

# 유저-상품 적합도 기반의 정확한 광고효과 계산 방안

## Accurate Ad-Effect Estimation Method based on Relevance between User and Item

홍석진, 고윤용, 김상욱\*, 박계환+  
한양대학교, (주)청년+

Hong suk-jin, Ko yun-yong, Kim sang-wook\*,  
Park gye-hwan+  
Hanyang University, CHUNGNYEON Inc.+

### 요약

최근 소셜 네트워킹 서비스(SNS)의 급격한 성장과 함께, SNS를 대상으로 상품 마케팅을 하는 기업(광고주)들이 증가하고 있다. 이에 따라 SNS에서 상품을 효과적으로 광고할 수 있는 광고 대행 유저들을 광고주에게 추천해주는 서비스들이 등장하였다. 하지만 위와 같은 대부분의 서비스들은 단순히 유저의 이웃 수를 기반으로 유저의 광고 효과를 평가하기 때문에, 유저를 통해 단계적으로 파급되는 광고 효과는 고려하지 못한다는 한계를 가지고 있다. 위와 같은 문제를 해결하기 위해, 본 논문은 영향력 최대화 (Influence maximization) 연구 분야의 기술을 활용하여, (1) 유저를 통해 단계적으로 파급되는 광고 효과를 고려하는 광고효과 최대화 방안을 제안한다. 또한 보다 정확하게 광고효과를 평가하기 위해, (2) 광고 상품과 유저 사이의 적합도를 정의하여 광고 대행인 선출 과정에 적용하였다. 실제 데이터베이스를 이용한 실험을 통해 제안하는 광고 대행 유저 선출 방안이 전통적인 선출 방안들과 비교하여 광고 효과가 더 큰 유저들을 선출한다는 것을 입증하였다.

## I. 서론

최근 Facebook, Instagram, NAVER 블로그 등과 같은 소셜 네트워킹 서비스(SNS)를 이용하는 유저들의 급증으로 인해, SNS를 대상으로 상품을 광고하고자 하는 기업(광고주)들이 늘어나고 있다. 그 결과로 SNS에서 효과적으로 상품을 광고할 수 있는 광고 대행 유저들을 추천해주는 서비스들이 생겨났다. 위와 같은 서비스들은 유저의 팔로워(이웃) 수, 방문자 수 등 유저가 가지고 있는 일부 특징을 기반으로 유저의 광고효과를 평가하여, 광고대행 유저를 선출한다.

위와 같은 방식은 광고효과를 평가하기 위해, 각 유저의 특성만을 이용할 뿐, SNS에서 유저들 간 단계적으로 파급되는 잠재적인 광고효과는 고려하지 못한다. 결과적으로 유저의 광고효과를 올바르게 평가하기 어려워, 최적의 광고대행 유저를 선출하지 못한다.

위와 같은 문제를 해결하기 위해, 본 연구는 영향력 최

대화 (Influence maximization) 연구 분야의 기술을 활용한다. 영향력 최대화란 소셜 네트워크와 선출하고자 하는 유저의 수  $k$ 가 주어졌을 때, 영향력을 최대화하는  $k$ 명의 유저를 선출하는 문제로 (1)과 같이 정의된다.

$$S^* = \operatorname{argmax}_{S \subseteq V} \sigma(S) \quad (1)$$

위 식에서  $\sigma(S)$ 는 네트워크를 통해 파급된 유저의 총 영향력을 의미한다.

본 연구는 영향력 최대화 기술을 광고대행 유저 선출 문제에 적용하여, SNS에서 유저들 간 단계적으로 파급되는 광고효과를 고려하는 (1) 광고효과 최대화 방안을 제안한다. 또한, 유저 간 파급되는 광고효과를 보다 정확하게 평가하기 위해, (2) 광고 상품과 유사 간 적합도를 정의하고 계산 및 적용 방안을 제안한다.

## II. 제안 방안

### 1. 광고효과 최대화 방안 및 파급 모델

본 연구는 영향력 최대화 기술을 활용하여 광고대행 유저를 선출하는 광고효과 최대화 방안을 제안한다. 광고효과 최대화 방안의 최적해를 구하는 것은 영향력 최대화 문제와 마찬가지로 NP-난해 문제이다[1]. 이 문제를 근사적으로 해결하기 위해, 본 연구는 영향력 최대화 문제의 해결방안 중 하나인 그리디 (greedy) 알고리즘을 이용한다[2]. 그리디 알고리즘은 매 단계 광고효과 한계 이득(marginal gain)이 최대인 유저를 하나씩 선출하여 총  $k$ 명의 유저를 선출하는 방안이다. 유저  $u$ 의 광고

\* : 교신저자

본 논문은 중소기업청에서 지원하는 2017년도 산학협력기술평가사업(No. C0511678)의 연구수행으로 인한 결과물임을 밝힙니다.

이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보·컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(No. NRF-2017M3C4A7069440). "본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성 지원사업의 연구결과로 수행되었음"(IPTP-2018-2013-1-00881)

효과 한계 이득은 (2)와 같이 표현한다.

$$\text{Marginal gain}(S,u) = \sigma(SUu) - \sigma(S) \quad (2)$$

그리고 유저의 광고효과를 평가하기 위해서는 광고효과가 네트워크에서 어떻게 파급되는지를 설명해주는 모델이 필요하며 이러한 모델을 파급 모델이라 한다. 대표적인 파급 모델로 Independent cascade (IC) 모델이 있다. IC 모델은 네트워크의 간선 (edge)의 가중치를 1/(팔로워의 수)로 할당한다.

