

01

2021 vol. 601

저널리즘의 품격

줄어들던 TV와 PC 이용, 코로나19로 감쪽 받은

30주년 맞은 <PD수첩>

댓글 작업 흔적, 추적 가능하다

유해 콘텐츠 규제·디지털 광고 독과점 방지에 열인하는 영국 정부

한국언론진흥재단
Korea Press Foundation



신문과 방송

NEWSPAPER & BROADCASTING



댓글 작업 흔적, 추적 가능하다 완벽한 댓글 정책은 없다... 사회 전체가 함께 풀어갈 숙제

정지완 / KAIST 전산학부 박사과정

포털은 뉴스나 정보를 찾기 위한 관문으로 누구나 한 번쯤은 거치는 공간이다. 이용자가 많은 만큼 포털 뉴스의 댓글 조작 논란은 뜨거운 감자였다. 사람들은 어떤 작업을 통해 댓글을 조작할까, 그리고 이를 해결하기 위한 해법은 무엇인지 알아본다. 편집자 주



2013년 7월 네이버는 뉴스 댓글을 순공감순(공감-비공감)으로 정렬하기 시작했다. 공감이 높을수록 상단에 노출시키되, 비공감이 많은 만큼 감점하는 방식이다. 순공감순 정렬은 네이버만의 독특한 정책이 아니었다. 다음도 이렇게 댓글을 정렬했다. 야후를 비롯한 세계 각국의 주요 포털도 대부분 비슷한 방식으로 댓글을 정렬했다.

순공감순 정렬은 효과적인 댓글 운영 정책으로 꼽히기도 했다.²⁾ 포털에는 매일 수십 수백만 건의 댓글이 달린다. 관리자가 검토하는 데는 한계가 있다. 이용자들이 클릭한 공감, 비공감은 관리자 역할을 대신했다. 많은 이용자는 좋은 댓글에 공감을 늘

려 상단으로 올리는 한편, 악성 댓글에는 비공감을 눌러 노출도를 낮췄다.

2015년 12월 네이버는 호감순(공감-3×비공감)으로 정렬 기준을 바꿨다. 비공감의 비중이 커졌기 때문에 악성 댓글을 더 효율적으로 걸러낼 수 있었지만, 단지 반대 의견의 댓글을 보이지 않게 만들기 위해 악용될 수도 있었다.

1) 이 글의 주요 내용은 다음 논문의 일부를 풀이한 것이다. Jeong, J., Kang, J., & Moon, S., <Identifying and Quantifying Coordinated Manipulation of Upvotes and Downvotes in Naver News Comments>, Proceedings of the International AAAI Conference on Web and Social Media, 14, pp.303-314, 2020.

2) World Editors Forum, <Online Comment Moderation: Emerging Best Practices>, World Association of Newspapers and News Publishers, 2013.

이 시기에 강남역 살인사건, 국정농단 의혹 등 굵직한 사건들이 터져 나왔다. 이는 여초-남초, 진보-보수 커뮤니티 사이의 집단적 댓글싸움으로 번지기도 했다. 구체적 지시사항이 담긴 '댓글 모니터 요원 매뉴얼'이 유출되는 일도 있었다. 뿐만 아니다. 매크로를 이용한 여론 조작 사건이 밝혀지기도 했다.

흔히들 '댓글 조작'이라 말하는 이 작업의 핵심은 댓글 자체보다는 공감 표시에 있다. 여러 계정을 동원해 단시간에 공감과 비공감을 퍼부어 댓글의 순위를 뒤엎거나 인기도를 왜곡하는 것이다.

△집단적인 댓글 작업은 어떤 흔적을 남길까?
△그런 흔적이 남아 있는 기사는 얼마나 많을까?

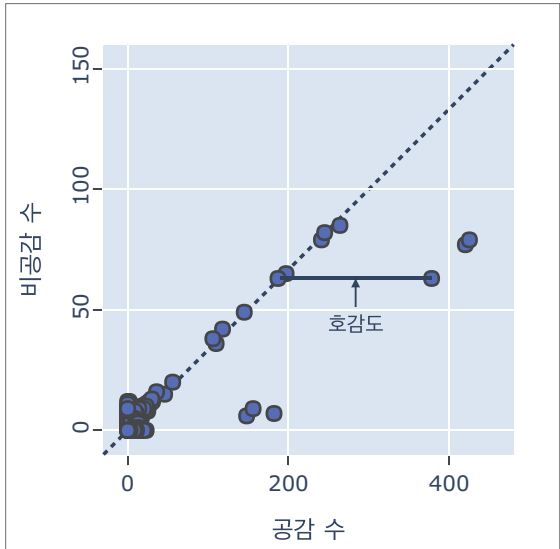
이에 답하기 위해 네이버 뉴스의 댓글란을 분석했다. 몇 개의 온라인 커뮤니티에서 '좌표 찍힌 기사'를 수집해 댓글 작업의 흔적을 추론했고, 이를 바탕으로 네이버 뉴스 전체에서 그 흔적을 찾았다. 이 글을 통해 그 결과의 일부를 공유한다.

