Novelty Detection on Web-server Log Dataset

웹서버 로그 데이터의 이상상태 탐지 기법

AD20216801

김재민(Jae-Min Kim)

Novelty Detection on Web-server Log Dataset

웹서버 로그 데이터의 이상상태 탐지 기법

- Journal of the Korea Institute of Information and Communication Engineering
- 한국정보통신학회논문지 Vol. 23, No. 10: 1311~1319, Oct. 2019

- Hwaseong Lee1* · Ki Su Kim2 1*Senior Researcher, Agency of Defense and Development, Daejeon 34186, Korea 2Researcher, Agency of Defense and Development, Daejeon 34186, Korea

- 이화성1* 정보보호학과 공학박사 , 국방과학연구소
- 김기수2 컴퓨터공학과 공학석사 , 국방과학연구소

웹 해킹

웹환경에서 개인 정보 유출이나 시스템 장애 등을 목표로 하는 외부 해킹의 공격 .

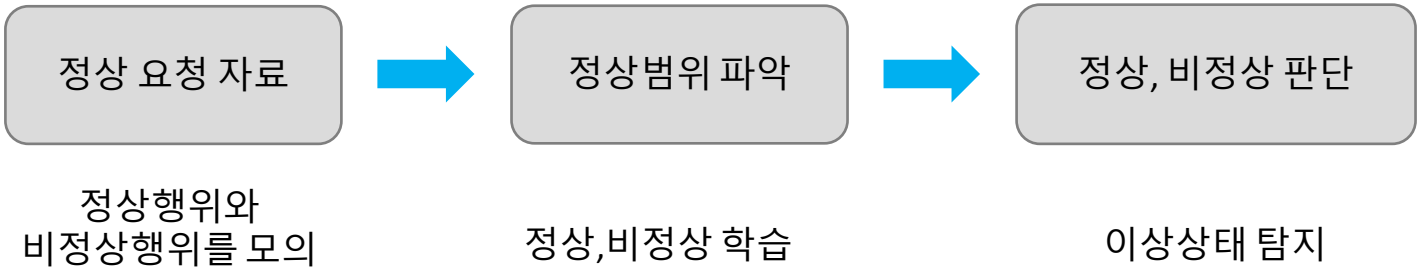
2017년 발표된 웹 어플리케이션 보안 위협 랭킹

- 1위 SQL Injection
- 3위 XSS(cross-site scripting)

사이버 공격 탐지 기술: 시그니처 기반 분석

제안

이상상태 탐지기법(Novelty Detection)은 정상적인 요청으로 정상상태 모델을 생성한 후 정상적인 요청에서 벗어난 형태의 요청이 발생할 경우 이상상태로 탐지하는 기법이다.



관련 연구 논문

J. Liang, W. Zhao, and W. Ye : Anomaly-Based Web Attack Detection: A Deep Learning Approach

비정상을 탐지하기 위해 딥러닝 모델 기반 비정상 탐지 방법을 제안

- URL의 절대경로와 쿼리 파라미터를 파싱하고,
- LSTM(Long Short-term Memory) / GRU (Gated Recurrent Unit)를 이용한 RNN(Recurrent Neural Network) 모델의 입력 데이터로 사용하여 절대 경로와 쿼리 파라미터의 정상 패턴을 학습.
- 학습된 RNN 모델의 결과는 Multi layer Perceptron 모델의 입력으로 활용하여 웹 요청의 비정상을 탐지.

장점: 학습된 모델은 좋은 성능을 보여줌.

단점: GET 요청 방식만 고려한다는 점.
실제운용상황에서 정상요청과 비정상요청 데이터 학습하는 딥러닝 모델, 운용상황이라 분리가 필요함

J. Liang, W. Zhao, and W. Ye, “Anomaly-Based Web Attack Detection: A Deep Learning Approach,” the VI International Conference on Network, Communication and Computing. ACM, pp. 80-85, 2017.

관련 연구 논문

H. Mac, D. Truong, L. Nguyen, H. A. Tran, and D. Tran: “Detecting Attacks on Web Applications using Autoencoder,”

Regularized Deep Autoencoder를 사용한 비정상 탐지 방법을 제안.

- URL을 요청 방식, 절대 경로, 쿼리 파라미터로 구분하기 위해 토큰화(Tokenizing)하고 문자를 아스키 코드와 일치하는 숫자로 치환하고 정규화.
- 전처리 단계에서 생성한 벡터로 오토인코더(Autoencoder)를 학습.
- 학습된 오터인코더 모델의 결과는 실제 데이터와 모델이 예측한 데이터간의 오차 오류 값으로 이 값의 분포 내 임계값(Threshold)을 기준으로 정상과 비 정상을 식별.

