

논문 2015-52-11-6

SNS에서 콘텐츠 오염자 탐지를 위한 개선된 특징 추출 방법

(Improved Feature Extraction Method for the Contents Polluter
Detection in Social Networking Service)

한진섭*, 박병준*

(Jin Seop Han and Byung Joon Park[©])

요약

인터넷의 발달과 스마트폰 등과 같은 휴대기기 보급의 확산으로 트위터, 페이스북과 같은 SNS 사용자의 수가 증가하고 있다. 그리고 이와 함께 상품 광고, 비방 및 성인 콘텐츠 등을 게재함으로써 SNS를 오염시키는 콘텐츠 오염 문제 또한 점차 커지고 있다. 따라서 본 논문은 SNS에서의 콘텐츠 오염자를 탐지하기 위한 개선된 콘텐츠 오염자의 특징 추출 방법을 제안한다. 특히, 본 논문은 콘텐츠 오염자의 예측 및 분류 단계에서 새로운 사용자 데이터의 특징 값을 효율적으로 추출하기 위하여 전체 데이터를 대상으로 하는 일괄 처리 방식이 아니라 데이터 증가분만을 고려하는 점진적 접근 방법에 기초한 콘텐츠 오염자 특징 추출 방법을 제안한다. 그리고 제안한 방법이 일괄 처리한 방법과 비교해서 분류 정확도는 유지되고 시간 효율성이 향상되는 것을 실험을 통해 비교 평가한다.

Abstract

The number of users of SNS such as Twitter and Facebook increases due to the development of internet and the spread of supply of mobile devices such as smart phone. Moreover, there are also an increasing number of content pollution problems that pollute SNS by posting a product advertisement, defamatory comment and adult contents, and so on. This paper proposes an improved method of extracting the feature of content polluter for detecting a content polluter in SNS. In particular, this paper presents a method of extracting the feature of content polluter on the basis of incremental approach that considers only increment in data, not batch processing system of entire data in order to efficiently extract the feature value of new user data at the stage of predicting and classifying a content polluter. And it comparatively assesses whether the proposed method maintains classification accuracy and improves time efficiency in comparison with batch processing method through experiment.

Keywords : SNS, 콘텐츠 오염자 탐지, 점진적 처리, 특징 추출

I. 서론

SNS(Social Networking Service)는 가깝게 아는 이

* 정회원, 광운대학교 컴퓨터학과
(Department of Computer Science, Kwangwoon University)

© Corresponding Author(E-mail: bjpark@kw.ac.kr)

※ 이 논문은 2014년도 광운대학교 교내학술연구비 지원에 의해 연구되었음

Received ; August 17, 2015 Revised ; October 12, 2015

Accepted ; November 2, 2015

는 물론 서로 알지 못하는 임의의 사람들과 관계를 맺고 정보를 공유하는 인터넷 서비스이다. SNS는 간단한 가입 절차를 거쳐서 빠르게 계정을 생성하고 손쉽게 다수 사용자와 관계를 맺어서 메시지를 공유할 수 있기 때문에, 악의적 목적을 가지고 관계를 요청하고 메시지를 게재함으로써 SNS를 오염시키는 스팸 행위가 빈번하게 일어날 수 있다. 이러한 콘텐츠 오염자 탐지를 위해서 대부분의 연구들은 인공지능 기계 학습의 지도 학습(Supervised Learning) 방법을 활용한다. 지도 학습을 이용한 SNS 콘텐츠 오염자 탐지 방법은 훈련 데이

터로부터 오염자와 비 오염자를 구분할 수 있는 특징(Feature)을 식별하고 그 값을 추출하여 분류기 모델을 생성하고, 학습한 분류기 모델을 통해서 콘텐츠 오염자를 분류 예측하는 것이다. 이 때 오염자와 비 오염자를 효과적으로 구분할 수 있는 특징을 식별하고 그 값을 추출하는 것은 분류 성능에 중요한 영향을 미친다. 특히 분류 예측 단계에서 임의의 사용자의 특징 값을 추출하기 위하여 매번 전체 데이터를 대상으로 하는 것은 비효율적인 처리 방식이다. 따라서 본 논문은 SNS에서 콘텐츠 오염자 탐지에 필요한 특징 값을 추출하는데 있어서 데이터의 증가분만을 고려하여 특징을 추출하여 분류 예측하는 점진적 특징 추출 방법을 제안한다.