영점댓글이란



[그림 1]은 좌표 찍힌 기사 한 건에 달린 모든 댓글에 대한 공감·비공감 산포도다. 각 점이 하나의 댓글을 나타내며, 가로축은 해당 댓글이 받은 공감 수, 세로축은 비공감 수를 나타낸다. 이 기사가 올라온 당시 댓글은 호감순으로 정렬됐다. 이 점수가 0이 되는 지점을 점선으로 표시했다. 이 대각선에서 댓글을 나타내는 점까지의 수평거리가 그 댓글의 호감도이다.

점들의 분포를 살펴보자. 우선 일반적인 댓글란의 특징이 보인다. 소수의 댓글이 대각선에서 오른쪽으로 멀리 떨어져 있다. 높은 호감도를 가진 베스트 댓글들이다. 수많은 댓글이 원점 근처에 몰려 있다. 대부분의 사용자는 기사별 최상단 몇 개의 댓글



[그림 1] 좌표 찍힌 댓글란의 공감·비공감 산포도 <출처 - 필자 제공>

만 읽기 때문에³⁾ 나머지 댓글은 많은 공감 표시를 받지 못한 채 묻힌다.

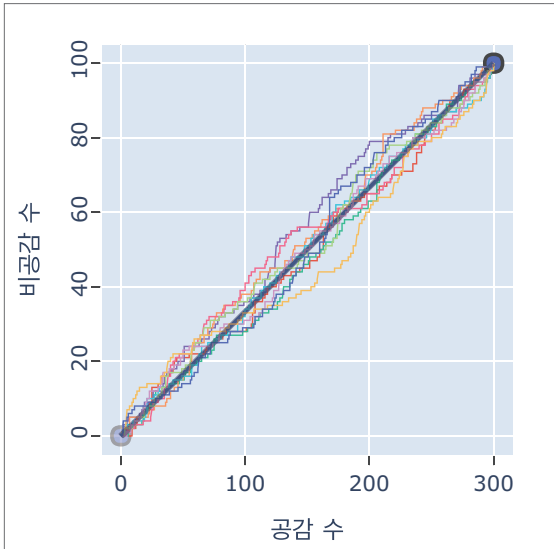
그런데 이 좌표 찍힌 기사에는 독특한 댓글들이 보인다. 여남은 개의 댓글이 0점을 나타내는 대각선에 줄지어 있다. 공감도 많이 받았지만 비공감도 많이 받아 0에 가까운 호감도를 갖게 된 것이다. 이런 댓글을 '영점댓글'이라 부르자. 많은 공감 표시를 받은 영점댓글, 좌표 찍힌 기사에서는 흔하게 나타났다. 하지만 대부분의 기사에는 존재하지 않았다. 영점댓글이 집단적 공감 표시의 흔적일까? 그렇다면 그 이유를 어떻게 설명할 수 있을까?

작업의 흔적일까

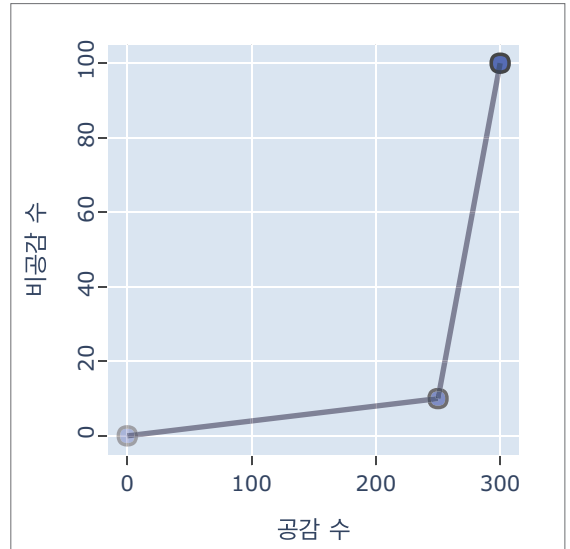


작업이 벌어지면 특정 성향의 독자들이 단시간에 표적 기사에 몰려와 공감, 비공감을 클릭한다. 반면 일반적인 상황에서는 독자들이 성향에 따라 멍치지 않고 개별적으로 댓글 활동을 할 것이다.

