

온라인 토론의 댓글 응답 구조를 이용한 사용자 특성 분석

User Characterization from Replying Comment Structures in Online Discussion

김성환, 탁해성, 조환규
부산대학교 전기전자컴퓨터공학과

Sung-Hwan Kim(sunghwan@pusan.ac.kr), Haesung Tak(tok33@pusan.ac.kr),
Hwan-Gue Cho(hgcho@pusan.ac.kr)

요약

온라인 커뮤니티에서는 사용자들이 댓글을 이용하여 다양한 주제에 대한 의견과 감정을 교환한다. 댓글을 통한 의사소통은 신속하고 편리하지만 때로는 이러한 무게감이 덜한 특성이 사용자로 하여금 무례하고 공격적인 언사를 쉽게 행하도록 만들어 분쟁을 쉽게 유발하기도 한다. 따라서 이러한 분쟁을 미리 예측하고 대응하기 위하여 사용자들의 특성을 미리 파악하고 분류하는 작업이 중요하다. 본 논문에서는 이러한 사용자의 특성이 해당 사용자들이 참여한 발제글을 나타내는 댓글트리의 구조적인 특징에 나타난다는 가정을 바탕으로 댓글트리 구조를 서술하기 위한 여러 가지 정량적인 지표를 제안한다. 제안하는 정량 지표들의 분포를 통하여 발제글 작성자 및 댓글 작성자에 따른 지표의 분포를 살펴보고, 추가적으로 관리자에 의하여 경고를 받은 사용자들을 분류하는 실험을 통하여 제안하는 구조적 지표의 효과성을 보인다.

■ 중심어 : | 인터넷 커뮤니티 | 웹기반 토론 | 댓글 트리 | 사용자 분류 |

Abstract

In online communities, users use comments to exchange their opinions and feelings on various subjects. Communication based on comments is quick and convenient, but sometimes this light-weight characteristic makes users use impolite and aggressive words, which leads to an online conflict. Therefore, it is important to analyze and classify users according to their characteristics in order to predict and take action for this kind of troubles. In this paper, we present several quantitative measures for describing the structures of comments trees based on the assumption that the user characteristics be observed as a form of some structural feature in comment trees of articles in which they posted comments. We examine the distribution of the proposed measures over article posters and commenters, and in addition, we show the effectiveness of the presented structural features by conducting experiments to classify users who have received warnings of the administrator from benign users.

■ keyword : | Internet Community | Web-based Discussion | Comment Tree | User Classification |

1. 연구동기

온라인 커뮤니티에서는 다양한 사용자들이 시공간적

인 제약을 받지 않고 여러 주제에 대하여 동시다발적으로 의견과 감정을 교환한다[1]. 특히 많은 수의 온라인 사이트들이 게시판 형태의 의사소통 도구를 제공하는

* 이 논문은 부산대학교 자유과제 학술연구비(2년)에 의하여 연구되었음

접수일자 : 2018년 09월 18일

수정일자 : 2018년 10월 23일

심사완료일 : 2018년 10월 24일

교신저자 : 조환규, email : hgcho@pusan.ac.kr

양한 방법을 통해 독립적으로 연구되어왔다. 특히 트위터와 같은 소셜 네트워크 서비스에서 사용자들이 어떻게 군집화 되고 의견이 전파되는지에 대한 연구가 많이 이루어져 왔다[4]. Reddit[5]이나 Slashdot[1][6]과 같은 특정 커뮤니티의 특징에 대하여 분석을 수행한 연구도 있다. 온라인 토론에 대한 분석과 관련해서는 주로 자연어 처리 기법을 이용하여 의견의 대립[7]이나 특정 주제에 대한 입장[8]을 분류하는 기법이 소개되어 왔으며, 의견부분과 사실부분을 탐지함으로써 차후 논쟁 분석에 활용할 수 있도록 한 사례도 있다[9]. 토론 과정에서 정상적인 토론을 방해하는 반사회적 성향의 사용자에 대한 분석도 각각으로 이루어져왔다. 특히 반사회적 성향을 보이는 이유와 이와 연관되어 커뮤니티의 반응이 사용자의 행동을 어떠한 영향을 미치는지에 대한 연구가 있었다[10][11]. 본 논문과 가장 관련이 깊은 연구는 사용자가 댓글을 남기는 특성을 이용하여 사용자를 분류하는 연구이다[12]. 이 연구에서는 댓글의 개수, 활동 시간대, 평균 단어 및 문장 수, 대소문자 비율, URL 개수와 더불어 같은 글에 댓글을 작성한 다른 사용자와의 관계망상에서의 해당 사용자 노드의 차수와 최대 클리크(maximal cliques)에 포함되는 횟수 등을 특징으로 이용하여 기계학습을 통해 비정상사용자를 분류해내는 작업을 수행하였다. 그러나 이 연구에서는 계층적 댓글 시스템으로부터 얻을 수 있는 사용자간의 밀접한 상호작용을 고려하지 않았다는 한계점이 있다.

3. 댓글 트리 구조

본 장에서는 본 논문에서 사용할 댓글과 관련한 구조와 용어에 관하여 서술한다. 우선 발제글과 그에 달린 댓글들은 하나의 트리 $T=(V,E)$ 로 나타낼 수 있다. 발제글과 댓글들은 노드로 표현이 되며, 특히 발제글은 트리의 루트노드가 된다. 각각의 노드는 그에 해당하는 발제글 또는 댓글을 작성한 사용자에게 대한 식별자를 레이블로 가지고 있으며, 노드 v 에 대응하는 사용자를 $\lambda(v)$ 로 표현한다. 노드 간에 연결된 에지는 대응하는 댓글(또는 발제글)간의 답글 관계를 나타낸다. 어떤 댓글노드 u 에 대하여 댓글 v 가 달렸다면 트리 상에 에지

(u,v) 를 연결한다. 이 때 u 는 v 의 부모노드가 되며 $u=p(v)$ 로 표기하도록 한다. 루트노드는 r 로 표기를 하고, 노드의 깊이는 루트노드까지의 경로를 구성하는 에지의 수로 정의하며 $\delta(v)$ 라고 표기한다. 예를 들어, $\delta(r)=0$ 이고, 루트노드(발제글노드)에 바로 달린 자식 댓글노드 v 에 대해서는 $\delta(v)=1$ 이며 이들 노드는 특별히 Depth-1 노드(댓글)이라 명칭 한다. 자식노드가 없는 말단노드의 집합은 L 로 표기를 한다.