본 논문의 구성은 II장에서 점진적 마이닝과 기존 SNS 콘텐츠 오염자 탐지 관련 연구에 대해 기술하고, III장에서 제안한 점진적 특징 추출 방법을 기술한다. 그리고 IV장에서 제안한 방법을 적용하여 수행한 실험 결과를 기술하고, V장에서 결론을 맺는다.

II. 관련 연구

데이터 마이닝은 대부분의 경우 대용량의 데이터를 처리하며 패턴을 찾기 위하여 매번 전체 데이터를 대상으로 마이닝을 한다는 것은 비효율적인 방법이다. 점진적 마이닝은 순차적 패턴을 찾는 연관 규칙(Association Rule) 마이닝에 주로 적용된 방법^[4-5, 7]이며, 데이터의 패턴 및 규칙을 생성하는데 있어서 데이터 전체를 대상으로 하지 않고 데이터 증가분만을 대상으로 마이닝 하는 방법이다^[11]. 점진적 마이닝 방법은 동적인 SNS 네트워크에서 임의의 노드의 중심도를 측정하는 매개 중심성(Betweenness Centrality)^[8]과 근접 중심성(Closeness Centrality)^[9] 연구 및 같은 성향의 노드 집단을 발견하는 연구^[12]에 적용된 사례가 있다.

일반 웹의 스파머 탐지^[20] 및 SNS 보안^[21] 연구와 함께 SNS 콘텐츠 오염자 탐지 연구 또한 다양한 방법으로 수행되었다. 먼저 LDA(Latent Dirichlet Allocation) 알고리즘을 이용하여 트위터 메시지의 신뢰도를 측정하여 콘텐츠 오염자를 탐지하는 연구^[1, 3]가 있으며 페이스북에서 부당하게 획득한 페이지 'Like'를 탐지하는 CopyCatch 시스템^[6]이 제안되었다. 친구 관계를 그래프 구조로 표현하여 지역 군집 계수(Local clustering coefficient) 및 상호 링크 비율 등의 특징을 식별한 연

구^[22]가 있다.

그리고 메시지, 사용자 및 사용자 간의 관계를 고려하여 스파머 탐지 프레임워크를 제안하는 연구^[16]가 있고, 페이스북에서 스팸 메시지를 탐지하는 Defensios 시스템이 제안되었다^[19]. 이것은 악의적 메시지와 URL을 중심으로 탐지한다. 그리고 트위터에서 단어와 URL을 고려하여 논리회귀 분석 기법 기반의 스팸 탐지하는 시스템이 제안되었다^[17]. 또한 트위터에서 친구 수, 친구 요청 수 및 클러스터링 기법을 통해 메시지들 간의 연관구조를 분석해서 정상 메시지와 구별되는 특성을 식별하는 방법^[13]이 제안 되었다. 관계가 없는 사용자에게로의 메시지 전송 횟수 등의 특징을 분석한 스팸 분류 방법^[14]이 있으며 기존에 비해 높은 정확도 성능을 보이거나 상대적으로 적은 수의 데이터를 이용하여 대용량의 데이터를 사용했을 때 그 성능과 효율성을 보장하기 어렵다. 마지막으로 소셜 허니팟은 비교적 많은 트위터 사용자의 특징을 제안하고 기존 대비 많은 데이터를 이용하여 콘텐츠 오염자 탐지 실험을 수행하여 높은 분류 정확도(98.42%)의 분류기 모델을 제안한다^[18]. 표 1과 같이 기존 연구들의 특징, 실험 데이터 및 학습 알고리즘을 비교하였다.

본 논문은 기존 대비 높은 성능을 보인 소셜 허니팟을 참고하여 제안한 방법을 실험하였으며 소셜 허니팟의 자세한 설명은 다음과 같다.

소셜 허니팟은 학습에 필요한 훈련 데이터를 수집하기 위하여 다수의 소프트웨어 에이전트 계정(허니팟)을

표 1. 기존 SNS 오염자 탐지 연구
Table 1. Existing work on SNS polluter detection.

