Journal of KIIT. Vol. 24, No. 8, pp. 105-118, Aug. 31, 2026. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2026.24.8.105

뉴스 추천의 확증 편향을 완화하는 댐핑 요소 기반 반편향 하이브리드 재순위화 알고리즘

현정석*, 박찬정**

Damping Factor-based Debiasing Hybrid Reranking Algorithm for Mitigating Confirmation Bias in News Recommendation

Jung Suk Hyun*, Chan Jung Park**

이 논문은 2025학년도 제주대학교 교원성과지원사업에 의하여 연구되었음

요 약

기존의 편향 완화 연구는 모델 재학습이나 계산 복잡도가 높은 재순위화 기법에 의존해 온 한계가 있었다. 이에 본 연구는 뉴스 추천 시스템의 확증 편향을 완화하는 그래프 기반 하이브리드 재순위화 알고리즘 DDHR(Damping-Debiased Hybrid Reranking)을 제안하였다. MIND 데이터셋을 분석한 결과는 카테고리 편향은 주로 후보 선정 단계에서 비롯되며, 다량 사용자일수록 선택 편향이 강화되는 것으로 나타났다. DDHR은 그래프 확산 알고리즘의 리셋 벡터를 반편향 방향으로 재설계하여, 단일 댐핑 요소로 편향 완화와 추천 정확도 간의 상충 관계를 제어하는 구조를 갖는다. 본 연구는 편향 형성 과정을 규명함과 동시에, 재학습 없이 적용 가능한 실용적 후처리 완화 기법을 제시했다는 데 의의가 있다.

Abstract

Prior research on bias mitigation has been limited by its reliance on model retraining or re-ranking techniques with high computational complexity. To address this limitation, this paper proposes Damping-Debiased Hybrid Reranking (DDHR), a graph-based hybrid re-ranking algorithm that mitigates confirmation bias in news recommendation systems. Analysis of the MIND dataset showed that category-level bias primarily originates at the candidate selection stage, and that selection bias intensifies among heavy users. DDHR redesigns the reset vector of a graph diffusion algorithm in a debiasing direction, enabling a single damping factor to control the trade-off between bias mitigation and recommendation accuracy. This paper is significant in that it elucidates the process by which bias forms, while also presenting a practical post-processing mitigation technique that can be applied without retraining.

Keywords

confirmation bias, news recommendation system, MIND dataset, graph-based reranking, recommendation algorithm

* 제주대학교 경영정보학과 교수
- ORCID: https://orcid.org/0000-0002-2334-0126
** 제주대학교 컴퓨터교육과 교수(교신저자)
- ORCID: https://orcid.org/0000-0002-3146-2342

· Received: Jul. 21, 2026, Revised: Aug. 11, 2026, Accepted: Aug. 14, 2026

· Corresponding Author: Chan Jung Park
Dept. of Computer Education, Jeju National University, Jejudaehak-ro 102 63243, South Korea
Tel.: +82-64-654-3296, Email: cjpark@jejunu.ac.kr

Ⅰ. 서 론

최근까지도 뉴스 추천 시스템은 사용자의 정보 소비 편의성을 극대화하는 방향으로 발전해 왔다[1][2]. 그러나 이러한 발전은 동시에 사용자가 선호하는 성향의 뉴스만 지속적으로 노출시키는 필터 버블(Filter bubble) 현상과 이것이 사용자의 인지적 편향과 결합하여 나타나는 확증 편향(Confirmation bias)을 심화시킨다는 비판을 받았다[3]-[5]. 확증 편향은 본래 개인이 자신의 기존 신념과 일치하는 정보를 선택적으로 수용하는 인지적 현상을 가리키지만, 추천 시스템이라는 알고리즘 매개 환경에서는 이 개인적 성향이 시스템이 만들어내는 구조적 편향과 뒤섞여 나타난다[6][7]. 따라서 사용자가 특정 카테고리의 뉴스에 편중되어 소비하는 현상을 관찰하더라도, 이것이 사용자 개인의 능동적 선택에 기인한 것인지, 아니면 애초에 추천 시스템이 편중된 후보군만을 제시했기 때문인지를 구분하기 어렵다는 근본적인 문제가 존재한다[8][9].

확증 편향과 필터 버블을 다룬 선행연구들은 추천 결과의 카테고리 편중도나 다양성 지표를 측정하는 데 그칠 뿐, 그 편향이 시스템의 후보 선정단계와 사용자의 실제 선택 단계 중 어디서 비롯되었는지 분리하여 정량화하지 않는다[10][11]. 이 때문에 추천 시스템이 편향되어 있다는 진단은 가능하지만, 개선하려면 시스템 설계를 개선해야 할지 혹은 사용자의 선택 패턴에 개입해야 할지 구체적인 방향을 정하기 어렵다.

또한, PageRank[12] 메커니즘을 추천 시스템에 이식한 대표적 연구로서 ItemRank[13]를 비롯한 Personalized PageRank 계열 알고리즘[14]은 개인화 점수를 계산할 때에 리셋 벡터(Reset vector)를 사용자의 과거 소비 이력 그대로 반영한다. 이 경우 댐핑 요소(Damping factor)가 커질수록 사용자의 기존 선호와 편향이 오히려 증폭되는 구조적 한계가 나타난다. 즉, 그래프 확산 메커니즘 자체는 정확도 향상에는 효과적이지만, 다양성이나 편향 완화라는 목적에는 그대로 적용하기 어려운 문제가 대두된다. 다양성 확보를 위한 후처리 재순위화(Post-hoc re-ranking) 기법으로 흔히 활용되는 행렬식 점 과정(DPP, Determinantal Point Process) 계열 방법[15] 역시 계산복잡도가 높고 커널 설계가 도메인에 민감하며, 편향 완화의 정도를 직관적인 파라미터 하나로 통제하기 어렵다는 제약이 있다.

본 연구는 MIND(Microsoft News Dataset)[16][17]의 사용자 클릭 로그를 활용하여 뉴스 추천 서비스 내 확증 편향의 메커니즘을 데이터 기반으로 검증하고, 이를 완화할 수 있는 알고리즘을 제안한다. 또한 본 연구의 분석 대상 알고리즘으로는 학계와 산업계에서 공통으로 검증된 대표적인 뉴스 추천 모델인 멀티 헤드 셀프 어텐션을 활용한 인공신경망 기반 뉴스 추천(NRMS, Neural News Recommendation with Multi-Head Self-Attention)[17]을 기반으로 한다. 이 모델은 선행연구에서 공통 벤치마크로 널리 활용되어 왔으며, 후보 뉴스 집합에 대한 클릭 확률을 예측하여 상위 항목을 추천한다.

기존의 MIND 데이터셋을 활용한 다수의 선행연구에서 NRMS 이외에 어텐션 기반 멀티뷰 학습을 활용한 인공신경망 기반 뉴스 추천 알고리즘[18]과 장기 및 단기 사용자 표현을 활용한 인공신경망 기반 뉴스 추천 알고리즘[19]도 함께 비교하였으나, 세 모델 간의 정확도 차이는 뚜렷한 우열 없이 논문마다 상이하게 보고되고 있음을 확인하였다[20]-[22]. 이에 본 연구는 세 모델을 모두 학습시켜 편향을 비교하는 대신, 선행연구에서 가장 널리 쓰이는 베이스라인이자 세 모델 중 구조가 가장 단순하여 그래프 기반 재순위화 기법의 효과를 가장 명확하게 관찰할 수 있는 NRMS를 대표 모델로 선정하여, DDHR(Damping-Debiased Hybrid Reranking)과의 비교실험을 수행하였다.

