소셜 네트워크 서비스 기반 마이닝을 이용한 실시간 랜섬웨어 위험도 분석 시스템 설계

나재호, 김미희*
한경대학교 컴퓨터공학과
e-mail : magrejh@hknu.ac.kr, mhkim@hknu.ac.kr*

Design of a Real-time Risk Analysis System for Ransomware Using Mining based on Social Network Service

Jaeho Na, Mihui Kim*
Department of Computer Science & Engineering, Computer System Institute, Hankyong National University (*Corresponding author)

요 약

본 논문에서는 소셜 네트워크 서비스 중 트위터를 마이닝하여 실시간으로 랜섬웨어 위험도 분석을 하는 시스템을 설계한다. 이를 위해 2017 년 5 월 12 일에 가장 피해가 컸던 워너크라이 랜섬웨어를 중심으로 5 월 10 일에서 20 일 사이의 트윗 데이터를 마이닝하고, 기존 시스템인 구글 트렌드와의 유사성을 비교 실험하여 트윗 데이터의 가치를 확인한다. 마지막으로 제안하는 시스템에 대한 향후 연구주제를 제시한다.

1. 서론

랜섬웨어는 악성 소프트웨어의 일종으로서 사용자의 시스템을 감염시켜 파일을 암호화하고 공격자는 복호화 키에 대한 돈을 요구한다. 전세계적으로 랜섬웨어의 위협은 증가하고 있다. 2017 년에는 병원, 극장 등의 사회간접자본을 마비시키는 사례도 발생하였다[1][2]. 현재 국내에 랜섬웨어에 위험도를 실시간으로 알려주는 언론 매체나 보안 사이트가 없다. 대부분 피해 이후에 결과에 대해서 알려주는 정도에 그친다.

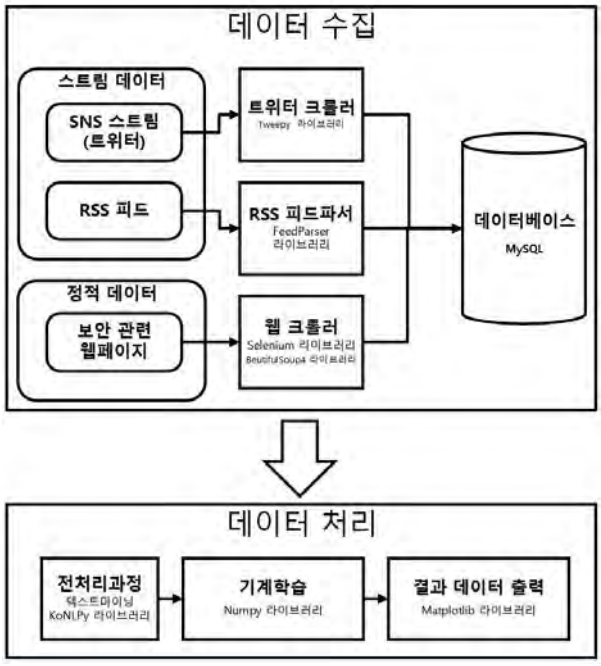
소셜 네트워크 서비스를 지진과 같은 재난 예측 및 분석에 사용한 연구가 많이 있다[3]. 그리고 2016 년에는 소셜 네트워크 서비스를 실시간으로 활용한 지진을 감지해주는 시스템이 이슈가 됐던 사례가 있다[4]. 랜섬웨어도 하나의 재난으로 본다면 랜섬웨어에 대한 위협을 실시간으로 분석해 주는 시스템이 필요할 것이며, 소셜 네트워크 서비스를 활용한다면 실시간으로 랜섬웨어에 대한 위협을 감지하여 대비하는데 사용할 수 있을 것이다.

이에 본 논문에서는 소셜 네트워크 서비스기반 마이닝을 이용한 실시간 랜섬웨어 위험도 분석 시스템을 제안하고자 한다.

2. 제안 시스템

2.1 시스템 개요

본 논문에서 제안하는 시스템의 개요는 그림 1 과 같다. 본 시스템은 웹에서 데이터를 수집해 데이터베이스로 저장하는 데이터 수집 부분과 수집된 데이터의 분석을 용이하게 하고 결과를 출력하는 데이터 처리 두 부분으로 구성된다.