저자	특징 (Feature)	실험 데이터	학습 알고리즘
Ito 외 ^[2]	LDA 신뢰도	2000개 트위터 메시지	RandomForest
Chao Yang 외 ^[22]	그래프 구조 특징을 비롯한 24개 특징	5000개 비 스팸 메시지, 500개 스팸	RandomForest 외
Hongyu Gao 외 ^[13]	메시지 게재 시간 간격 등 6개 특징	217,802개 페이스북 메시지, 467,390개 트위터 메시지	SVM
Maarten Bosma 외 ^[14]	메시지 간 관계성, 지자, 사용자 스팸 기록 등	28,998개 스팸 메시지, 13,188개 비 스팸 메시지, 9,491명 사용자	Naive Bayes
Jonghyuk Song 외 ^[16]	사용자 간의 관계 중심의 특징 식별 기반, 총 11개 특징	-55000개 메시지 -1100명 사용자	BayesNet 외
Lee, K. 외 ^[18]	계정의 길이 등 총 19개 특징	-5,643,297개 메시지 -41,499명 사용자	RandomForest 외

인터넷에 배치하고, 친구 요청을 시도한 계정을 수집한다. 소셜 허니팟 환경은 허니팟 계정 간에 한해서 서로 메시지 송수신 및 친구 요청이 가능하기 때문에, 허니팟 계정에 친구 요청 한 비 허니팟 계정을 콘텐츠 오염자 및 스팸으로 간주한다. 그리고 유도한 계정들의 사용자 프로필, 사용자 사이의 관계, 메시지, 그리고 사용자 이력 정보 등을 분석하여 총 19개 특징을 식별하였다. 식별한 특징은 다음과 같다:

- **프로필 특징:** 계정 길이, 소개 글 길이, 계정 사용 기간
- **관계 특징:** 친구요청(Following) 수, 친구(Follower) 수, 친구요청 수와 친구 수의 비율, 쌍방향 친구 비율, 쌍방향 친구요청 비율, 친구 계정(Numeric ID)의 표준편차, 친구요청 계정(Numeric ID)의 표준편차
- **메시지 특징:** 메시지 수, 하루 평균 메시지 수, 메시지 별 URL 수, 메시지 별 고유(중복 제외) URL 수, 메시지 별 @ 수, 메시지 별 고유(중복 제외) @ 개수, 메시지 유사도(Cosine Similarity) 평균, 메시지 압축률
- **사용자 이력 특징:** 일정 기간 일별 친구요청 수의 평균 변화율

위와 같은 소셜 허니팟을 비롯한 기존 연구는 분류 예측 단계에서 임의의 사용자의 특징 값을 추출하기 위하여 전체 데이터를 대상으로 추출하는 일괄 처리 방식이다. 본 논문은 매번 전체 데이터를 일괄 처리하지 않고, 새롭게 축적된 데이터만을 이용하여 특징 값을 추출하여 효율적으로 콘텐츠 오염자를 분류할 수 있는 점진적 특징 추출 방법을 제안한다.

III. 점진적 특징 추출

본 논문은 그림 1과 같이 점진적 마이닝 접근 방법에 기초하여 분류 예측 단계에서 특징 값을 추출하여 콘텐츠 오염자를 분류 및 탐지하는 방법을 제안한다.

임의의 사용자 데이터는 시간이 지남에 따라 누적된다. 따라서 이전 추출한 특징 값 및 지식과 새로 추가된 데이터의 증가분만을 고려하는 점진적 처리를 통해서 특징 값을 추출한 후, 학습 단계에서 기 생성한 분류기

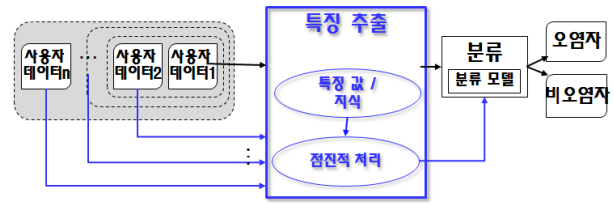


그림 1. 점진적 특징 추출

Fig. 1. Incremental feature extraction.