본 연구가 제안하는 DDHR 편향 완화 알고리즘은 ItemRank의 그래프 확산 메커니즘을 기반으로 설계되었다. ItemRank는 아이템 간 상관 그래프 상에서 PageRank 방식의 확산과 감쇠 메커니즘을 적용하여 사용자별 선호도를 효율적으로 계산한다. 이러한 구조적 특성은 기존 추천 모델을 재학습하지 않고도 후처리 방식으로 편향 완화 기법을 적용할 수 있는 기틀을 제공하며, 실제 서비스 환경에서의 적용 가능성을 높이는 장점이 있다.

본 연구는 먼저 확증 편향의 원인을 추천 시스템

이 후보군을 구성하는 단계에서 발생하는 시스템 편향과 사용자가 노출된 후보 중 실제로 클릭을 선택하는 단계에서 발생하는 사용자 선택 편향으로 분리하여 분석하는 절차를 제안한다. 이러한 분석 절차를 바탕으로, 본 연구는 NRMS 모델을 실제로 학습시켜 각 모델의 구조가 만들어내는 카테고리 편향의 정도를 정량적으로 비교하여 평가한다. NRMS는 뉴스 제목만을 입력 정보로 활용하는 대표적인 뉴스 추천 모델로서 본 연구에서는 편향 분석을 위한 벤치마크 기준점으로 사용한다[17].

본 연구는 기존 알고리즘에서 발생하는 편향을 효과적으로 제어하기 위해, 기존 추천 모델의 재학습 없이 후처리 방식으로 적용 가능한 그래프 기반 하이브리드 재순위화 알고리즘인 DDHR을 제안한다. DDHR의 핵심적 차별성은 리셋 벡터의 설계 방식에 있다. 기존 ItemRank 계열의 알고리즘은 사용자의 과거 소비 이력을 리셋 벡터에 원형 그대로 반영하여 기존 선호를 강화하고 편향을 고착화하는 한계가 있었다. 반면 DDHR은 사용자의 기존 카테고리 편중과 반대 방향으로 작용하는 반편향(Debiasing) 리셋 벡터를 설계하여, 사용자가 과거에 상대적으로 적게 소비하거나 노출 빈도가 낮았던 카테고리에 더 높은 확률적 가중치를 부여한다. 이러한 설계를 통해, 댐핑 요소는 단순히 그래프 탐색의 확산 범위를 조절하는 하이퍼파라미터에 그치지 않는다. 즉, 댐핑 요소는 사용자의 기존 선호를 보존하는 것과 새로운 카테고리로 얼마나 확장할 것인지 균형을 조율하는 핵심 제어 변수로 작용한다. 최종적으로 DDHR은 딥러닝 기반 추천 모델이 산출한 관련성 점수와 그래프 기반 확산 점수를 가중 결합하는 하이브리드 재순위화 구조를 채택함으로써, 기존 추천 모델의 구조를 변경하거나 추가적인 재학습을 수행하지 않고도 즉각 적용이 가능한 실용성과 확장성을 제공한다.

II. 관련 연구

2.1 뉴스 추천과 MIND 데이터셋

MIND 데이터셋은 대규모 실제 사용자의 클릭 로그와 풍부한 텍스트 메타데이터를 함께 제공한다는 점에서 뉴스 추천 연구 분야에서 독보적인 벤치마크로 자리매김해 왔다[16][23]. 그러나 이 데이터셋을 활용한 대부분의 선행연구는 AUC(Area Under the ROC Curve), MRR(Mean Reciprocal Rank), nDCG(Normalized Discounted Cumulative Gain)와 같은 정확도(Accuracy) 지표를 중심으로 모델의 예측 성능을 평가하는데 치중되어 왔으며, 데이터셋이 담고 있는 사용자의 시계열적 클릭 행동 자체를 분석한 연구는 드물다[23][24].

본 연구는 이와 같은 통상적인 활용 방식에서 벗어나, MIND 데이터셋에 내재된 실제 사용자 로그를 다각도로 분석함으로써 확증 편향의 형성 과정을 종단적으로 추적하는 새로운 분석 프레임워크의 기초 데이터로 이를 재정의하였다. 즉, 본 연구는 MIND 데이터셋을 단순히 추천 알고리즘의 성능 평가를 위한 도구에 국한하지 않고, 추천 시스템 환경 내에서 편향이 발현되고 고착화되는 동적 메커니즘을 규명하기 위한 관측 자료로 재해석한다는 점에서 기존 연구와 차별성을 갖는다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 MIND 데이터셋의 특성과 선행연구를 검토하고, 확증 편향의 형성 과정을 추적하는 관측 자료로 재해석한다. Ⅲ장에서는 MINDsmall_train 로그를 통해 편향의 상대적 크기와 종단적 강화 양상을 분석한다. Ⅳ장에서는 반편향 리셋 벡터를 설계한 DDHR의 이론적 구조와 차별성을 기술한다. Ⅴ장에서는 NRMS를 베이스라인으로 DDHR 적용 전후의 정확도·편향 지표를 비교하고, 댐핑 요소와 결합가중치에 따른 트레이드오프를 검증한다. Ⅵ장에서는 연구 결과를 종합하고 연구의 한계와 향후 방향을 제시한다.

2.2 추천 시스템에서의 확증 편향 및 완화 연구

추천 시스템 환경에서 필터 버블과 확증 편향은 사용자가 자신의 기존 선호와 일치하는 콘텐츠에만 지속적으로 노출됨으로써 정보 소비의 편중이 심화되는 현상으로 정의된다[4]. 이러한 문제를 완화하기 위해 선행연구들은 주로 다양성 확보 및 반편향

화 기법을 제안해 왔다[25][26]. 대표적으로 모델 내부의 손실 함수에 다양성 규제항을 추가하거나, 임베딩 학습 단계에서 편향 요소를 직접 배제하는 방식이 활용되었다[26]. 그러나 이러한 모델 내재적 접근법은 두 가지 근본적인 한계를 지닌다. 첫째, 손실 함수 수정이나 임베딩 단계의 개입은 모델의 전면적인 재학습을 처음부터 요구하므로, 이미 실제 서비스에 배포된 대규모 추천 시스템에 실시간으로 적용하기에는 막대한 연산 비용과 실무적 제약이 수반된다. 둘째, 기존 연구들은 편향의 발생 원인을 사용자의 내재적 선택 편향과 시스템의 구조적 편향으로 구별하지 않은 채 추천 모델 전반에 획일적으로 개입한다는 한계가 있다. 이로 인해 편향 완화 과정에서 추천 정확도와 다양성 간의 과도한 상충 관계(트레이드오프)가 발생하여 추천의 효율성을 저해시킨다. 이러한 한계는 편향의 근본적인 원인에 대한 정밀한 진단 없이 완화 기법을 일률적으로 적용하는 데서 기인한다. 따라서 본 연구는 시스템 편향과 사용자 선택 편향을 독립적으로 분리하여 진단하는 분석 프레임워크를 제안함으로써 이러한 한계를 극복하고자 한다.

2.3 그래프 기반 추천 알고리즘

PageRank와 ItemRank를 비롯한 그래프 기반 추천 알고리즘은 사용자와 아이템 간의 관계를 그래프 구조 위에서 전파하는 방식으로 추천을 수행한다는 점에서 딥러닝 기반 모델과 차별화된다. 이러한 그래프 기반 접근은 노드 간 연결 관계를 기반으로 추천 결과를 생성하므로 추천 과정의 설명 가능성이 높고, 사용자와 아이템 간의 구조적 관계를 명시적으로 반영할 수 있다는 강점을 지닌다. 또한 그래프 기반 알고리즘은 이미 계산된 추천 결과에 후처리 방식으로 재순위화를 적용하기 용이한 구조를 제공하며, 그래프 확산 과정이 댐핑 요소와 리셋 벡터라는 명시적인 수학적 요소를 통해 제어된다. 본 연구는 이러한 구조적 특성에 주목하여 추천 결과의 확산 범위와 방향을 체계적으로 조절할 수 있는 근거를 제공한다. 특히 댐핑 요소와 리셋 벡터를 적절히 설계할 경우에 대규모 딥러닝 추천 모델을 재학습하지 않고도 추천 정확도와 편향 완화 간의 균형을 효과적으로 조율할 수 있다. 이는 기존 연구에서 한계로 지적되어 온 막대한 재학습 비용 문제와 편향 완화를 위해 추천 정확도를 희생해야 하는 문제를 동시에 극복하는 대안이 된다. 이에 본 연구는 그래프 기반 알고리즘의 특성을 활용하여 기존 추천 결과를 후처리 단계에서 재순위화함으로써 확증 편향을 완화하는 방법을 제안한다.

