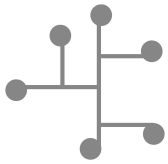


트윗을 이용한 그래프와 시계열 기반 사이버 공격 예측

Predicting of Cyber Attacks based on Graphs and Time-Series Using Tweets

이준하(Junha Lee)[1] 한보영(Boyoung Han)[2] 권혁윤(Hyuk-Yoon Kwon)[3]

요 약

본 논문에서는 트윗을 이용한 그래프와 시계열 기반 사이버 공격 예측 방법을 제안한다. 본 논문은 사이버 공격과 연관된 사용자들 간의 관계를 표현한 그래프 데이터와 시간에 따른 그들의 활동을 표현한 시계열 데이터를 동시에 활용하여 사이버 공격을 예측한 최초의 연구이다. 먼저, 사이버 공격과 연관된 키워드가 포함된 트윗을 자주 작성하는 사용자들을 대상으로 사용자들 간의 관계를 그래프로 모델링한다. 이를 기반으로 사용자 간 그래프 상의 인접성을 고려한 효과적인 클러스터링 기준을 선별한다. 다음으로, 사이버 공격에 대한 뉴스 기사를 기반으로 트윗 사용자들의 트윗 작성 추이를 시계열로 분석한다. 이때, 사용자 그룹 간 사이버 공격 전후의 트윗 작성 빈도를 분석함으로써, 그래프 분석에서 선별된 클러스터링 기준이 사이버 공격 예측에 효과적임을 보인다. 2013-2018년 사이에 발생한 총 58건의 사이버 공격에 대한 분석을 수행한다. 구체적으로, 사이버 공격과 연관된 트윗을 자주 작성하는 1,000명의 사용자 중에서 임의로 선정된 100명의 사용자 그룹이 작성한 트윗에 비하여 그래프 상의 인접성에 따른 클러스터링 기준에 따라 선별된 100명의 사용자 그룹이 작성한 트윗이 사이버 공격 예측에 최대 약 18%만큼 더 효과적임을 보인다.

주제어: 사이버 공격 예측, 그래프 분석, 시계열 분석

1 서울과학기술대학교 글로벌융합산업공학과 IT Management 전공, 학부생, 공동 제1저자
2 서울과학기술대학교 글로벌융합산업공학과 IT Management 전공, 학부생, 공동 제1저자
3 서울과학기술대학교 글로벌융합산업공학과 IT Management 전공, 조교수, 교신저자
+ 이 연구는 서울과학기술대학교 교내연구비의 지원으로 수행되었습니다.
+ 논문접수: 2019년 11월 7일, 심사완료: 2019년 12월 4일, 게재승인: 2019년 12월 11일.

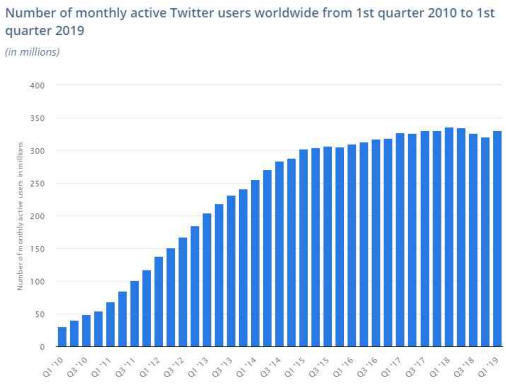
Abstract

In this paper, we propose a method for predicting of cyber attacks based on graphs and time series using tweets. This paper is the first research effort that utilizes both graph data representing the relationship between users who are related to cyber attacks and time-series data representing their activity according to the time. First, we perform a graph-based modeling for describing the relationship between users who frequently write tweets related to cyber attacks. Based on the graph, we identify the effective clustering criteria between users by taking into account their proximity on the graph. Next, we analyze the trend that the users write tweets in a time series based on news articles about cyber attacks. Here, by analyzing the frequency of tweets according to the groups of tweet users before and after cyber attacks, we show that the clustering criteria identified by the graph analysis is effective. We perform the analysis on a total of 58 cyber attacks between 2013 and 2018. Specifically, out of a total of 1,000 users who frequently write tweets related to cyber attacks, we show that tweets written by a group of 100 users, who are selected according to the proximity on the graph, are more effective for predicting cyber attacks by up to 18% compared to the tweets written by a group of 100 users, who are randomly selected.

Keywords: predicting cyber attacks, graph analysis, time-series analysis

1. 개 요

트위터(Twitter)는 2006년 개발된 소셜 네트워크 서비스(SNS)로서 약 1억 9천만 명 이상의 사용자들을 보유하고 있으며, 트위터에는 하루에 약 5,500만 개의 새로운 트윗(Tweets)이 생성된다[16]. 그림 1은 2010년부터 2018년까지 트위터 사용자의 증가 추이를 보여준다.



[그림 1] 트위터 사용자의 증가 추이[4]

트위터는 트위터 사용자들의 트윗이 실시간으로 반영되기 때문에, 트위터를 통해 다양한 유형의 데이터를 수집함으로써 특정 사회 현상이나 사건을 빠르게 예측하는데 유용하게 사용될 수 있다[17]. [13]에서는 트윗 데이터를 이용하여 미국에서 발생하는 질병 활성 정도를 추적하는 연구를 수행하였다. [14]에서는 트윗으로부터 주식 시장의 분위기를 분석하여 주식 가격을 예측하는 연구를 수행하였다. [15]에서는 트윗 분석을 통해 대중들의 시위를 예측하는 연구를 수행하였다.

