

형태소 분석을 통한 악성 댓글 필터링 방안

A Filtering Method of Malicious Comments Through Morpheme Analysis

하예람*, 천준석**, 왕인서**, 박민옥***, 우균**
사상경찰서*, 부산대학교 정보융합공학과**, (주)유시스***

Yeram Ha(hayeram@hanmail.net)*, Junseok Cheon(jscheon@pusan.ac.kr)**,
Inseo Wang(inseowang@pusan.ac.kr)**, Minuk Park(mupark3028@gmail.com)***,
Gyun Woo(woogyun@pusan.ac.kr)**

요약

인터넷 기사에 대한 댓글은 토론 및 소통 등 긍정적 효과도 있지만, 악성 댓글은 사람의 목숨을 앗아가는 정도로 심각한 문제이다. 이런 측면에서 악성 댓글을 자동으로 검출하는 것은 중요한 문제이다. 그러나 현재 채택되고 있는 금지어 바탕의 댓글 필터링 방안은, 특히 한글 댓글의 경우 효과적이지 못한 실정이다. 이 논문에서는 형태소 분석을 통해 비속어 형태소와 존대어 형태소를 구별함으로써 댓글 필터링을 수행하는 방법을 제안한다. 이들 두 형태소를 바탕으로 댓글 건전도를 계산하고, 또한 건전도를 바탕으로 한 댓글 충격량 계산 방법을 여러 제안한다. 악성 댓글을 대상으로 한 실험을 수행한 결과, 포털에서 수행되고 있는 클린봇보다 재현율이 37.93% 포인트 개선되었으며 F-지수 값은 47.66 포인트 개선된 것으로 나타났다. 이 결과에 따르면, 형태소 분석을 기초로 한 새로운 필터링 방법이 금지어를 바탕으로 한 방법의 좋은 대안이 될 수 있을 것으로 기대된다.

■ 중심어 : | 악성 댓글 | 형태소 분석 | 댓글 건전도 | 댓글 충격량 |

Abstract

Even though the replying comments on Internet articles have positive effects on discussions and communications, the malicious comments are still the source of problems even driving people to death. Automatic detection of malicious comments is important in this respect. However, the current filtering method of the malicious comments, based on forbidden words, is not so effective, especially for the replying comments written in Korean. This paper proposes a new filtering approach based on morpheme analysis, identifying coarse and polite morphemes. Based on these two groups of morphemes, the soundness of comments can be calculated. Further, this paper proposes various impact measures for comments, based on the soundness. According to the experiments on malicious comments, one of the impact measures is effective for detecting malicious comments. Comparing our method with the clean-bot of a portal site, the recall is enhanced by 37.93% point and F-measure is also enhanced up to 47.66 points. According to this result, it is highly expected that the new filtering method based on morpheme analysis can be a promising alternative to those based on forbidden words.

■ keyword : | Malicious Comments | Morpheme Analysis | Soundness of Comments | Impact of Comments |

* 이 과제는 부산대학교 기본연구지원사업(2년)에 의하여 연구되었음.

접수일자 : 2021년 07월 22일
수정일자 : 2021년 08월 24일

심사완료일 : 2021년 08월 24일
교신저자 : 우균, e-mail : woogyun@pusan.ac.kr

I. 서론

인터넷 뉴스의 댓글은 참여자들의 소통 및 토론의 장을 마련한다는 측면에서는 긍정적이지만, 악성 댓글은 사회적인 문제가 되고 있다. 정치 기사에서는 자신이 지지하지 않는 정당에 대해서는 수많은 악성 댓글이 생성되는 것을 쉽게 찾아볼 수 있다. 이러한 악성 댓글은 소중한 인명을 앗아가기도 한다. 특히 연예 기사와 스포츠 기사에서는 명예훼손 및 비하 댓글로 인해, 2020년부터는 연예 및 스포츠 관련 기사에는 댓글 서비스를 중단하는 사태에 이르렀다[1]. 그러나 정치 및 사회 등 일반 기사의 경우에는 여전히 악성 댓글이 큰 문제가 되고 있다.

악성 댓글이 이렇게 빈번하게 발생하고 있는 원인 중 하나는 작성자가 누구인지 확인이 힘든 익명성 때문이다. 익명성이란 “상대방의 정체에 잘 드러나지 않는 상태를 지칭하며, 커뮤니케이션 당사자의 정체 확인이 어려운 정도”를 말한다[2]. 댓글을 작성하는 이가 누구인지 알 수 없는 사이버공간에서 댓글 작성자는 아무런 죄의식이나 일탈 의식 없이 누군가를 마음껏 비난할 수 있다. 그 결과, 개개인들끼리의 사회적 실재감이 낮아지게 되고 이로 인해 언어적인 마찰이 발생하여 가해자, 피해자 모두에게 큰 정신적 피해를 끼치게 된다. 이러한 익명성을 해결하기 위해 실명제를 도입하려는 움직임이 보였으나 표현의 자유와 언론의 자유 등과 같은 기본권을 침해한다는 이유로 위헌 판결이 내려져 실행되지 못하였다[3].

악성 댓글 문제를 해결하기 위한 여러 노력이 시행되고 있지만 빠른 효과를 기대하기 어렵기 때문에, 악성 댓글 피해자가 직접 댓글 작성자를 고소하는 법적 소송이 증가하고 있다. 대검찰청에 접수되는 사이버 명예훼손 사건도 매년 크게 증가하고 있다[4]. 그러나 법적 소송 역시 많은 시간과 비용이 소요되므로 악성 댓글 문제의 빠른 해결책이라고 보기 어렵다.

하지만 이러한 방법은 사후 처방에 불과하기 때문에 피해를 사전에 막기 위한 예방책을 마련하는 것이 더 낫다. 포털 사이트 또는 뉴스를 다루는 웹사이트는 방송통신위원회가 제시한 SafeNet 기준을 참고하여 댓글 규제 정책을 수립하였다[5]. 또한, 댓글을 읽는 사람들

이 직접 신고할 수 있게 하는 기능을 추가하는 등 악성 댓글에 대한 대응 방안들을 마련하고 있다. 자동 악성 댓글 검출 또한 이러한 방안 중 하나로 제시되고 있다.

자동 악성 댓글 검출 방법은 주로 댓글에 욕설이나 특정 어휘가 포함되면 제재하는 방식, 즉 금칙어 방식을 취하고 있는데, 이는 쉽게 벗어날 수 있으므로 실효가 높지 않은 것으로 판단된다. 작성자는 시스템을 우회하기 위해 변형된 형태를 사용하기도 하고 매일 새로운 형태의 악성 금칙어가 생성되기 때문이다. 따라서 현재 작성된 인터넷 뉴스 댓글 상황을 바탕으로 댓글 작성 시 실시간으로 필터링 작업을 할 수 있는 방법이 필요하다.

이 논문은 뉴스 사이트의 악성 댓글을 필터링하기 위해 댓글의 속성 및 어조의 특성을 파악하여 비속어의 강도에 따른 심층적인 비속어 필터링 방법을 제안한다. 먼저 뉴스 포털(네이버 뉴스)의 기사별 댓글을 수집하여 댓글 길이별로 악성 댓글 분포를 조사하였다. 또한, 비속어 및 존대어 형태소를 바탕으로 한 댓글 건전도 개념을 확장하여 댓글 충격량을 제시하였다. 실제 악성 댓글을 대상으로 건전도 및 충격량 실험을 통해 악성 댓글 검출에 효과적인 기준을 선정하였다.