III. MIND 데이터셋 기반 확증 편향 분석

3.1 분석의 개요

본 연구는 MIND의 MINDsmall_train 데이터를 분석하였다. 분석 데이터셋의 구체적인 명세는 표 1과 같다. 본 연구에서 사용된 핵심 지표는 다음 네 가지이다. top_category_ ratio는 클릭 또는 노출된 뉴스 중 가장 높은 빈도를 차지하는 카테고리의 비율로서 이 값이 클수록 특정 카테고리에 대한 편중이 큰 것을 의미한다. 정규화 섀넌 엔트로피(Normalized Shannon Entropy)는 카테고리 분포의 다양성을 0~1 범위로 정규화한 값이다[27]. ratio_gap은 clicked_top_ratio에서 exposed_top_ ratio를 뺀 값으로서 추천 노출에 비해 실제 클릭 행동의 편중 정도를 이용하여 사용자의 선택적 편향이 얼마나 작용했는지 측정한다. 젠센 섀넌 정보량(JSD, Jensen Shannon Divergence)는 두 카테고리 분포 간의 수학적 거리를 단일 수치로 요약하여 나타낸다[28].

표 1. 분석에서 사용한 MINDsmall_train 데이터셋 개요
Table 1. Overview of the MINDsmall_train dataset used in the analysis

Item	Details
behavior.tsv size	156,965 rows (impression logs) / 50,000 users
news.tsv size	17 category types (news, sports, finance, lifestyle, etc.)
Key components	impression_id, user_id, time, history (past click history), impressions (impression-click labels)
observation period	Approx. Nov 9, 2019 – Nov 14, 2019

본 연구는 사용자의 과거 클릭 이력(History)과 누적 노출 개수(Impressions)를 기준으로 다량(Heavy) 사용자를 정의하였다. 표본 추출을 위해 절대 임계값 기준(history≥50회 및 impressions≥100회, n=3,246)과 백분위수 기준(상위 10% history≥42회 및 impressions≥283회, n=2,376)의 두 가지 방식을 검토하였다. 최종 분석 대상은 데이터의 편향을 줄이고 데이터 분포 특성을 반영할 수 있는 백분위수 기준(전체 사용자의 4.75%에 해당)으로 선정하였다. 분석 대상 전체 사용자의 평균 클릭 이력은 18.5회, 평균 노출 수는 116.9회, 평균 방문 횟수는 3.1회로 나타났다.

3.2 사용자 집단별 뉴스 다양성 비교

본 연구는 기존 연구에서는 밝히지 못한 사용자 집단을 구별하여 뉴스 선택의 다양성 비교를 실시하였다. 즉, 다량 사용자와 소량(Light) 사용자가 이용한 전체 뉴스(과거 클릭 이력 및 당일 실제 클릭 뉴스)를 대상으로 카테고리 분포의 entropy와 top_category_ratio를 비교 분석하였다. 소량 사용자의 평균 entropy는 0.8438, 다량 사용자는 0.7617로 나타났으며, top_category_ratio는 각각 0.4207과 0.3758로 나타났다. 두 집단 간의 통계적 차이를 검증하기 위해 Mann-Whitney U 검정[29]을 실시한 결과, 다량 사용자의 entropy가 소량 사용자에 비해 통계적으로 유의미하게 낮음이 확인되었다(U=28,950,218, p<0.001). 즉, 다량 사용자가 소량 사용자보다 상대적으로 더 특정 뉴스 카테고리에 집중하여 클릭하는 경향이 있는 것으로 나타났다.

또한, 다량 사용자 2,376명 개개인의 노출 뉴스와 실제 클릭 뉴스의 카테고리 편중도를 비교 분석하였다. 그 결과는 표 2와 같고 동일 집단 내 대응표본의 차이를 검증하기 위해 Wilcoxon 부호순위 검정[30]을 수행한 결과, 실제 클릭 뉴스의 편중도가 노출 뉴스의 편중도에 비해 통계적으로 유의미하게 높은 것으로 나타났다(W=2,554,509, p<0.001). 이는 다량 사용자가 노출된 뉴스 환경에 머무르지 않고 실제로 뉴스를 선택하고 클릭하는 과정에서 특정 카테고리에 더 집중하는 사용자의 자발적 선택 편향이 작용하고 있음을 시사한다.

표 2. 다량 사용자의 뉴스 노출 대비 실제 클릭 카테고리 편중도 비교(n=2,376)
Table 2. Comparison of actual clicked category bias relative to news impressions among heavy users(n=2,376)

Metric	Mean	SD
Exposure concentration ratio (shown_top_ratio)	0.2829	0.0882
Click concentration ratio (clicked_top_ratio)	0.3946	0.1445
ratio_gap	0.1117	0.1287
entropy_gap	0.0144	0.1035

3.3 추천 단계별 카테고리 분포 분석

본 연구는 추천 단계별 뉴스 소비의 편향성 추이를 살펴보기 위해 분석 단계를 다음과 같이 구체화하여 카테고리 분포를 비교하였다. 분석 대상은 첫째, 전체 뉴스(News.tsv) 풀(Pool), 둘째, 사용자에게 노출된(Exposed) 추천 뉴스, 셋째 사용자가 최종 소비한 실제 클릭(Clicked) 뉴스이다. 이와 같은 3단계 추천 파이프라인 모델링은 추천 시스템의 단계별 필터링 과정에서 편향 전이 현상을 실증한 기존의 분석 프레임워크[23]에 기반하여 재구성되었다.

데이터 분석 결과, 뉴스 카테고리별 분포에서 스포츠 카테고리의 비중이 가장 큰 폭으로 변동하였다. 전체 뉴스에서 28.29%를 차지하던 스포츠 카테고리는 노출 단계에서 10.13%로서 -18.16%p 급감하였다. 이와 달리 라이프스타일은 4.83%에서 11.22%로, 엔터테인먼트는 1.14%에서 5.99%로 증가함으로써 소프트 뉴스 카테고리의 비중이 노출 단계에서 확대되는 것으로 나타났다. top_category_ratio는 전체 뉴스에서 0.3076, 노출 후보군에서 0.2723, 실제 클릭 뉴스가 0.2937로 나타났다. 이는 추천의 단계적 이행 과정에서 카테고리 분포가 단계별로 재구성되었음을 의미한다.

jsd 지표를 통해 각 단계 간 분포의 거리를 측정한 결과, 알고리즘 기반의 후보군 선정단계(Pool-exposed)의 거리는 0.2836, 사용자 선택 단계(Exposed-clicked)의 거리는 0.0931, 전체 뉴스와 실제 클릭 뉴스 간(Pool-clicked) 거리는 0.2711로 나타났다. 즉, 뉴스가 얼마나 편향되게 노출되느냐를 결정하는 데는 추천 시스템이 어떤 뉴스를 후보로 골라주느냐가 사용자가 그중에서 무엇을 클릭하느냐

보다 약 3배 더 큰 영향을 미쳤다. 결국, 대부분의 편향은 사용자가 확증 편향으로 스스로 뭔가를 선택하기도 전에, 추천 시스템이 후보 뉴스를 미리 걸러내고 구성하는 단계에서 이미 정해진다.