본 논문에서는 트위터에서 수집할 수 있는 다양한 유형의 데이터를 분석하여 사이버 공격을 예측하고자 한다. 그림 2는 트위터에서 수집할 수 있는 텍스트 유형 데이터인 트윗의 내용과 사이버 공격과의 연관성을 보여주는 예시이다. 그림 2(a)는 니카라과에 대한 사이버 공격을 예고하는 실제 트윗이며, 그림 2(b)는 트윗이 예고한 날짜에 실제로 발생한 사이버 공격을 다룬 뉴스 기사이다. 이러한 예시를 통해 트윗의 내용과 사이버 공격 사이에 연관성이 있으며, 트윗 분석을 통하여 사이버 공격을 예측할 수 있음을 확인할 수 있다.

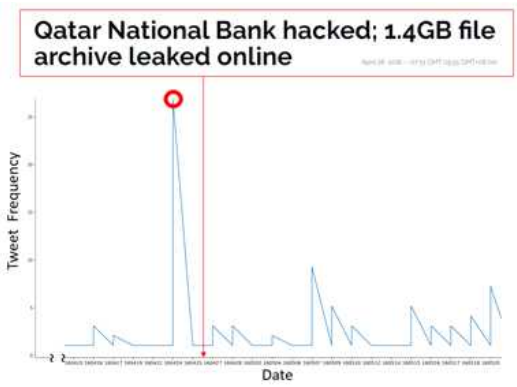
(a) 사이버 공격을 예고하는 트윗.

(b) 사이버 공격을 다룬 뉴스 기사.

[그림 2] 트윗 내용과 사이버 공격 간의 연관성을 보여주는 사례.

4 https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

그림 3은 2016년 4월 26일 카타르 국립 은행을 대상으로 발생한 사이버 공격에 대해 트윗 빈도수와 사이버 공격 간에 연관성이 있음을 보여주는 간단한 분석 결과이다. 사이버 공격과 연관된 약 50명의 사용자들이 2016년 4월 한 달 간 작성한 1,416건의 트윗 빈도수에 대하여 시계열 분석을 수행한 결과, 실제 사이버 공격이 발생한 날인 2016년 4월 26일의 전날인 4월 25일의 트윗 빈도수가 다른 날에 비하여 매우 높음을 알 수 있다. 이러한 분석 예시를 통하여 특정 집단내의 사용자가 작성한 트윗 빈도수는 사이버 공격과의 연관성을 판단할 수 있는 중요 요소임을 확인할 수 있다.



[그림 3] 트윗 빈도 수와 사이버 공격 간의 연관성 분석 예시.

본 논문에서는 트윗을 이용한 그래프와 시계열 기반 사이버 공격 예측 방법을 제안한다. 본 논문은 사이버 공격과 연관된 사용자들 간의 그래프 데이터와 시간에 따른 그들의 활동을 표현한 시계열 데이터를 동시에 활용하여 사이버 공격을 예측한 최초의 연구이다. 먼저, 사이버 공격과 연관된 키워드가 포함된 트윗을 자주 작성하는 것을 기준으로 선별된 1,000명의 사용자를 대상으로 해당 사용자들의 전체 관계 데이터(즉, 사용자들 간의 following, follower 관계)를 수집하고 사용자들 간의 관계를 그래프로 모델링한다. 이 그래프를 기반으로 그래프 상의 인접성을 고려한 사용자 간 클러스터링 효과를 검증한다. 구체적으로, 1,000명의 사용자 중 임의로 선별된 100명의 사용자들 간의 클러스터링과 그래프 상의 인접성을 고려한 기준에 따른 100명의 사용자들 간의 클러스터링 정도를 비교한다. 비교 결과, 그래프 상의 인접성을 고려한 경우, 100명의 사용자 모두 같은 클러스터로 클러스터링 되는 반면, 임의로 선별한 경우, 36명의 사용자만이 같은 그룹으로 클러스터링 됨을 보인다. 다음으로, 사이버 공격에 대한 뉴스 기사를 기반으로 트윗 사용자들의 트윗 작성 추이를 시계열로 분석한다. 트윗 사용자들의 사이버 공격 전후의 트윗 작성 빈도를 분석함으로써, 그래프 분석을 통하여 선별된 사용자 간 클러스터링 기준이 사이버 공격 예측에 효과적임을 보인다. 구체적으로, 2013-2018년 사이에 발생한 총 58건의 사이버 공격에 대한 분석을 통하여 1,000명 중 임의로 선별된 100명의 사용자 그룹이 작성한 트윗에 비하여 그래프 상의 인접성을 고려한 클러스터링 기준에 따라 선별된 100명의 사용자 그룹이 작성한 트윗이 사이버 공격 예측에 최대 약 18%만큼 더 효과적임을 보인다.

본 논문에서 트위터 데이터를 이용하여 사이버 공격을 예측하기 위한 과정은 다음과 같다.

1. 데이터 수집

먼저, 트윗으로부터 대상이 되는 트위터 사용자에 대한 관계 데이터 및 트윗 데이터를 수집한다. 수집되는 트윗 데이터의 유형에는 트윗 내의 텍스트 데이터와 사용자 간의 following/follower 데이터이다. 또한, 실제 사이버 공격에 대한 발생 날짜, 지역, 대상 등 구체적인 정보를 확인할 수 있는 뉴스 데이터를 수집한다.

2. 그래프 분석

본 논문에서는 트위터 사용자들 간의 관계망을 그래프 데이터베이스에 저장하여 관리한다. 그래프 데이터베이스로부터 그래프 상의 인접성을 기준으로 트위터 사용자들 간의 관계를 분석함으로써 사이버 공격과 연관된 사용자 그룹을 효과적으로 식별할 수 있는 기준을 선별한다.

3. 시계열 분석

그래프 분석에서 선별된 클러스터링 기준을 사용하여 시간에 따라 특정 그룹에 속한 사용자들의 활동을 시계열 분석한다.

4. 사이버 공격 예측 모델 생성

사이버 공격과 특정 그룹에 속한 사용자들의 트윗 활동 간의 연관성 분석을 위한 모델을 생성한다.

5. 예측 모델 검증

사이버 공격 별로 사이버 공격 예측 모델을 적용하고, 사이버 공격 예측의 정확도를 측정한다.