(그림 1) 시스템 개요도

2.2 데이터수집

2.2.1 스트림 데이터

본 시스템에서는 소셜 네트워크 서비스 중에서 트위터(twitter.com)를 사용한다. 연구가 이뤄진 시점에서 트위터는 140 자 미만의 텍스트 위주의 개인 블

로그이다. 또한 충분히 많은 사용자에 의해 널리 사용되고 있다. 따라서 마이닝의 대상이 되는 소셜 네트워크 서비스 중에서 표본 데이터로서 충분한 가치가 있고 분석이 상대적으로 용이하다. 트위터에서 제공하는 퍼블릭 스트림에 대해서 특정 키워드로 필터링을 하고 그 결과를 마이닝 한다. 사용하고자 하는 라이브러리는 Tweepy 이다[5].

국내와 해외의 랜섬웨어 피해사례에서 이용된 랜섬웨어의 종류는 상이했다[6]. 따라서 국내 실정에 맞는 시스템이 필요하다. RSS 피드는 안랩에서 제공하는 보안 뉴스 제목을 제공해 주는 피드를 사용한다[7]. RSS FeedParser 라이브러리를 사용해 피드를 주기적으로 수집하여 랜섬웨어 관련 키워드를 추출하여 사용한다.

2.2.2 정적 데이터

Selenium 라이브러리와 PhantomJS 드라이버를 이용하여 웹페이지를 동적으로 크롤링한다[8][9]. 그 후 수집된 자료를 BeautifulSoup4 라이브러리를 이용해 스크래이핑한다[10]. 스크래이핑된 자료는 MySQL Connector 를 이용하여 미리 만들어 놓은 데이터베이스 테이블에 저장한다. 크롤링은 각 페이지마다 그리고 크롤링이 실시되는 시점에 따라 다른 크롤링 방법을 요구하기도 한다. 따라서 특정하기 힘들지만 여러 보안관련 페이지(예, Rancert 보안뉴스 페이지, Rancert 랜섬웨어 피해사례 페이지 등)에서 수행되며, RSS 피드와 마찬가지로 국내 실정에 맞는 시스템 개발을 위해 이용자가 충분히 많은 국내사이트에서 랜섬웨어 관련 키워드를 주기적으로 수집한다.

2.3 데이터처리

2.3.1 전처리과정

모든 데이터는 자연어 형식으로 구성 되어 있으며 표준어 형식을 지키지 않은 데이터가 많다. 이 때의 노이즈를 최소화 하기 위해 KoNLPy 라이브러리를 사용하여 비표준어 데이터를 정규화하고, 형태소를 분리하여 명사를 추출하는 과정을 수행한다[11].

2.3.2 기계학습

기계학습을 활용하는 목적은 첫째, 빅데이터의 예측분석을 하기 위함, 둘째, 노이즈가 많은 복잡한 문제를 해결하기 위함에 있다. 측정 결과로 퍼블릭 스트림을 통해서 일일 7 천개 이상의 트윗이 만들어진다. 트윗 중에 스팸성인 데이터와 너무 짧아서 분석하기 어려운 경우가 많다. 따라서 기계학습을 이용해 효율적으로 방대하고 노이즈가 많은 트윗 빅데이터를 사용하고자 하며, 특히 노이즈 처리에 유리한 딥러닝을 통해 구현하고자 한다. 기본적인 벡터 연산을 효율적으로 수행하기 위해 Numpy 라이브러리를 사용하고자 한다[13].

2.3.3 데이터출력

Matplotlib 라이브러리를 사용하여 자료를 쉽게 그래프와 같은 형태로 시각화 하고자 한다[14].