4. 데이터 수집 및 통계적 특성

분석을 위하여 인터넷 커뮤니티 중 하나인 PGR21의 자유게시판에서 2015년부터 2017년까지 3년간 작성된 발제글 17,516개와 그에 달린 댓글 975,500개를 수집하였다. 이를 위하여 파이썬 스크립트를 작성하여 게시판 목록상에 나타난 링크 중 해당 기간 내에 작성된 모든 발제글의 본문 페이지를 다운로드하여 파싱한 후 댓글 트리 정보를 추출하였다. 분석 대상으로서 이 사이트를 선택한 이유는 다음과 같다. 우선 이 커뮤니티에서는 사용자들이 계층적 댓글 시스템을 적극적으로 이용하여 토론을 수행한다. 또한 발제글의 생성률이 다른 사이트에 비하여 낮은데, 이는 최신 발제글들이 첫 페이지에 길게 머무르게 되어 사용자들이 발제글에 참여하여 충분히 의견을 나눌 수 있도록 한다. 짧은 시간 내에 대량의 발제글이 생성되는 커뮤니티에서는 발제글들이 금방 게시판 상의 다음 페이지로 넘어가기 때문에 극소수의 발제글을 제외하고는 많은 수의 사용자들이 발제글로 유입되기 힘들고 사용자들로부터의 관심도가 급격하게 떨어져 지속적인 참여와 토론을 어렵게 만든다. 또한 이 커뮤니티는 다른 사이트에 비하여 다소 엄격한 토론 규칙이 있으며, 규칙을 어기는 경우 운영자가 댓글을 차단하고 필요에 따라 사용자에게 벌점을 부여한다. 이와 관련하여, 1491개의 발제글에 걸쳐 있는 4,158개의 차단된 댓글 정보 역시 수집하였다. 또한, 사이트의 인터페이스에서는 댓글의 깊이가 제한되지만 최대 깊이에 있는 댓글들은 답글을 다는 상대방을 지칭하고 있어 이를 토대로 계층적 구조를 재구성하였다.

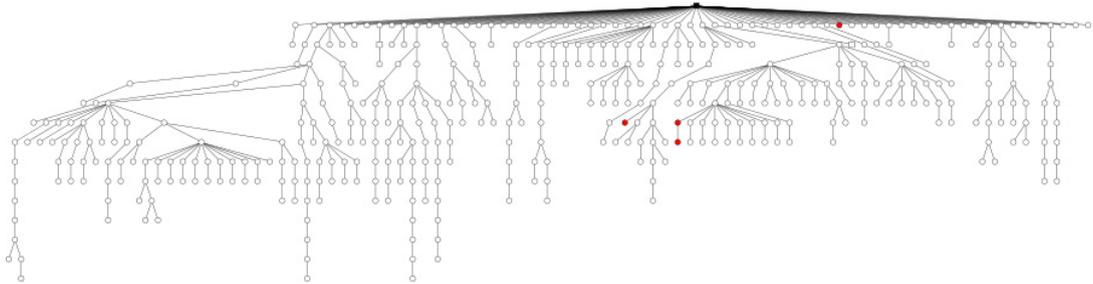


그림 2. 성차별과 관련된 발제글의 댓글 트리 예시
(댓글수=455, 사용자수=92, 최대깊이=14, 최대너비=Depth1댓글수=64, 빨간색 노드는 차단된 댓글)

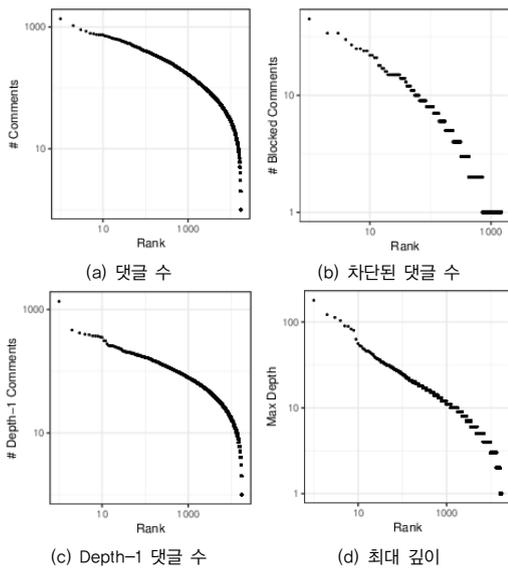


그림 3. 분석 데이터 특성. 발제글 별 댓글, 차단댓글, Depth-1 댓글, 최대 깊이를 내림차순으로 나타내었다.

[그림 2]은 성차별과 관련된 주제로 작성된 실제 데이터 상의 발제글의 댓글 트리를 나타낸다. 총 92명의 사용자가 455개의 댓글을 주고받으며 열띤 토론을 벌였고, 그 중 4개의 댓글은 비방 등을 이유로 차단당했다.

[그림 3]은 분석 대상이 되는 데이터에 대한 특성을 나타낸다. 발제글의 댓글수와 차단된 댓글 수, Depth-1 댓글 수, 최대 깊이를 순위에 따라서 내림차순으로 나타낸 결과이다. 가장 많은 댓글이 작성된 발제글은 댓글이 1,362개이며, 깊이는 최대 179까지 내려가기도 하였다.