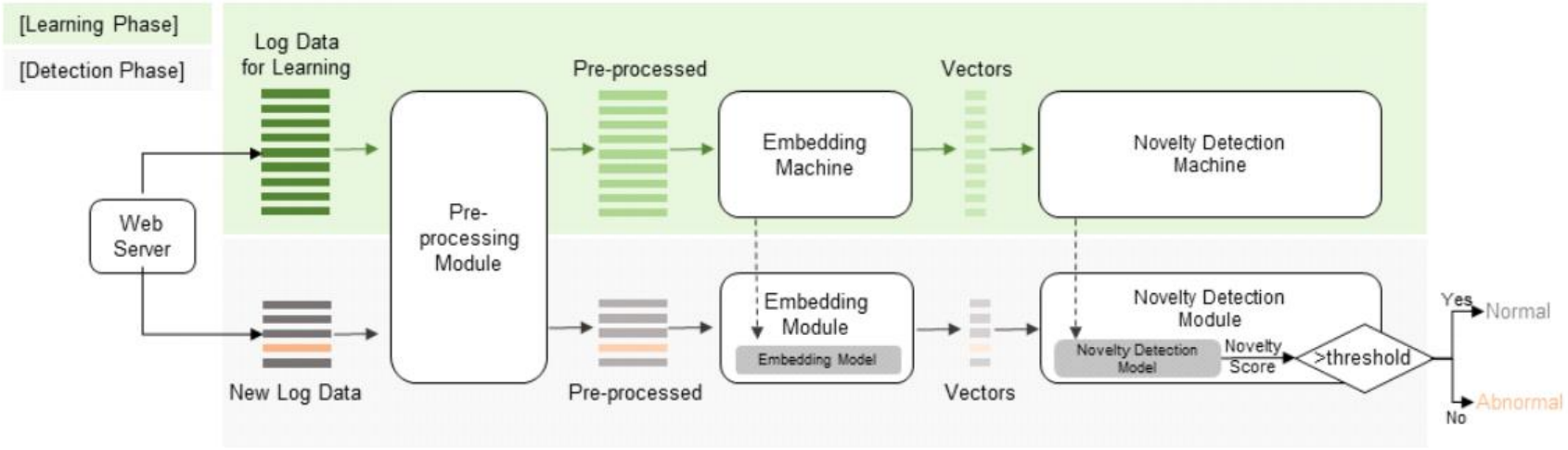
장점: 학습된 모델은 매우 좋은 성능을 보여줌.

단점: URL 경로의 파라미터들 간의 순서는 배제하고 학습하기 때문에 비정상적인 위치 또는 방법으로 접근한 정상행위나, 고도화된 공격의 맥락은 놓칠 수 있다.

H. Mac, D. Truong, L. Nguyen, H. A. Tran, and D. Tran, “Detecting Attacks on Web Applications using Autoencoder,” the 9th Internationa Symposium on Information and Communication Technology, Viet Nam, pp. 416-421, 2018.

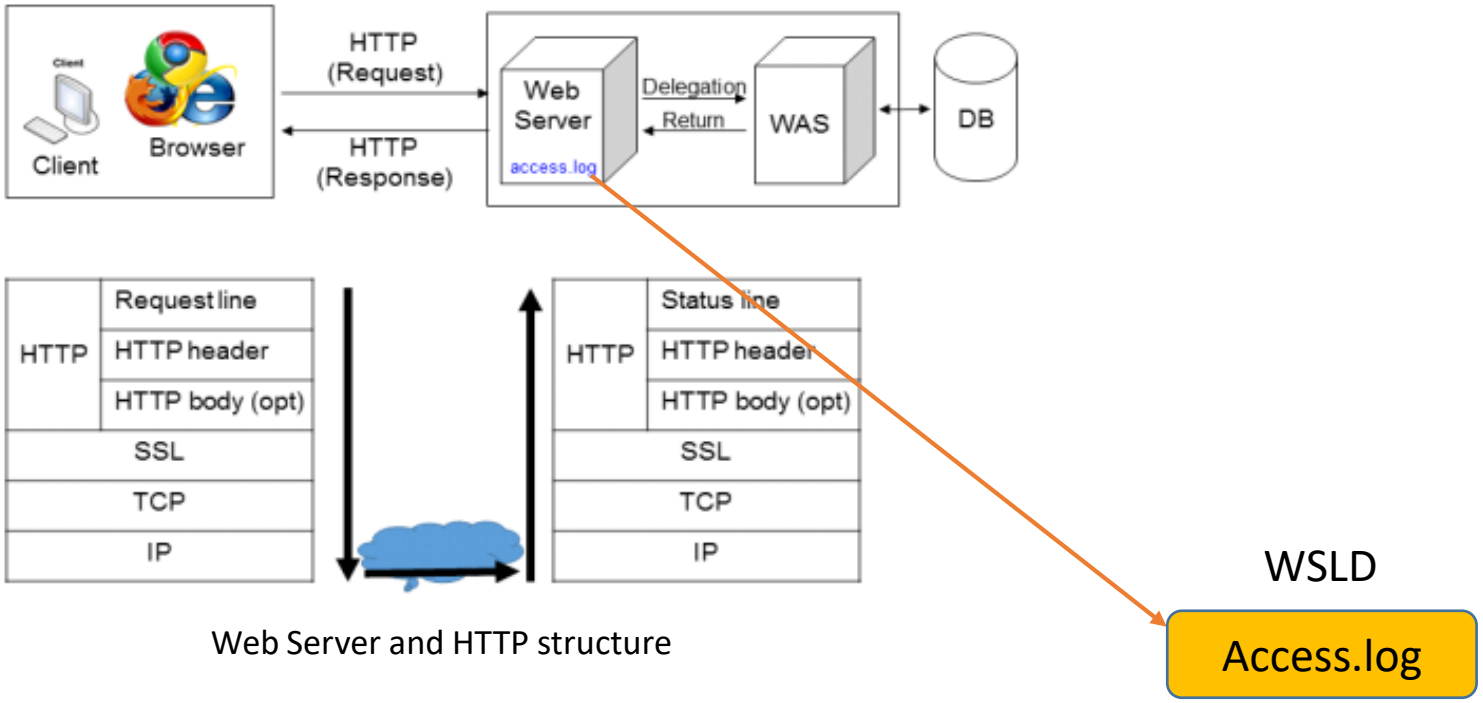
웹서버 로그 기반 이상상태 탐지 기법

이상상태 탐지 기법은 그림과 같이 학습단계(Learning Phase)와 탐지단계(Novelty Detection Phase)로 나뉜다



The architecture of the web novelty detection (웹 이상치 감지 아키텍처)