2절에서는 악성 댓글에 대한 연구와 댓글을 크롤링 방법에 대해 소개한다. 3절에서는 제안 시스템 설계와 구현에 대해 설명한다. 4절에서는 제안 시스템을 이용한 실험과 평가 결과에 대해 소개한다. 5절에서 제안 시스템의 한계점과 해결책을 논의한 후, 6장에서 결론을 맺는다.

II. 관련연구

악성 댓글의 필터링 기법이 채택되고 있음에도 불구하고 악성 댓글은 여전히 문제가 되고 있는 상황이다. 문제의 원인을 파악하기 위해 이 절에서는 여러 선행 연구들을 소개하고 각 연구에서 제시하고 있는 이론적인 관점 및 악성 댓글 관련 요인을 살펴본다.

댓글이 많이 달릴수록 기사는 주목할만하고 중요한 것으로 인식될 수 있다. 그 정도로 댓글은 큰 영향을 미치고 있으나 많은 수의 댓글이 달릴 경우, 전체적인 내

용 및 성향을 파악하기가 힘들다. 그래서 내용의 유사도에 따라 여러 집단으로 분류하고 시각화하는 연구[6]가 있다. 이 연구에서는 이용자들이 댓글 성향 파악을 통해 원하는 댓글에 접근할 수 있도록 하며 스팸 및 무의미한 댓글에 대한 속성을 제시하고 있다.

한국어와 영어의 인터넷 뉴스 댓글을 비교하여 악성 댓글의 특성을 파악한 연구[7]도 있는데, 이 연구에서는 언어 분석 프로그램인 LIWC(Linguistic Inquiry and WordCount)[8]와 KLIWC(Korean Linguistic Inquiry and Word Count)[9]의 비교로 악성 댓글의 형태소를 분석하여 일반 댓글보다 악성 댓글이 형태소가 더 적으며 일시적인 감정 표출로 작성되는 면에서 악성 댓글에는 감정적, 정서적 단어가 일반 댓글보다 많이 포함되어 있음을 확인하였다.

형태소 분석기를 통해 추출한 단어를 바탕으로, 문서 내부 및 문서 간의 출현 빈도, 구체적으로 TF-IDF(Term Frequency-Inverse Document Frequency)[10-12] 값에 따른 가중치를 계산하여 악성 댓글과 일반 댓글을 분류한 연구[13]가 있다. 이 연구에서는 의미있는 품사들을 선별하여 가중치를 구하였으며 SVM(support vector machine)[14]을 통해 악성 댓글을 분류하였다.

댓글에서 자질 추출을 통해 FFP(Feature Frequency Profile)[15] 방법을 적용하여 악성 댓글을 필터링하는 연구[16]도 있다. 우리말의 경우에 단어를 변형하여 필터링을 회피하는 경우가 있는데, 이런 경우에도 초성은 대부분 같다. 이 연구에서는 이러한 특징을 기반으로 하여 빈도수가 높은 글자는 초성으로 변경한 후 자질을 추출하였다. SVM과 랜덤 포레스트(random forest)를 적용한 결과 SVM이 성능이 우수함을 확인하였다.

글자 형태의 특수문자, 숫자 조합을 한글 정규화 단계를 거친 후 감성 사전으로 구축하고 SVM을 통해 악성 댓글을 추출하는 연구[17]도 있다. 이 연구에서는 특수문자와 숫자를 활용한 변형 비속어를 정규화하여 필터링 가능한 형태로 만든 다음, 악성 댓글을 추출한다.

이들 연구는 기계적인 처리에 따른 필터링으로 실제 댓글을 읽는 독자가 받는 충격을 고려하지 않은 방법이다. 비속어가 포함된 댓글을 찾아내서 필터링하는 기본

적인 방법 외에 실제 독자들에게 충격을 주는 심각성이 높은 악성 댓글에 대한 필터링 또한 연구가 필요하다.

III. 시스템 설계 및 구현

이 논문에서는 이전 연구와 달리, 댓글의 어조를 파악하고 이를 활용하여 악성 댓글 필터링을 수행하고자 한다. 댓글의 어조를 비속어, 존대어, 예사어로 구분하고 댓글 길이를 분석한다. 형태소 분석을 통해 추출한 비속어 및 존대어 형태소를 바탕으로 댓글 건전도 및 충격량을 분석한다. 논문에서 제안하는 비속어 댓글 필터링 시스템의 구조는 [그림 1]과 같다.

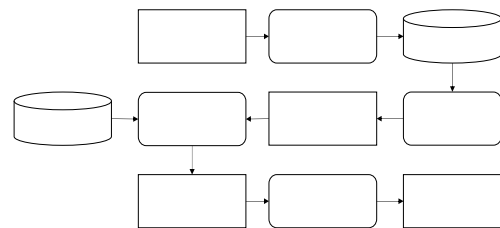


그림 1. 비속어 댓글 필터링 시스템

비속어 댓글 필터링 시스템에서 뉴스 포털(네이버 뉴스)의 첫 페이지를 통해 분야별로 주요 뉴스를 수집한다. 실시간으로 자동 갱신되는 주요 뉴스를 대상으로 자체 제작한 웹 크롤러를 이용하여 기사 댓글을 수집한 후 SQLite 데이터베이스 파일로 저장하였다. 길이 분석을 통해 평균 길이 이상의 댓글만을 선택하여 형태소를 분석한 후, 비속어 목록 데이터베이스를 참고하여 댓글의 비속어 및 존대어를 분석한다.

1. 댓글 수집

네이버 뉴스¹의 첫 페이지는 정치, 경제, 사회, 생활/문화, 세계, IT/과학으로 기사가 나뉘어서 실시간으로 게재된다. 섹션별 뉴스는 네이버의 AiRS 추천을 통해 구성된 뉴스로 사람들이 관심 있는 주제의 뉴스를 인공

1 여론집중도조사위원회의 '2016~2018 뉴스 이용집중도 조사 결과'에 따르면, 인터넷 뉴스 이용 집중도의 경우 네이버, 다음, 네이트 포털 점유율이 2018년에 88.4%이며[18], 세 곳의 포털 중에서 뉴스 첫 페이지에 분야별 주요 뉴스가 출력되는 곳은 네이버이므로 네이버를 크롤링 대상으로 정하였다.

지능을 통해 실시간으로 분석하여 제공한다. 주요 뉴스에는 댓글이 달릴 확률이 높으므로, 주요 뉴스를 실시간으로 계속 크롤링하여 댓글을 수집하였다.

웹 크롤링은 데이터 수집, 데이터 추출, 데이터 보존 단계로 구성된다. 데이터 수집 단계에서는 웹 사이트의 HTML 등의 데이터를 수집한다. 하지만 HTML에는 필요한 문장 데이터만이 아닌 태그 등의 데이터도 들어있기 때문에 필요한 것만 추출하는 작업이 필요하다. 따라서 다음 작업인 데이터 추출 단계가 진행된다. 여기에서는 복잡한 구조의 HTML 데이터를 분석하여 필요한 정보만 추출한 후 데이터베이스에 저장한다.