모델을 통하여 오염자와 비 오염자를 분류하는 것이다. 본 논문은 먼저 기존 소셜 허니팟의 19개 특징과 본 논문의 이전 연구에서 제안한 특징^[2, 10]을 분석하여 표 2와 같이 비 점진적 처리와 점진적 처리를 분류한다. 분류 기준은 시간이 지남에 따라 누적되는 데이터로부터 특징 값을 계산해야 하는 경우와 기존 데이터 추출 과정에서 특징 값 및 지식 정보 활용이 가능한 특징의 경우 점진적 처리로 분류하였다.

비 점진적 처리로 분류한 특징들 대부분은 트위터 API의 실행을 통해서 값을 얻었을 때 마이닝에 적합한 특징 값을 추출할 수 있다. 또한 ‘메시지 압축률’은 원본 메시지 크기와 압축한 메시지 크기의 비율을 의미하며, 전체 데이터를 대상으로 특징 값을 계산했을 때, 정확한 값을 추출할 수 있기 때문에 비 점진적 처리로 분류하는 것이 효과적이다.

점진적 처리로 분류한 특징은 대부분 메시지와 사용자 이력 관련 특징이다. 본 논문은 이 특징 값들을 점진적으로 추출하기 위하여 다음과 같이 수식을 정의하여 점진적으로 처리한다.

표 2. 비 점진적 처리와 점진적 처리

Table 2. Non-incremental process and incremental process.

비 점진적(13개)	점진적 처리(12개)
<ul style="list-style-type: none"> • 계정 길이, 소개 글 길이, 계정 사용 기간 • 친구요청 수, 친구 수, 친구요청 수와 친구 수의 비율 • 쌍방향 친구 비율, 쌍방향 친구요청 비율 • 친구 계정의 표준편차, 친구요청 계정의 표준편차 • 메시지 수, 하루 평균 메시지 수 • 메시지 압축률 	<ul style="list-style-type: none"> • 메시지 별 URL 수 및 고유 URL 수 • 메시지 별 @ 수 및 고유 @ 수 • 메시지 별 RT, 해시태그, OH 및 HT 수 • 메시지 유사도 평균 • 일정 기간 일별 친구요청 수의 평균 변화율 • 일정 기간 일별 메시지 수의 평균 변화율 및 표준편차

• 메시지 별 URL 수

$$URL(x,y) = \frac{\text{기존 값} \times \text{기존메시지수} + x}{y} \quad (1)$$

수식 (1)은 메시지에서 URL이 포함된 비율 특징 값을 점진적으로 처리하는 수식이며 비 오염자에 비해서 오염자 메시지에서 더 높은 값이 나타나는 경향이 있다. 기존 메시지 데이터의 스캔 없이 이전 계산한 URL 비율(기존 값)에 기존 메시지 수를 곱하여 기존 URL 수를 구하고 여기에 추가된 데이터의 URL 수(x)를 더한 값을 전체 메시지 수(y)로 나누어 줌으로써 점진적으로 특징 값을 추출한다. 메시지 별 URL 수 외 3개(고유 URL 수, 메시지 별 @ 수 및 고유 @ 수) 특징 역시 같은 방식으로 수식을 정의하여 처리한다. 본 논문은 이전 연구[2]에 더해서 총 4개의 특징(메시지 별 RT, 해시태그, OH 및 HT 수)을 추가하여 같은 방식으로 수식을 정의하여 실험하였다.

• 메시지 유사도 평균

$$MS(x,y) = \frac{\text{기존 값} \times \text{기존메시지수} C_2 + x}{y C_2} \quad (2)$$

수식 (2)는 사용자가 게재한 메시지에서 모든 메시지 쌍의 유사도 평균이며 오염자 메시지에서 더 큰 값이 나타나는 경향이 있다. 소셜 허니팟 연구[18]의 수식을 개선하여 기존 유사도 평균 값(기존 값)에 기존 메시지 수의 조합을 곱하여 기 유사도의 합을 구하고 여기에 추가 유사도의 합(x)을 더한 후 총 메시지 수(y)의 조합으로 나누어서 점진적 처리가 가능하도록 정의한다.