5. 발제글 작성자 분류를 위한 특징 추출

본 장에서는 발제글 작성자들의 특성을 분석하기 위하여 각각의 발제글과 그에 달린 댓글들이 구성하는 댓글 트리의 전체적인 구조적인 특징을 나타내기 위한 정량적인 지표를 제안하고, 제안한 지표가 어떠한 분포 특성을 나타내는지를 살펴보고자 한다. 각 구조적 특징은 [표 1]에 요약되어 있다.

표 1. 발제글 작성자 분류를 위한 특징

특징	설명
댓글 수	$ V -1$
참여자 수	$\{\lambda(v) : v \in V - \{r\}\}$
발제자의 댓글 수	$\{v \in V - \{r\} : \lambda(v) = x\}$
Depth-1 댓글 수	$\{v \in V : p(v) = r\}$
최대 깊이	$\max_{v \in V} \{\delta(v)\}$
말단 노드 평균 깊이	$\frac{1}{ L } \sum_{v \in L} \delta(v)$
말단 노드 평균 너비	$ L E / \left(\sum_{v \in L} \delta(v) \right)$

5.1 댓글 트리의 구조적 특성 지표

먼저 발제글 작성자의 특성을 분석하기 위하여 개별 발제글에 해당하는 댓글트리의 구조적인 특성을 나타내기 위한 지표를 기술한다. 개별 댓글트리의 구조적인 특성을 정량적인 지표로 나타낼 수 있다면, 각각의 사용자가 게시한 발제글들의 댓글 트리 지표에 대한 합계 및 평균값 등을 이용하여 해당 사용자가 발제글 작성자로서 어떠한 특성을 지니고 있는지를 알 수 있다.

발제글의 속성 중 가장 자명한 것은 댓글 수와 참여

자 수이다. 댓글 수가 많을수록 해당 발제글에서 사용자들이 의견 표명을 많이 하였다는 것을 의미한다. 댓글이 많은데 참여자 수가 적은 경우는 소수의 참여자들이 열띤 논쟁을 벌이고 있다는 것을 의미하며, 댓글 수와 참여자 수가 거의 같다면 사실상 각 사용자가 큰 의미가 없는 일회성 댓글만을 작성했다는 것으로 해당 발제글에는 논쟁할 만한 주제가 크게 없는 것으로 볼 수 있다.

발제글을 작성한 발제자가 자신의 글에 얼마나 많은 댓글을 남겼는지도 발제글의 특성을 분석하기 위한 중요한 지표가 될 수 있다. 어떤 발제자는 논란이 되는 주제의 발제글을 자주 올리면서도 자신은 참여하지 않는 경우가 있는가 하면, 어떤 경우에는 발제자 자신이 적극적으로 개입하여 논쟁을 이끌어 나가는 경우도 있다. 전자는 주로 주요 이슈에 대한 최신 동향을 해당 커뮤니티로 옮기는 콘텐츠 생산자가 주 역할인 반면, 후자는 보다 적극적으로 의견을 개진하고 전파하는 경향이 더 크다고 할 수 있다. 그렇다고 해서 발제자가 개입한 횟수가 많을수록 반드시 그 발제글에 발제자를 중심으로 논쟁이 심각하게 일어난다고는 단언할 수 없다. 예를 들어 어떤 정보를 공유하기 위한 글에서 발제자의 댓글이 많으면 다른 사용자들의 호기심을 유발하는 콘텐츠를 생산하고 질문에 성실히 응답한다고 해석할 수도 있다.

댓글 트리의 깊이와 너비도 발제글의 특성을 분석하기 위한 중요한 지표가 된다. 댓글 트리의 깊이는 루트 노드에서 말단 노드에 이르는 가장 긴 경로의 길이로 사용자들 간에 댓글 교환이 얼마나 길게 지속되었는지를 나타낸다. 논쟁이 벌어진 경우 댓글 트리의 깊이가 깊을수록 사용자들 간에 의견공방이 더욱 치열하게 일어났다는 것을 의미하기도 한다. 댓글 트리의 너비는 다양한 방법으로 정의할 수 있다. [1][6]에서와 같이 같은 깊이를 가지는 댓글들의 최대 개수로 정의할 수도 있으나, 본 논문에서는 부모노드가 발제글 노드인 댓글들의 개수로 정의하였다. 발제글에 직접 달린 댓글을 발제글이 내재하고 있는 주제 중 일부에 대하여 의견을 표명함으로써 해당 세부 주제에 대한 토론을 개시한다는 의미가 있다. 또한 데이터 상에서 확인했을 때 발제

글에 직접 달린 댓글의 수와 같은 높이에서의 댓글 수가 최대인 값의 차이가 실질적으로 존재하지 않았다.

마지막으로 말단노드들의 평균 깊이와 평균 너비가 있다. 한 댓글에 댓글이 두 개 이상 달린다는 것은 해당 댓글의 세부 주제가 다시 그 자식노드 수만큼 분화한다고 볼 수 있는데, 이러한 관점에서 볼 때 말단노드는 최종적으로 분화된 세부주제들 중 하나의 세부주제에 대한 토론이 종결된 시점을 나타낸다. 다시 말해 말단노드의 평균 깊이는 각각의 세부주제들에 대하여 평균적으로 얼마나 심도 있게 토론이 이루어졌는지를 의미하며, 평균 너비는 발제글의 주제로부터 각 세부주제로 얼마나 다양하게 분화했는지를 나타내는 지표이다. 평균 너비에 대하여 보다 구체적으로 설명하면 다음과 같다. 우선 루트 노드로부터 말단노드 v 에 이르는 경로의 길이는 정의된 바에 따라 $\delta(v)$ 와 같다. 이를 모든 말단노드에 대하여 합한 값 $\sum_{v \in L} \delta(v)$ 을 트리의 에지 수 $|E|$ 로 나누게 되면 루트 노드에서 말단 노드까지의 경로들이 공유하는 평균 에지 수가 나오게 된다. 이 값이 높을수록 루트-말단 경로 간에 공유하는 에지의 수가 많다는 것을 의미한다. 이 값에 역수를 취하면 전체 에지의 비중을 1로 두었을 때, 각 루트-말단 경로가 에지를 나누어 쓰는 비율이 된다. 만약 모든 루트-말단 경로가 에지를 전혀 공유하지 않는다면 각 경로가 에지를 온전히 사용하게 되므로 이 값은 1이 된다. 그러나 경로들 간에 공유하는 에지가 많아질수록 이 값은 $1/|L|$ 에 가까워진다. 여기에 전체 말단 노드의 수를 곱하게 되면 분기 시점으로 보정된 트리의 너비를 얻을 수 있게 된다.