웹서버 로그 데이터셋(Web-server Log Dataset) 수집

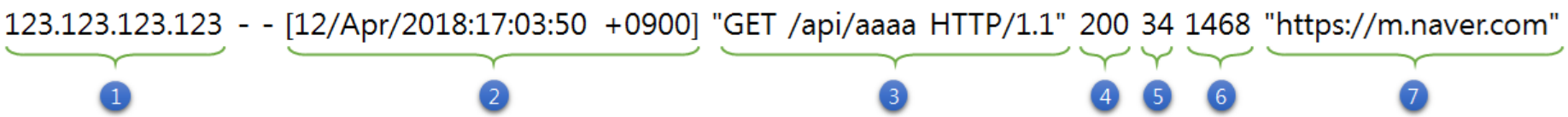


Web Server and HTTP structure

Apache HTTP server의 Access log는 서버가 처리하는 모든 요청을 기록한다.

Access.log 구조

```
123.123.123.123 - - [12/Apr/2018:17:03:50 +0900] "GET /api/aaaa HTTP/1.1" 200 34
1468 "https://m.naver.com" "Mozilla/5.0 (iPhone; CPU iPhone OS 11_3 like Mac OS X)
AppleWebKit/605.1.15 (KHTML, like Gecko) Mobile/15E216 NAVER(inapp; search; 580;
8.6.3; 7)"
```



1. 원격 호스트 IP 주소(요청자)
2. 요청 시간
3. 'GET' 메서드를 사용하고 '/api/aaaa'라는 경로에 'HTTP/1.1'의 프로토콜로 요청
4. HTTP 상태 코드(1**, 2**, 3**, 4**)
5. HTTP 헤더를 제외한 전송 바이트 수
6. 요청을 처리하는 데 걸린 시간(ms)
7. 리퍼러(referrer), 요청이 위치한 주소

A list of fields in web-server log (웹서버 로그 필드 정보 목록)

Field Name	Desc
method	Request method to a web server (GET, POST etc)
URL	Path to access to HTTP server, including parameters(operator, operand and variables) on GET method
status	Status code that the server sends back to the client, including success, failure, redirection etc
http_version	http version information
referrer	Address of webpage which is linked to the resource being requested. Web-server can see where the request originated (Optional)
user_agent	Software agent that is acting on behalf of a client, such as web browser [ex:Mozilla/5.0 (compatible; MSIE 10.0; Windows NT 6.1; WOW64; Trident/6.0)]

Access.log 수집 행위 모의

정상행위 모의

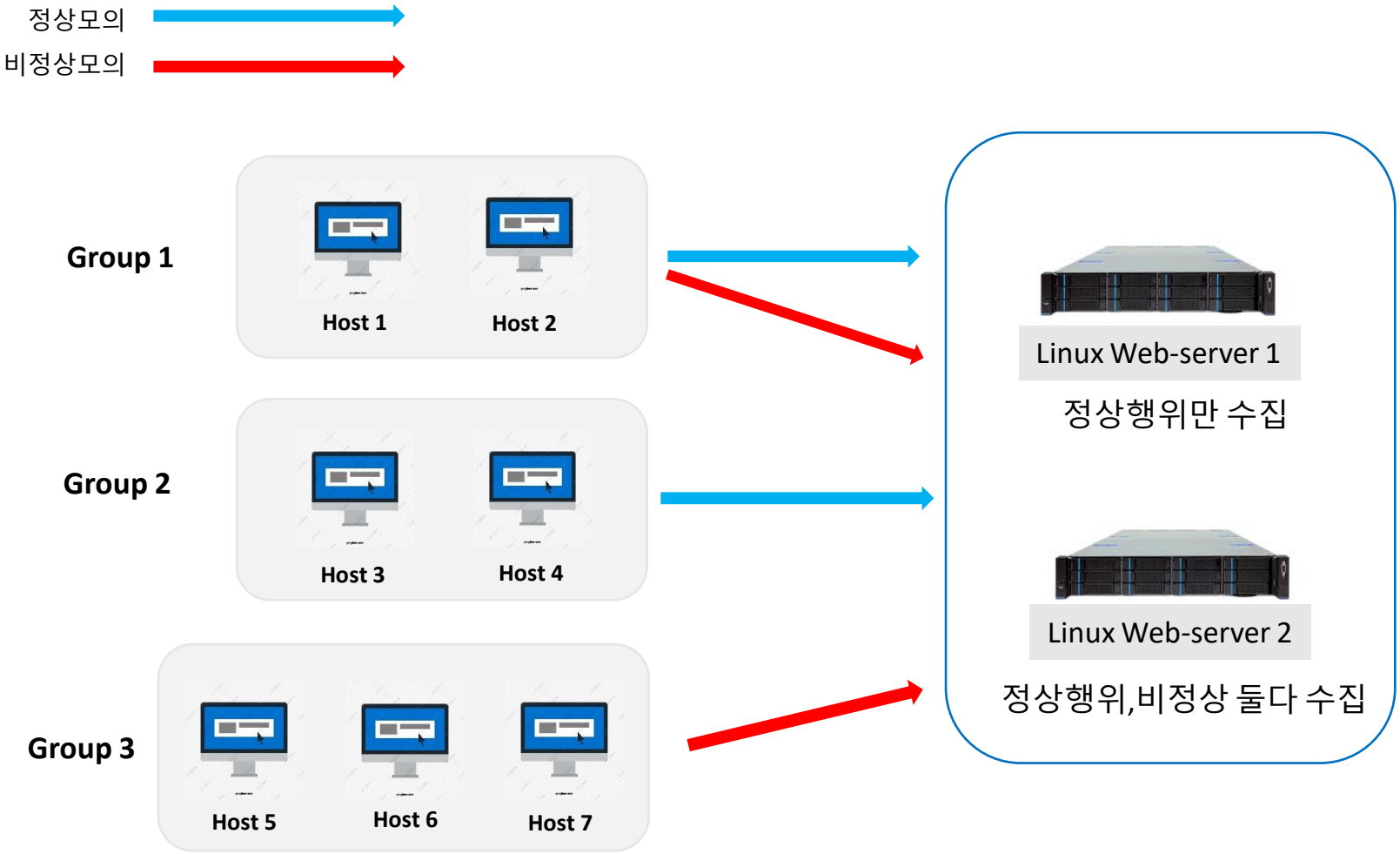
- 사용자가 일반적으로 웹서버에 접속하는 방식 행위를 모의.
- GET과 POST에 해당되는 행위를 모의