Python의 웹 크롤링 라이브러리로는 Requests[19], BeautifulSoup[20], Selenium[21]이 있다. Request는 HTML 데이터 수집에 사용되며 BeautifulSoup은 수집한 데이터로부터 필요한 정보를 추출하는 데 사용된다. Selenium은 데이터 수집뿐만 아니라 데이터 추출도 가능하지만 동작이 느리기 때문에, 우리는 Selenium으로 데이터 수집 후에 BeautifulSoup으로 데이터를 추출하였다.

네이버에서는 자체적으로 악성 댓글을 감지하여 숨기는 클린봇이라는 기능이 있는데, 크롤러 구현 시에는 이 기능의 설정 및 해제를 통해 클린봇의 효과도 확인하였다. 수집 댓글은 중복 여부를 확인 후, 중복되지 않은 댓글만 SQLite DB파일로 저장하였다.

2. 길이 분석

길이 분석 도구는 수집된 댓글을 길이를 기준으로 분류한다. 크롤링을 통해 하루 동안에 수집되는 댓글 개수는 기사당 10,000개가 훌쩍 넘으므로, 댓글 필터링 시스템의 속도 향상이 중요하다. 이전 연구[22]와 마찬가지로, 악성 댓글을 수집하기 위해 평균 길이보다 긴 댓글을 수집하였다. 길이는 공백을 제외한 글자 수로 산출하였다.

3. 형태소 분석

댓글을 분석하기 위해 KoNLPy(Korean Natural Language Processing in Python)[23]를 활용하였는데, 형태소 분석기로는 꼬꼬마(KKMA)[24]를 사용하였다. 꼬꼬마 형태소 분석기는 다른 도구에 비해 속도는

느리지만 품사 분류가 자세하다는 장점이 있다. 꼬꼬마 형태소 분석기에서는 한글 형태소의 품사를 체언, 용언, 관형사, 부사, 감탄사, 조사, 어미, 접사, 어근, 부호, 한글 이외로 총 11가지 종류로 나누고, 세부적으로 최대 55가지로 구분하고 있다.

꼬꼬마 형태소 분석기를 통해 댓글을 분석하면 품사 태그와 형태소의 쌍으로 이루어진 배열을 구할 수 있다. 댓글의 어조를 판단하기 위해, NNG(보통 명사), VV(동사), VA(형용사), MAG(일반 부사) 외에 ‘-놈’을 나타내는 NNB(일반 의존 명사)와 ‘-님니다’를 나타내는 EFN(평서형 종결 어미), NP(대명사), IC(감탄사), JKS(주격 조사), JKM(목적격 조사), EPH(존칭 선어말 어미), EPT(시제 선어말 어미), EPP(공손 선어말 어미), EFQ(의문형 종결 어미), EFO(명령형 종결 어미), EFA(청유형 종결 어미), EFI(감탄형 종결 어미), EFR(존칭형 종결 어미), ECD(의존적 연결 어미), XSN(명사 파생 접미사), XR(어근)으로 21개의 품사를 유효 품사로 정의하였다.

형태소 분석 도구에서는 비속어 목록 데이터도 참고한다. 비속어 데이터 목록으로는 게임산업진흥원의 ‘게임언어 건전화 지침서 연구’[25]에 사용된 데이터를 기반으로 하여 최신화한 자료를 사용하였다.

4. 댓글 건전도 및 충격량 분석

댓글이 악성 댓글임을 확신할 수 있는 요인은 댓글 내 비속어의 유무이다. 2002년 정보통신윤리위원회에서 규정한 사이버 언어폭력의 유형에는 욕설, 비방(명예훼손), 도배, 성적 욕설, 유언비어가 있으며, 이 중 비속어 형태로 추출할 수 있는 유형은 욕설, 성적 비방을 의미하는 단어다[26].

비속어와 달리 존대어는 상대방을 높이고 존중하는 의미를 담고 있으므로 댓글을 건전하게 만든다. 따라서 댓글이 건전하다고 느낄 수 있는 요인 중 하나가 된다. 존대어는 EF(종결 어미) 내 ‘님니다’, ‘님니까’, ‘십시오’, ‘님시다’ 등의 높임 종결어미, 유효 품사 내 존대어를 나타내는 품사인 EPH(존칭 선어말 어미), EPP(공손 선어말 어미), EFR(존칭형 종결 어미)가 있으며 NNG(보통 명사)에서 존대의 의미를 가진 “진지”, “성함”, “연세”, “덕”, “부인”, “생신”, “말씀” 등의 단어를 반영하였다.

건전도 분석 도구에서는 꼬꼬마 형태소 분석기를 통해 얻어진 형태소 분석된 댓글과 형태소 분석된 비속어를 이용하여 댓글의 건전도(W)를 계산한다. 댓글의 건전도는 전체 형태소 개수 중 비속어 형태소 개수와 존대어 형태소 개수에 대한 연산으로 구한다. 건전도 계산식은 식 1과 같이 정의하였다[27].

$$W = \sum W_R \times \frac{M_R}{M_V} - \sum W_C \times \frac{M_C}{M_V} \quad (1)$$

where $M_V = M_R + M_C + M_P$

식 1에서 W_R 은 존대어 가중치, W_C 는 비속어 가중치로서, 각 가중치 값으로 0이상 1이하의 조합을 실험한 결과, 정규분포가 가장 정규분포에 가까운 $W_R = W_C = 1$ 을 사용하였다. M_R , M_C , M_P 는 각각 존대어, 비속어, 예사어 형태소 개수이며, M_V 는 이들의 합, 즉 전체 형태소 개수이다. 이들 개수는 당연히 중복을 허용한 집합의 원소 수로서, 즉 출현 횟수를 의미한다.

댓글 건전도 값이 높을수록 댓글이 악성 댓글일 가능성이 작아지며, 댓글 건전도 값이 낮을수록 댓글이 악성 댓글일 가능성이 커진다. 건전도 분석 도구를 통해 나오게 되는 값이 비속어 댓글이므로 이 댓글을 필터링하면 악성 댓글을 줄일 수 있다.

댓글 건전도는 비속어 및 존대어의 형태소 출현 횟수만을 고려하기 때문에, 댓글 길이와 무관한 지수이다. 하지만 같은 건전도의 댓글이라고 하더라도 그 위력은 다를 수 있다. 짧은 댓글은 비속어만 사용해서 짧게 작성되었기 때문에, 기분은 나쁘지만 독자의 충격은 적을 수 있다. 하지만 긴 댓글의 경우 비속어뿐만 아니라 기사 대상이 좋아하지 않는 내용이 추가로 들어있기 때문에 짧은 댓글보다 더욱 위력적일 수 있다. 이를 고려하여 댓글 충격량(CI) 지수를 고안하였다.

$$CI_1 = M_C \times (M_V - M_C) \times M_I \quad (2)$$

$$CI_2 = M_C \times (M_V - M_C) \times M_N \quad (3)$$

$$CI_3 = W \times L \quad (4)$$

$$CI_4 = L \quad (5)$$

식 2와 3에서 M_V 은 식 1과 같이 유효 형태소 수, M_I

은 무효 형태소(외국어 등 분석 불가 형태소) 수, M_N 은 비형태소(문장 부호 등) 수이며 식 4와 5에서 L 은 댓글 길이를 나타낸다. 식 2와 식 3은 형태소 종류로 구별되는 직육면체의 부피를 충격량으로 산출한 것이며, 식 4는 댓글 건전도와 댓글 길이로 산출된 넓이를, 식 5는 단순히 댓글 길이만을 댓글 충격량으로 산출한 것이다.