T	T_a	T_b	T_c	T_d
댓글수	9	5	5	5
최대깊이	3	3	3	3
Depth-1	3	1	3	1
평균깊이	3,000	3,000	1,667	2,667
평균너비	3,000	1,667	3,000	1,874

그림 4. 댓글 트리 구조적 지표의 예

평균 너비 값이 낮다는 것은 공유하는 에지 비율이 높다는 것이며 이는 댓글 작성자의 관점에서 봤을 때 해당 세부주제들의 공통주제가 있고, 시간적으로 나중에 가야 그 공통주제로부터 분화가 되었다는 것을 의미한다. 반대로 평균 너비 값이 높다는 것은 공통 경로의 비율이 낮으며, 다시 말해 토론이 게시된 시점으로부터 얼마 지나지 않아 세부주제들이 분화되었다는 것을 의미하며 그만큼 발제글의 자체가 여러 개의 세부주제들을 복합적으로 지니고 있다고 볼 수 있다.

[그림 4]는 댓글 트리의 구조적 특성을 나타내는 제안지표 중 댓글 수, 최대 깊이, Depth-1 댓글 수, 평균 깊이, 평균 너비를 나타낸 것이다. 트리 T_a 는 3개의 독립적인 세부주제에 대한 토론이 각각 3번의 댓글교환을 통해 일어난 것을 의미하며, T_b 는 하나의 세부 주제에 대한 토론이 2번의 댓글교환을 통해 일어나다가 마지막에 다시 3개의 추가적인 세부 주제로 분화되는 것을 나타내고 있다. T_c 는 처음에는 3개의 세부주제 토론이 시작되었지만 그중 하나만 추가적인 댓글교환이 일어난 상태를 나타낸다. 댓글 트리 T_d 는 기존에 주로 사용되는 댓글수나 최대깊이와 같은 지표로는 T_b 와 구분이 잘 안되지만 말단노드 평균깊이와 평균너비로는 구분이 가능한 예를 보여주고 있다.

5.2 발제글 작성자에 따른 특성 지표의 분포

각 발제글 작성자별 특성 지표의 분포가 [그림 5]에 나타나 있다. [그림 5-(a)]는 각 사용자가 작성하여 게시한 발제글의 수이고, 나머지는 댓글트리들의 구조적 지표값의 평균값을 발제글 작성자별로 나타낸 결과이다. 모든 그래프는 지표값을 내림차순으로 정렬하여 양좌표축을 로그 스케일로 나타내었다.

발제글을 최소 1개 이상 작성한 사용자는 총 2492명이었으며, 상위 10위에서 1000위 사이의 값에 대하여 로그-로그 그래프 상에서 선형회귀를 수행하여 각 그래프에 그 기울기 계수 k 를 표시하였다. 일반적으로 지프(Zipf) 법칙과 같은 특성을 지닌 데이터에서 최상위와 최하위 데이터가 선형성에서 다소 벗어나는 경향이 있는데 본 논문의 데이터도 이러한 성향이 관찰되어 이

를 반영하여 선형성이 뚜렷한 부분에 대해서만 회귀를 수행한 것이다. 전반적으로 로그-로그 그래프 상에서 선형성을 보이는 구간이 존재하며, 최하위 값들은 급격하게 감소하는 성향을 보이는 것을 확인하였다.

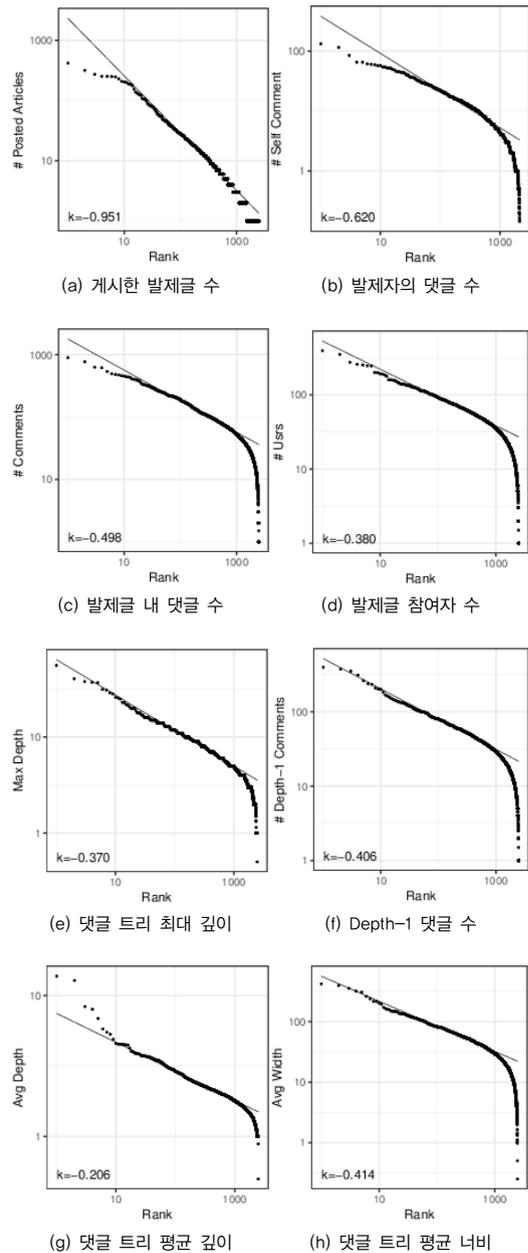


그림 5. 발제글 작성자 별 구조적 특성 지표 값의 분포

6. 댓글 작성자 분류를 위한 특징 추출

본 장에서는 댓글 작성자들의 특성을 분석하기 위한 방법으로서 각각의 사용자가 작성한 댓글들이 지니는 댓글 트리 상에서의 구조적인 특징에 관하여 서술한다.