비정상행위 모의

- 사용자가 악의적인 목적으로 수행
- 사이버 공격 , 단순 오용 등 정상행위에 서 벗어난 행위 전반을 발생.

- 정적 공격: 존재하지 않은 리소스에 대한 요청 , 임의의 폴더나 파일에 무작 위로 접근을 시도.
- 동적 공격: 웹서버 취약점을 악용한 공격, 인자를 수정한 공격(Command injection, SQL injection, XSS)
- 단순 오용 등: 웹 어플리케이션의 정상적인 행위 를 따르지 않은 시도.

Access.log 수집 행위 모의



데이터셋 라벨링

- 라인별로 라벨(정상, 비정상)을 부여
- 호스트IP를 기준으로 정상행위와 비정상행위 대상 파악이 가능

정상 51.222.253.7 - - [13/Mar/2022:04:08:02 +0900] "GET /robots.txt HTTP/1.1" 200 21 "-" "Mozilla/5.0 (compatible; AhrefsBot/7.0…

비정상 114.119.159.122 - - [13/Mar/2022:09:02:18 +0900] "GET /read.php3?aid=1635385996715664073 HTTP/1.1" 500 658199 "-" "…

정상 114.119.156.185 - - [13/Mar/2022:08:18:59 +0900] "GET /js/jssor.slider.min.js HTTP/1.1" 200 50627 "http://tig0204.kr/index.php" …

정상 114.119.156.185 - - [13/Mar/2022:08:18:59 +0900] "GET /js/jquery.scrollbox.js HTTP/1.1" 200 8668 "http://tig0204.kr/index.php" …

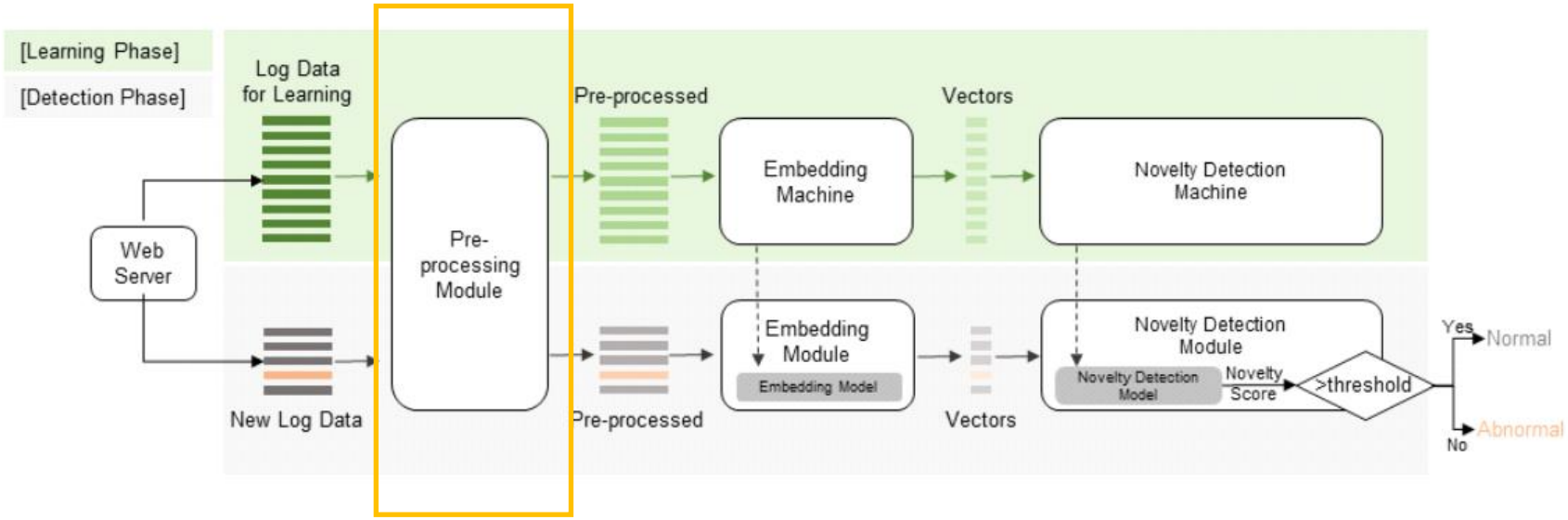
비정상 119.207.72.84 - - [13/Mar/2022:12:22:40 +0900] "GET /sky/log/ HTTP/1.1" 307 235 "http://tig0204.kr/index.php" "Mozilla/5.0 …

정상 114.119.156.185 - - [13/Mar/2022:08:18:59 +0900] "GET /js/jquery-1.11.3.min.js HTTP/1.1" 200 95957 "http://tig0204.kr/index.php...