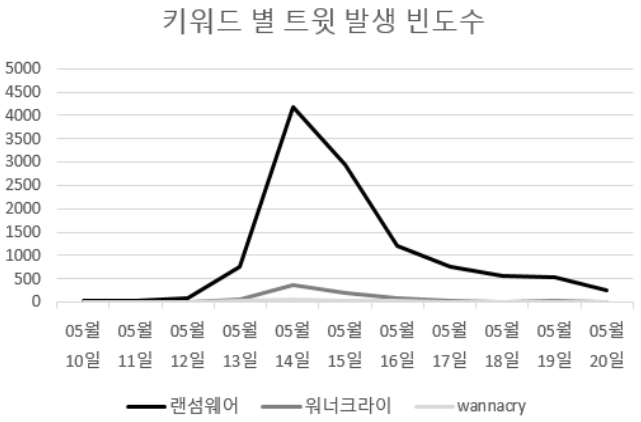
3. 모의실험

3.1 모의실험의 목적과 개요

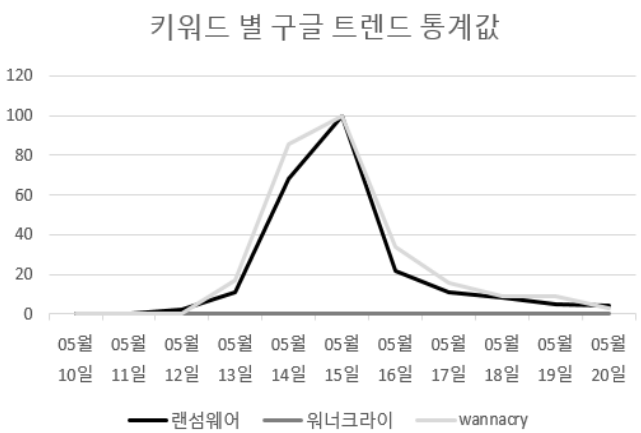
우선 트위터 소셜 네트워크 서비스의 데이터가 랜섬웨어 위험도 분석을 위해 가치가 있는지를 확인하기 위해 실험한다. 워너크라이 랜섬웨어의 피해가 가장 컸던 2017 년 5 월 12 일을 중심으로 하여 5 월 10 일부터 20 일 사이의 트위터의 데이터를 수집하였다. 구글 트렌드는 다양한 분야에서 통계 데이터로 활용되고 있다. 이에 구글 트렌드가 제공하는 통계 결과와 유사한 결과가 나오는지 확인해 보았다.

트위터에서 제공하는 API 는 최근 7 일내의 트윗만을 제공한다[15]. 모의 테스트는 2017 년 5 월 12 일을 중심으로 5 월 10 일에서 20 일 사이의 자료를 사용하고자 한다. 따라서 공식적으로 제공되는 API 로는 자료 마이닝이 불가능 했다. 그래서 크롤러를 사용하여 트위터 검색 페이지를 크롤링하였다. 크롤러는 앞서 시스템 개요에서 소개한 크롤러를 사용하였다. 크롤러로 수집된 트윗 데이터는 웹페이지 태그가 포함된 정제되지 않은 형태이지만 트위터에서 제공하는 검색기능을 활용하여 충분히 필터링 된 통계자료를 낼 수 있었다.

3.1 실험 결과



(그림 2) 키워드 별 트윗 일일 발생 빈도수



(그림 3) 키워드 별 구글 트렌드 일일 통계 값

항목 : 구글 트렌드 값(백분율) / 트윗 수(백분율)

항목	랜섬웨어	워너크라이	Wannacry
10 일	0(0) / 27(0.2)	0(0) / 0(0)	0(0) / 1(0.5)
11 일	0(0) / 29(0.3)	0(0) / 0(0)	0(0) / 0 (0)
12 일	2(0.9) / 91(0.8)	0(0) / 0(0)	0(0) / 6 (3.1)
13 일	11(4.8) / 762(6.7)	0(0) / 73(8.9)	17(6.2) / 33 (16.8)
14 일	68(29.4) / 4193(36.9)	0(0) / 370(45.3)	86(31.4) / 63 (32.1)
15 일	100(43.3) / 2954(26)	0(0) / 207(25.3)	100(36.5) / 27 (13.8)
16 일	22(9.5) / 1208(10.6)	0(0) / 81(9.9)	34(12.4) / 20 (10.2)
17 일	11(4.8) / 772(6.8)	0(0) / 35(4.3)	16(5.8) / 13 (6.6)
18 일	8(3.5) / 554(4.9)	0(0) / 17(2.1)	9(3.3) / 12 (6.1)
19 일	5(2.2) / 538(4.7)	0(0) / 20(2.4)	9(3.3) / 16 (8.2)
20 일	4(1.7) / 244(2.1)	0(0) / 14(1.7)	3(1.1) / 5 (2.6)
합계	231(100) / 11372(100)	0(-) / 817(100)	274(100) / 196(100)

(표 1) 키워드 별 세부 수치

워너크라이 랜섬웨어에 피해사례가 발생한 시점은 5 월 12 일이며, 사회적으로 이슈가 된 시점은 13 일이다. 그림 2, 3 과 표 1 의 결과에서 공통적으로 확인 할 수 있는 점은 랜섬웨어 관련 이슈에 충분히 반응을 보인다는 점과 그 시기가 같다는 것이다. 구글 트렌드의 경우 13 일에서 15 일 사이에 전체의 77.4%의 비중을 가지며, 트위터의 경우 13 일에서 16 일 사이에 전체의 80.1%의 비중을 갖는다. 따라서 키워드별 일일 트윗 발생 빈도수 추이와 구글 트렌드의 일일 키워드 통계 값의 추이가 유사함을 확인할 수 있으며, 트위터 소셜 네트워크 서비스의 데이터가 랜섬웨어 위험도 분석을 위해 가치가 있다고 판정할 수 있다.