본 논문의 구성은 다음과 같다. 2장에서는 트위터 사용자 간의 관계를 기반으로 수행한 그래프 분석에 대해 설명한다. 3장에서는 트위터 사용자들의 활동을 시간에 따라 분석한 시계열 분석에 대해 설명한다. 4장에서는 관련 연구를 설명하고, 5장에서는 본 논문의 결론을 설명한다.

2. 그래프 기반 트위터 사용자들의 관계 분석

2.1 사용자 관계 데이터 수집

- **데이터 수집 방법**

Recorded Future[5]는 다양한 외부 소스로부터 보안에 위협이 될 수 있는 데이터를 수집 및 분석하여 실시간으로 위협 정보를 제공하는 종합 위협 인텔리전스 서비스를 제공한다. 본 논문에서는 Recorded Future로부터 "exploit"이라는 단어를 언급한 트윗을 작성한 전체 사용자인 1,935명을 수집한다. 그림 4에서 보는 바와 같이 "exploit"이라는 단어는 사이버 공격에 사용되는 취약점을 표현하기 위하여 자주 사용된다. 본 논문에서는 1,935명의 사용자 중 "exploit"이 포함된 트윗의 개수에 따라 정렬한 상위 1,000명의 사용자 중에서 트위터 계정 사용을 중지하였거나, 비공개로 설정한 계정을 제외한 919명의 사용자에 대해 분석한다. 그들의 관계를 분석하기 위하여 그들이 Following하는 사용자들과 그들을 Follower하는 사용자들의 트위터 ID를 수집한다. 해당 데이터는 Tweepy[6]에서 제공하는 API를 사용하여 사용자 별 제한 없이 수집이 가능하다. Following하는 사용자에 대해서는 Friends_ids(), Follower하는 사용자에 대해서는 Follower_ids()를 사용하여 수집하였다. 본 논문에서는 919명의 사용자를 *사이버 공격과 연관된 사용자*라고 정의하고, 그들의 Following 및 Follower에 해당하는 트위터 사용자들을 *사이버 공격과 잠재적으로 연관된 사용자*라고 정의한다.

[그림4] 사이버 공격에 사용된 취약점을 표현한 트윗의 예.

- **데이터 수집 결과**

사이버 공격과 연관된 사용자 중 100명과 일반 트위터 사용자 100명을 대상으로 그들의 Following 및 Follower 사용자들에 대한 데이터 수집 결과는 표 1과 같다.

5 https://recordedfuture.com/

6 https://www.tweepy.org/

[표 1] 사용자 관계 데이터 수집 결과.

사용자 구분	노드의 수	엣지의 수
사이버 공격과 연관된 사용자(100명)	195,932	388,319
일반 트위터 사용자(100명)	414,800	476,555

표 1에서 노드(Node)의 수는 각 100명의 사용자와 그들이 Following하는 또는 그들을 Follower하는 사용자의 수를 더한 숫자이며, 엣지(Edge)의 수는 사용자들 간 관계의 총 개수이다. 사이버 공격과 연관된 사용자에 대한 노드와 엣지의 비율은 약 1:1.98인 반면, 일반 트위터 사용자에 대한 노드와 엣지의 비율은 약 1:1.15에 불과하다. 이는 사이버 공격과 연관된 사용자와 사이버 공격과 잠재적으로 연관된 사용자들 간의 관계가 일반 사용자들의 관계에 비하여 매우 긴밀함을 알 수 있다.

2.2 그래프 데이터베이스

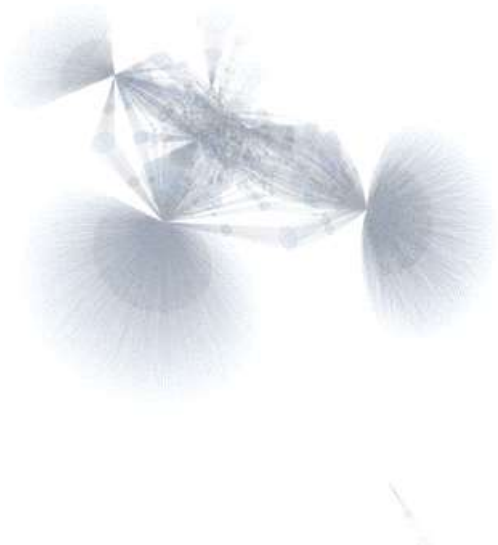
본 논문에서는 사이버 공격과 연관된 사용자와 사이버 공격과 잠재적으로 연관된 사용자 간의 관계 분석을 위한 그래프 데이터베이스로 Neo4J[7]를 사용한다. Neo4J는 그래프 데이터를 저장하고 관리하기 위한 그래프 데이터베이스로서, 노드(Node), 엣지(Edge), 속성(Property)으로 그래프를 구성한다. 이때, 속성은 노드와 엣지 각각에 정의될 수 있다. 본 논문에서 노드의 속성은 사이버 공격과 연관된 사용자와 사이버 공격과 잠재적으로 연관된 사용자를 정의하기 위하여 사용한다. 엣지의 속성은 사용하지 않음으로써 엣지는 노드 간의 관계 유무만을 표현하도록 한다.

2.3 사용자 관계 데이터 시각화

그림 5는 사이버 공격과 연관된 사용자의 데이터를 Neo4J에 저장하고 데이터를 시각화한 결과이다. 그림 5(a)는 919명의 사이버 공격과 연관된 사용자들만으로 이루어진 그래프를 시각화한 결과이며, 그림 5(b)는 사이버 공격과 연관된 사용자와 이들로부터 연결된 사이버 공격과 잠재적으로 연관된 사용자들이 포함된 총 913,760명의 사용자들 중 그래프에 추가된 순서에 따른 상위 100,000명의 관계를 그래프로 시각화한 결과이다. 사이버 공격과 연관된 사용자는 붉은색 노드, 사이버 공격과 잠재적으로 연관된 사용자는 푸른색 노드로 표현하였다[8]. 그림에서 보는 바와 같이 사이버 공격과 연관된 사용자들 간에는 긴밀한 관계를 확인할 수 있고, 이를 확장한 사이버 공격과 연관된 사용자 및 사이버 공격과 잠재적으로 연관된 사용자들 간에는 사용자들 간의 그룹이 존재함을 확인할 수 있다.