정상 114.119.156.185 - - [13/Mar/2022:08:18:59 +0900] "GET /js/printThis.js HTTP/1.1" 200 13098 "http://tig0204.kr/index.php...

정상 114.119.156.185 - - [13/Mar/2022:08:18:59 +0900] "GET /js/tweenmax.js HTTP/1.1" 200 189934 "http://tig0204.kr/index.php....

데이터 전처리



The architecture of the web novelty detection (웹 이상치 감지 아키텍처)

데이터 전처리 방법

207.46.13.166 - - [13/Mar/2022:16:57:12 +0900] "GET /main_program/221?search=sea HTTP/1.1" 200 278171 "-" "Mozilla/5.0 (compatible; Barkrowler/0.9; +https://babbar.tech/crawler)"

1. 각 “/” 와 파라미터에 공백으로 처리해서 문장화

URL : /main_program/221?search=sea ⟹ main_program 221 search sea

2. 공백을 제거해서 단어(Word)화

http_referrer, user_ agent:

"Mozilla/5.0 (compatible; Barkrowler/0.9; +https://babbar.tech/crawler)"

"Mozilla/5.0(compatible;Barkrowler/0.9;+https://babbar.tech/crawler)"

3. IP: 임베딩시 문서 ID (Documentation ID)로 사용

4. status(상태코드) : 2**는 성공, 4**은 범주형 변수이므로 수치화 데이터로 변환

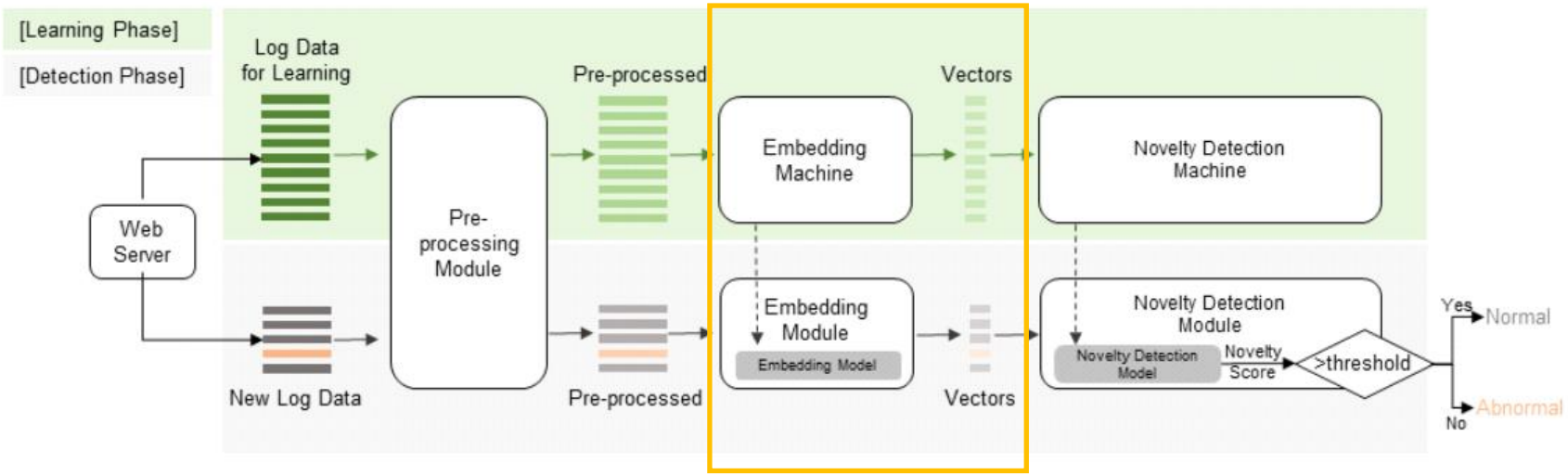
데이터 전처리 항목

Field Name	Desc
method	Request method to a web server (GET, POST etc)
URL	Path to access to HTTP server, including parameters(operator, operand and variables) on GET method
status	Status code that the server sends back to the client, including success, failure, redirection etc
http_version	http version information
referrer	Address of webpage which is linked to the resource being requested. Web-server can see where the request originated (Optional)
user_agent	Software agent that is acting on behalf of a client, such as web browser [ex:Mozilla/5.0 (compatible; MSIE 10.0; Windows NT 6.1; WOW64; Trident/6.0)]

·전처리 대상 항목 Ⅰ: URL

·전처리 대상 항목 Ⅱ: method, URL, status, http_version, referer, user-agent

임베딩



The architecture of the web novelty detection (웹 이상치 감지 아키텍처)

URL

method, URL, status, http_version, referer, user-agent

Doc2Vec Autoencoder

임베딩 기법

1. Doc2Vec

Word2Vec의 방법론을 응용하여 하나 이상의 문장으로 구성된 문서 단위의 객체 내 내재된 맥락 정보를 학습하여 수치화된 벡터를 생성하는 비지도 학습 모델.