표 2. 댓글 작성자 분류를 위한 특징

특징	설명
댓글 수	$V_x = \{v \in V - \{r\} : \lambda(v) = x\}$
댓글 받은 수	$\{v \in V : \lambda(p(v)) = x\}$
Depth-1 댓글 수	$\{v \in V_x : p(v) = r\}$
말단 댓글 수	$\{v \in V_x : C(v) = \phi\}$
부모 댓글의 서로 다른 사용자 수	$\{\lambda(u) : v \in V_x, u = p(v), u \neq r\}$
자식 댓글의 서로 다른 사용자 수	$\{\lambda(v) : v \in V, \lambda(p(v)) = x, p(v) \neq r\}$
댓글 최대 깊이	$\max_{v \in V_x} \delta(v)$
댓글 간 최대 길이	$\max\{\delta(v) - \delta(u) : u, v \in V_x, u \in \text{ancestor}(v)\}$

6.1 발제글 내 댓글 작성자에 대한 구조적 특성 지표

댓글 트리 내에서 특정 사용자 x 가 작성한 댓글에 대한 노드들의 집합을 V_x 라고 하자. 본 절에서는 주어진 발제글에 대한 댓글 트리 $T = (V, E)$ 내의 각각의 댓글 작성자 $x \in \{\lambda(v) : v \in V - \{r\}\}$ 에 대하여 해당 사용자가 작성한 댓글들의 집합 V_x 이 지니는 구조적인 특성을 나타내기 위한 지표들을 살펴보고자 한다.

우선 가장 간단한 지표 중의 하나는 해당 사용자가 댓글을 얼마나 많이 작성했고, 얼마나 많은 대댓글을 받았는지에 대한 것이다. 댓글을 많이 작성했을수록 그 사용자의 참여도가 높다고 할 수 있으며, 많은 대댓글을 받을수록 다른 사용자의 이목을 쉽게 끄는 사용자라고 할 수 있다.

부모가 발제글 노드인 댓글, 즉 Depth-1 댓글을 얼마나 작성했는지도 중요한 지표가 될 수 있다. 이는 발제글 내에서 해당 사용자가 세부주제에 대한 토론을 얼마나 많이 게시했는지를 알려준다. 또한 사용자가 작성한 말단 댓글, 즉 대댓글을 받지 못한 댓글의 수는 이 사용

자가 얼마나 많은 세부토론을 마무리 지었는지를 알려준다. 예를 들어 어떤 사용자의 말단 댓글 수가 다른 사용자에게 비하여 아주 높다면, 이 사용자는 일반적으로 상대방이 더 이상 자신의 댓글에 대댓글을 달지 않을 때까지 반박을 절대 멈추지 않는 열띤 논쟁꾼이라고 볼 수도 있다.

부모 댓글과 자식 댓글의 작성자 수도 중요한 정보를 지니고 있다. 예를 들어 어떤 발제글 내에서 한 사용자가 작성한 댓글의 부모 댓글의 작성자가 모두 동일하다면, 이 사용자는 특정 사용자만 노려서 반박 댓글을 작성하고 있는 것이라고 볼 수 있다. 한편 부모 댓글이나 자식 댓글의 서로 다른 작성자 수가 매우 많다면 이 사용자는 좋은 의미이든 나쁜 의미이든 간에 많은 사람의 관심을 불러일으키거나 그렇게 하려고 노력하고 있다는 것을 나타낸다.

댓글의 최대 깊이는 이 사용자가 얼마나 오랫동안 세부 주제에 대한 토론에 참여하는지를 의미한다. 만약 댓글을 작성하지 않다가 말단노드 부근에서 댓글을 작성했다고 하더라도 이 사용자는 이전 부모 노드들의 흐름을 읽으면서 따라왔다는 의미이기 때문에 최대 깊이 자체가 의미를 가질 수 있다. 이와 비슷하게 댓글 간 최대 길이는 루트노드로부터 말단노드에 이르는 경로 상에서 이 사용자가 작성한 댓글 중 루트노드에 가장 가까운 댓글과 말단노드에 가장 가까운 댓글 사이의 거리를 의미하며, 이는 이 사용자가 자신이 직접적으로 개입한 토론에 대하여 마지막까지 충실히 토론에 참여하는지에 대한 지표로 활용될 수 있다.

6.2 댓글 작성자에 따른 특성 지표의 분포

댓글 작성자의 특성분포는 크게 전체 댓글 수와 같이 게시판 규모에서의 합계치와 각 참여 발제글 당 평균 댓글수와 같은 평균치, 두 그룹으로 나눌 수 있다. 먼저 전체 댓글 수, 답글 받은 수, Depth-1 댓글 수, 말단 댓글 수, 참여 발제글 수와 더불어 차단된 댓글 수에 따른 사용자 수를 [그림 6]에 로그-로그 그래프로 나타내었다. 예를 들어 통틀어 1개의 댓글을 작성한 사용자의 수는 [그림 6-(a)]의 x 좌표 값이 가장 작은 점에 해당하며, 이는 1,080명에 달한다.