3) 김선호·오세욱, <포털 뉴스서비스 및 댓글에 대한 인터넷 이용자 인식조사>, 미디어 이슈, 4(5), 2018.



[그림 2] 댓글 작업이 없을 때 영점댓글의 성장 <출처 - 필자 제공>



[그림 3] 댓글 작업이 있을 때 영점댓글의 성장 <출처 - 필자 제공>

위 가정을 영점댓글에 적용해 보자. 독자들이 개별적으로 활동했다면 특정 댓글을 좋아하는 독자와 싫어하는 독자들이 뒤섞여 기사를 보러 온다. 그렇다면 댓글이 받는 공감과 비공감도 뒤섞여 클릭될 것이다. [그림 2]는 공감 수가 높은 영점댓글에 대해 이 과정을 10회 시뮬레이션한 결과이다. 대부분의 경우 호감도는 항상 0점 근처에 머문다. 호감도가 높았던 적이 없는데 베스트 댓글에 오를 수 있었을까? 상단에 노출되지 못했다면 어떻게 많은 공감 표시를 받을 수 있었을까?

하지만 댓글 작업이 있었다면 영점댓글의 성장은 쉽게 설명될 수 있다. [그림 3]에서 영점댓글은 먼저 우호적인 독자들에게 많은 공감을 받아 상단에 오른다. 이후 반대하는 독자들에게 많은 비공감을 받아 내려간다.

비공감을 받다가 0점 부근에 도달하면 원점 근처의 수많은 댓글과 뒤섞여 묻힌다. 한번 묻힌 댓글은 수 페이지를 넘겨야만 보이기 때문에 쉽게 노출되지 않는다. 더 이상 공감도 비공감도 받기 어렵다. 그렇기 때문에 영점댓글은 다시 올라가지도, 더 밑으로 내려가지도 못하고 그 자리에 묻힌다. 이 과정

이 여러 댓글에 일어나면, 그 댓글들은 [그림 1]에서처럼 ‘0점’을 나타내는 대각선 위에 늘어서게 된다.

영점댓글은 공감과 비공감이 몰려서 클릭됐다는 것을 암시할 뿐, 공감과 비공감 중 어느 쪽이 작업 인지는 알려 주지 않는다. 작업으로 올려놓은 댓글을 대중이 끌어내린 것일 수도 있다. 반대로 대중이 만든 베스트 댓글을 작업으로 묻어 버린 것일 수도 있다. 아니면 두 집단이 부딪혔던 댓글 싸움의 현상일 수도 있다. 각 영점댓글이 이 가운데 어떤 과정을 거쳐 만들어진 것인지는 단정하기 어렵다. 하지만 그 과정에 집단행동의 영향이 있었음은 유력하다.

영점댓글, 그 내용은?



네이버 뉴스 전체에서 찾아본 영점댓글은 온라인 커뮤니티에서 관찰된 것보다 훨씬 다양한 주제의 기사에 광범위하게 나타났다. 작업됐을 거라고는 예상하지 못했던 날씨 기사에도 남아 있었다.

한 방송사 일기예보에는 “어리고 보송한 남자 기캐(기상캐스터)의 날씨예보는 언제쯤 볼 수 있을까…”

라는 댓글이 달렸다. 이 댓글은 117건의 공감과 44건의 비공감을 받아 최종 호감도 -15점이 됐다. 앞서 말했듯 이것만으로는 밀어올린 쪽이 집단인지 끌어내린 쪽이 집단인지 판단할 수 없다. 하지만 성별 갈등과 관련된 집단행동이 있었음을 짐작할 수 있다.

또 다른 일기예보에 달린 “무슨 맑고 쾌적이나, 주말까지 미세먼지 대박이구만”이라는 내용의 댓글은 194건의 공감과 67건의 비공감을 받았다. 최종 호감도는 -7점이다. 해당 기사의 본문엔 미세먼지에 관한 언급은 없었다. 댓글을 이용해 미세먼지에 대한 불만을 부추기려는 작업, 혹은 댓글을 통해 표출된 부정적 여론을 감추려는 작업이 있었던 것으로 보인다.