IV. 실험 및 평가

이 절에서는 제안 시스템을 이용한 실험 결과를 소개한다. 먼저 댓글 길이에 따른 악성 댓글 분포를 조사하였다. 어조별 평균 댓글 길이 및 다음으로 건전도가 낮은 악성 댓글은 주로 어느 분야에서 많이 생성되는지를 확인하였다. 앞서 제시한 건전도 및 충격량의 상관관계 분석을 실시하였다. 또한 네이버 클린봇과 제안 시스템의 성능을 비교하였다.

1. 댓글 길이별 악성 댓글 분포 조사

여기에서는 댓글 길이별로 악성 댓글 분포도가 차이가 있는지를 확인하였다. 앞에서 우리는 댓글 길이가 길면 악성 댓글이 있을 확률이 높다고 가정하였다. 실험을 위해 2021년 6월 1일 하루 동안 업로드되는 네이버 기사를 분야별로 댓글을 수집하였다. 수집 결과, 총 7,241개의 댓글을 수집하였는데, 수집한 댓글의 길이별로 세 분류로 나눈 후, 비속어 댓글 비율을 확인하였다. 네이버에는 댓글이 공백을 제외하고 최대 300자가 입력되지만, 수집한 댓글 분석 결과 평균 길이는 41.71자인 것을 확인하였다. 이 값을 기준으로 하여 댓글을 짧은 댓글, 중간 댓글, 긴 댓글을 같은 수만큼 세 그룹으로 분류한 후 실험을 진행하였다. 이 실험에서 댓글의 악성 여부는 비속어 포함 여부로 판단하였다. [표 1]은 댓글 길이에 따른 악성 댓글 분포 실험 결과를 나타낸 것이다. 실험 결과, 하루 동안 수집한 7,241개의 댓글 중 10.22%인 740개의 댓글이 악성 댓글임을 확인하였다. 총 댓글을 댓글 길이별로 나눠서 계산하면 짧은 댓글은 5.88%가, 보통 댓글은 9.16%가, 긴 댓글은 14.58%가 악성 댓글인 것을 확인하였다. 즉, 댓글 길이가 길어질수록 악성 댓글 확률이 높아졌다.

표 1. 댓글 길이에 따른 분야별 악성 댓글 비율

| 구분 | 정치 | | | 경제 | | | 사회 | | |
|----|-------|-------|-------|------|-------|-------|------|------|-------|
| | 짧음 | 보통 | 김 | 짧음 | 보통 | 김 | 짧음 | 보통 | 김 |
| 전체 | 1,018 | 1,018 | 1,018 | 329 | 329 | 329 | 500 | 498 | 500 |
| 악성 | 70 | 106 | 196 | 18 | 36 | 40 | 28 | 47 | 50 |
| 비율 | 6.88 | 10.41 | 19.25 | 5.47 | 10.94 | 12.16 | 5.60 | 9.44 | 10.00 |

| 구분 | 생활 | | | 세계 | | | IT/과학 | | |
|----|------|------|------|------|-------|-------|-------|------|------|
| | 짧음 | 보통 | 김 | 짧음 | 보통 | 김 | 짧음 | 보통 | 김 |
| 전체 | 172 | 173 | 172 | 253 | 252 | 253 | 142 | 143 | 5142 |
| 악성 | 6 | 14 | 17 | 11 | 28 | 35 | 9 | 7 | 14 |
| 비율 | 3.49 | 8.09 | 9.88 | 4.35 | 11.11 | 13.83 | 6.34 | 4.90 | 9.86 |

위의 실험을 살펴보면, 모든 분야의 기사에서 긴 댓글에 악성 댓글이 많은 것을 확인하였다. 하지만 보통 댓글도 긴 댓글 못지않게 악성 댓글이 많았다. 따라서 댓글 평균 길이인 41.7자 이상인 댓글에 대해서 어조별 댓글 평균 길이는 [표 2]와 같다.

표 2. 어조별 평균 댓글 길이(평균 길이 이상의 댓글 대상)

| 구분 | 정치 | 경제 | 사회 | 생활 | 세계 | IT/과학 | 평균 |
|-----|--------|-------|-------|--------|-------|--------|-------|
| 비속어 | 89.86 | 96.61 | 83.00 | 94.05 | 97.39 | 107.61 | 94.75 |
| 예사어 | 83.30 | 91.08 | 77.94 | 72.78 | 80.25 | 81.90 | 81.21 |
| 존대어 | 101.08 | 96.28 | 91.00 | 107.45 | 86.36 | 81.90 | 94.01 |

[표 2]를 보면, 예사어 댓글보다 비속어 및 존대어 댓글의 평균 길이가 다소 긴 것을 확인할 수 있다.

다음으로 하루 동안 수집한 네이버 뉴스 댓글의 건전도 분포를 분석하였다. 하루 동안 수집된 댓글 중 평균 길이 이상의 댓글은 정치 1,018개, 경제 325개, 사회 513개, 생활 178개, 세계 238개, IT/과학 133개였는데 이들 분포는 [그림 2]와 같다. 표준편차는 정치 0.023166, 경제 0.019978, 사회 0.022359, 생활 0.024594, 세계 0.027912, IT/과학 0.011812인 것으로 나타났다.

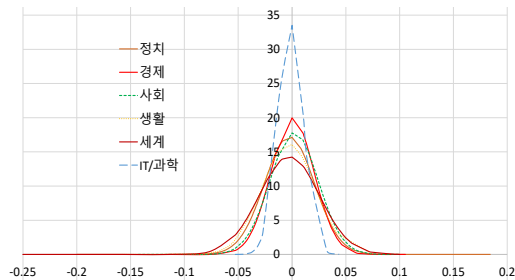


그림 2. 기사 분야별 댓글 건전도에 대한 정규분포 그래프

[그림 2]에 따르면 각 분야별 분포는 단봉형(unimodal)으로 나타났는데, 최빈값의 x 좌표는 정치 -0.00241, 경제 0.00042, 사회 0.00191, 생활 -0.00305, 세계 -0.00254, IT/과학 -0.00149으로 나타났다. COVID-19로 인해 그 래프 분포가 대부분 왼쪽으로 이동된 것으로 생각된다.

다음으로 분야별 악성 댓글 편향성을 조사하였다. 6개 분야의 악성 댓글 중 건전도가 낮은 순으로 100개를 추출한 후 이를 분석한 결과 [그림 3]과 같은 결과를 얻었다.