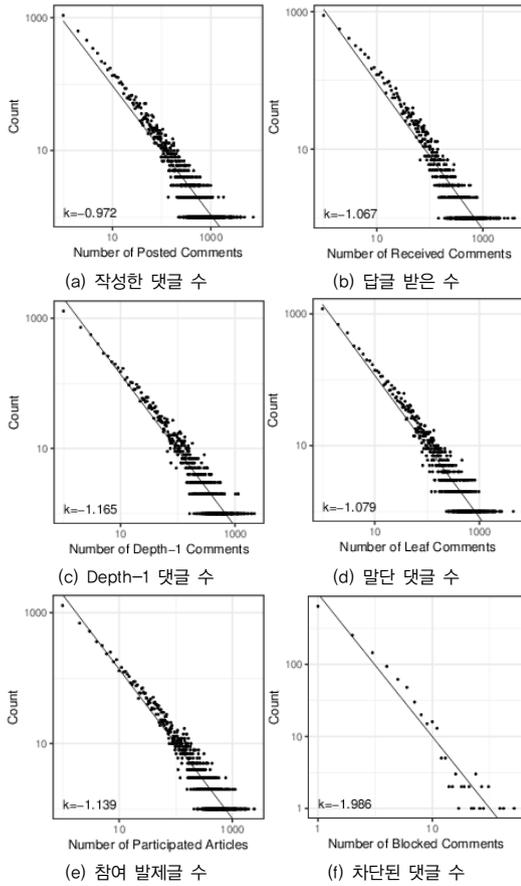


그림 6. 댓글 작성자 별 구조적 특성 지표 값의 분포

이들 지표는 로그-로그 그래프 상에서 강한 선형성을 띠는 것을 알 수 있다. 이는 각 지표들이 멱법칙 (Power Law)를 따른다는 것을 알 수 있으며, 선형회귀 결과 계산된 계수(직선의 기울기)를 각 그래프 좌측 하단에 표시하였다.

[그림 7]은 각 댓글 작성자가 자신이 참여한 발제글 당 평균적으로 작성한 댓글 수, 받은 댓글 수, Depth-1 댓글 수, 말단 댓글 수, 부모 및 자식 댓글의 서로 다른 작성자 수, 댓글 깊이 및 댓글 간 최대 길이를 내림차순으로 정렬하여 로그-로그 그래프 상에 순서대로 나타내었다. 발제글 작성자의 지표와 마찬가지로 최상위 값과 최하위값을 제외하고 10위에서 1,000위 사이의 값들에 대하여 선형회귀를 수행한 후 그 계수를 각 그래프마다 표시하였다.

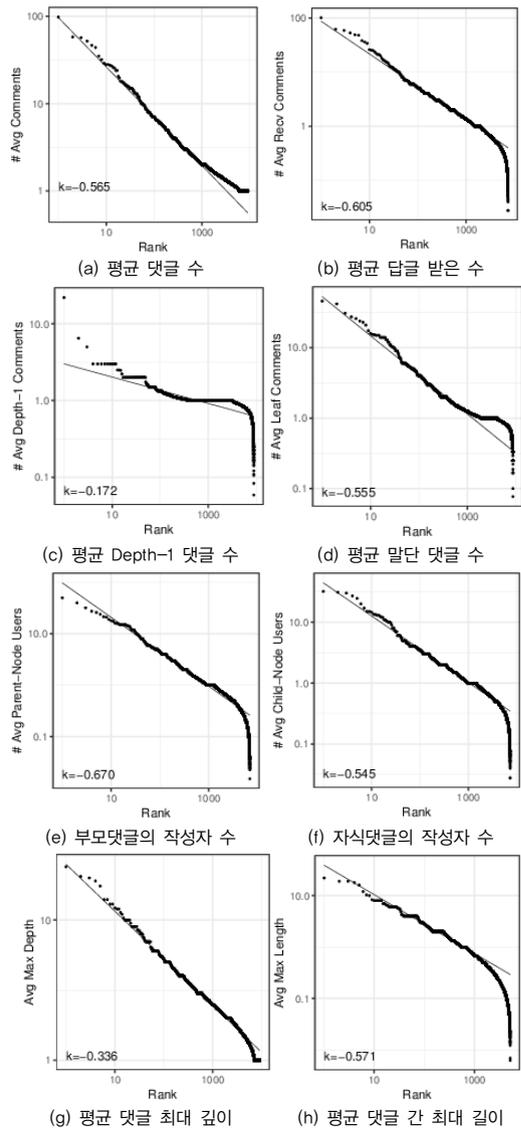


그림 7. 댓글 작성자 별 구조적 특성 지표 평균값의 분포

특히 평균 댓글 수와 평균 댓글 최대 깊이는 데이터 전반에 걸쳐서 강한 선형성을 나타낸다는 사실이 직관적으로 확인이 되고, 다른 지표들은 모두 비슷한 분포를 보이나, 평균 Depth-1 댓글 수는 1.0 부근에 상대적으로 많은 수의 사용자들이 있는데, 이는 해당 사용자들이 발제글에 직접 댓글을 단 후, 추가적으로 토론을 더 하고 싶으면 해당 댓글에 답글을 다는 식의 참여를 하는 경향이 있다는 것을 확인할 수 있는 부분이다.

7. 실험 및 결과

실험을 위하여 관리자로부터 차단된 댓글 수를 기준으로 발제글 게시자 및 댓글 작성자를 각각 두 그룹으로 분할하였다. 우선 발제글 게시자를 살펴보면, 만약 어떤 발제글 게시자가 작성한 발제글 중 차단된 댓글이 발생한 발제글의 수가 많다면, 이 발제글 게시자가 주로 게시하는 내용에 분란의 소지가 많다는 것을 의미한다. 따라서 자신이 작성한 발제글 중 θ 개 이상의 발제글에 차단된 댓글이 하나 이상씩 존재하는 발제글 게시자들과 차단된 댓글이 하나도 존재하지 않는 발제글만 작성한 발제글 게시자(일반 사용자)를 분류하는 실험을 수행한다. 마찬가지로, 댓글 작성자의 경우 θ 개 이상의 발제글에서 자신의 댓글이 차단된 댓글 작성자와, 차단된 댓글 한 번도 당하지 않은 댓글 작성자(일반 사용자)를 분류하는 실험을 수행한다. θ 값에 따라 해당되는 발제글 게시자와 댓글 작성자의 수는 [표 3]과 같다. 각 θ 값에 대한 대상 사용자 및 이와 같은 수의 일반 사용자를 혼합하여 데이터 집합을 구성하는데, 일반 사용자를 선택하는 과정에서 전체 일반 사용자 집합으로부터 무작위로 선택할 수도 있으나(Random User로 표시), 실제로 분쟁을 유발시키는 사용자는 매우 활동적인 사용자일 것이므로 참여 게시글, 댓글 수 등을 기준으로 정렬한 후 상위 θ 에 속한 사용자들만을 추려내어 이용할 수도 있다(Active User로 표시). 이렇게 구성된 데이터 상에서 SVM(Support Vector Machine)을 사용하여 학습을 시킨다. 본 실험에서는 R에서 제공하는 e1071 패키지에서 제공하는 SVM 구현을 사용하였으며, 커널은 가우시안 커널을 사용하였다. 또한, 교차검증을 위하여 Leave-one-out 방식을 이용한다. 즉, 구성된 사용자 집합으로부터 1명을 제거한 후 나머지에 대하여 학습을 수행한 후 해당 사용자를 올바르게 분류하는지를 전체 사용자에 대하여 평가하여 평균 정확도를 계산한다.