3.4 재방문 사용자의 종단 분석

이 절에서는 재방문 사용자(21,750명)를 대상으로 종단 분석을 실시하였다. 그 결과, 방문 순서에 따라 ratio_gap이 증가하는 사용자가 61.3%였으며 평균 기울기는 0.01919로 통계적으로 유의하였다(t=25.667, p<0.0001). 이 증가가 알고리즘적 강화 때문인지 검증하기 위해 시차 변수 회귀와 Granger 인과 검정[31]을 수행하였으나, 과거 클릭이 이후 이번 노출에 미치는 영향은 유의하지 않았고(p=0.625) Granger 검정도 양방향 모두 5.3%로 우연 수준에 그쳤다. 다만 이는 알고리즘적 강화가 없다는 확정적 증거가 아니라, 사용자별 관측치가 짧아(평균 5~8회) 검정력이 부족했을 가능성이 크다는 점을 함께 고려해야 한다. 카테고리 일관성 지표는 평균 0.48로, 사용자들이 매 방문 특정 카테고리에 쏠리되 그 대상 카테고리는 방문마다 자주 바뀌는 유동적 패턴을 시사한다. 결론적으로, 표 3과 같이 시스템 편향이 사용자 선택 편향보다 약 3배 크며, 사용자 선택 편향은 통계적으로 유의하나 동일 카테고리로의 고착화라기보다 방문 단위의 유동적 편중에 가깝다.

표 3. 선행연구와 본 분석의 결과 비교
Table 3. Comparison of Results with Previous Studies

Item	Vrijenhoek (2023)	Present analysis
System vs. user bias	Reported bias at candidate selection stage	Confirmed: system bias > user bias (jsd 0.284 vs. 0.093)
Heavy vs. light users	Not addressed	Newly examined. No clear difference (results vary by aggregation method)
Bias change over time	Not feasible with short-term data	Significant increasing trend in longitudinal analysis
Causal direction of bias	Not addressed	Newly examined. No clear causal signal (lagged regression, Granger test)

IV. DDHR 알고리즘 제안

4.1 DDHR의 기본 개념

그래프 기반 추천 알고리즘의 대표적 계열인 PageRank 및 그 변형 모델들은 그래프 위에서의 확산과 감쇠 방법을 통해 노드의 중요도를 계산한다. 일반화된 PageRank의 점수 벡터 R은 다음과 같은 고정점 방정식으로 정의된다.

$$R = d \cdot M \cdot R + (1-d) \cdot E \quad (1)$$

여기서 M은 그래프의 정규화된 전이행렬, E는 리셋 벡터(또는 텔레포트 벡터), d∈[0,1]는 댐핑 요소이다. 식 (1)에서 확인할 수 있듯이 댐핑 요소 d는 확산 과정에서 그래프의 링크 구조 M을 반영하는 신뢰도의 수준과 나머지 확률 질량 (1−d)을 리셋 벡터 E가 지정하는 방향으로 재분배하는 비중을 결정하는 핵심 파라미터로 기능한다. 따라서 E의 설계는 전체 확산 과정의 방향성을 결정짓는 핵심 요인이 된다. 표준적인 Personalized PageRank와 이를 추천 시스템에 이식한 ItemRank는 리셋 벡터 E를 사용자의 과거 소비 이력에 기반하여 설정한다. 이 경우 댐핑 요소 d가 커질수록 사용자의 기존 선호가 그래프 확산 과정을 통해 그대로 증폭되는 구조를 취하게 되는데, 이는 추천의 정확도 관점에서는 효과적일 수 있으나 확증 편향을 완화하기보다 오히려 강화할 위험을 내포하고 있다.

4.2 알고리즘

본 연구는 리셋 벡터의 설계 원리를 근본적으로 반전시킨 DDHR 알고리즘을 제안한다. DDHR은 사용자의 기존 카테고리 소비 분포와 반대 방향을 향하도록 리셋 벡터 E를 재설계하며, 구체적으로는 사용자가 상대적으로 적게 노출되었거나 소비하지 않은 카테고리에 더 높은 확률 질량을 배정하는 반편향 리셋 벡터를 도입한다.

이러한 설계 하에서 댐핑 요소 d는 더 이상 단순히 링크 구조에 대한 신뢰도만을 의미하지 않는다.

d가 1에 가까울수록 확산은 그래프의 연결 구조, 즉 기존 소비 패턴과 상관된 항목들을 따라 이루어지는 반면, d가 0에 가까울수록 확산은 반편향 리셋 벡터가 지정하는 다양성 방향으로 재분배된다. 즉 DDHR에서 댐핑 요소는 추천의 정확도(기존 선호 패턴의 유지)와 편향 완화(새로운 다양성 방향으로의 이동) 사이의 균형을 명시적으로 조절할 수 있는 하나의 통제 가능한 파라미터로 재정의된다.

마지막으로, DDHR은 그래프 확산을 독립적인 추천 모델로 운용하는 대신, 사전 학습된 신경망 기반 추천 모델인 NRMS가 산출한 관련성 점수와 그래프 확산 점수를 가중합하는 후처리 재순위화 구조를 채택한다. 이러한 접근 방식은 추천 모델을 처음부터 재학습하는 데 수반되는 계산 비용과 시간적 부담을 덜 뿐만 아니라, 기존에 운영 중인 추천 시스템에 편향 완화 메커니즘을 별도의 플러그인 형태로 손쉽게 적용할 수 있도록 한다.

d는 그래프 확산 과정에서 인접 링크를 따라 이동할 확률과 반편향 방향으로 점프할 확률을 제어하는 역할을 한다. alpha는 최종 후처리 재순위화 단계에서 신경망 기반 추천 모델의 정확도와 그래프 기반 다양성 간의 상충 관계를 독립적으로 조절한다. 이들 두 파라미터를 독립적으로 스윕(Sweep)하면, 편향-정확도 트레이드오프를 2차원 곡면상에 표현할 수 있다. DDHR의 구체적인 알고리즘은 알고리즘 1과 같다.

DDHR은 ItemRank와 동일한 PageRank 기반 그래프 확산 메커니즘과 댐핑 구조를 유지하되, 사용자의 기존 선호를 증폭하던 리셋 벡터의 방향성을 반전시킴으로써 성능 극대화 중심의 알고리즘을 편향 완화 목적으로 재목적화하였다. 추천 다양성 확보를 위해 널리 활용되는 DPP 기반 재순위화 기법은 높은 계산 복잡도와 도메인에 민감한 커널 설계로 인해 편향 완화 수준을 직관적으로 제어하기 어렵다. 반면, DDHR은 댐핑 요소 d와 결합 가중치 alpha라는 두 개의 저차원 스칼라 파라미터만으로 편향-정확도 트레이드오프를 체계적으로 조절할 수 있어 실험적 통제와 결과 해석이 용이하다.

실용적 측면에서도 DDHR은 ItemRank의 계산 효율성을 그대로 계승하여 대규모 코퍼스에도 효율적으로 적용할 수 있다. 더욱이 신규 추천 모델을 재학습할 필요 없이 사전 학습된 NRMS의 예측 점수에 후처리 재순위화 계층으로 결합할 수 있으므로 기배포된 추천 시스템에 대한 적용성과 확장성이 높다. 아울러 로그 기반의 확증 편향 분석에서 활용된 지표체계(entropy, gini 계수, jsd, ratio_gap)를 알고리즘 평가에 동일하게 적용함으로써 데이터 현상 분석부터 알고리즘 설계 및 성능 평가에 이르기까지 하나의 일관된 프레임워크 내에서 유기적으로 연계하였다.