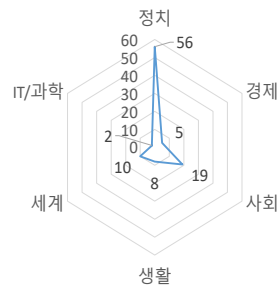


그림 3. 건전도 하위 100개의 댓글에 대한 분야별 분포

건전도 하위 댓글은 대부분 정치 분야에 나타나는 것을 확인할 수 있다. 2위는 사회 분야, 3위는 세계 분야였다. 다음으로 악성 댓글 길이별 건전도를 확인해 보았다. [그림 4]는 분야별 악성 댓글 길이와 건전도를 계산한 값을 바탕으로 생성된 그래프이다. 대부분의 악성 댓글은 120자 내로 작성되는 것을 확인하였으며, 길게 작성한 악성 댓글의 경우에도, 순수하게 비속어만으로 작성된 댓글은 없는 것으로 확인되었다.

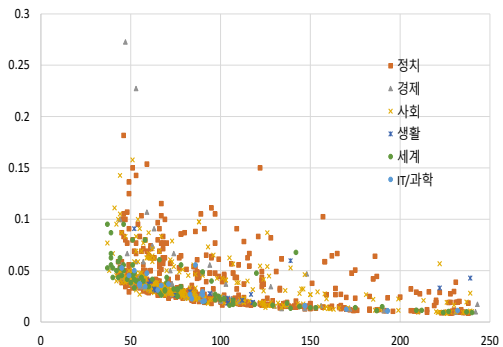


그림 4. 악성 댓글 길이별 건전도 분포 그래프

다음 실험은 댓글 충격량 실험이다. [표 3]에서 볼 수 있는 것처럼 건전도가 같은 댓글이라고 하더라도 충격량이 다른 것으로 판단된다.

표 3. 댓글 길이는 차이가 나지만 건전도가 같은 댓글 예

| 분야 | ID | 댓글 내용 | 건전도 |
|-------|------|--|--------|
| 정치 | 102 | 내로남불죄받 짐승들 | -0.143 |
| | 1439 | 문재인 양아치새~휘~ 전화받을 필요없다. 또 쇼할려고 저러는거지. 국민의 힘을 무시하고 이간질하려고 호시탐탐 노리는 놈인데... 문재인 넌적한 썬판때기나 후려갈겨라. | |
| 경제 | 227 | 멍청한 쓰레기들이 세금도 존나 올렸다 | -0.111 |
| | 588 | 공산주의빨갱이들은 집가진자 싫어하지요 ! 자본주의는사회악인 돼지추중세력 더불어빨갱이들 ! | |
| 사회 | 383 | 저교사 웬지 일베충일거 같은 느낌이다 | -0.091 |
| | 729 | 저런 개 쓰레기도 선생이라고... 학생들이 불쌍하다.그 놈의 세월호는 언제까지 우려먹는 거냐? 저런 선생 쉐귀 놈은 OOO버려야 한다. | |
| 생활 | 87 | 분명 중공파장농들짓이다 (조선족포함) | -0.125 |
| | 80 | 중국농들은 챔피언것도 모르는 미개한 도둑놈 다지들거래 우기면 다된다 생각하는 지저분한놈들 | |
| 세계 | 224 | 썩바리3기들이 다 그렇지 모 | -0.111 |
| | 321 | 원숭이족속 쪽국은 올림픽 열지못한다.그리고 경제 대공황이나 맞아서 쪽국 빈국으로 떨어져라 | |
| IT/과학 | 157 | 존나 잘자는데? 오히려 조용하면 잠이 더 안 옴 | -0.091 |
| | 55 | 문기놈은 지원은 못해줄망정 주장 잡아놓고 방해만 놓고 있으니, 이렇게 기업 잡는 나라 세상에 없다. 파이팅 삼성 천원문기놈 | |

충격량 실험을 위해 먼저 2021년 6월 10일 목요일부터 2021년 6월 16일 수요일까지 일주일 동안 네이 버 댓글을 수집하였다. 분야별로 수집한 전체 댓글 수와 악성 댓글 수는 [표 4]와 같다.

표 4. 댓글 충격량 계산을 위해 수집한 악성 댓글(건전도 0 미만)

| 구분 | 정치 | 경제 | 사회 | 생활 | 세계 | IT/과학 | 합계 |
|----------|--------|--------|--------|-------|-------|-------|--------|
| 전체 댓글 | 42,895 | 10,249 | 18,486 | 3,453 | 9,441 | 2,873 | 87,397 |
| 악성 댓글 | 5,454 | 981 | 2,312 | 225 | 922 | 193 | 10,087 |
| 악성 댓글 비율 | 12.71 | 9.57 | 12.51 | 6.52 | 9.77 | 6.72 | 11.54 |

댓글 충격량을 계산하기 위해 먼저 [표 5]와 같이 악성 댓글 정보를 분석하였다. 수집한 악성 댓글 수가 만개 이상이므로 건전도가 같은 댓글 사례만 선정하여 기타 정보를 요약하였다.

표 5. 악성 댓글 정보 분석 예

| 분야 | ID | 댓글 길이 | 건전도 | 유효 형태소 | 무효 형태소 | 비속어 형태소 | 비형태소 |
|----|------|-------|--------|--------|--------|---------|------|
| 정치 | 102 | 9 | -0.143 | 7 | 1 | 1 | 0 |
| | 1439 | 74 | -0.143 | 28 | 21 | 4 | 7 |
| 경제 | 227 | 16 | -0.111 | 9 | 3 | 1 | 0 |

| | | | | | | | |
|-------|-----|----|--------|----|----|---|---|
| 사회 | 588 | 42 | -0.111 | 18 | 7 | 2 | 2 |
| | 383 | 17 | -0.091 | 17 | 11 | 1 | 4 |
| | 729 | 59 | -0.091 | 22 | 19 | 2 | 6 |
| 생활 | 87 | 18 | -0.125 | 8 | 4 | 1 | 2 |
| | 80 | 40 | -0.125 | 17 | 14 | 2 | 1 |
| 세계 | 224 | 12 | -0.111 | 9 | 1 | 1 | 0 |
| | 321 | 41 | -0.111 | 18 | 6 | 2 | 2 |
| IT/과학 | 157 | 19 | -0.091 | 11 | 4 | 1 | 1 |
| | 55 | 52 | -0.091 | 22 | 15 | 2 | 3 |

[표 5]에서 ID는 악성 댓글 고유 번호이며, 댓글 길이는 공백을 제외한 글자 수이다. 유효 형태소는 건전도를 구할 때 사용하는 21 종류의 유효 형태소를 가리킨다. 무효 형태소는 KKMA가 구분하는 형태소 중에서 유효 형태소를 제외한 것으로 총 35 종류의 형태소이다. 비 형태소는 형태소가 아닌 것으로 이모티콘, 자음, 모음, 특수문자, 숫자 등이 있다. 댓글 충격량(식 2~5)을 산출한 결과는 [표 6]과 같다.