(a) 사이버 공격과 연관된 사용자들 간의 관계.

7 https://neo4j.com/



(b) 사이버 공격과 연관된 사용자 및 사이버 공격과 잠재적으로 연관된 사용자들 간의 관계.

[그림 5] Neo4J를 이용한 사용자 관계 데이터 시각화.

2.4 사용자 클러스터링 기준

본 논문은 919명의 사이버 공격과 연관된 사용자들 중 사용자 클러스터링 기준에 따른 상위 100명에 대해서 분석을 수행한다. 본 논문에서 사용한 기준은 다음과 같다.

1. Random

919명의 사용자들 중 무작위로 100명을 추출한다.

2. Betweenness Centrality (BC)

그래프 내의 임의의 모든 두 노드 s와 t간의 최소 경로의 총 개수 σ_{st}와 최소 경로에 특정 노드 N이 포함되는 개수인 $\sigma_{st}(N)$를 구하여 식 (1)의 값이 높은 순으로 정렬한다[18][22].

$$BC(N) = \sum_{s \neq N \neq t} \frac{\sigma_{st}(N)}{\sigma_{st}} \qquad (1)$$

3. Closeness Centrality (CC)

특정 노드 N과 그래프 내의 임의의 모든 노드 s간의 경로의 길이 $d(s, N)$를 구하여 식 (2)의 값이 높은 순서대로 정렬한다[19][22].

$$CC(N) = \frac{1}{\sum d(s, N)} \qquad (2)$$

4. Recorded Future Rank (RF)

Recorded Future로부터 “exploit”이라는 단어를 언급한 트윗을 작성한 전체 사용자인 1,935명을 수집하고, 이 중 “exploit”이 포함된 트윗의 개수가 높은 순서대로 정렬한다.

본 논문에서는 919명의 사용자들에 대해 위와 같이 4가지 기준을 각각 적용하여 사용자들을 정렬하고, 각 기준의 상위 100명에 해당하는 사용자들을 하나의 그룹으로 편성하여 분석을 수행한다. 4가지의 사용자 클러스터링 기준 중, Betweenness Centrality와 Closeness Centrality는 Neo4J에서 지원하는 그래프 알고리즘이다[21].

2.5 사용자 클러스터링 기준 적용 결과

각 클러스터링 기준에 따라 상위 100명에 대해 비교한 결과는 표 2와 같다. 4가지의 사용자 클러스터링 기준을 적용시켰을 때, Random, BC, CC, RF 각 그룹에 대해 하나의 기준에만 포함되는 사용자는 각각 86명, 54명, 19명, 15명이다. 또한, BC와 CC의 교집합은 80명으로서 두 기준에 의해 클러스터링된 사용자는 상당 부분 일치함을 알 수 있다.

2.6 사용자 그룹 클러스터링

본 논문에서는 각 클러스터링 기준 별 상위 100명에 해당하는 사용자들에 대해 세부 그룹에 대한 클러스터링을 수행한다. 이를 위하여 Neo4J에서 제공하는 Community Detection Algorithm 중 Label Propagation Algorithm을 사용한다[20].

[표 2] 클러스터링 기준에 따른 100명에 대한 비교 결과.

집합	사용자 수
Random − (BC ∪ CC ∪ RF)	86
BC − (Random ∪ CC ∪ RF)	54
CC − (BC ∪ Random ∪ RF)	19
RF − (BC ∪ CC ∪ Random)	15
Random ∩ RF	9
BC ∩ CC	80
BC ∩ RF	46
CC ∩ RF	42
Random ∩ BC	10
Random ∩ CC	9

[표 3] 4가지 사용자 클러스터링 기준에 따른 클러스터링 결과.

클러스터링 기준	Cluster1	Cluster2	Cluster3	Cluster1에 대한 밀집도
일반사용자	36	4	4...	36%
Random	65	1	1	65%
RF	95	1	1	95%
BC	100	−	−	100%
CC	100	−	−	100%

4가지 사용자 클러스터링 기준에 대해 Label Propagation Algorithm을 적용시킨 결과는 표 3과 같다.

각 기준에 따른 클러스터링 결과 Cluster1에 포함되는 비율은 일반 사용자의 경우 36%, Random의 경우 65%, RF는 95%, BC와 CC는 100%임을 알 수 있다. 4가지 사용자 클러스터링 기준 모두 일반 사용자들의 경우 보다 밀집도가 높음을 알 수 있으며, 그 중에서도 BC와 CC는 100%의 밀집도를 보여 다른 기준들보다 서로 연관성이 높은 사용자들로 클러스터링 시켜주는 기준임을 알 수 있다.

3. 사이버 공격에 대한 트윗 사용자 활동량의 시계열 분석

3.1 트윗 데이터 수집

- **데이터 수집 방법**

본 연구에서는 사이버 공격과 연관된 사용자들이 작성한 트윗을 수집하기 위하여 Scrapy[9]를 이용하여 트윗 데이터를 수집하였다. 사이버 공격과 연관된 사용자가 직접 작성한 트윗 중 분석 대상인 2013년에서 2018년 사이에 작성된 트윗을 수집 대상으로 지정하였다.

- **데이터 수집 결과**

919명의 사이버 공격과 연관된 사용자들이 작성한 트윗에 대해 Scarpy를 이용하여 수집한 결과는 표 4와 같다. 수집한 데이터의 크기는 약 5.4GB이며, 트윗의 총 개수는 약 2,100만 건, 사용자 별 평균 트윗 수는 약 23,000건 이다.