알고리즘 1. DDHR 알고리즘
Algorithm 1. DDHR algorithm

```
Input:  candidate set C, neural scorer s(·), co-click
graph adjacency A, user history H, category map cat(·),
damping factor d, mixing weight alpha, top-K size K
Output: re-ranked list R

 1:  M <- A . diag(colsum(A))^-1
 2:                  // column-normalized transition
matrix
 3:  P_hist <- category distribution of {cat(v) : v in H}
 4:  for each news v in V do
 5:      w[v] <- 1 - P_hist[cat(v)]
 6:                          // favor under-seen
categories
 7:  end for
 8:  e <- w / sum(w)    // debiased reset vector
 9:  r <- e
10:  repeat
11:      r_new <- d · M · r + (1 - d) · e
12:      if ||r_new - r||_1 < tol then break
13:      r <- r_new
14:  until max_iter reached
15:  for each c in C do
16:      score_n[c] <- normalize(s(c))
17:      score_g[c] <- normalize(r[c])
18:      score[c]   <- alpha . score_n[c] + (1 - alpha)
19:                    · score_g[c]
20:  end for
21:  R <- top-K items of C sorted by score in
descending
22:  order
23:  return R
```

V. 성능 평가

5.1 실험 설계

본 연구는 DDHR의 편향 완화 효과를 검증하기 위해 사전 학습된 NRMS를 베이스라인 모델로 설정하고, ① NRMS 단독 추천 결과와 ② NRMS의 예측 점수에 DDHR의 그래프 확산 점수를 결합한 후처리 재순위화 추천(NRMS+DDHR) 결과를 비교 분석하였다. 이를 위해 MINDsmall_train 데이터셋의 사용자 클릭 이력에 슬라이딩 윈도우 알고리즘(윈도우 크기 ③을 적용하여 동시 클릭(co-click)된 뉴스 노드 간의 관계를 반영한 공동 클릭 그래프(노드 51,282개, 엣지 1,502,554개)를 구축하였다.

그래프 확산 점수는 개인화 PageRank 방식의 반복 계산을 통해 산출하였다. 반복 계산은 연속된 두 반복 간 점수 벡터의 L1 노름 차이가 수렴 임계값(tolerance) 10^{-6} 미만이 될 때 종료하도록 설정하였으며, 최대 반복 횟수는 50회로 제한하여 수렴이 지연되는 경우에도 계산이 무한히 지속되지 않도록 하였다. 슬라이딩 윈도우 크기는 세션 기반 추천 연구에서 일반적으로 사용되는 범위(2~5)를 참고하여 3으로 설정하였다. 하지만, 슬라이딩 윈도우 크기가 결과에 미치는 영향을 확인하기 위해, 윈도우 크기 5로 재구성한 그래프(노드 51,282개, 엣지 2,279,707개)를 대상으로 동일한 조건(d=0.7, alpha=0.7)에서 1,000건 표본에 대해 재검증하였다. 그 결과는 표 4와 같이 정확도 지표 모두 윈도우 크기 3의 결과와 상대 변화율 0.15% 이내로 거의 동일하게 나타나, 본 연구의 윈도우 크기 선택이 핵심 결론에 실질적인 영향을 미치지 않음을 확인하였다. 이후 분석은 윈도우 크기 3을 기반으로 실행하였다.

표 4. 윈도우 크기에 따른 지표 격차 비교(n=1,000)
Table 4. Comparison of metric differences by window size (n=1,000)

Metric	Window size		Difference
	3	5	
group_AUC	0.643192	0.643203	+0.000011
mean_MRR	0.357277	0.357495	+0.000218
nDCG@5	0.339405	0.339494	+0.000090
nDCG@10	0.403254	0.403309	+0.000055
top_ratio	0.211338	0.211639	+0.000301
entropy	0.863752	0.863791	+0.000039
gini	0.462634	0.462589	−0.000045

반편향 리셋 벡터는 각 사용자의 과거 클릭 카테고리 분포와 역방향으로 가중치를 부여하는 방식으로 생성하였다. 본 실험에서 제시하는 결과는 d=0.7, alpha=0.7 조건에서 MINDsmall_dev 데이터셋의 impression 로그 전체 73,152건을 대상으로 산출한 것이다.

5.2 정확도 평가

전체 테스트 데이터셋(n=73,152)을 기준으로 산출한 정확도 지표는 표 5와 같다. 정확도 평가에는 뉴스 추천 연구에서 표준적으로 사용되는 네 가지 지표를 활용하였다. group_AUC는 사용자별 impression 단위로 클릭/비클릭 뉴스를 얼마나 잘 구분하여 순위를 매기는지를 나타내며, mean_MRR(Mean Reciprocal Rank)은 실제 클릭된 뉴스가 추천 순위에서 차지한 위치의 역수를 평균한 값으로 최초 클릭 항목의 순위를 반영한다. nDCG@k(Normalized Discounted Cumulative Gain at k)는 상위 k개 추천 항목 내에서 클릭된 뉴스들의 배치 순서를 종합적으로 반영하는 지표로, 본 연구에서는 k=5와 k=10을 사용하였다.

표 5. NRMS 단독 대비 NRMS+DDHR의 추천 정확도 지표 비교(n=73,152)
Table 5. Comparison of recommendation accuracy metrics: NRMS vs. NRMS+DDHR(n=73,152)

Metric	NRMS	NRMS+DDHR	Difference
group_AUC	0.6609	0.6466	−0.0143
mean_MRR	0.3201	0.354	+0.0339
nDCG@5	0.3526	0.3369	−0.0157
nDCG@10	0.415	0.3994	−0.0156

group_AUC는 2.16%, nDCG@5와 nDCG@10은 각각 4.45%, 3.76% 소폭 하락하였으나, mean_MRR은 오히려 10.59% 상승하였다. mean_MRR의 상승은 DDHR이 관련성 점수에 그래프 다양성 점수를 결합하는 과정에서 추천 순위 전체가 재배열되며, 최초 클릭 항목의 순위만을 반영하는 MRR의 특성상 이러한 재배열이 오히려 유리하게 작용했을 가능성을 시사한다.

이러한 해석은 표 6의 편향 완화 결과와 함께 놓고 볼 때 보다 구체화될 수 있다. DDHR 적용 시 top_category_ratio, entropy, gini 계수가 모두 일관되게 개선되었다는 것은, 재순위화 과정에서 사용자의 기존 소비 패턴과 다른 카테고리의 뉴스가 상위권으로 이동하는 방향으로 순위가 재배열되었음을 시사한다. mean_MRR은 실제 클릭된 항목 하나의 순위 이동에 민감하게 반응하는 지표인 반면, group AUC와 nDCG는 클릭/비클릭 전체 후보 간의 상대적 순서를 종합적으로 평가하므로, 편향 완화를 위한 재배열 과정에서 발생한 비클릭 항목들 간의 순서 변화까지 함께 반영되어 상대적으로 하락한 것으로 해석된다. 다만 이는 지표의 정의적 특성에 근거한 해석이며, 개별 impression 단위에서 클릭 항목의 순위가 실제로 어떻게 이동하였는지를 직접 추적하는 분석은 수행하지 못하였다는 한계가 있다. 또한 mean_MRR이 상승한 반면 group AUC와 nDCG가 하락한 현상에 대해서는 지표의 정의적 차이에 근거한 해석을 제시하는데 그쳤으며, 클릭 항목의 원래 순위 구간별 재배열 양상을 직접 분해하거나 대표 사례를 통해 검증하는 심층 분석은 수행하지 못하였다.