## 2. 상품-유저 간 적합도 (relevance)

광고효과는 광고하고자 하는 상품과 각 유저들의 적합도에 따라 달라질 수 있다. 예를 들어, 어떤 유저가 상품에 대한 광고를 접했을 때, 해당 상품이 자신에게 적합하다면 광고에 영향을 받을 가능성이 크지만, 그렇지 않다면 해당 광고의 영향을 거의 받지 않을 것이다. 그러므로 본 연구는 유저와 상품 간 적합도를 정의하고, 유저와 상품 간 적합도를 반영한 광고효과 계산 방안을 제안한다. 유저 A의 상품 X에 대한  $\sigma(A, X)$ 를 (3)과 같이 정의한다.

$$\sigma(A, X) = 1 + \sum_{B \in O_A} \text{rel}(B, X) \cdot \frac{1}{|I_B|} \cdot \sigma(B, X) \quad (3)$$

위 식의  $\text{rel}(B, X)$ 는 유저 B와 상품 X의 적합도로 유저 B의 대표 단어와 상품 X의 대표단어 간 유사도를 이용하여 극대 매칭을 찾고, 찾은 극대 매칭 단어쌍의 유사도와 유저 B의 대표단어 빈도를 이용하여 기중평균을 계산하며 그 수식은 (4)와 같다. 이때 단어 간 유사도는 학습된 Word2Vec 모델의 단어 간 유사도를 사용하였다[3].

$$\text{rel}(B, X) = \frac{\sum_{(bw, xw) \in m^*} \text{sim}(bw, xw) \cdot n(bw)}{\sum_{(bw, xw) \in m^*} n(bw)} \quad (4)$$

## III. 성능 평가

### 1. 실험 데이터

본 연구는 NAVER 블로그 네트워크를 크롤링하여 실세계의 소셜 네트워크 데이터를 구축하였다. 미리 선정된 유저 30명에서 시작하여, 각 유저로부터 2 단계 (2-hop)까지의 이웃 유저들에 대한 팔로워 정보, 1단계 (1-hop)까지의 이웃 유저들에 대한 최신 10개 게시글을 수집하였다. 표 1은 수집한 데이터의 요약 정보를 보여준다.

표 1. NAVER 블로그 네트워크

# Users (유저)	# Degree (이웃 쌍)	# Documents (게시글)
2,336,953	11,590,716	253,660

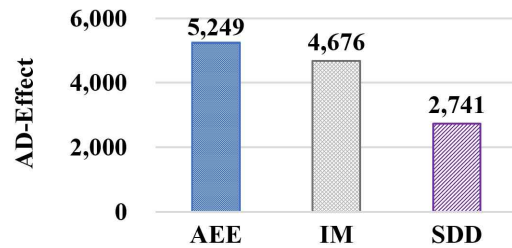
### 2. 실험 방법

본 연구는 다음 3가지의 방안을 비교 분석하였다: (1) 제안 방안 (AEE): 본 논문에서 제안하는 방안; (2) 기존

영향력 최대화 방안 (IM): 영향력 최대화 문제를 위한 방안으로, 유저-상품 간 적합도를 고려하지 않는 그리디 알고리즘; (3) Single Degree Discount (SDD): 영향력 최대화 분야의 해결방안 중 하나로, 유저의 이웃 수를 기반으로 광고대행 유저를 선출하되, 다음 대행 유저 선출 시, 기 선출된 유저의 간선을 제거하는 방안.

비교 방안들의 광고효과를 비교하기 위해, 우리는 각 방안으로 10명의 광고대행 유저 집합을 선출하여 선출된 유저 집합의 광고효과를 측정하였다. 유저 집합의 광고효과는 광고대행 유저의 광고에 영향을 받은 유저의 수를 의미한다.

### 3. 실험 결과



▶▶ 그림 1. 각 방안에 의한 광고효과

그림 1은 각 비교방안으로 선출한 10명의 광고대행 유저의 광고효과를 보여준다. x축은 각 비교 방안을 나타내며, y축은 광고효과를 나타낸다. 제안 방안 (AEE)으로 선출한 광고대행 유저 집합이 가장 큰 광고효과를 보였으며, IM, SDD와 비교하여 각각 12%, 91% 만큼 더 큰 광고효과를 갖는 유저를 선출하였다. 이러한 실험 결과는 제안 방안(AEE)이 유저-상품 간 적합도를 반영하여 유저들의 광고효과를 더욱 정확하게 계산했다는 것을 의미한다.

## IV. 결론

본 논문은 (1) 광고효과 최대화 방안을 제안하고, 상품에 대한 유저의 광고 효과를 보다 정확하게 평가하기 위해, (2) 유저-상품 간 적합도를 정의하고 계산 방안을 제안하였다. 실 세계 데이터를 이용한 실험을 통해 제안하는 방안이 전통적인 선출 방안들과 비교하여 광고 효과가 더 큰 유저들을 선출한다는 것을 입증하였다.

## ■ 참고 문헌 ■

- [1] D. Kempe, J. Kleinberg and E. Tardos, "Maximizing the Spread of Influence through a Social Network," ACM SIGKDD, pp. 137-146, 2003.
- [2] Y. Ko, D. Chae and S. Kim, "Accurate Path-based Methods for Influence Maximization in Social Networks," ACM WWW, pp. 59-60, 2016.
- [3] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," eprint arXiv, 1301.3781, 2013