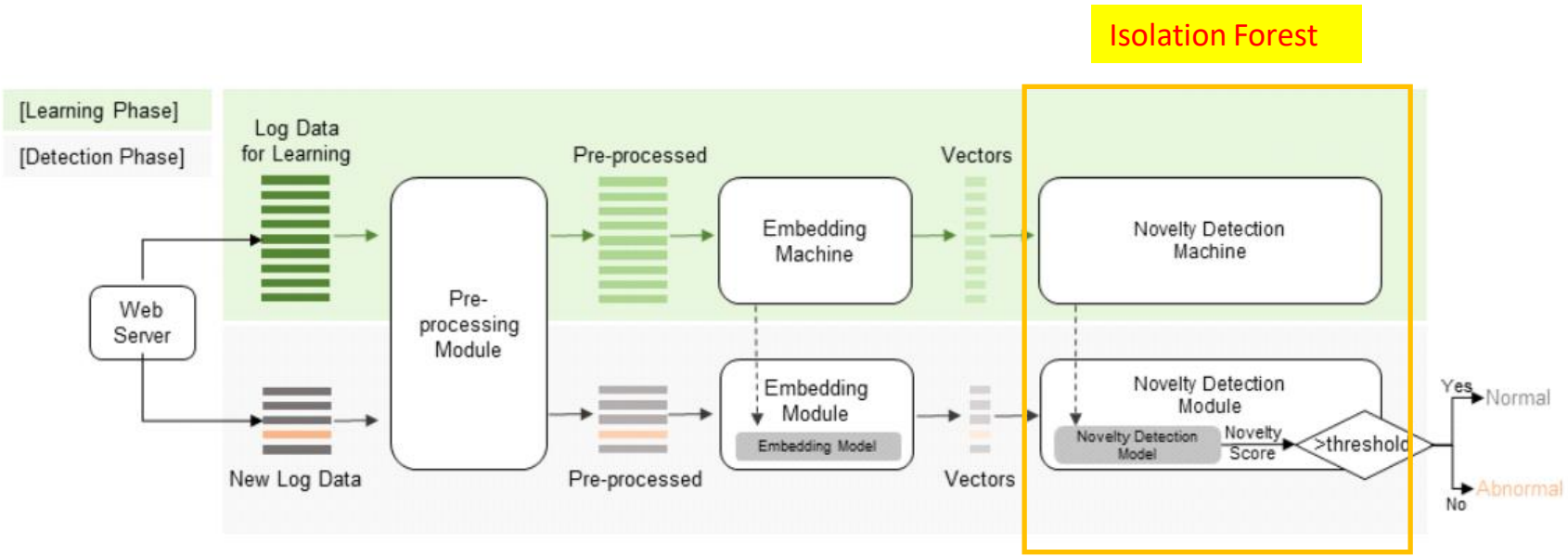
Word2Vec의 경우 임의의 벡터 공간에 각 단어에 대응되는 벡터를 생성하여 단어 간 의미론적 연관관계를 효과적으로 표현할 수 있지만, 하나 이상의 단어로 구성된 문장에 내재된 맥락 정보를 충분히 표현 하지 못한다는 단점이 있다.

Doc2Vec은 Word2Vec으로 생성된 단어 벡터와 문서 단위의 벡터를 연관시키는 학습으로 N개의 단어로 구성된 문서에서 1개의 문서 벡터와 N-1개의 단어들로 남은 하나의 단어를 예측하는 방식으로 학습한다.

2. Autoencoder

- 신경망을 이용해서 입력으로부터 계산되는 출력이 입력값과 비슷해지도록 학습 하는 기법
- 입력값을 중간층으로 인코딩한 후 다시 입력값과 같은 차원으로 디코딩하고 그 오차를 줄이는 방향으로 역전파 하는 기법

이상상태 탐지 방법



The architecture of the web novelty detection (웹 이상치 감지 아키텍처)

학습단계: 정상 , 비정상을 판별하기 위한 임계치를 계산

탐지단계: 로그의 한줄 라인을 읽고 그 라인을 학습 모델에 넣어서 임계치화 비교하여 정상인지 비정상인지만 파악

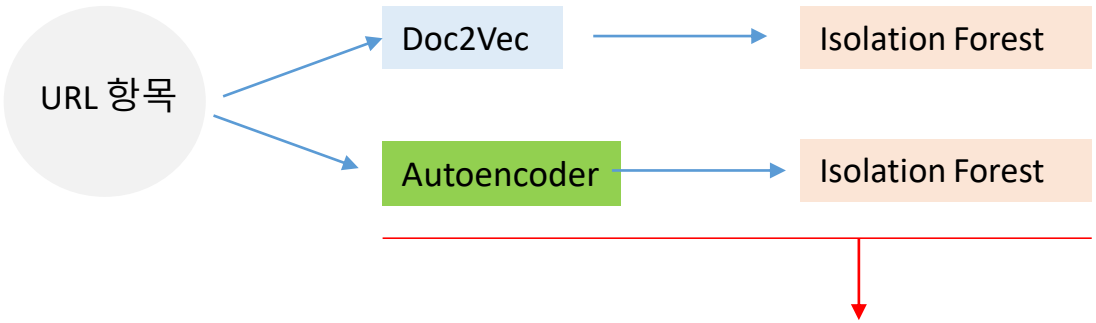
Isolation Forest

- Isolation Forest는 비정상을 탐지하는 모델 중 하나로 일반적으로 알려진 비정상 탐지 모델과는 다른 접근법을 제안한다.
- 일반적인 비정상 탐지 모델은 정상 상태의 프로파일을 생성하고 이 프로파일을 기준으로 정상과 비정상을 구별한다.
- Isolation Forest는 이상치 (Outlier) 데이터의 특성을 고려한 모델로 이상치 데이터는 정상치 데이터에 비해 그 양이 적으며, 정상 데이터와는 다른 특성을 갖는다는 전제를 가지는 모델이다
- 학습을 위해서 전체 데이터 셋 중 일부를 샘플링하여 하나 이상의 트리 모델을 생성하는 앙상블 기법
- 각 트리는 랜덤 하게 데이터를 구분하며 데이터 셋의 모든 데이터가 단말 노드에 의해 구별된 유일한 데이터로 고립될 때까지 랜덤으로 분기를 반복한다.
- 하나의 트리가 완성되었을 때 이상치 데이터는 트리를 탐색하는 경로가 정상 데이터에 비해 경로 길이가 짧으며
- 정상과 비정상 사이의 탐색 경로 길이를 임계값으로 정상과 비정상을 구분한다.