표 6. NRMS 단독 대비 NRMS+DDHR의 카테고리 편향 지표 비교(n=73,152)

Table 6. Comparison of category bias Metrics: NRMS vs. NRMS+DDHR(n=73,152)

Metric	NRMS	NRMS+DDHR	Difference
top_category_ratio	0.232755	0.213634	-0.019121
entropy	0.831618	0.861664	0.030046
gini	0.506251	0.463688	-0.042563

5.3 편향 완화 평가

전체 MINDsmall_dev 데이터셋(n=73,152)을 대상으로 생성한 Top-5 추천 리스트의 카테고리 분포를 비교한 결과는 표 6과 같다. DDHR을 적용한 재순위화 추천(NRMS+DDHR)은 NRMS 단독 추천에 비해서 세 가지 평가 지표 모두 편향이 완화되는 경향을 보였다. 지표별 상대적 변화율을 살펴보면, top_category_ratio는 8.2% 감소하였고, entropy는 3.6% 증가하였으며, gini 계수는 8.4% 감소하였다.

top_category_ratio는 특정 카테고리에 대한 추천 집중도를 나타내고, entropy는 추천 카테고리 분포의 다양성을 나타내고, gini 계수는 추천 분포의 불균형 정도를 각각 측정하는 지표이다. 서로 상이한 메커니즘을 갖는 세 지표에서 모두 일관된 개선이 이루어진 것은 DDHR의 반편향 리셋 벡터가 추천 결과의 카테고리 분포를 보다 균형 있게 재구성하는 데 실질적으로 기여하고 있음을 시사한다.

5.4 댐핑 요소 및 결합 가중치 조절 분석

이전 절에서 DDHR은 의도한 정확도-편향 트레이드오프를 명확하게 실현하였다. 정확도 지표는 group_AUC 와 nDCG 계열의 소폭 하락에도 불구하고 mean_MRR이 오히려 상승하는 등, 정확도를 일방적으로 희생시키는 구조가 아님을 확인하였고, 세 편향 지표(top_category_ratio, entropy, gini) 모두 일관되게 편향 완화 방향을 보였다. 단, 앞선 정확도 및 편향 평가는 d=0.7, alpha=0.7 단일 조건에서의 결과였다. 이번 절에서는 이 두 파라미터의 조합에 따라 트레이드오프가 어떻게 달라지는지 체계적으로 탐색하였다.

표 7은 데이터 수 1,000개의 표본을 대상으로 댐핑 요소 d ∈ {0.3, 0.5, 0.7, 0.85, 0.95}와 결합가중치 alpha ∈ {0.5, 0.7, 0.9}의 15가지 조합에 대해 정확도(group_AUC)와 다양성(Entropy)의 변화를 동시에 분석하였다. 실험 결과, alpha가 트레이드오프를 결정하는 주된 요인이며 댐핑 요소는 각 alpha 수준 내에서 미세 조정 역할을 하는 것으로 나타났다. alpha=0.5, 0.7, 0.9의 세 구간은 서로 뚜렷이 분리된 군집을 이루었으며, 동일한 alpha 내에서 댐핑 요소에 의한 변동폭은 alpha가 커질수록 (0.5→0.9) 감소하였다. 이는 결합가중치가 클수록 그래프 확산 점수의 반영 비중이 작아지면서 댐핑 요소의 영향력이 함께 희석되기 때문으로 해석된다. 15개 조합 중 14개는 entropy와 group_AUC 사이에 명확한 단조적 트레이드오프를 보였으나, d=0.95, alpha=0.9 조합은 d=0.85, alpha=0.9 조합에 비해 group_AUC(0.6558 대 0.6563)와 entropy(0.8426 대 0.8431) 모두 낮게 나타나 예외적 지점으로 확인되었다.

표 7. 댐핑 요소와 가중치별 정확성과 다양성 비교 (n=1,000)
Table 7. Comparison of accuracy and diversity across damping factors and weights (n=1,000)

d	a	group AUC	nDCG@5	top_category_ratio	entropy
0.3	0.5	0.6001	0.2939	0.1938	0.8954
0.3	0.7	0.6395	0.3391	0.2061	0.8701
0.3	0.9	0.6551	0.3509	0.2274	0.8455
0.5	0.5	0.6066	0.2957	0.1903	0.8951
0.5	0.7	0.6416	0.3383	0.2089	0.8665
0.5	0.9	0.6553	0.3515	0.2282	0.8455
0.7	0.5	0.6115	0.2967	0.1918	0.8891
0.7	0.7	0.6432	0.3394	0.2113	0.8638
0.7	0.9	0.656	0.3523	0.2292	0.8439
0.85	0.5	0.6148	0.2953	0.1976	0.8841
0.85	0.7	0.6453	0.3406	0.2109	0.8635
0.85	0.9	0.6563	0.3529	0.2298	0.8431
0.95	0.5	0.6185	0.2907	0.2046	0.8776
0.95	0.7	0.6464	0.341	0.2121	0.8605
0.95	0.9	0.6558	0.3534	0.2296	0.8426

1,000건 표본의 예비 탐색에서 d=0.95, alpha=0.9 조합이 d=0.85, alpha=0.9 조합에 비해 정확도와 다양성이 모두 낮은 예외적 현상이 관찰되어, 이 현상의 재현성을 확인하기 위해 10,000건 표본으로 동일한 15개 조합을 재검증하였다. 표 8에서는 데이터셋의 개수를 늘려 10,000개를 대상으로 댐핑 요소 $d \in \{0.3, 0.5, 0.7, 0.85, 0.95\}$와 결합가중치 $alpha \in \{0.5, 0.7, 0.9\}$의 15가지 조합에 대해 정확도(group_AUC)와 다양성(Entropy)을 동시에 측정하였고 이를 그림 1과 같이 시각화하였다.

이전 분석과 동일하게 alpha가 정확도-다양성 트레이드오프의 주된 결정 요인으로 작용하였으며, 댐핑 요소 d는 각 alpha 수준 내에서 미세 조정 역할을 하는 것으로 나타났다. alpha 수준별 group_AUC 변동폭은 alpha=0.5에서 0.0256이었던 반면 alpha=0.9에서는 0.0013으로 크게 줄어, 결합가중치가 클수록 그래프 확산 점수의 영향력이 희석됨을 확인하였다. 15개 조합 중 14개는 entropy와 group_AUC 사이에 완전한 단조적 트레이드오프를 보였으나, d=0.95, alpha=0.9 조합은 d=0.85, alpha=0.9 조합에 비해 group_AUC (0.6614 대 0.6615)와 entropy(0.8392 대 0.8401) 모두 낮게 나타났다.

표 8. 댐핑 요소와 가중치별 정확성과 다양성 비교 (n=10,000)
Table 8. Comparison of accuracy and diversity across damping factors and weights(n=10,000)

d	a	group AUC	nDCG@5	top_category_ratio	entropy
0.3	0.5	0.6033	0.2951	0.2014	0.8920
0.3	0.7	0.6430	0.3367	0.2096	0.8670
0.3	0.9	0.6602	0.3523	0.2251	0.8434
0.5	0.5	0.6120	0.2970	0.1992	0.8922
0.5	0.7	0.6467	0.3382	0.2096	0.8645
0.5	0.9	0.6609	0.3534	0.2258	0.8422
0.7	0.5	0.6186	0.2960	0.1997	0.8891
0.7	0.7	0.6490	0.3388	0.2108	0.8630
0.7	0.9	0.6614	0.3540	0.2263	0.8411
0.85	0.5	0.6234	0.2959	0.2049	0.8835
0.85	0.7	0.6509	0.3409	0.2107	0.8618
0.85	0.9	0.6615	0.3539	0.2266	0.8401
0.95	0.5	0.6288	0.2962	0.2129	0.8764
0.95	0.7	0.6520	0.3412	0.2124	0.8597
0.95	0.9	0.6614	0.3543	0.2265	0.8392

이 현상은 1,000건 예비 표본에서도 동일하게 관찰되었으며 10,000건 표본에서 재현됨으로써, d가 지나치게 높은 극단값에서는 그래프 확산이 협소한 구조로 수렴하여 오히려 추천 순위에 잡음을 더할 수 있다는 가설을 뒷받침하는 안정적인 근거를 확보하였다. 전체 조합 중 d=0.95, alpha=0.7 조합이 정확도 손실(-1.35%)과 다양성 개선(+3.38%)의 균형이 가장 우수한 절충 지점으로 확인되었다.