[표 4] 트윗 데이터 수집 결과.

사용자 수	총 트윗	사용자 별 평균 트윗	데이터 크기
919명	21,161,766개	약 23,027개	5.4GB

3.2 뉴스 데이터 수집

- **데이터 수집 방법**

본 논문에서는 구글 뉴스를 이용하여 실제로 발생한 사이버 공격에 대한 뉴스 기사를 수집하였다. 이를 위하여 “hit by cyberattack”이라는 키워드로 검색하여 2013년~2018년 사이에 발생한 사이버 공격에 대한 뉴스 기사들을 수집하였다.

- **데이터 수집 결과**

본 논문에서 수집한 뉴스 데이터의 연도 별 결과는 표 5와 같다. 2013년부터 2018년까지 수

9 https://scrapy.org/

집한 뉴스 데이터의 총 개수는 58개이며, 각 연도별로 평균 약 9.67개의 뉴스 데이터를 수집하였다. 뉴스가 작성된 날짜와 실제 사이버 공격이 발생한 날짜는 차이가 있을 수 있기 때문에, 수집한 뉴스를 기반으로 실제 사이버 공격이 발생한 날짜를 추출하였다.

[표 5] 연도별 뉴스 데이터 개수.

연도	뉴스 데이터 개수
2013	12
2014	10
2015	9
2016	8
2017	8
2018	11
총합	58

3.3 사이버 공격과 관련된 트윗 추출

본 연구에서는 사용자가 작성한 트윗 중 사이버 공격과 연관된 트윗 만을 추출하여 분석에 사용하기 위하여 Recorded Future에서 자체적으로 관리하는 사이버 공격과 관련된 키워드들의 집합인 “industrial term”을 *RF Keyword Set*으로 정의하여 사용하였다. *RF Keyword Set*에 포함된 키워드의 집합은 총 639개이며, 키워드들의 예시는 다음과 같다. “Internet-security flaw, via, PoC, Exploit, RCE, Online Security Blog, CVE, Flash, Sandworm, Shellshock, Neutrino, Samba, Stagefright, Bin”.

표 6는 2장에서 식별한 사용자 클러스터링 기준 별로 2013년부터 2018년 사이에 작성된 사용자들의 전체 트윗을 사이버 공격과의 연관성에 따라 추출한 결과를 보여준다. 이때, 트윗과 사이버 공격과의 연관성 여부는 트윗 내에 RF Keyword Set에 포함된 키워드가 포함되는지의 여부로 판단하였다.

[표 6] 사용자 클러스터링 기준 별 사이버 공격과 연관된 트윗 추출 비율.

사용자 클러스터링 기준	전체 트윗 개수	사이버 공격과 연관된 트윗 개수	사이버 공격과 연관된 트윗의 비율
Random	1,041,022	128,529	12.35%
RF	1,171,956	52,922	4.52%
BC (CC)	1,526,580	303,276	19.87%

표2에 따르면 BC와 CC에 공통으로 속한 사용자는 80명인데 본 연구에서의 분석 대상인 2013년에서 2018년 사이에는 이 80명 이외에 BC와 CC 각각에만 속한 사용자들의 활동은 감지되지 않았다. 이에 따라 분석 기간 내에 두 그룹에 속한 사용자들은 정확히 일치하였으며, 향후 두 그룹을 동일한 그룹으로 간주하여 분석을 진행한다.

표 6을 통해, BC (CC) 기준에 따라 클러스터링된 사용자들은 다른 사용자 클러스터링 기준에 따른 사용자들에 비하여 사이버 공격과 관련된 트윗을 더욱 많이 작성함을 알 수 있다.

3.4 뉴스 데이터 기반 사이버 공격 선별

- **중요도가 높은 사건**

본 논문에서는 사이버 공격의 중요도에 따른 트윗 활동 경향의 연관성을 분석하기 위하여 전체 사이버 공격 중 뉴스 데이터 기반으로 중요도 높은 사건을 선별한다. 이를 위하여 사이버 공격별 작성된 뉴스 기사의 횟수를 중요도로 정의하고, 연도 별 평균 중요도 이상의 중요도를 가지는 사이버 공격을 *중요도가 높은 사건*이라 정의한다. 중요도가 높은 사이버 공격의 예시는 그림 6과 같다. 이와 같이 중요도가 높은 사이버 공격은 다양한 매체에서 뉴스 기사를 작성하였음을 알 수 있다. 즉, 그림 6에서와 같이 특정 사이버 공격에 대해 6개의 기사가 작성되었을 경우, 해당 사이버 공격의 중요도는 6이다. 사이버 공격의 중요도를 기준으로 2013년에서 2018년 사이에 발생한 사이버 공격에 대해 선별한 결과는 표 7과 같다.

[그림 6] 중요도가 높은 사이버 공격의 예시.

[표 7] 사이버 공격을 중요도가 높은 정도를 기준으로 선별한 결과.

연도	총 사건 수	평균 중요도	중요도가 높은 사건의 수
2013	12	2.67	5
2014	10	2.6	4
2015	9	2.78	4
2016	8	2.25	2
2017	8	3.25	4
2018	11	3.45	4
총합	58	2.83	23

2013년부터 2018년 사이에 발생한 사이버 공격들 중 뉴스 데이터를 기반으로 연도별 평균 중요도보다 높은 중요도를 보인 사이버 공격의 수는 총 23건이다.