• 친구요청 수의 평균 변화율

$$Followig(x) = \sqrt{\frac{\text{기존 값}^2 \times (\text{기간}-2) + |\text{전날요청수} - x|}{\text{기간}-1}} \quad (3)$$

수식 (3)은 일정 기간에 요청한 친구요청 수의 평균 변화율을 계산한 수식이며, 오염자 메시지에서 더 큰 값이 나타나는 경향이 있다. 소셜 허니팟의 수식을 고려해서 기존 변화율(기존 값)에 기간, 금일 요청 수(x) 및 전날 요청 수를 적용하여 점진적 처리가 가능하도록 정의한다.

• 메시지 수의 평균 변화율

$$ACNM(x) = \frac{\text{기존 값} \times (\text{기간}-2) + |\text{전날메시지수} - x|}{\text{기간}-1} \quad (4)$$

‘메시지 수의 평균 변화율’은 본 논문의 이전 연구[10]에서 제안한 특징이며 일정 기간에 사용자가 게재한 메시지 수의 평균 변화율을 계산한다. 오염자 메시지에서 더 큰 값이 나타나는 경향이 있다. 수식 (4)과 같이 기존 변화율(기존 값)에 기간, 금일 메시지 수(x) 및 전날 메시지 수를 적용하여 점진적 처리가 가능하도록 정의한다.

• 메시지 수의 표준편차

$$SDNM(x) = \sqrt{\frac{1}{\text{기간}} ((\text{기간}-1)(\text{기존 값}^2 + (\frac{\text{기메시지합}}{\text{기간}-1})^2 + x^2) - (\frac{\text{기메시지합} + x}{\text{기간}})^2)} \quad (5)$$

‘메시지 수의 표준편차’는 일정 기간에 사용자가 게재한 메시지 수의 표준편차를 의미한다. 수식 (5)과 같이 기간, 기존 표준편차 값(기존 값), 기존 메시지 수 합 및 금일 메시지 수(x)를 적용하여 점진적 처리가 가능하도록 정의한다. 이 수식을 유도한 과정은 그림 2와 같다.

(Msg는 일일 메시지 수)

$$\text{기존 값} = \sqrt{\frac{1}{\text{기간}-1} (\sum_{i=1}^{n-1} \text{Msg}_i^2) - (\frac{\text{기존메시지수합}}{\text{기간}-1})^2}$$

$$\sum_{i=1}^{n-1} \text{Msg}_i^2 = (\text{기간}-1)(\text{기존 값}^2 + (\frac{\text{기존메시지수합}}{\text{기간}-1})^2)$$

$$\text{구하려는 값} = \sqrt{\frac{1}{\text{기간}} (\sum_{i=1}^n \text{Msg}_i^2) - (\frac{\text{기존메시지수합} + \text{금일메시지수}}{\text{기간}})^2}$$

최종 유도식

$$\sqrt{\frac{1}{\text{기간}} ((\text{기간}-1)(\text{기존 값}^2 + (\frac{\text{기존메시지수합}}{\text{기간}-1})^2) + \text{금일메시지수}^2) - (\frac{\text{기존메시지수합} + \text{금일메시지수}}{\text{기간}})^2}$$

그림 2. 메시지 게재 수의 표준 편 수식 유도
Fig. 2. Derivation formula for the standard deviation of the number of messages.

위와 같이, 본 논문은 점진적 처리가 가능한 특징을 분석하고 이를 처리하기 위한 수식을 정의하여 특징 값을 추출함으로써 기존 보다 효율적인 콘텐츠 오염자 탐지 방법을 제안한다.

다음 장에서 제안한 방법을 검증하기 위하여 수행한 실험 결과를 기술한다.

IV. 실험 결과

이번 장은 제안한 방법을 적용한 실험 및 평가를 기

술한다. 실험 및 평가에 앞서서 먼저 실험 데이터의 구성에 대해서 기술한다.

1. 실험 데이터

제안한 점진적 특징 추출 실험은 기존 소셜 허니팟과 동일한 데이터^[18]를 사용한다. 소셜 허니팟은 트위터에서의 정상 사용자 19,276명(46.4%)과 콘텐츠 오염자 22,223명(53.6%) 그리고, 메시지 수는 각각 3,263,238개와 2,380,059개로 구성되어 있다.