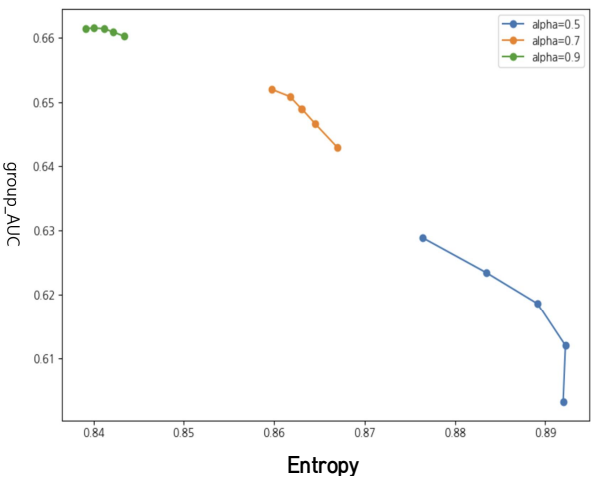

그림 1. 정확성과 다양성 시각화 비교(n=10,000)
Fig. 1. Visualization comparison of accuracy and diversity (n=10,000)

VI. 결 론

본 연구는 MIND의 사용자 클릭 로그를 분석하여 뉴스 추천 서비스 내 확증편향을 검증하고, 이를 완화하는 그래프 확산 기반 재순위화 알고리즘 DDHR을 제안 및 검증하였다. 분석 결과에서 후보군 구성 단계의 시스템 편향이 사용자의 선택 편향보다 약 3배 크게 나타나 초기 필터 버블이 지배적임을 확인하였다. 아울러 종단분석에서는 방문이 거듭될수록 선택 편향이 완만히 강화되는 유의한 추세가 관찰되었으나 인과검정에서는 뚜렷한 신호를 식별하지 못하였다.

이를 바탕으로 사용자의 소비 성향과 반대되는 반편향 리셋 벡터를 도입한 DDHR을 NRMS에 적용하여 전체 테스트 데이터셋(n=73,152)으로 검증한 결과, top_category_ratio(-8.24%), entropy(+3.62%), gini 계수(-8.41%) 등 편향 지표가 일관되게 개선되었고, 이는 1,000건 예비 결과와도 유사하게 재현되어 안정적인 효과임이 확인되었다. 정확도는 group_AUC와 nDCG가 소폭 하락한 반면, mean_MRR은 오히려 10.59% 상승하여, 정확도를 일방적으로 희생시키지 않는 트레이드오프임을 보였다. 나아가 댐핑 요소와 결합가중치의 격자 탐색을 통해, 결합가중치가 트레이드오프의 주된 결정 요인이며 댐핑 요소는 미세 조정 역할을 함을 확인하였다. 다만 댐핑 요소가 극단적으로 높은 특정 지점(d=0.95, alpha=0.9)에서는 정확도와 다양성이 동시에 저하되는 예외 현상이 두 표본 규모에서 일관되게 재현되어, 댐핑 요소상승이 항상 유리하지는 않음을 시사하였다.

본 연구의 시사점은 다음과 같다. 먼저 확증 편향을 단일한 현상으로 다루던 기존 연구와 달리, 본 연구는 편향의 원인을 시스템 요인과 사용자 요인으로 분해하여 각각의 상대적 크기를 정량화함으로써 편향이 형성되는 다층적 구조를 규명하였다. 아울러 로그 기반 관찰 분석만으로는 검증하기 어려웠던 인과적 질문에 대해, 댐핑 요소와 결합가중치를 직접 조작하는 통제된 개입 실험을 통해 실증적 근거를 제시함으로써 관찰연구의 한계를 보완하였다. 나아가 DDHR은 기존 추천 모델의 재학습 없이 후처리 방식으로 적용 가능하며, 두 개의 저차원 파라미터만으로 정확도-편향 트레이드오프를 체계적으로 통제할 수 있다는 점에서 실무적 활용성이 높다.

본 연구의 한계점은 다음과 같다. 본 연구에 사용한 데이터셋에서 관찰 기간이 5~6일로 짧고 대분류 카테고리 수준에 그쳐 장기적이고 세부적 필터 버블은 다루지 못하였다. 또한, 다량 사용자와 소량 사용자 간 편향 비교 분석은 집계 방식에 따라 결과가 상반되는 것으로 나타났다. DDHR 검증은 데이터 정합성 문제로 NRMS 단일 모델에 한정되었다. 파라미터 탐색도 15개 조합에 그쳤고 통계적 유의성 검정을 수행하지 못하였다. 추후에는 지표 차이의 통계적 검증, 다양한 추천 알고리즘으로의 확장, 댐핑 요소를 사용자별로 자동 조절하는 적응형 메커니즘 개발이 필요하다. 또한 클릭 항목의 원래 순위 구간에 따라 DDHR 적용 후 순위가 어떻게 변화하는지를 정량적으로 분해하고, 대표적인 사례에 대한 정성적 분석을 병행함으로써 mean_MRR 상승의 메커니즘을 규명할 필요가 있다.

References

[1] K. Han, "Personalized News Recommendation and Simulation based on Improved Collaborative Filtering Algorithm", Complexity, Vol. 2020, No. 1, Art. no. 8834908, pp. 1-12, Oct. 2020. https://doi.org/10.1155/2020/8834908.

[2] S. Lee and S. W. Kang, "User Understanding and Perceptions of News Recommendation Algorithms : Relationships with Attitude-Consistent News Exposure, News Trust, and News-Seeking Behavior", Korean Journal of Journalism & Communication, Vol. 68, No. 1, pp. 348-385, Feb. 2024. https://doi.org/10.20879/kjjcs.2024.68.1.010.

[3] E. Pariser, "The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think", Penguin, May 2011.

[4] S. Flaxman, S. Goel, and J. M. Rao, "Filter Bubbles, Echo Chambers, and Online News Consumption", Public Opinion Quarterly, Vol. 80,

No. S1, pp. 298-320, Mar. 2016. https://doi.org/10.1093/poq/nfw006.

[5] A. R. Arguedas, C. Robertson, R. Fletcher, and R. Nielsen, "Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review", Reuters Institute for the Study of Journalism, pp. 1-42, Jan. 2022. https://doi.org/10.60625/risj-etxj-7k60.

[6] H. Kang, S. T. Han, B. Jung, and Y. Shin, "On the Algorithm of Recommendation System for Personalization", Journal of The Korean Data Analysis Society Vol. 6, No. 4, pp. 1043-1049, Aug. 2004.

[7] J. S. Lee, "Selective Exposure and Bias by Recommendation Systems", KISO Journal, Vol. 51, pp. 52-55, Jun. 2023.

[8] A. J. Chaney, B. M. Stewart, and B. E. Engelhardt, "How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility", Proceedings of the 12th ACM Conference on Recommender Systems, pp. 224-232, Vancouver, British Columbia, Canada, Oct. 2018. https://doi.org/10.1145/3240323.3240370.

[9] B. Edizel, F. Bonchi, S. Hajian, A. Panisson, and T. Tassa, "FaiRecSys: Mitigating Algorithmic Bias in Recommender Systems", International Journal of Data Science and Analytics, Vol. 9, No. 2, pp. 197-213, Mar. 2020. https://doi.org/10.1007/s41060-019-00181-5.