우선 발제글 게시자 분류를 위한 비교실험 및 그 결과를 분석한다. 이와 관련된 기존 연구는 없으므로 본 논문에서 제시한 특징들의 서로 다른 조합을 [표 4]와 같이 구성하여 학습 및 평가를 수행하였다. [그림 8-(a)] 및 (b)에 실험결과가 나타나 있다. 일반 사용자

표 3. θ 에 따른 분류 대상 사용자 수

(a) 발제글 게시자							
θ	1	2	3	4	5	6	7
n	649	234	133	86	57	42	28

(b) 댓글 작성자								
θ	2	3	4	5	6	7	8	9
n	648	379	239	156	102	67	40	27

표 4. 발제글 게시자 분류를 위한 특징 집합

특징집합	사용한 구조적 특성 지표
FS1	발제글 수, 평균 댓글 수, 평균 사용자 수
FS2	FS1+평균길이, 평균너비
FS3	FS2+최대길이, 발제자댓글수, Depth1댓글수

집합을 무작위로 추출한 경우에는 일반 사용자에 해당하는 발제글 작성자들이 올린 발제글 수나 인기가 분쟁이 일어난 발제글을 올린 사용자에 비하여 턱없이 낮기 때문에 가장 간단한 특징 집합인 FS1이 가장 높은 성능을 보이는 것이 관찰되었다. 이에 반하여 일반 사용자 집합을 활동적인 사용자들로 구성된 실험결과에서는 $\theta < 3$ 으로 작거나, 반대로 큰 $\theta \geq 7$ 의 경우에 전체 특성 지표를 모두 조합함으로써 성능 향상을 얻을 수 있음을 확인하였다. 이는 $4 \leq \theta \leq 6$ 에 해당하는 발제글 게시자들 중에 특이한 행동양식을 보이는 사용자가 포함되어 있어 학습에 방해가 된 것일 가능성을 나타내기도 한다.

댓글 작성자 분류를 위해서는 기존의 관련 연구[12]에서 사용된 특징을 사용하여 학습한 결과와 6장에서 기술한 구조적 특성 지표를 사용하여 학습한 결과에 대한 비교 분석을 수행하였다. [그림 8-(c)]와 [그림 8-(d)]에 기존 기법[12]에 대한 결과는 BASE로, 제안 기법은 OURS로 표기하여 나타내었다.

무작위로 추출한 일반사용자와의 분류실험에서는 낮은 $\theta \leq 5$ 에서는 제안 기법이 우수하나 이 값이 높아질수록 기존 기법과의 성능 차이가 줄어들거나 역전되는 결과를 얻었는데, 기존 기법의 특징들이 사용자의 활동성을 중심으로 기술되었기 때문이다. 반대로 활동적인 사용자를 일반사용자로 구성된 분류실험에서는 [그림 8-(d)]와 같이 제안 기법이 항상 우수한 성능을 보인다.

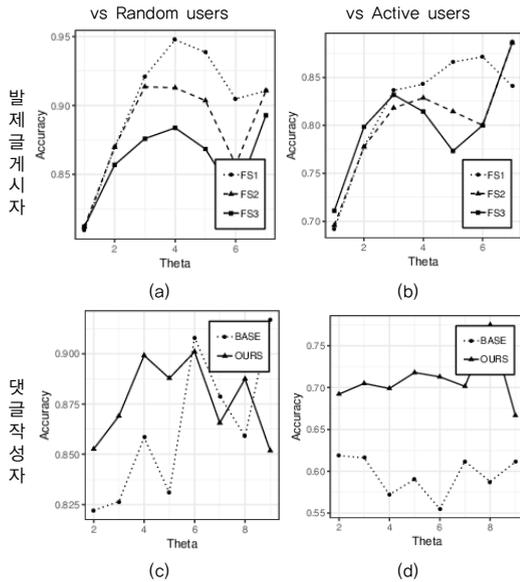


그림 8. 사용자 분류 실험 결과

SVM 이외에 의사결정트리인 C5.0를 사용하여 동일한 실험을 수행하였으나, 전반적인 결과는 큰 차이가 없었다. 발제글 게시자 분류에서는 활동적인 사용자 내에서 낮은 θ 값의 경우 모든 특징 조합을 사용한 경우에 우수한 성능을 보였으며, 댓글 작성자 분류는 무작위 사용자 집합과의 실험에서 낮은 θ 값에서 높은 정확도를 보였고, 활동적인 사용자와의 실험에서는 마찬가지로 모든 경우 기존 기법에 비하여 우수하였다.