위와 같은 소셜 허니팟 데이터를 가지고 제안하는 방법을 실험하기 위하여 원본 데이터를 재구성한다. 정상 사용자와 콘텐츠 오염자가 동일하게 수집된 기간인 2009년 10월 31일부터 2009년 11월 29일(30일) 데이터를 일별로 재구성하여 점진적 특징 추출 실험을 수행한다. 결과적으로 30일 동안에 2,523,219개의 정상 사용자 메시지와 78,847개의 콘텐츠 오염자 메시지가 본 논문의 실험에 사용되었다.

2. 실험 및 평가

실험 내용은 기존 일괄 처리와 III장에서 제안한 점진적 처리 방법을 적용하여 특징 값을 추출했을 때, 두 방법이 동일한 정확도 결과가 나오면서 점진적 처리가 시간 효율성 측면에서 얼마나 향상되는지 비교 평가한다. 실험은 윈도우즈7 운영체제에서 C언어를 이용하여 구현하였으며 시스템 사양은 Intel(R) Core(TM) i7-2600k 3.40 GHz CPU 및 8GB RAM 이다. 각각의 방법으로 분류기 모델 생성 후, 정확도 성능 측정은 Weka toolkit 버전 3.7.6^[15]을 사용하였다. 점진적 처리로 분류한 특징을 일괄 처리와 점진적 처리를 하여 정확도를 측정할 결과, 99.427%로 같은 결과가 나온다. 그리고 30일 기간 동안 일별로 재구성한 실험 데이터를 가지고 점진적 처리로 분류한 특징들을 일괄 처리 및 점진적 처리로 특징 값을 추출했을 때, 추출 시간을 측정하였다. 표 3과 같이 30일 간 평균 시간이 점진적 처리가 일괄 처리에 비해서 평균적으로 시간 단축이 된 것을 알 수 있다.

‘메시지 유사도 평균’은 그 특징 값을 추출하는데 있어서 많은 시간이 소요되는 것으로써 점진적 처리로 하였을 때, 약 18분 30초가 단축되어 특히 많은 시간 단축이 된 것을 알 수 있다. 그림 3 및 그림 4는 이전 연구 [2]에 더해서 점진적 처리로 분류한 메시지 수 표준편차와 메시지 별 RT, 해시태그, OH 및 HT 수 특징들의

표 3. 30일간 평균 특징 추출 시간

Table 3. Average time of the feature extraction over 30 days.

특징	처리 방식	일괄 처리 시간 (분:초)	점진적 처리 시간 (분:초)
메시지 별 URL 수와 고유 URL 수, 메시지 별 @ 수와 고유 @ 수		00:1.777	00:0.192
메시지 유사도 평균		19:52.688	01:32.760
친구 요청 수의 평균 변화율		00:0.077	00:0.035
메시지 수의 평균 변화율		00:1.225	00:0.109
메시지 수 표준편차		00:1.228	00:0.088
메시지 별 RT, 해시태그, OH 및 HT 수		00:1.699	00:0.252

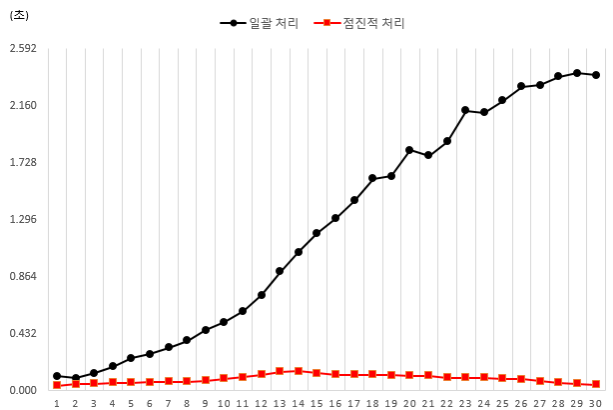


그림 3. 메시지 수 표준편차의 추출 시간

Fig. 3. Extraction time of the standard deviation of the number of messages

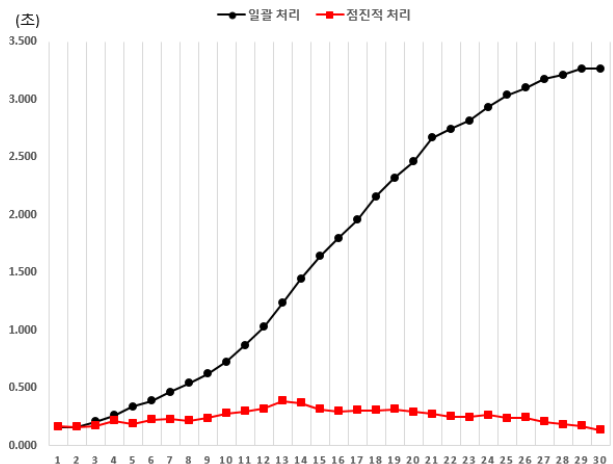


그림 4. 메시지 별 RT, 해시태그, OH 및 HT 수의 추출 시간

Fig. 4. Extraction time of the number of RT, Hashtag, OH and HT in a tweet.