4. 결론 및 향후 연구

모의실험의 결과로 트위터에서 랜섬웨어와 관련된 키워드가 포함된 트윗의 발생빈도가 랜섬웨어 관심도 분석을 위해 유의미한 결과를 가짐을 확인할 수 있었다. 이 때의 분석은 현재의 관심도를 실시간으로 분석하고, 가까운 미래의 관심도를 예측하는 것을 포함한다. 이는 실시간으로 현재 유행하는 랜섬웨어에 위험도를 분석할 수 있음을 의미하지만 유행 정도를 예측한 결과가 유의미하게 피해예방을 위해 사용될 수 있는지는 추가적인 연구가 필요할 것이다.

제안한 시스템에서 RSS 피드나 여러 보안관련 웹사이트를 크롤링하는 목적은 랜섬웨어와 관련된 키워드와 관련된 지표를 찾기 위함이 있었다. 그러나 이를 위해 구체적인 방법이 제시되지는 않았는데, 랜섬웨어와 관련된 키워드와 지표를 실시간이고 직접적으로 제공해주는 사이트가 마땅하지 않았기 때문이다. 이를 해결하기 위해 두 가지 방향을 제시하고자 한다. 첫 째로, 크롤링을 하지 않아도 되도록 공공 API 를 제공하는 인프라가 구축되는 것이다. 공공 API 인프라가 마련된다면 상대적으로 더 공신력이 있고, 크롤러와 같은 방법보다 훨씬 안정적으로 관련 지표와 키워드를 공급받을 수 있을 것이다. 두 번째로, 트윗을 활용하여 새로이 생겨나는 랜섬웨어에 대한 키워드를 얻는 것이다. 트위터에서 지진이 발생한 곳의 지명을 얻으려는 연구가 있었다[16]. 이와 유사한 방법으로 랜섬웨어와 관련된 키워드를 얻는 연구가 있다면 유효할 것이다.

향후, 제안하는 시스템에서 사용하는 기계학습의 회귀모델에 대한 상세한 연구가 필요하다. 트윗 데이터를 직접적으로 입력으로 사용하는 방법도 가능할 것이다. 하지만 트윗 데이터를 그대로 사용한다면 노이즈가 많은 데이터고 그 학습과정이 어떤 과정을 거치는지 명시적이지 않다. 성능의 최적화를 위해서 그리고 학습과정을 보다 명시적으로 분석을 위해서 위험성과 관련된 지표를 활용한다면 더 좋을 것이다.

5. Acknowledgement

“이 논문은 2015 년도 정부(교육부)의 재원으로 한국연구재단 기초연구사업의 지원을 받아 수행된 연구임(No. 2015R1D1A1A01057362)”. 교신저자 김미희.

참고문헌

[1] 김효섭, “동시다발 랜섬웨어 공격 100 개 국가 피해, 병원・기업 마비”, 연합뉴스, 2017.05.13, https://www.youtube.com/watch?v=C0jFu0iP3OI
[2] 윤동진, “CGV "최다 50 개 극장 광고서버 랜섬웨어 감염 복구작업 중”, 연합뉴스, 2017.05.15, http://www.yonhapnews.co.kr/bulletin/2017/05/15/02000000000AKR20170515060800005.HTML,
[3] Takeshi Sakaki, Makoto Okazaki, Yutaka Matsuo, “Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors,” 19th International Conference on World Wide Web, pp.851-860, 2010.
[4] 김수빈, “'지진희알림'은 정말로 기상청보다 빨랐다”, 허핑턴포스트코리아, 2016.09.21.
[5] 트위터 개발자 문서. https://dev.twitter.com/docs
[6] 한국랜섬웨어 침해대응 센터, “2017 랜섬웨어 침해분석 보고서”, 2017.02.02.
[7] 안랩 RSS 피드, http://www.ahnlab.com/kr/site/etc/rss.do
[8] Selenium, http://www.seleniumhq.org/docs/
[9] PhantomJS, http://phantomjs.org/documentation/
[10] BeautifulSoup4, https://www.crummy.com/software/BeautifulSoup/bs4/doc
[11] KoNLPy: 파이썬 한국어 NLP, http://konlpy-ko.readthedocs.io/ko/v0.4.4/#
[12] 사이토 고키, “밑바닥부터 시작하는 딥러닝”, 한빛미디어, 2017.01.03.
[13] Numpy, https://docs.scipy.org/doc/
[14] Matplotlib, https://matplotlib.org/2.0.2/index.html
[15] 트위터 개발자 문서. https://dev.twitter.com/docs
[16] 하현수, 황병연, “트위터를 활용한 실시간 이벤트 탐지에서의 재난 키워드 필터링과 지명 검출 기법”, 정보처리학회논문지/소프트웨어 및 데이터 공학, 제 5 권 제 7 호, 2016.07.