- **뉴스 기사와 인접한 사건**

본 논문에서 수집한 사이버 공격을 다룬 뉴스 기사를 분석해보면 실제 사이버 공격이 발생한 날짜와 뉴스 기사가 나온 날짜 간의 차이를 확인할 수 있다. 본 논문에서는 이 기간의 차이가 클수록 트위터에서 사용자들의 활동을 측정하는데 오차가 발생할 가능성이 크다고 가정하고, 이러한 가정을 확인하기 위하여 사이버 공격이 발생한 날짜와 해당 사이버 공격을 다룬 뉴스 기사가 작성된 날짜의 차이가 하루 이하인 사이버 공격들을 선별하여 *뉴스 기사와 인접한 사건*이라 정의하였다. 이에 따라 2013년에서 2018 사이에 발생한 사이버 공격에 대해 뉴스 기사와 인접한 사건을 선별하였으며, 그 결과는 표 8과 같다.

[표 8] 뉴스 기사와 인접한 사건을 선별한 결과.

연도	총 기사 수	뉴스 기사와 인접한 사건
2013	12	7
2014	10	7
2015	9	7
2016	8	3
2017	8	3
2018	11	3
총합	58	30

2013년부터 2018년 사이에 발생한 사이버 공격들 중 뉴스 기사와 인접한 사건의 수는 총 30건이다.

3.5 시계열 분석 기준

본 논문에서는 사이버 공격과 연관된 사용자들의 트윗 활동으로부터 실제 사이버 공격을 예측하기 위하여 “사이버 공격 시점을 기준으로 특정 기간 동안 사이버 공격과 연관된 사용자들의 평균 트윗 활동량이 그 해의 평균 활동량보다 높다”라는 기준을 사용한다. 이에 따라 사이버 공격이 발생한 날짜를 전후로 사이버 공격과 연관된 트윗 활동을 측정하는 세 가지 기준은 표 9와 같다.

그림 7(a)는 특정 사이버 공격에 대한 뉴스 기사를 보여주고, 그림 7(b)는 해당 사이버 공격이 발생한 날짜를 전후로 기준 3에 따라 특정 사용자 그룹 내의 사용자들이 작성한 날짜 별 트윗 빈도수를 표현한 그래프를 보여준다.

그림 7에서 알 수 있듯이, 사이버 공격이 발생하고 이틀 뒤인 2014년 10월 22일에 발행된 뉴스 기사를 통해 2014년 10월 20일에 사이버 공격이 발생했음을 알 수 있다. 그림 6(b)에서 검은색 선은 각 날짜 별 트윗 빈도수, 빨간색 선은 사이버 공격이 발생한 사건을 전후로 기준 별

[표 9] 사이버 공격과 연관된 트윗 활동의 측정 기준.

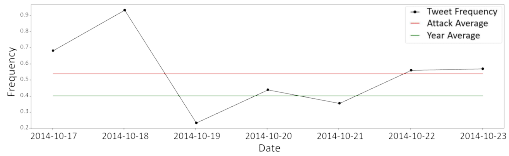
구분	기준	트윗 활동 측정 기간
기준 1	사건 전 7일 + 사건 당일 + 사건 후 7일	15일
기준 2	사건 전 5일 + 사건 당일 + 사건 후 5일	11일
기준 3	사건 전 3일 + 사건 당일 + 사건 후 3일	7일


The Washington Post

Apple's iCloud service suffers cyber-attack in China, putting passwords in peril

Oct. 22, 2014 at 5:56 a.m. GMT+9

(a) 2014년 10월 20일에 발생한 사이버 공격에 대한 뉴스 기사.



(b) 2014년 10월 20일에 발생한 사이버 공격에 대하여 날짜 별 트윗 빈도수 그래프.

[그림 7] 사이버 공격이 발생한 날짜를 전후로 날짜 별 트윗 빈도수를 표현한 그래프.

평균 트윗 빈도수, 초록색 선은 해당 사이버 공격이 발생한 연도의 평균 트윗 빈도수를 나타낸다. 또한, X축은 날짜, Y축은 정규화 과정을 거친 트윗의 빈도수를 나타낸다. 트윗 빈도수의 정규화는 식 (3)에 의해 수행된다. x는 특정 날짜의 트윗 빈도수를 의미하여, $x_{\max}$는 사이버 공격이 발생한 연도의 일일 트윗 빈도수 중 최고 빈도수, $x_{\min}$은 최저 빈도수를 의미한다.

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \qquad (3)$$

본 논문에서는 사이버 공격과 트윗 간의 연관성을 판단하는 기준을 식 (4)와 같이 정의한다. 사이버 공격이 발생한 사건을 전후로 각 기준 P와 사이버 공격과 연관된 트윗을 작성하는 사용자들을 클러스터링하는 기준 U가 주어졌을 때, 해당 기간 동안 U내의 사용자에 의해 작성된 트윗은 $Tweet(P, U)$로 정의한다. 또한, 해당 사이버 공격이 발생한 연도 Y 내에서 U 내의 사용자에 의해 작성된 트윗은 $Tweet(Y, U)$로 정의한다. 즉, 그림 7과 같은 시계열 그래프에서 빨간색 선이 초록색 선보다 위쪽에 위치할 경우, 해당 사이버 공격은 사이버 공격과 연관된 사용자들의 트윗 활동과 연관성이 있는 것으로 판단한다.

$$AVG(Tweet(Y, U)) < AVG(Tweet(P, U)) \qquad (4)$$

3.6 시계열 분석 결과

표 10은 기준1에 대하여 2장에서 식별한 사용자 클러스터링 기준에 따라 선별된 100명의 사용자 그룹에 대하여 식 (4)를 적용하여 사이버 공격과 트윗 활동 간의 연관성을 확인한 결과이다. 이때, 사용자 그룹에서 전체 트윗 중 3.3절에서 제시한 RF Keyword Set에 따라 사이버 공격과 연관된 트윗만을 분석한 경우는 "_Filter"라는 postfix를 붙여서 표현하였다. 사이버 공격과 연관된 사용자를 클러스터링하는 3개의 기준과 사이버 공격과 연관된 트윗 여부에 따른 2개의 기준을 적용한 총 6가지 기준에 대해, BC가 다른 사용자 클러스터링 기준에 비하여 높은 정확도를 보였으며, 전체적으로 사이버 공격과 연관된 트윗만 분석한 경우의 정확도가 전체 트윗을 분석한 경우에 비하여 높았다. 구체적으로, BC_Filter의 경우, 평균적으로 다른 기준에 비하여 2.54% ~ 11.41%의 정확도 향상을 보였다.

