가: KR
- 회사: AS23576 NAVER Cloud Corp.
- 도메인이 사용하는 시간대 : Asia/Seoul
- 도메인이 네임이 처음 등록된 날짜: 1997-09-12
- 도메인 네임이 등록이 만료되는 날짜: 2032-09-11
- 마지막으로 도메인 네임 정보가 수정된 날짜: 2023-06-29T06:40:26Z
- 도메인의 추정 나이: 9797
- 도메인 등록 대행사 이름: gabia
- 도메인 등록 대행사 ID: 244

네임서버 정보

- 네임 서버 호스트 이름 목록: ['ns1.naver.com', 'ns2.naver.com']

동작 과정(hash 결과창)

LKP

File Analysis Result

Basic properties

분석 ID: MmI0MGM5OGVkMGY3YTFkM2IwOTFhM2U4MzUzMTMyZGM6MTcyMTU2ODA1NA==
파일 크기: 289792 Byte
SHA-256: badf4752413cb0cbdc03fb95820ca167f0cdc63b597ccdb5ef43111180e088b0
MD5: 2b40c98ed0f7a1d3b091a3e8353132dc
SHA-1: df79c86fdd11b9ccb89148458e509f879c72566c

Analysis Result

악성: 0
의심스러운: 0
검출되지 않음: 71
무해함: 0
시간 초과: 2
확인된 시간 초과: 0
검사 실패: 1
지원되지 않는 유형: 4

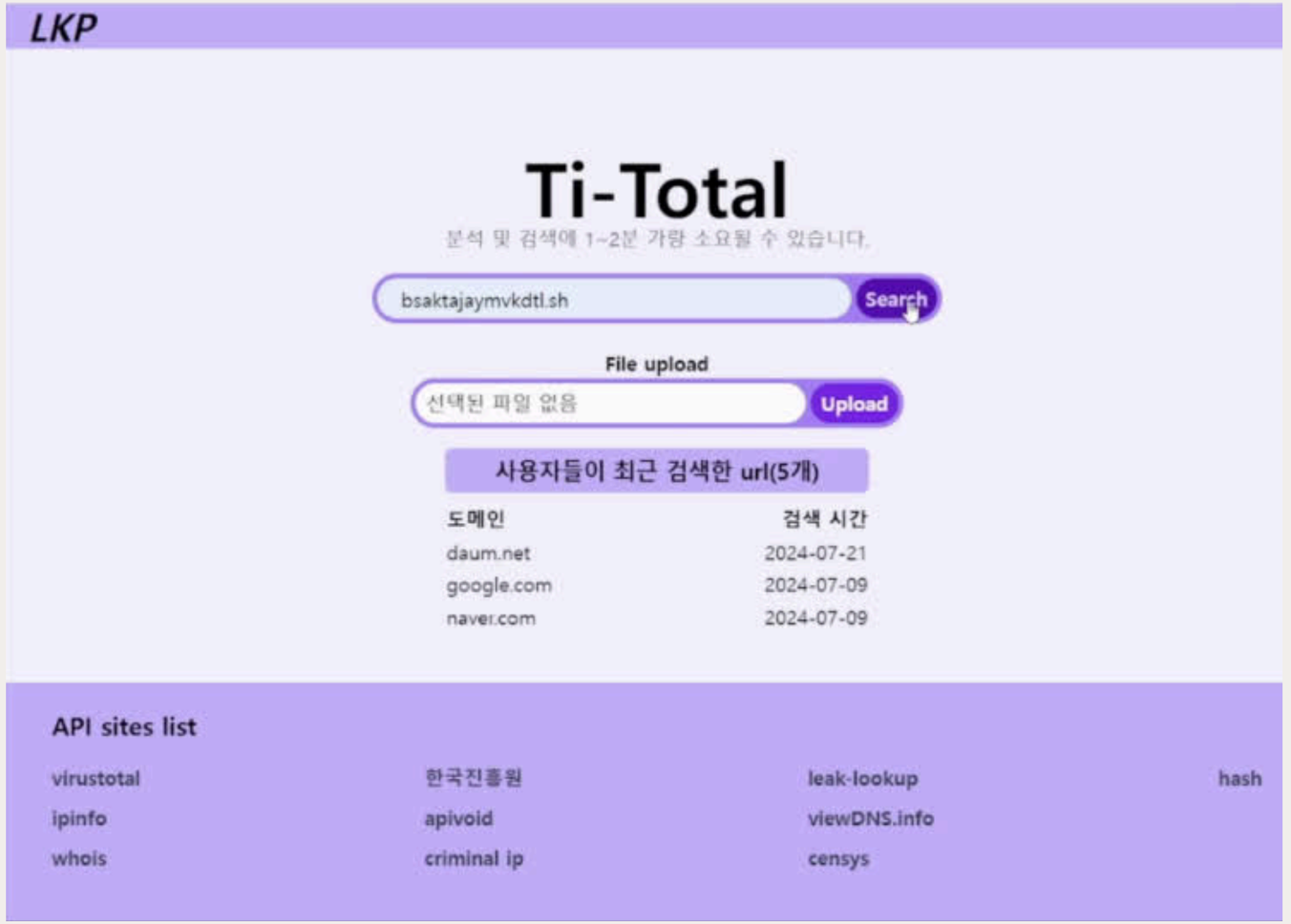
▼ 엔진별 결과

Bkav: undetected
Lionic: undetected
Elastic: undetected

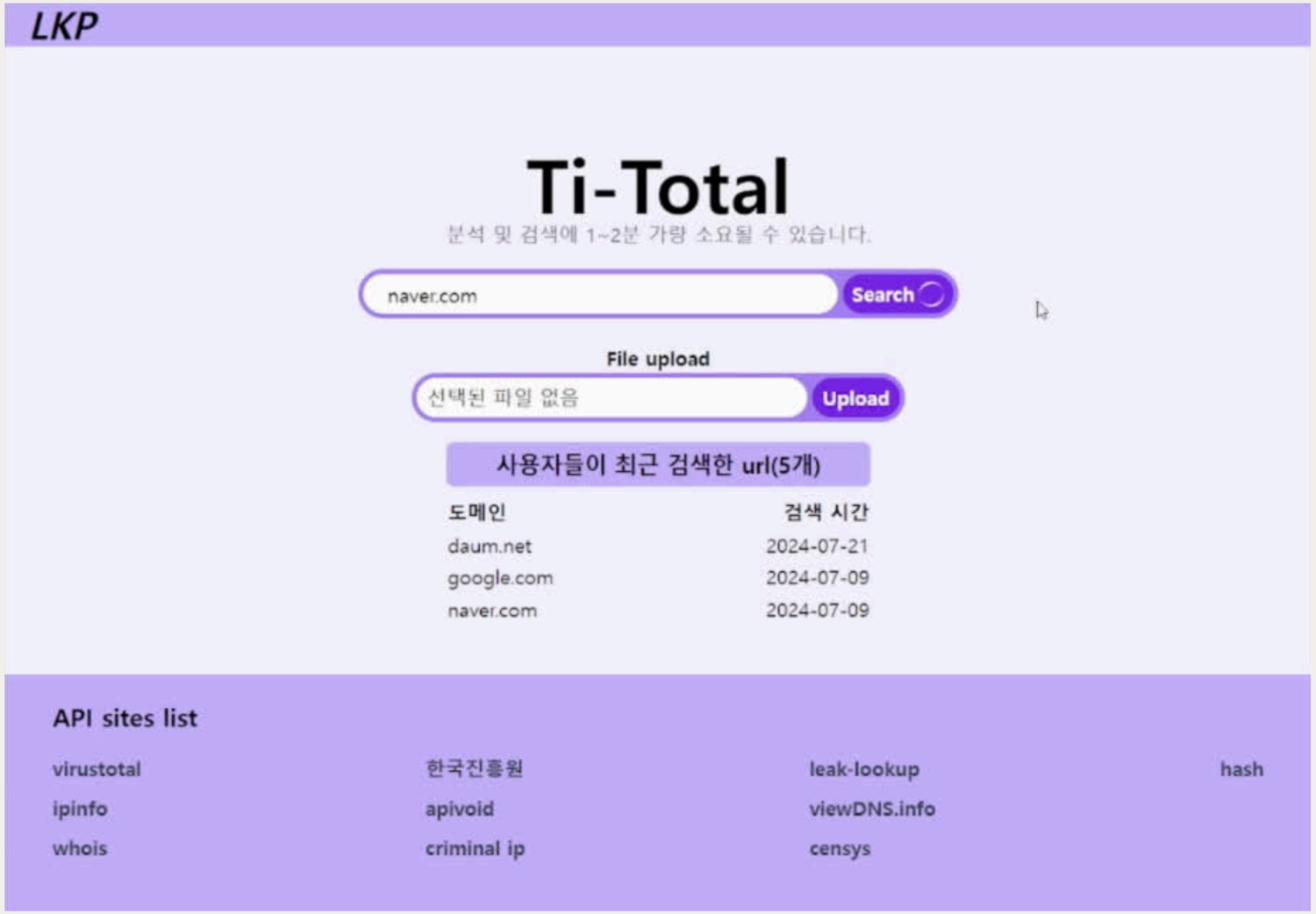
기대 효과

다양한 Ti 정보를 통합하여, 각종 도메인에 대해 알고 싶은 사람들에게
다양한 정보를 제공할 수 있다. 이로 인하여, 기존에 존재하던
검색 엔진에 일일히 들어 가는 소요를 줄여, 불편함을 최소화 할 수 있다.