영점댓글 중에는 명백한 악성 댓글도 있었지만 그렇지 않은 댓글도 많았다. 몇몇 영점댓글은 오히려 수준 높은 내용을 담고 있기도 했다. 공감, 비공감이 댓글의 질에 따라 클릭된 것이 아니라, 댓글 여론을 주도하기 위한 도구로 사용됐음을 알 수 있다.

영점댓글로 댓글 작업을 측정할 수 있다?



댓글을 더 높은 공감 수에 도달시키기 위해서는 더 많은 계정이 필요하다. 마찬가지로 더 높은 공감



**댓글 정책 변경 이후
영점기사의 숫자는 크게
줄었다. 댓글 조작이 없어진
것일까? 아닐 것이다.
조작하려는 집단이 있는 한,
막고 뚫는 과정은 끝없이
반복될 것이다.**

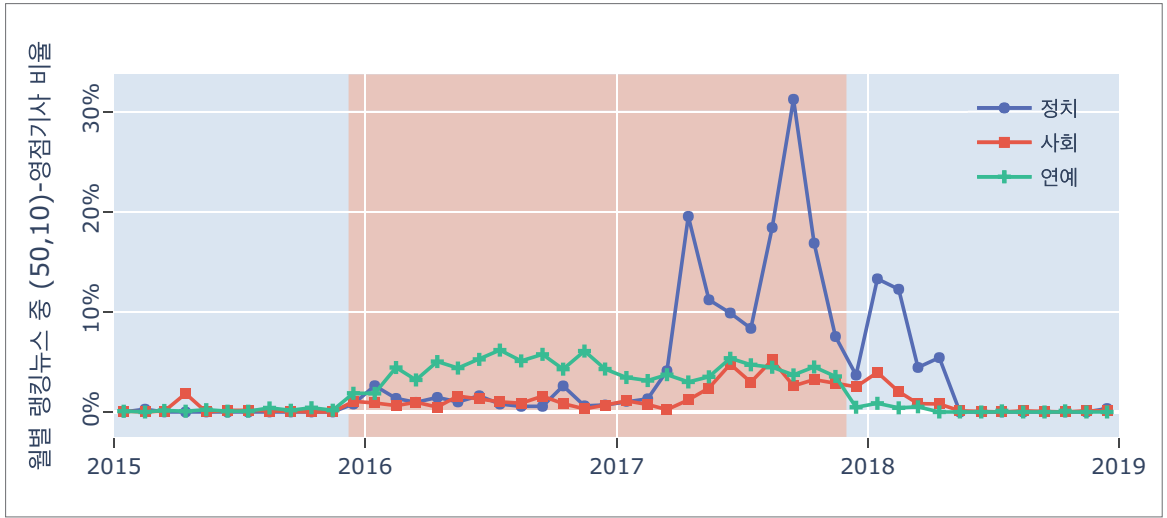
수의 댓글을 끌어내릴 때 더 많은 계정이 필요하다. 따라서 공감 수가 높은 영점댓글이 만들어지기 위해선 더 큰 집단이 동원돼야 한다. 몇몇 영점댓글은 작업 없이 우연히 만들어졌을지도 모른다. 하지만 한 기사에 여러 개의 영점댓글이 존재한다면 우연으로 치부하기는 어렵다.

즉, 한 기사에 공감 수가 높은 영점댓글이 많을수록 더 확실한 집단행동의 흔적이라고 볼 수 있다. 공감 수가 k 이상인 영점댓글이 n 개 이상 존재하는 기사를 ‘(k,n)-영점기사’라고 정의했다.⁴⁾ 이 두 값을 조절함으로써 크고 작은 규모의 댓글 작업을 추정해 볼 수 있다.

이 글에서는 (50,10)-영점기사, 즉 50건 이상의 공감을 받은 영점댓글이 10건 이상 있는 기사를 세어 봤다. 2013년 7월부터 2019년 6월까지 네이버 뉴스 전체 기사 중 11,500여 건의 기사가 이에 해당했다. 그중 5,100여 건은 섹션별 랭킹뉴스(일별 최다 조회 기사 30건)에 속했다. 정치, 연예, 사회 섹션의 기사가 가장 많았다.