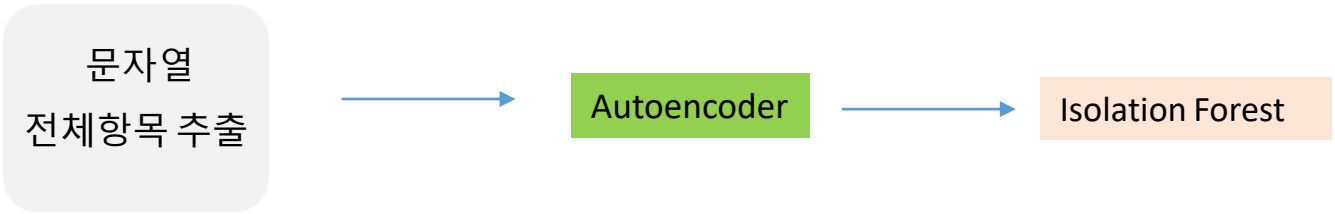
실험 및 평가

1. 임베딩 기법별 이상상태 탐지 실험

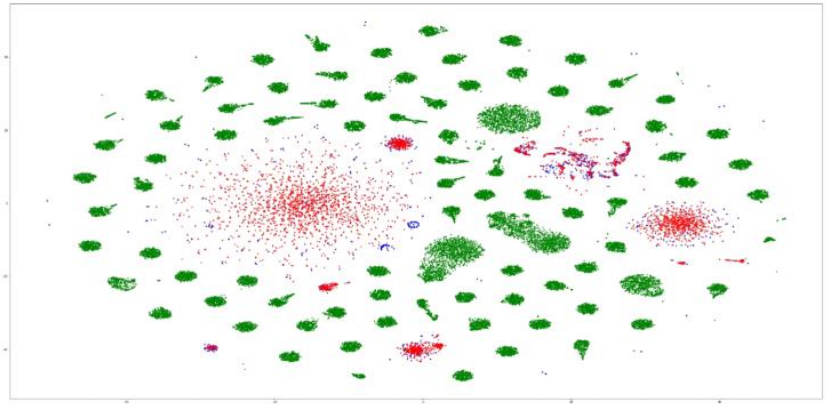
Isolation Forest
앙상블 트리 개수는 100, 이상치 기준은 0.1



2. 추출항목별 이상상태 탐지 실험



t-SNE visualization (t-SNE 시각화)