표 6. 악성 댓글에 대한 충격량 계산 결과

| 분야 | ID | CI_1 | CI_2 | CI_3 | CI_4 |
|-------|------|--------|--------|---------|--------|
| 정치 | 102 | 6 | 0 | -1.286 | 9 |
| | 1439 | 2016 | 672 | -10.571 | 74 |
| 경제 | 227 | 24 | 0 | -1.778 | 16 |
| | 588 | 224 | 64 | -4.667 | 42 |
| 사회 | 383 | 30 | 0 | -1.545 | 17 |
| | 729 | 760 | 240 | -5.364 | 59 |
| 생활 | 87 | 28 | 14 | -2.250 | 18 |
| | 80 | 420 | 30 | -4.706 | 40 |
| 세계 | 224 | 8 | 0 | -1.333 | 12 |
| | 321 | 192 | 64 | -4.556 | 41 |
| IT/과학 | 157 | 40 | 10 | -1.727 | 19 |
| | 55 | 600 | 120 | -4.727 | 52 |

CI_2 (식 3)의 경우에는 비형태소(M_N) 값이 0인 댓글로 인해 0으로 산출되는 경우가 있었다.

2. 상관관계 분석

이 절에서는 실제 악성 댓글의 악성 정도와 건전도 및 충격량 간의 상관관계를 분석한다. 이전 실험을 통해 댓글 건전도 외에도 댓글 크기도 악성 댓글을 판별에 영향을 줄 수 있음을 확인하였다. 이 절 상관관계 실험을 통해 악성 댓글 검출에 더 효과적인 지표를 선정하고자 한다.

진리 집합을 결정하기 위해 우리는 [표 4]에서 수집한 정치 분야 댓글 42,895개 중에서 100개를 무작위로 추출하여 사람이 댓글을 평가하게 하였다. 편의를 위해 평가 항목을 매우 기분 좋음, 기분 좋음, 조금 기분 좋

음, 조금 기분 나쁨, 기분 나쁨, 매우 기분 나쁨의 6개로 제공하였다. 이후 각 항목을 건전도 -1, -0.6, -0.2, 0.2, 0.6, 1로 변경하였으며, 식 6을 사용해서 이를 정규화하였다.

$$N_Q = \frac{\sum_{i \in I} (N_{(i,Q)} + s_i)}{|I|} \text{ where } N_{(i,Q)} = \frac{V_i - \mu_i}{s_i} \quad (6)$$

식 6은 사람이 평가한 것을 정규화한 값(N_Q)을 구하기 위한 식을 나타낸 것이다. $N_{(i,Q)}$ 는 한 사람($i \in I$)의 평가 결과(Q) 100개를 정규화한 것으로, 개인이 평가한 문항별 값인 V_i 에서 개인별 평균값 μ_i 를 뺀 후, 이를 표준편차인 s_i 로 나눈 값이다. 산출된 $N_{(i,Q)}$ 에 s_i 를 더하여 건전도가 낮은 댓글만 추출한 후, 사람 수 $|I|$ 로 나누어 정규화한다. N_Q 를 바탕으로 진리 집합을 구성한 결과, 100개 중 58개의 악성 댓글 집합을 구성하였다.

진리 집합의 N_Q (standard)와 건전도(soundness) 및 댓글 충격량(CI) 간의 상관 관계를 분석하였다. 피어슨 상관계수 분석 결과는 [그림 5]와 같다.

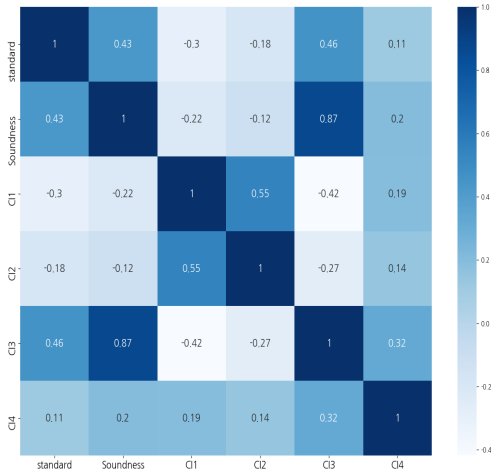


그림 5. 판별 결과에 대한 피어슨 상관 분석

[그림 5]를 살펴보면, 5개의 변수 사이의 상관계수 값으로 -0.42에서 0.87 사이의 값을 얻은 것을 확인할 수 있다. 실험 결과 댓글 충격량 CI_3 (식 4)과 N_Q 사이에 뚜렷한 양의 상관관계가 있음을 확인할 수 있다.

상관관계의 산점도 시각화 결과는 [그림 6]과 같다.

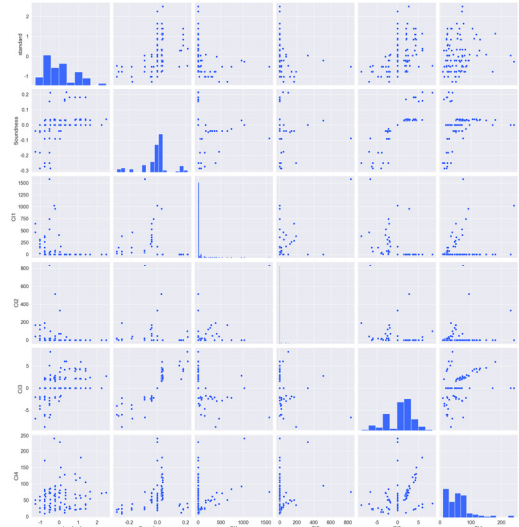


그림 6. 판별 결과에 대한 변수들의 상관관계 시각화

[그림 6]에서도 N_Q 와 CI_3 사이에 상관관계가 높음을 확인할 수 있다.

3. 검출 성능 평가

네이버에서는 악성 댓글을 줄이기 위해서 지난 2019년 11월에 악성 댓글을 필터링하는 클린봇 기능을 추가하였다. 클린봇이란 악성 댓글에 대한 자동 숨김 기능이다. 클린봇을 활성화하면, 악성 댓글 대신에 '클린봇이 부적절한 표현을 감지한 댓글입니다.'라는 메시지가 출력된다.

제안 시스템의 성능을 측정하기 위해, 진리 집합(58개의 악성 댓글)을 바탕으로 네이버 클린봇과의 비교 실험을 수행하였다. 실험 결과는 [표 7] 및 [표 8]과 같다.

표 7. 판별 결과와 논문 시스템 결과 비교표

| 제안 시스템 | | 검출 결과 | | 재현율 |
|--------|-------|--------|-------|-------|
| | | 악성 댓글 | 일반 댓글 | |
| 판별 결과 | 악성 댓글 | 27 | 31 | 46.55 |
| | 일반 댓글 | 0 | 42 | |
| 정밀도 | | 100.00 | | |

표 8. 판별 결과와 네이버 클린봇 결과 비교표

| 네이버 클린봇 | | 검출 결과 | | 재현율 |
|---------|-------|--------|-------|------|
| | | 악성 댓글 | 일반 댓글 | |
| 판별 결과 | 악성 댓글 | 5 | 53 | 8.62 |
| | 일반 댓글 | 0 | 42 | |
| 정밀도 | | 100.00 | | |

두 시스템의 검출 능력을 평가한 결과는 [표 9]와 같다.