[그림 4]는 이 세 섹션에 대해 월별 랭킹뉴스 기사 중 ‘(50,10)-영점기사’의 비율을 그린 것이다. 2015년 이전과 2019년 이후는 그 수가 많지 않아 생략했다. 처음 순공감순으로 댓글을 정렬하던 시기에는 영점기사가 거의 없었다. 2015년 4월 사회 섹션에서 조금 치솟는데, 주로 세월호 1주기와 관련된 내용이었다. 2015년 12월 호감순 정렬이 도입됐다. 이후 약 1년 동안 영점기사는 연예 섹션에서 가장 많이 보인다. 팬덤 문화였다. 2017년 10월 국정농단 의혹 보도와 함께 정치 섹션 기사가 작업되기 시작한 이후 탄핵, 경선, 대선을 거치며 폭발적으로 증가한다. 특히 2017년 9월에는 정치 섹션 랭킹뉴스 기사 중 무려 30%가 넘는 비율이 ‘(50,10)-영점기사’에 해당했다. 2017년 12월 네이버는 다시 댓글 정렬 기준을 순공감순으로 바꾸었다. 하지만 정

4) 본 연구에서는 ‘호감도 $\leq 0.1 \times$ 공감 수’를 만족하는 댓글을 영점댓글로 정의했다. 순공감순 정렬에 대해서는 ‘순공감도 $\leq 0.1 \times$ 공감 수’를 기준으로 했다.



[그림 4] 월별 랭킹 뉴스 기사 중 '(50,10)-영점기사'의 비율 <출처 - 필자 제공>

치 섹션의 영점기사 비율은 이전만큼 줄어들지 않았다. 아이스하키 단일팀 논란, 드루킹 사건 등 관련 기사에 작업의 흔적이 남아 있었다.

2018년 4월 네이버 뉴스는 정치 섹션의 댓글을 정렬하지 않기로 했다. 또한 사용자들의 전반적인 공감 표시 활동을 크게 제한하기 시작했다. 이후 몇 번의 추가적인 댓글 정책 개선 작업이 이뤄졌다. 사실상 전과 같은 규모의 댓글 작업을 펼칠 수 없게 됐고, 영점기사의 숫자도 크게 줄어들게 됐다.

댓글 조작, 해결된 것일까



일련의 댓글 정책 변경 이후 영점기사의 숫자는 크게 줄었다. 댓글 조작이 없어진 것일까? 아닐 것이다. 순공감순 정렬이 처음 도입됐던 때도 그랬다. 순공감순 정렬은 악성 댓글에 대한 대응책이기도 했지만, 같은 논지의 댓글을 반복적으로 작성해 항상 최신순 정렬의 상단에 위치시키는 ‘댓글 밀어내기’ 작업 또한 효과적으로 막아냈다. 하지만 이내 공감표시를 이용한 작업이 나타났다. 조작하려는 집단이 있는 한, 막고 뚫는 과정은 끝없이 반복될 것이다.

댓글 조작 사태는 이용자 인식에도 영향을 끼쳤다. 드루킹 사건을 통해 많은 국민이 댓글 여론의 조작 가능성에 대해 알게 됐다. 이후의 설문조사에 따르면 이전에 비해 댓글로 여론을 유추하는 경향 역시 하락한 것으로 보인다.^{5) 6)} 댓글을 믿지 않으면 문제가 해결되는 걸까? 댓글에 대한 불신이 독자들의 자주적인 속의로 이어진다면 그렇다고 볼 수 있다. 하지만 자신과 다른 의견은 모두 ‘조작된 것’으로 치부해 버리는 편협한 사고가 만연해질 수도 있다.

댓글란은 공론과 표현의 자유, 재미, 정보성 등 여러 가치가 얽힌 공간이다. 어떤 댓글 정책도 모든 면에서 완벽할 수는 없다. 댓글란의 운영에 어떤 가치를 우선으로 뒀야 할지, 그것을 어떻게 댓글 정책을 통해 추구할 수 있을지, 이용자의 입장에서 그것이 어떻게 받아들여질지는 포털뿐만 아닌 사회 전체가 함께 고민해야 할 숙제일 것이다. 🗑️

5) 김선호·오세욱·최민재, <댓글 문화 분석>, 미디어 이슈, 2(10), 2016.

6) 김선호·오세욱, <포털 뉴스서비스 및 댓글에 대한 인터넷 이용자 인식조사>, 미디어 이슈, 4(5), 2018.