URL-then-Doc2Vec

green: normal data for learning,

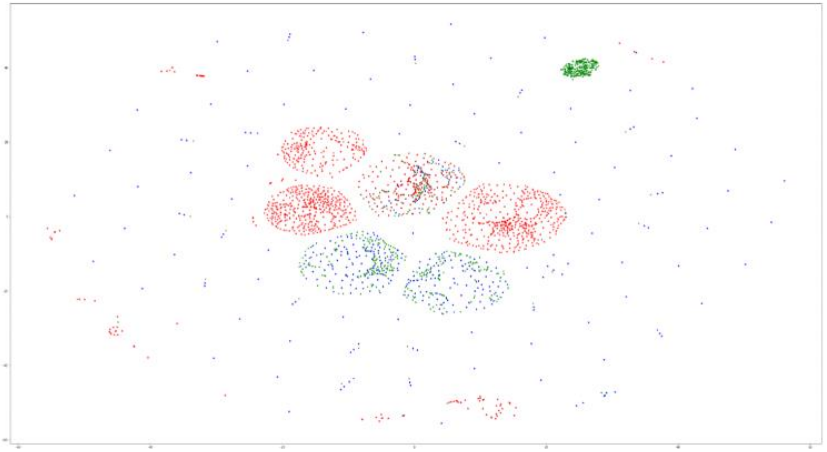
blue: normal data for validation,

red: abnormal data for validation

녹색: 학습용 정상 데이터

파란색: 검증용 정상 데이터

빨간색: 검증용 비정상 데이터



URL-then-Autoencoder

green: normal data for learning,

blue: normal data for validation,

red: abnormal data for validation

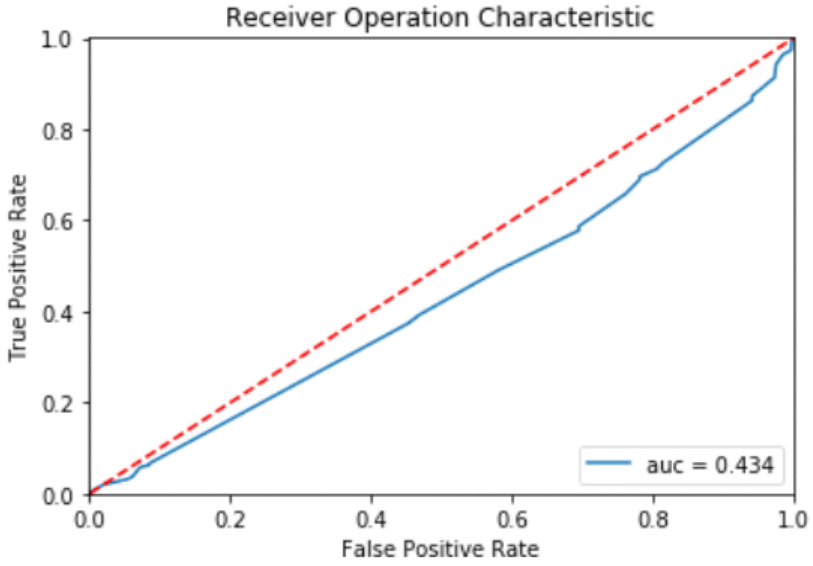
녹색: 학습용 정상 데이터

파란색: 검증용 정상 데이터

빨간색: 검증용 비정상 데이터

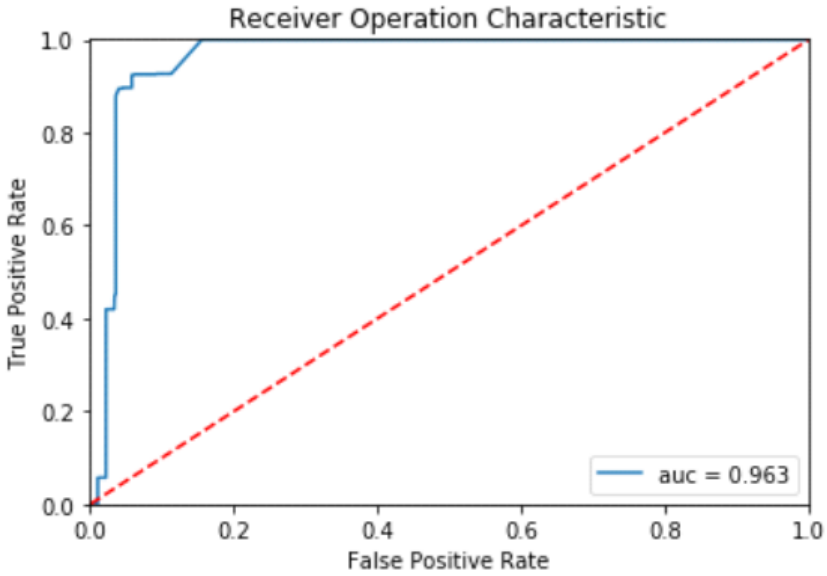
Receiver operating characteristics, ROC 그래프

앙상블 트리 개수는 100, 이상치 기준은 0.1



ROC cuve of 'URL-Doc2Vec-IF' experiment

FPR (False positive rate)오탐률: 상승



ROC cuve of 'URL-Autoencoder -IF' experiment

FPR (False positive rate)오탐률: 대략 0.16

		ACC	Precision	AUROC
URL only	Doc2Vec	0.513	0.558	0.459
	Autoencoder	0.913	0.902	0.963
		정확도	정밀도	성능평가

ROC curve of ‘Strings-Autoencoder -IF’ experiment

앙상블 트리 개수는 100, 이상치 기준은 0.1

Table. 4 Comparison of novelty detection on the WSLD 분류성능 평가 비교

		ACC	Precision	AUROC
URL only	Doc2Vec	0.513	0.558	0.459
	Autoencoder	0.913	0.902	0.963
All strings	Autoencoder	0.946 ↑	0.907 ↑	0.983 ↑
		정확도	정밀도	성능평가

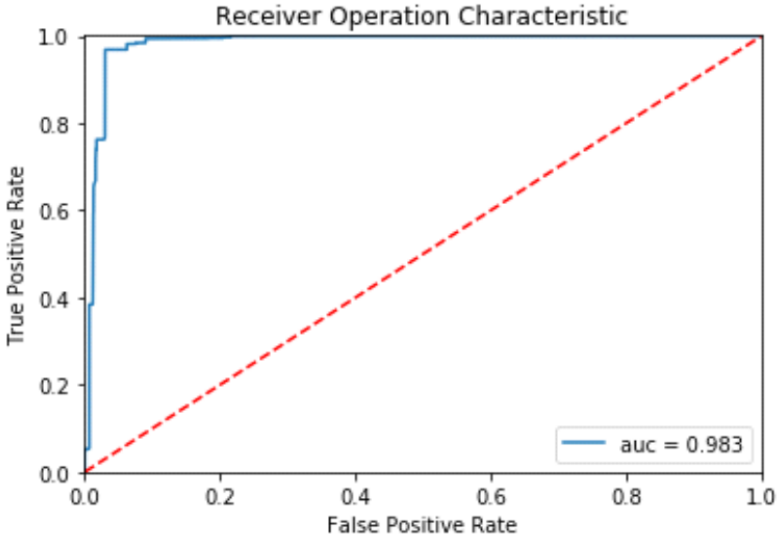


Fig. 7 ROC curve of ‘Strings-Autoencoder -IF’ experiment

FPR (False positive rate)오탐률: 대략 0.102

결론

Doc2Vec		Autoencoder
AUROC 0.459	<	AUROC 0.963

웹서버 로그 도메인에서는 Doc2Vec보다 오터인코더가 더 적합한 임베딩 기법임을 확인

문자열 전체항목 추출 → Autoencoder → Isolation Forest

URL 단일항목보다 문자열 전체 항목을 임베딩 후 Isolation Forest를 진행할 경우 오탐률은 동일한 수준으로 유지하면서 정탐률이 높이는 이상 상태 탐지 모델이 생성됨을 확인.

감사합니다.