표 9. 제안 시스템과 네이버 클린봇의 악성 댓글 검출 성능 평가

| 평가 지표 | 네이버 클린봇 | 제안 시스템 | 증감 |
|--------|---------|--------|--------|
| 재현율(%) | 8.62 | 46.55 | +37.93 |
| 정밀도(%) | 100.00 | 100.00 | 0 |
| F1 지수 | 15.87 | 63.53 | +47.66 |

재현율(recall)과 정밀도(precision), F1 지수를 비교하였는데, 제안 시스템의 재현율과 F1 지수가 크게 향상된 것으로 나타났다. [표 9]의 성능 평가 결과, 제안 시스템의 재현율은 46.55로서 네이버 클린봇에 비해 37.93% 포인트 증가하였으며 5.4배 정도 개선되었다. 정확도가 높을수록 양성오류(false positive) 수가 줄어든다. 다음으로 정밀도는 둘 다 100인 것을 확인할 수 있는데, 이 값이 높을수록 음성오류(false negative)가 줄어든다. 마지막으로 F1 지수의 경우 논문 시스템은 63.53이고, 네이버 클린봇은 15.87로 4배가 넘는 차이를 보이는 것을 확인할 수 있다. 재현율과 정밀도는 상호보완(trade-off) 관계이므로 F1 지수가 높을수록 두 값이 균형 있게 높은 것으로 판단할 수 있다.

V. 토 의

제안 시스템은 비속어 및 존재어 형태소 분석을 기초로 하고 있다. 그러나 모든 언어가 그렇듯이 비속어도 새로 생기기도 하고 사라지기도 한다. 비속어 데이터베이스의 최신화는 제안 시스템에서 매우 중요하다. 앞서 언급한 대로 현재 시스템은 게임산업진흥원의 비속어 목록[25]을 사용하고 있으나 최신화가 중단된 상태이다. 다행스럽게도 네이버와 카카오톡에서 공동으로 유해어 DB를 구축하여 공개할 예정이므로[28] 이를 활용한다면 제안 시스템의 성능을 더욱 높일 수 있을 것이라고 판단된다.

[표 10]은 [표 7]에서 실험한 내용 중에서 제안 시스템이 찾아내지 못한 댓글 중의 일부를 나타낸다. 6107번 댓글의 경우에는 띄어쓰기, 특수문자, 영어를 삽입하여 의도적으로 비속어를 걸러내지 못하도록 하였다. 다음으로 16296번 댓글의 경우에는 '더듬어 민주당'이라는 단어가 최근 추가된 비속어인데, 이를 추가하지 않아 건전도가 양수 값이 나온 것을 알 수 있다

한편, 모든 악성 댓글에 반드시 비속어가 나타나는 것은 아니다. 댓글에 비속어를 사용하지 않더라도 문맥상 악성 댓글인 경우가 존재하는데, [표 10]에서 28511번 댓글의 경우에는 비속어를 하나도 사용하지 않았지만, 기사 대상의 입장에서는 매우 기분이 나빠지는 악성 댓글이다. 그리고 16153번 댓글에서는 '인간'이라는 단어가 문맥상으로 비속어처럼 사용된 것을 확인할 수 있다. 이러한 예는 단순히 형태소 분석만으로는 악성 댓글인지 여부를 파악하기 힘들기 때문에 이러한 악성 댓글을 발견해내기 위한 연구가 추가로 필요하다.

표 10. 제안 시스템에서 찾아내지 못한 악성 댓글 목록

| ID | 댓글 내용 | 건전도 | 총격량 | 판별 |
|-------|---|-------|-------|--------|
| 6107 | 민주당께 좌 2 발 !뿔 뿔 G 이 r 똥을 싸라 장모 이미 법원판결 나온지가 언젠데 아직도 그거 물고 늘어지고 있네 징역6개월인가 벌써 판결 난거 가지고 치들끼리 난리 | 0.028 | 1.861 | -0.217 |
| 16296 | 새천년 NHK 정신의 황금 멤버 송영길 우상호 이종걸 김원석이 아직도 건재한 더듬어 만진당 아닌가!!! 심지어 송영길은 더듬어 만진당 대표!!! ㅋㅋㅋㅋ ㅋㅋㅋㅋ 국민의 힘은 더듬어 만진당하고는 확실하게 달라야 한다!! 피해 호소인 3인방 진선미 남인순 고민정도 건재한 더듬어 만진당!!! 이준석과 국민의 힘 여성 의원들에 대한 기대가 크다!!! 말 실수 조심하고 겸손한 자세로 화이팅~~~!!! | 0.032 | 4.129 | -0.265 |
| 28511 | 이 아저씨 참이상하네..군대는 갔다오셨나요? 면제 받고 ..학생들 교육시키는 교사라는 분이 말을 참참게하네...이런 분들은 교사이기이전에 인간 교육을 받아야 하겠네요.. | 0.029 | 2.229 | -0.217 |
| 16153 | 도와줄 인간도 아니고, 지혜를 구할 인간도 아님 | 0.000 | 0.000 | -0.770 |

VI. 결론

이 논문에서는 온라인 포털 사이트의 뉴스 기사에 대한 악성 댓글을 필터링하는 방법에 대해 제안하고 실험 결과를 제시하였다. 구체적으로 악성 댓글을 필터링하

기 위한 시스템을 제안하고 이 시스템을 기반으로 여러 가지 실험을 수행하였다.

먼저 우리는 댓글 길이별로 악성 댓글 분포도가 차이가 있는지 확인하였는데, 일반적으로 비속어를 사용하면 글의 길이가 짧아지는 경향이 있음에도 불구하고 악성 댓글이 일반 댓글보다 평균 댓글 길이가 긴 것을 확인하였다. 이는 악성 댓글은 비속어뿐만 아니라 기사 대상이 불쾌해질 만한 내용을 공들여 작성하기 때문인 것으로 판단된다.

이러한 특성을 반영하기 위해, 우리는 댓글 건전도 개념을 확장하여 댓글 충격량을 정의하였다. 같은 건전도의 악성 댓글이라 하더라도 내용에 따라 독자가 받는 충격량이 다를 수 있으므로, 건전도에 댓글의 규모를 반영하여 네 가지 충격량을 공식화하였다. 실험 결과에 따르면 충격량 C_3 가 실제 악성 댓글의 악성 정도와 상관관계가 가장 높은 것으로 나타났다.

포털 사이트의 클린봇과의 성능 비교 실험을 수행한 결과 제안 시스템이 클린봇에 비해 4배 정도로 성능이 높은 것으로 나타났다. 진리 집합을 구성하기 위해 100개의 무작위 정치 기사 댓글을 추출하여 사람이 직접 악성도를 평가하였다. 이를 기반으로 논문에서 제안하는 시스템과 네이버 클린봇에 100개의 무작위 정치 기사 댓글을 넣어서 출력 결과를 비교해보았다. 실험 결과, 재현율이 37.93% 포인트 개선되었으며 F-지수 값은 47.66 포인트 개선된 것으로 나타났다.

향후에는 댓글 필터링 시스템의 속도 향상에 관한 연구를 진행하려고 한다. 우리가 제안한 시스템은 댓글에서 존재어와 비속어를 찾아내는 데 꼬꼬마 형태소 분석기를 사용한다. 꼬꼬마 형태소 분석기는 결과가 정확하고 비속어 및 존재어 관련 품사를 다양하게 지원한다는 장점이 있지만, 형태소 분석기의 특징이 결과가 정확해 질수록 속도가 느려진다는 단점이 있다. 시스템 성능을 향상시키기 위해 형태소 분석 부하를 최소한으로 줄여야 하는데, 이는 향후 연구로 생각하고 있다.