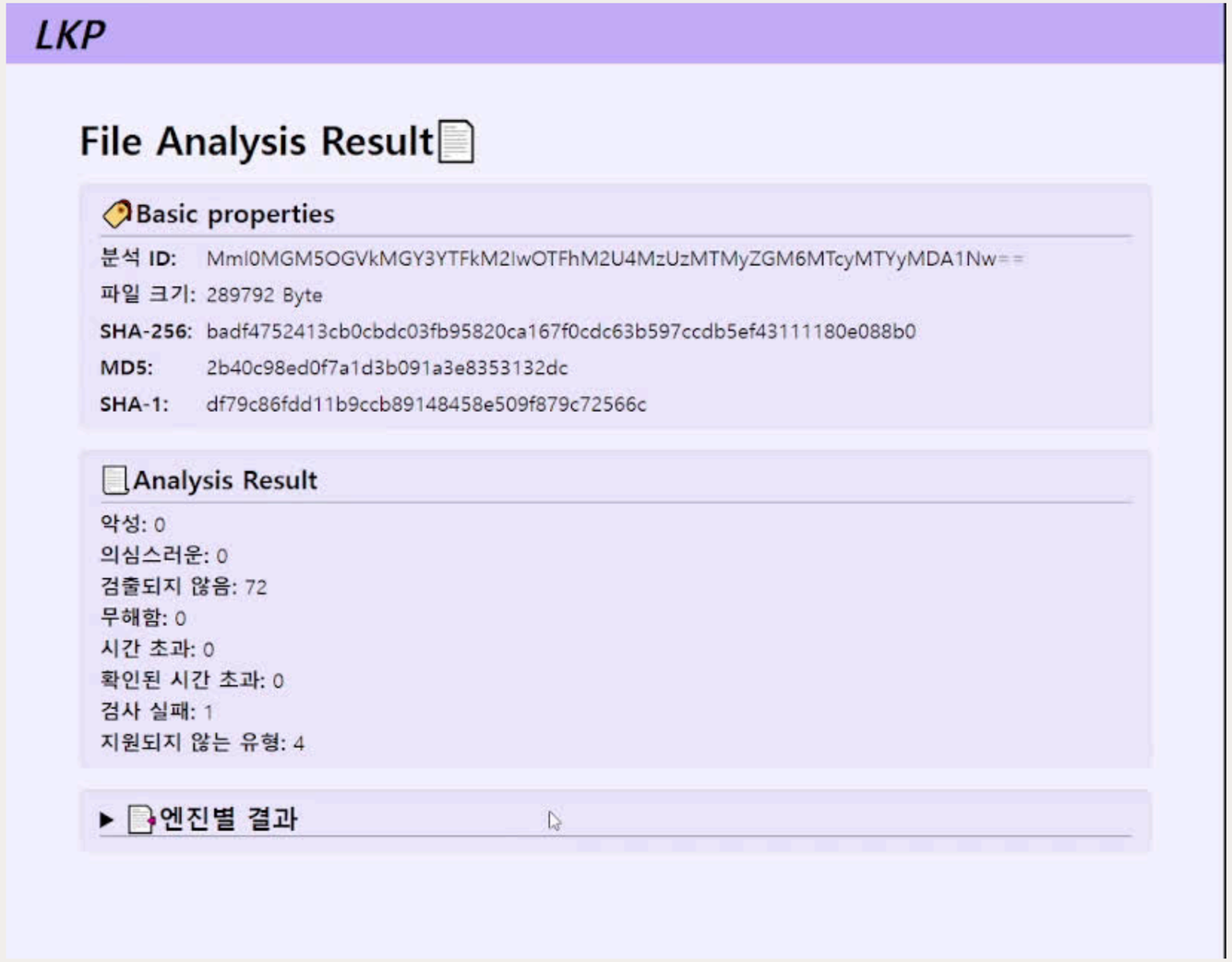
시연영상(DGA 탐지)


LKP
Ti-Total
분석 및 검색에 1~2분 가량 소요될 수 있습니다.
bsaktajaymvkdtl.sh
Search
File upload
선택된 파일 없음
Upload
사용자들이 최근 검색한 url(5개)
도메인
검색 시간
daum.net
2024-07-21
google.com
2024-07-09
naver.com
2024-07-09
API sites list
virustotal
한국진흥원
leak-lookup
hash
ipinfo
apivoid
viewDNS.info
whois
criminal ip
censys

시연영상(도메인 결과)


LKP
Ti-Total
분석 및 검색에 1~2분 가량 소요될 수 있습니다.
naver.com
Search
File upload
선택된 파일 없음
Upload
사용자들이 최근 검색한 url(5개)
도메인
검색 시간
daum.net
2024-07-21
google.com
2024-07-09
naver.com
2024-07-09
API sites list
virustotal
한국진흥원
leak-lookup
hash
ipinfo
apivoid
viewDNS.info
whois
criminal ip
censys

시연영상(hash 결과)


LKP
File Analysis Result
Basic properties
분석 ID: MmI0MGM5OGVkMGY3YTFkM2IwOTFhM2U4MzUzMTMyZGM6MTcyMTYyMDA1Nw==
파일 크기: 289792 Byte
SHA-256: badf4752413cb0cbdc03fb95820ca167f0cdc63b597ccdb5ef43111180e088b0
MD5: 2b40c98ed0f7a1d3b091a3e8353132dc
SHA-1: df79c86fdd11b9ccb89148458e509f879c72566c
Analysis Result
악성: 0
의심스러운: 0
검출되지 않음: 72
무해함: 0
시간 초과: 0
확인된 시간 초과: 0
검사 실패: 1
지원되지 않는 유형: 4
엔진별 결과

THANK YOU