추가 실험으로서 일반 사용자를 경고를 한 번도 받지 않은 모든 사용자(발제글 게시자 1842명, 댓글 작성자 7712명)으로 두고 $\theta = 2$ 에 따른 사용자 검출 실험을 수행하였다. 마찬가지로 가우시안 커널을 사용한 SVM을 이용해 학습하였으며, 이번에는 10-fold 교차검증을 수행한 뒤 Precision과 Recall을 이용하여 F1 score를 측정하였다. 발제글게시자 분류실험에서는 모든 특징을 조합한 FS3가 73.1%의 성능을 보이며 FS1, FS2 조합을 각각 1.7%p, 1.5%p차로 우세함을 보였고, 댓글작성자와 분류실험에서는 제안 기법의 성능이 47.3%, 기존기법이 43.8%로 측정되었으며, 제안기법의 Precision과 Recall이 기존기법에 비하여 각각 3.1%p, 2.8%p 높음을 확인할 수 있었다.

8. 결론

본 논문에서는 댓글로 의사소통을 하며 의견을 주고 받는 온라인 커뮤니티에서 사용자의 특징이 그 사용자가 속한 발제글의 댓글 트리가 가지는 구조적인 특성으로 나타난다는 가정 하에 댓글 트리를 기반으로 발제글 게시자 및 댓글 작성자 분류를 위한 구조적 지표들을 제안하였으며, 실험을 통해 밝혀낸 사실은 다음과 같다.

1. 제안한 구조적 지표들이 로그-로그 그래프 상에서 선형성을 가지는 것을 관찰함을 통하여 해당 지표에 멱법칙의 특성이 나타나는 것을 확인하였다.
2. 제안한 구조적 지표들은 활동적인 사용자들 사이에서 관리자의 개입을 많이 받은 사용자를 분류해 내는 데에 적합하다는 사실을 실험을 통해 확인하였다. 즉, 관리자의 개입을 많이 받은 사용자들이 댓글 트리의 구조적인 특성에 유의미한 영향을 미친다는 사실을 확인할 수 있었다.
3. 구조적인 특성을 이용한 사용자 분류 작업은 발제글 게시자에 비하여 개별적인 댓글 사용자에게 더욱 효과적임을 실험적으로 확인하였다. 특히 댓글 작성자 분류실험에서는 기존 기법에 비하여 최대 18.8%p의 성능 향상을 보였다.

참고 문헌

- [1] S. Gonzalez-Bailon, A. Kaltenbrunner, and R. E. Banchs, "The Structure of Political Discussion Networks: A Model for the Analysis of Online Deliberation," *Journal of Information Technology*, Vol.25, No.2, pp.230-243, 2010.
- [2] J. Cheng, M. Bernstein, C. Danescu-Niculescu-Mizil, and J. Leskovec, "Anyone Can Become a Troll: Cause of Trolling Behavior in Online Discussions," in *Proc. ACM CSCW*, pp.1217-1230, 2017.
- [3] E. E. Buckels, P. D. Trapnell, and D. L. Paulhus, "Trolls Just Want to Have Fun," *Personality and Individual Differences*, Vol.67, pp.97-102,

2014.

[4] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito, "Measuring Political Polarization: Twitter Shows the Two Sides of Venezuela," *Chaos*, Vol.25, No.033114, 2015.

[5] P. Singer, F. Flöck, C. Meinhart, E. Zeitfogel, and M. Strohmaier, "Evolution of Reddit: From the Front Page of the Internet to a Self-Referential Community," In Proc. WWW, pp.517-522, 2014.

[6] V. Gómez, A. Kaltenbrunner, and V. López, "Statistical Analysis of the Social Network and Discussion Threads in Slashdot," in Proc. WWW, pp.645-654, 2008

[7] K. Allen, G. Carenini, and R. T. Ng, "Detecting Disagreement in Conversations using Pseudo-Monologic Rhetorical Structure," in Proc. EMNLP, pp.1160-1180, 2014.

[8] M. Qui, Y. Sim, N. A. Smith, and J. Jiang, "Modeling User Arguments, Interactions, and Attributes for Stance Prediction in Online Debate Forum," In Proc. SDM, pp.855-863, 2015.

[9] P. Biyani, S. Bhatia, C. Caragea, and P. Mitra, "Using Non-Lexical Features for Identifying Factual and Opinionative Threads in Online Forums," *Knowledge-Based Systems*, Vol.69, pp.170-178, 2014.

[10] J. Cheng, C. Danescu-Niculescu-Mizil, and J. Leskovec, "How Community Feedback Shapes User Behavior," In Proc. ICWSM, pp.41-50, 2014.

[11] B. A. Coles and M. West, "Trolling the Trolls: Online Forum Users Constructions of the Nature and Properties of Trolling," *Computers in Human Behavior*, Vol.60, pp.233-244, 2016.

[12] T. C. Li, J. Gharibshah, E. E. Papalexakis, and M. Faloutsos, "TrollSpot: Detecting Misbehavior in Commenting Platforms," in Proc. IEEE/ACM ASONAM, pp.171-175, 2017.

저 자 소 개

김 성 환(Sung-Hwan Kim)

정회원



- 2011년 : 부산대학교 정보컴퓨터공학과(공학사)
- 2013년 : 부산대학교 컴퓨터공학과(공학석사)
- 2018년 : 부산대학교 전기전자 컴퓨터공학과(공학박사)
- 2018년 ~ 현재 : 부산대학교 IT기반융합산업창업인력양성사업단 연수연구원
- <관심분야> : 데이터마이닝, 알고리즘

탁 해 성(Haesung Tak)

정회원



- 2013년 : 부산대학교 정보컴퓨터공학과(공학사)
- 2015년 : 부산대학교 전기전자 컴퓨터공학과(공학석사)
- 2015년 ~ 현재 : 부산대학교 전기전자컴퓨터공학과 박사과정
- <관심분야> : 데이터마이닝, 데이터시각화

조 환 규(Hwan-Gue Cho)

정회원



- 1984년 : 서울대학교 계산통계학과(이학사)
- 1986년 : KAIST대학원 전산학과(공학석사)
- 1990년 : KAIST대학원 전산학과(공학박사)
- 1990년 ~ 현재 : 부산대학교 전기컴퓨터공학부 교수
- <관심분야> : 알고리즘, 그래프이론, 생물정보학