일괄 처리 및 점진적 처리 추출 시간을 비교한 것이다.

그림과 같이, 매일 데이터는 축적되어 증가하며 매번 전체 데이터를 대상으로 하는 일괄 처리는 시간이 지남에 따라서 특징 값 추출 시간이 증가하는 것을 알 수 있다. 반면에 점진적 처리는 일별 증가분의 데이터를 가지고 특징 값을 추출하기 때문에 시간이 지남에 따라서 그 소요 시간의 변화가 작은 것을 알 수 있다.

V. 결 론

본 논문은 SNS에서 콘텐츠 오염자를 효율적으로 탐지하기 위한 점진적 특징 추출 방법을 제안하였다. 특히 분류 및 탐지 단계에서 임의의 사용자 데이터의 특징 값을 추출하는데 있어서 전체 데이터를 대상으로 일괄 처리하는 것이 아니라, 데이터의 증가분만을 고려하여 특징 값을 추출함으로써 점진적 처리 방법에 기초한 방법을 제안한다. 따라서 점진적 처리가 가능한 특징을 일정 기준에 따라 식별하고, 이 특징 값을 추출하기 위한 수식을 정의한다. 그리고 이 수식을 기반으로 일정 기간 동안 일괄 처리와 제안한 점진적 특징 추출 실험을 수행한 결과, 제안한 방법이 일괄 처리한 방법과 비교해서 분류 정확도 성능은 그대로 유지되고 시간 효율성은 향상되는 것을 검증한다.

REFERENCES

- [1] Ito, J., Song, J., Toda, H., Koike, Y., and Oyama, S., "Assessment of Tweet Credibility with LDA Features", Proc. of the 24th International Conference on World Wide Web, pp.953-958, 2015.
- [2] Jin Seop Han, Byung Joon Park, "Incremental Spammer Feature Extraction for the Spam Detection in Social Networking Service", *International Journal of Applied Engineering Research*, Vol. 10, No. 18, pp.39269-39273, 2015.
- [3] Tsoimon, Bayar, and Kyung-Soon Lee, "An event extraction model based on timeline and user analysis in latent dirichlet allocation", Proc. of the 37th international ACM SIGIR conference on Research & development in information retrieval. ACM, pp.1187-1190, 2014.
- [4] Sreedevi, M., Kumar, and G. V., "Parallel and Distributed Approach for Mining Closed Regular Patterns on Incremental Databases at User Thresholds", Proc. of the 2014 International Conference on Information and Communication Technology for Competitive Strategies, ACM, pp. 59-63, 2014.
- [5] Chen, Y. C., Weng, J. T. Y., Wang, J. Z., Chou, C. L., Huang, J. L., and Lee, S. Y., "Incrementally Mining Temporal Patterns in Interval-based Databases", *Data Science and Advanced Analytics (DSAA)*, pp. 304-311, 2014.
- [6] Beutel, A., Xu, W., Guruswami, V., Palow, C., and Faloutsos, C., "CopyCatch: stopping group attacks by spotting lockstep behavior in social networks", Proc. of the 22nd international conference on World Wide Web, pp.119-130, 2013.
- [7] Mehta, Gunjan, Deepa Sharma, and Ekta Chauhan, "Application of Incremental Mining and Apriori Algorithm on Library Transactional Database", *International Journal of Computer Applications*, pp. 73-78, 2013.
- [8] Miray Kas, Matthew Wachs, Kathleen M. Carley, and L. Richard Carley, "Incremental Algorithm for Updating Betweenness Centrality in Dynamically Growing Networks", Proc. of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ACM, pp. 33-40, 2013.
- [9] Miray Kas, Kathleen M. Carley, L. and Richard Carley, "Incremental Closeness Centrality for Dynamically Changing Social Networks", *Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 1250-1258, 2013.
- [10] Jin Seop Han, Byung Joon Park, "Efficient Detection of Content Polluters in Social Networks", *IT Convergence and Security 2012(LNEE)*, pp. 991-996, 2013.
- [11] Shah, Siddharth, N. C. Chauhan, and S. D. Bhandar, "Incremental Mining of Association Rules: A Survey", *International Journal of Computer Science and Information Technologies*, Vol. 3, no. 3, pp. 4071-4074, 2012.
- [12] Papadopoulos, S., Kompatsiaris, Y., Vakali, A., and Spyridonos, P., "Community detection in Social Media Performance and application considerations", *Data Mining and Knowledge Discovery* 24.(3), pp. 515-554, 2012.
- [13] Hongyu Gao, Yan Chen, Kathy Lee, Diana Palsetia, and Alok Choudhary, "Towards Online Spam Filtering in Social Networks", Proc. of 19th