[표 10] 기준 1을 적용한 시계열 분석 (단위: %).

사용자 클러스터링 기준	전체 사건	중요도가 높은 사건	뉴스와 인접한 사건	평균
Random	39.66	34.78	43.33	39.26
Random_Filter	41.38	**47.83**	50.00	46.40
RF	43.10	30.43	50.00	41.18
RF_Filter	34.48	34.78	43.33	37.53
BC	43.10	**47.83**	40.00	43.64
BC_Filter	**50.00**	43.48	**53.33**	**48.94**

표 11과 표 12는 기준 2, 기준 3에 대하여 동일하게 사이버 공격과 트윗 활동 간의 연관성을 분석한 결과이다. 표 10에서와 유사하게 BC가 다른 기준에 비하여 높은 정확도를 보였으며, 전체적으로 사이버 공격과 연관된 트윗만 추출한 경우의 정확도가 높았다. 구체적으로, 기준 2를 사용하였을 때 BC_Filter의 경우, 평균적으로 다른 사용자 클러스터링 기준에 비하여 4.14% ~ 16.19%의 정확도 향상을 보였다. 또한, 기준 3을 사용하였을 때 BC_Filter의 경우, 평균적으로 다른 사용자 클러스터링 기준에 비하여 4.18% ~ 18.21%의 정확도 향상을 보였다.

[표 11] 기준 2를 적용한 시계열 분석 (단위: %).

사용자 클러스터링 기준	전체	중요도가 높은 사건	뉴스와 인접한 사건	평균
Random	31.03	26.90	36.67	31.26
Random_Filter	39.66	**43.48**	46.67	43.27
RF	41.38	30.43	40.00	37.27
RF_Filter	34.48	34.78	40.00	36.42
BC	43.10	**43.48**	43.33	43.31
BC_Filter	**46.55**	39.13	**56.67**	**47.45**

4. 관련 연구

[2]에서는 특정 사이버 공격인 DDoS 후에 SNS상의 일반 사용자들의 반응을 자연어 처리 과정을 통해 분석하는 연구를 수행하였다. 특정 주제별로 일반 트위터 사용자들의 트윗을 라벨링

[표 12] 기준 3을 적용한 시계열 분석 (단위: %).

사용자 클러스터링 기준	전체	중요도가 높은 사건	뉴스와 인접한 사건	평균
Random	37.93	39.13	43.33	40.13
Random_Filter	43.10	**52.17**	50.00	48.43
RF	44.83	34.78	50.00	43.20
RF_Filter	32.76	30.43	40.00	34.40
BC	41.38	47.83	36.67	41.96
BC_Filter	**50.00**	47.83	**60.00**	**52.61**

하고, 각 주제에 해당하는 트윗의 빈도수 측정을 통해 사이버 공격을 예측하였다. [3]에서는 사이버 공격과 관련된 특정 키워드를 기준으로 온라인 상에서 사이버 공격의 심각성에 대해 언급한 글을 분석하여 사이버 공격을 예측하고자 하였다. 일반 사용자들의 트윗을 바탕으로 중요 소프트웨어의 취약점을 파악하였으며, Word Embedding을 통해 각 단어에 가중치를 부여하여 트윗 단위로 심각성을 측정하여 사이버 공격을 예측하였다. [4]에서는 머신러닝 기법을 적용하여 각종 웹 사이트에 등록된 글에 대한 분석을 수행하고 이로부터 사이버 공격을 예측하였다. 이때, 날짜 별 사이버 위협에 대한 글의 포스팅 수와 날짜 별 사이버 공격 빈도 수를 이용하였다.

이와 같은 연구들은 모두 SNS 및 웹 사이트 상의 데이터를 이용하여 사이버 공격을 예측하고자 하였다. 본 논문은 다음 측면에서 기존 연구와 차이가 있다. 본 논문에서 제안하는 방법은 사이버 공격과 연관된 사용자들의 트윗 활동을 시계열 분석을 통하여 사이버 공격을 예측하고자 한다. 이때, 트윗 사용자들을 그래프 기반 관계망 분석을 통해 사이버 공격과 연관된 사용자들을 효과적으로 클러스터링할 수 있는 기준을 선별하고 이 기준을 사용하여 시계열 분석에 활용한다.

5. 결 론

본 논문에서는 트윗을 이용한 그래프와 시계열 기반 사이버 공격 예측 방법을 제안하였다. 본 논문은 사이버 공격과 연관된 사용자들 간의 그래프 데이터와 시간에 따른 그들의 활동을 표현한 시계열 데이터를 동시에 활용하여 사이버 공격을 예측한 최초의 연구이다. 먼저, 사이버 공격과 연관된 키워드가 포함된 트윗을 자주 작성하는 사용자들을 대상으로 사용자들 간의 관계를 그래프로 모델링하였다. 이를 기반으로 사용자 간 그래프 상의 인접성을 고려한 효과적인 클러스터링 기준을 선별하였다. 다음으로, 사이버 공격에 대한 뉴스 기사를 기반으로 트윗 사용자들의 트윗 활동 변화를 시계열로 분석하였다.