[10] J. Chen, H. Dong, X. Wang, F. Feng, M. Wang, and X. He, "Bias and Debias in Recommender System: A Survey and Future Directions", ACM Transactions on Information Systems, Vol. 41, No. 3, pp. 1-39, Jul. 2023. https://doi.org/10.1145/3564284.

[11] Y. Zheng, C. Gao, X. Li, X. He, Y. Li, and D. Jin, "Disentangling User Interest and Conformity for Recommendation with Causal Embedding", Proceedings of the Web Conference 2021, Ljubljana Slovenia, pp. 2980-2991, Apr. 2021. https://doi.org/10.1145/3442381.3449788.

[12] D. F. Gleich, "PageRank Beyond the Web", SIAM Review, Vol. 57, No. 3, pp. 321-363, Sep. 2015. https://doi.org/10.1137/140976649.

[13] M. Gori, A. Pucci, V. Roma, and I. Siena, "Itemrank: A Random-Walk Based Scoring Algorithm for Recommender Engines", IJCAI 2007, Vol. 7, pp. 2766-2771, Jan. 2007. https://doi.org/10.1007/978-3-540-77485-3_8.

[14] S. Park, W. Lee, B. Choe, and S. G. Lee, "A Survey on Personalized PageRank Computation Algorithms", IEEE Access, Vol. 7, pp. 163049-163062, Nov. 2019. https://doi.org/10.1109/ACCESS.2019.2952653.

[15] A. Kulesza and B. Taskar, "Determinantal Point Processes for Machine Learning. Foundations and Trends® in Machine Learning", Vol. 5, No. 2-3, pp. 123-286, Dec. 2012. https://doi.org/10.1561/2200000044.

[16] F. Wu, Y. Qiao, J. H. Chen, Wu, C., T. Qi, J. Lian, D. Liu, X. Xie, J. Gao, W. Wu, and M. Zhou, "MIND: A Large-scale Dataset for News Recommendation", Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, pp. 2597-3606, Jul. 2020. https://doi.org/10.18653/v1/2020.acl-main.331.

[17] C. Wu, F. Wu, S. Ge, T. Qi, Y. Huang, and X. Xie, "Neural News Recommendation with Multi-head Self-attention", Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP), pp. 6389-6394, Hong Kong, China, Nov. 2019. https://doi.org/10.18653/v1/D19-1671.

[18] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Neural News Recommendation with Attentive Multi-view Learning", arXiv preprint arXiv:1907.05576, Jul. 2019. https://doi.org/10.48550/arXiv.1907.05576.

[19] M. An, F. Wu, C. Wu, K. Zhang, Z. Liu, and X. Xie, "Neural News Recommendation with Long-and Short-term User Representations",

Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, pp. 336-345, Jul. 2019. https://doi.org/10.18653/v1/P19-1033.

[20] Y. Liu, J. Zhang, Y. Dang, Y. Liang, Q. Liu, G. Guo, and X. Wang, "Cora: Collaborative Information Perception by Large Language Model's Weights for Recommendation", Proceedings of the AAAI Conference on Artificial Intelligence, Pennsylvania Convention Center, Vol. 39, No. 12, pp. 12246-12254, Feb.-Mar. 2025. https://doi.org/10.1609/aaai.v39i12.33334.

[21] R. Wang and W. Lu, "Modeling Multi-interest News Sequence for News Recommendation", arXiv preprint arXiv:2207.07331, Jul. 2022. https://doi.org/10.48550/arXiv.2207.07331.

[22] M. H. Nguyen, T. T. Nguyen, M. N. Ta, T. Le, and H. T. Nguyen, "Co-NAML-LSTUR: A Combined Model with Attentive Multi-view Learning and Long-and Short-Term User Representations for News Recommendation", International Conference on Multi-disciplinary Trends in Artificial Intelligence, pp. 106-119, Dec. 2025. https://doi.org/10.1007/978-981-95-4960-3_9.

[23] S. Vrijenhoek, "Do You MIND? Reflections on the MIND Dataset for Research on Diversity in News Recommendations", arXiv Preprint arXiv:2304.08253, Apr. 2023. https://doi.org/10.48550/arXiv.2304.0825.

[24] T. Qi, F. Wu, C. Wu, and Y. Huang, "Pp-rec: News Recommendation with Personalized User Interest and Time-aware News Popularity", Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Online, pp. 5457-5467, Aug. 2021. https://doi.org/10.18653/v1/2021.acl-long.424.

[25] W. Liu, Y. Xi, J. Qin, X. Dai, R. Tang, S. Li, and R. Zhang, "Personalized Diversification for Neural Re-ranking in Recommendation", 2023 IEEE 39th International Conference on Data Engineering (ICDE), Anaheim, CA, USA, pp. 802-815, Apr. 2023. https://doi.org/10.1109/ICDE55515.2023.00067.

[26] R. Jin and D. Li, "(Debiased) Contrastive Learning Loss for Recommendation (Technical Report)", arXiv preprint arXiv:2312.08517, Dec. 2023. https://doi.org/10.48550/arXiv.2312.08517.

[27] U. Kumar, V. Kumar, and J. N. Kapur, "Normalized Measures of Entropy", International Journal of General Systems, Vol. 12, No. 1, pp. 55-69, Feb. 1986. https://doi.org/10.1080/03081078608934927.

[28] M. L. Menéndez, J. A. Pardo, L. Pardo, and M. D. C. Pardo, "The Jensen-Shannon Divergence", Journal of the Franklin Institute, Vol. 334, No. 2, pp. 307-318, Mar. 1997. https://doi.org/10.1016/S0016-0032(96)00063-4.

[29] N. Nachar, "The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution", Tutorials in Quantitative Methods for Psychology, Vol. 4, No. 1, pp. 13-20, Mar. 2008. https://doi.org/10.20982/tqmp.04.1.p013.

[30] T. Y. Kim, C. Park, S. Kim, M. Kim, W. Lee, and Y. Kwon, "A Wilcoxon Signed-rank Test for Random Walk Hypothesis based on Slopes", Jouranl of the Korean Data and Information Science Society, Vol. 25, No. 6, pp. 1499-1506, Nov. 2014. https://doi.org/10.7465/jkdi.2014.25.6.1499.

[31] E. I. Dumitrescu and C. Hurlin, "Testing for Granger Non-causality in Heterogeneous Panels", Economic Modelling, Vol. 29, No. 4, pp. 1450-1460, Jul. 2012. https://doi.org/10.1016/j.econmod.2012.02.014.

저자소개

현 정 석 (Jung Suk Hyun)

1998년 2월 : 서강대학교 경영학과(경영학박사)
2024년 12월 : 발명 및 지식재산 교육 유공 특허청장 표창장 수상
2025년 2월 : 제주대학교 4단계 BK21사업 교육연구단 연구실적 우수교수상 수상
2002년 10월 ~ 현재 : 제주대학교 경영정보학과 교수
관심분야 : e-마케팅, 창의적 문제해결, 모순해결, 경영 데이터 마이닝

박 찬 정 (Chan Jung Park)

1988년 2월 : 서강대학교 전자계산학과(공학사)
1990년 2월 : KAIST 전산학과 (공학석사)
1998년 2월 :서강대학교 대학원 전자계산학과(공학박사)
1990년 3월 ~ 1994년 2월 : 한국통신 소프트웨어연구소 전임연구원
1998년 2월 ~ 1999년 9월 : 한국통신 멀티미디어연구소 전임연구원
2011년 1월 ~ 2011년 12월 : UC Berkeley 방문학자
2013년 4월 ~ 2015년 2월 : 제주대학교 교육과학연구소 소장
1999년 9월 ~ 현재 : 제주대학교 컴퓨터교육과 교수
관심분야 : 기술기반 교육, 에듀테크, 데이터마이닝, 생성형 점진적 프롬프팅 기반 교수학습 설계