- Network Distributed System Security (NDSS) Symposium, <http://www.internetsociety.org>, 2012.
- [14] Maarten Bosma, Edgar Meij, and Wouter Weerkamp, W., "A Framework for Unsupervised Spam Detection in Social Networking Sites", Proc. of European Conference on In-formation Retrieval (ECIR), pp. 364-375, 2012.
- [15] Eibe Frank, Available from <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>, 2012.
- [16] Jonghyuk Song, Sangho Lee and Jong Kim, "Spam filtering in twitter using sender-receiver relationship", Proc. of the 14th International Symposium on Recent Advances in Intrusion Detection (RAID), pp. 301-317, 2011.
- [17] KristoFer Beck, "Analyzing Tweets to Identify Malicious Messages", Proc. of Electro/Information Technology (EIT) IEEE International Conference, pp. 1-5, 2011.
- [18] Kyumin Lee, Brian David Eoff, and James Caverlee, "Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter", Proc. of International AAAI Conference on Weblogs and Social Media (ICWSM), pp. 185-192, 2011.
- [19] Saeed Abu-Nimeh, Thomas M. Chen, and Omar Alzubi, "Malicious and Spam Posts in Online Social Networks", *IEEE Computer Society*, Vol. 9, pp. 23-28, 2011.
- [20] Muhammad Atif Qureshi, Tae-Seob Yun, Jeong-Hoon Lee, Kyu-Young Whang, "Improving the Quality of Web Spam Filtering by Using Seed Refinement", *Journal of IEIE*, CI, Vol. 48, no. 6, pp. 123~139, 2011.11.
- [21] Seung-Hyun Seo, Taenam Cho, "Group Key Management Protocol for Secure Social Network Service", *Journal of IEIE*, CI, Vol. 48, no. 3, pp. 18~26, 2011.5.
- [22] Chao Yang, Robert Chandler Harkreader, and Guofei Gu, "Die free or live hard? Empirical evaluation and new design for fighting evolving twitter spammers", Recent Advances in Intrusion Detection (RAID), Vol. 6961, pp.318-337, 2011.

— 저 자 소 개 —



한진섭(정회원)

2001년 광운대학교 전자계산학과
학사 졸업(이학사).

2003년 광운대학교 컴퓨터과학과
석사 졸업(공학석사).

2003년~2005년 (주)한국컴퓨터.
기술연구소 연구원.

2005년~현재 광운대학교 컴퓨터과학과
박사 재학.

<주관심분야 : Artificial Intelligence, Data/Web
Mining>



박병준(정회원)

1984년 서울대학교 컴퓨터공학과
학사 졸업(공학사).

1988년 University of Minnesota,
Computer Science.
석사 졸업(공학석사).

1989년~1996년 US Army
Construction Engineering
Research Lab. Researcher

1997년 University of Illinois at
Urbana-Champaign, Computer Science
박사 졸업(공학박사).

1997년~1998년 Epic Systems Corp, Senior
Researcher.

1998년~2000년 SPSS Inc, Senior Researcher.

2000년~현재 광운대학교 컴퓨터소프트웨어학과
교수.

<주관심분야 : Artificial Intelligence, Data/Web
Mining, Knowledge-Based Systems>