본 논문의 주요 분석 결과는 다음과 같다. 먼저, 사용자 클러스터링 기준에 따른 그래프 기반 사용자들 간의 관계 분석 결과, 그래프 상에서 인접성을 고려한 Betweenness Centrality와 Closeness Centrality는 분석 대상인 100명의 사용자가 모두 같은 그룹에 클러스터링 됨을 보임으로써 사이버 공격과 연관된 사용자들의 연관성을 효과적으로 판단할 수 있는 기준임을 보였다. 또한, 사용자 클러스터링 기준에 따른 시계열 분석 결과, Betweenness Centrality와 Closeness Centrality내에 속한 사용자들이 작성한 트윗은 다른 그룹에 속한 사용자들에 비하여 사이버 공격 전후에 더욱 활발하게 활동함을 관찰하였다. 구체적으로, Betweenness Centrality와 Closeness Centrality내의 속한 사용자들이 작성한 트윗은 다른 그룹에 속한 사용자들이 작성한 트윗에 비하여 사이버 공격 예측에 최대 약 18%만큼 더 효과적임을 보였다.

6. 참고 문헌

[1] C. Sabottke, O. Suciu, T. Dumitras, “Vulnerability disclosure in the age of social media: Exploiting twitter for predicting real-world exploits,” USENIX Security, Vol. 15, 2015.

[2] N. Chambers, B. Fry, J. Mao, “Detecting denial-of-service attacks from social media Text: Applying NLP to computer security,” Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics Hum. Lang. Technol., pp. 1626-1635, Vol. 1, 2018.

[3] Shi Zong, Alan Ritter, Graham Mueller, and Evan Wright, “Analyzing the perceived severity of cybersecurity threats reported on social media,” arXiv preprint arXiv:1902.10680, 2019.

[4] P. Goyal, K. T. Hossain, A. Deb, N. Tavabi, N. Bartley, A. Abeliuk, E. Ferrara, K. Lerman, “Discovering Signals from Web Sources to Predict Cyber Attacks,” eprint arXiv:1806.03342, 2018.

[5] Ashok Deb, Kristina Lerman, and Emilio Ferrara. “Predicting Cyber-Events by Leveraging Hacker Sentiment,” Information, Vol. 9, No. 11, 2018.

[6] P. Filonov, F. Kitashov, A. Lavrentyev, “RNN-based early cyber-attack detection for the tennessee eastman process,” Proceedings of ICML Time Series Workshop, Aug. 2017.

[7] Yang Liu , Jing Zhang , Armin Sarabi , Mingyan Liu , Manish Karir , Michael Bailey, “Predicting Cyber Security Incidents Using Feature-Based Characterization of Network-Level Malicious Activities,” Proceedings of the 2015 ACM International Workshop on International Workshop on Security and Privacy Analytics, March 04-04, 2015, San Antonio, Texas, USA.

[8] Cipriano, C. et al., “NEXAT: History-based approach to predict attacker actions,” Proc. ACSAC, Orlando, FL, December 2011.

[9] B Munkhdorj and S. Yuji, “Cyber attack prediction using social data analysis,” Journal of High Speed Networks, Vol. 23, No. 2, pp. 109-135, 2017.

[10] Watters, P.A., McCombie, S., Layton, R., Pieprzyk, J., “Characterising and Predicting Cyber Attacks Using the Cyber Attacker Model Profile (CAMP),” Journal of Money Laundering Control, Vol. 15, No. 4, 2012.

[11] R. P. Khandpur, T. Ji, S. Jan, G. Wang, C.-T. Lu, N. Ramakrishnan, “Crowdsourcing cybersecurity: Cyber attack detection using social media,” Proceedings of the 2017 ACM on Conference on Information and Knowledge Management ser. CIKM ’17, pp. 1049-1057, 2017.

[12] Ahmet Okutan, Shanchieh Jay Yang, Katie McConky, “Forecasting cyber attacks with imbalanced data sets and different time granularities”, CoRR, vol. abs/1803.09560, 2018.

[13] Signorini A, Segre AM, Polgreen PM, “The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic,” PLoS One, Vol. 6, No. 5, e19467, May 2011.

[14] S Dey, “Stock Market Prediction Using Twitter Mood,” International Journal of Scientific & Engineering Research, Vol. 5, No. 5, p. 44 - 47, 2014.

[15] Bahrami, M., Findik, Y., Bozkaya, B., & Balcisoy, S, “Twitter Reveals: Using Twitter Analytics to Predict Public Protests,” arXiv preprint arXiv:1805.00358, 2018.

[16] Kumar, Shamanth, Fred Morstatter, and Huan Liu, “Crawling Twitter Data,” Twitter Data Analytics. Springer, New York, NY, 2014. 5-22.

[17] 강범일, 이재윤, “트위터 관련 연구에 대한 계량정보 학적 분석,” 정보관리학회지, Vol. 31, No. 3, pp. 293-311, 2014.

[18] https://neo4j.com/docs/graph-algorithms/current/labs-algorithms/betweenness-centrality/

[19] https://neo4j.com/docs/graph-algorithms/current/labs-algorithms/closeness-centrality/

[20] https://neo4j.com/docs/graph-algorithms/current/algorithms/label-propagation/

[21] https://neo4j.com/docs/graph-algorithms/3.5/

[22] Brandes, Ulrik, “A faster algorithm for betweenness centrality,” Journal of Mathematical Sociology, Vol. 25, No. 2, pp. 163-177, 2001.

이 준 하

2014년~현재 서울과학기술대학교 글로벌융합산업공학과 ITM 전공 학사과정

관심 분야 : 데이터베이스, 분산처리시스템, 데이터 분석

한 보 영

2015년~현재 서울과학기술대학교 글로벌융합산업공학과 ITM 전공 학사과정

관심 분야 : 데이터베이스, 정보보안, 알고리즘

권 혁 윤

2018년~현재 서울과학기술대학교 글로벌융합산업공학과 ITM 전공 조교수

2014년~2018년 국방부 연구원

2013년~2014년 KAIST 첨단정보기술연구센터 박사후연구원

2011년~2012년 Microsoft Research in Asia(Beijing) Research Intern

2013년 KAIST 전산학 박사

관심 분야 : 데이터베이스, 빅데이터 관리, 데이터 사이언스, 분산 처리 시스템