참 고 문 헌

[1] 연합뉴스, “네이버·카카오 스포츠뉴스 댓글 잠정 중

단...‘선수 고통 심각’(종합),” <https://www.yna.co.kr/view/AKR20200807067551017>.

[2] 박문서, “인터넷 익명성과 전자상거래,” 한국통상정보학회, Vol.4, No.2, pp.201-222, 2002.

[3] 헌법재판소, 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제44조의5 제1항 제2호 등 위헌확인, [http://law.go.kr/헌재결정례/\(2010헌마47,20120823\)](http://law.go.kr/헌재결정례/(2010헌마47,20120823)).

[4] 경찰청, 전체 사이버범죄 발생·검거 현황, <https://www.police.go.kr/www/open/public/public0204.jsp>.

[5] 방송통신심의위원회, “인터넷 내용 등급 서비스란?,” <http://www.safenet.net.kr/dstandard.do>.

[6] 이윤정, 지정훈, 우균, 조환규, “인터넷 게시물의 댓글 분석 및 시각화,” 한국콘텐츠학회논문지, 제9권, 제7호, pp.45-56, 2009.

[7] 김영일, 김영준, 김영진, 김경일, “악성 댓글의 특성: 한국어와 영어의 인터넷 뉴스 댓글 비교,” 한국콘텐츠학회논문지, 제19권, 제1호, pp.548-558, 2019.

[8] liwc.wpengine.com, Linguistic Inquiry and Word Count, <http://liwc.wpengine.com/>.

[9] 이창환, 심정미, 윤애선, “언어적 특성을 이용한 '심리학적 한국어 글분석 프로그램(KLIWC)' 개발 과정에 대한 고찰,” 한국인지과학회, 제16권, 제2호, pp.93-121, 2005.

[10] S. E. Robertson, *Term specificity*, letter to the editor], 1972.

[11] S. E. Robertson, “Documentation Note: Specificity and Weighted Retrieval,” *Journal of Documentation*, Vol.30, No.1, pp.41-46 1974.

[12] S. E. Robertson, “The probability ranking principle in information retrieval,” *Journal of Documentation*, Vol.33, No.4, pp.294-340, 1977.

[13] 김묘실, 강승식, “SVM을 이용한 악성 댓글 판별 시스템의 설계 및 구현,” 한국정보과학회 언어공학연구회 학술대회, pp. 285-289, 2006.

[14] Thorsten Joachims, SVM^{light}: Support Vector Machine, https://www.cs.cornell.edu/people/tj/svm_light/.

[15] G. E. Sims, S. Jun, G. A. Wu, and S. Kim, “Alignment-free genome comparison with feature frequency profiles (FFP) and optimal resolutions,” *Proceedings of the National Academy of Sciences*, Vol.106, No.8, pp.2677-2682, 2009.

[16] 김현정, 윤영미, 이병문, “향상된 FFP(Feature Frequency Profile)을 활용한 악성 댓글의 판별시스템,” 한국정보기술

학회논문지, 제9권, 제1호, pp.207-216, 2011.

- [17] 홍진주, *인터넷 악성댓글 탐지 기법*, 숭실대학교 소프트웨어특성화대학원, 석사학위논문, 2016.
- [18] 여론집중도조사위원회, *2016~2018 뉴스 이용집중도 조사 결과 발표*, 문화체육관광부, 2018.
- [19] K. Reitz, Requests: HTTP for Humans™, <https://docs.python-requests.org/en/master/>.
- [20] L. Richardson, Beautiful Soup Documentation, <https://www.crummy.com/software/BeautifulSoup/bs4/doc>.
- [21] J. Huggins, SeleniumHQ Browser Automation, <https://www.selenium.dev>.
- [22] 하예람, 박민욱, 우균, “비속어와 높임말의 분포를 고려한 인터넷 뉴스 댓글 분석,” 2019년 한국정보과학회종합학술대회, 제9권, 제1호, pp.509-513, 2019.
- [23] 박은정, 조성준, “KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지,” 한글 및 한국어 정보처리 학술대회 논문집, 제26회, 2014.
- [24] 꼬꼬마 형태소 분석기, <http://kkma.snu.ac.kr>.
- [25] 한국게임산업진흥원, *게임언어 건전화 지침서 연구 - 게임 내 채팅언어의 표준화를 중심으로*, 2008.
- [26] 안태형, “악성 댓글의 범위와 유형,” 우리말연구, 제32권, pp.109-131, 2013.
- [27] 하예람, *비속어 분포를 고려한 뉴스 댓글 필터링 방법*, 부산대학교, 석사학위논문, 2020.
- [28] 정현정, “네이버카카오 욕설 DB 합친다. 중소 커뮤니티에도 제공,” 전자신문, <https://m.etnews.com/20210718000047>.

저 자 소 개

하 예 램(Yeram Ha)

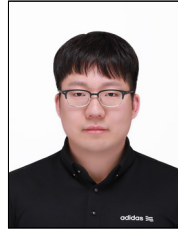
정희원



- 2009년 : 부산대학교 정보컴퓨터공학과(학사)
 - 2020년 : 부산대학교 전기전자컴퓨터공학과(석사)
 - 2007년 ~ 현재 : 경찰청 소속 국가직 공무원 재직중
- 〈관심분야〉 : 자연어 처리

천 준 석(Junseok Cheon)

정희원



- 2011년 : 동서대학교 컴퓨터공학과(학사)
- 2013년 : 부산대학교 전자전기컴퓨터공학과(석사)
- 2013년 ~ 현재 : 부산대학교 정보융합공학과(박사과정)

〈관심분야〉 : 프로그램 시각화, 데이터베이스, 블록 기반 프로그래밍, 함수형 언어

왕 인 서(Inseo Wang)

정희원



- 2021년 : 창원대학교 일어일문학과(학사)
- 2013년 ~ 현재 : 부산대학교 정보융합공학과(석사과정)

〈관심분야〉 : 자연어 처리, 기계학습, 인공지능

박 민 욱(Minuk Park)

정희원



- 1992년 : 1996 부산대학교 영어영문학과(학사)
- 1998년 : 부산대학교 인지과학협동과정(석사)
- 2000년 : 부산대학교 컴퓨터공학과(박사수료)
- 2001년 ~ 2015년 : 부산경상대학교 멀티미디어학과 부교수

■ 2017년 ~ 2020년 : 부산대학교 SW중심대학 산학협력교수
 ■ 2021년 ~ 현재 : (주)유시스 개발이사
 〈관심분야〉 : IT기반기술, 교육콘텐츠

우 균(Gyun Woo)

정희원



- 1991년 : KAIST 전산학(학사)
- 1993년 : KAIST 전산학(석사)
- 2000년 : KAIST 전산학(박사)
- 2000년 ~ 2004년 : 동아대학교 컴퓨터공학과 조교수
- 2004년 ~ 현재 : 부산대학교 정보컴퓨터공학부 교수

〈관심분야〉 : 프로그래밍 언어 및 컴파일러, 함수형 언어, 로봇 프로그래밍, 소프트웨어 메트릭, 프로그램 분석, 프로그램 시각화, 한글 프로그래밍 언어 개발, 자연어 처리