

Agent-Based Modelling Meets Generative AI in Social Network Simulations

<https://arxiv.org/html/2411.16031v1>

0. Introduction

- 온라인 소셜 네트워크에서 정보 확산, 영향력 동학, 커뮤니티 형성 등의 복잡한 사회적 현상을 모사하기 위해 Agent-Based Modelling(ABM)이 많이 활용되어 왔음.
- 그러나 ABM은 에이전트 간 상호작용 규칙이나 정보 흐름 역학을 수작업으로 설정해야 하며 이로 인해 실제 세계 적용가능성이 떨어지고 연구자 편향이 개입될 가능성이 있음.
- 최근 대형 언어모델(LLM)의 인간처럼 언어를 생성하고 추론하는 능력이 사회모델링에도 활용될 가능성이 제기됨.

1. Overview

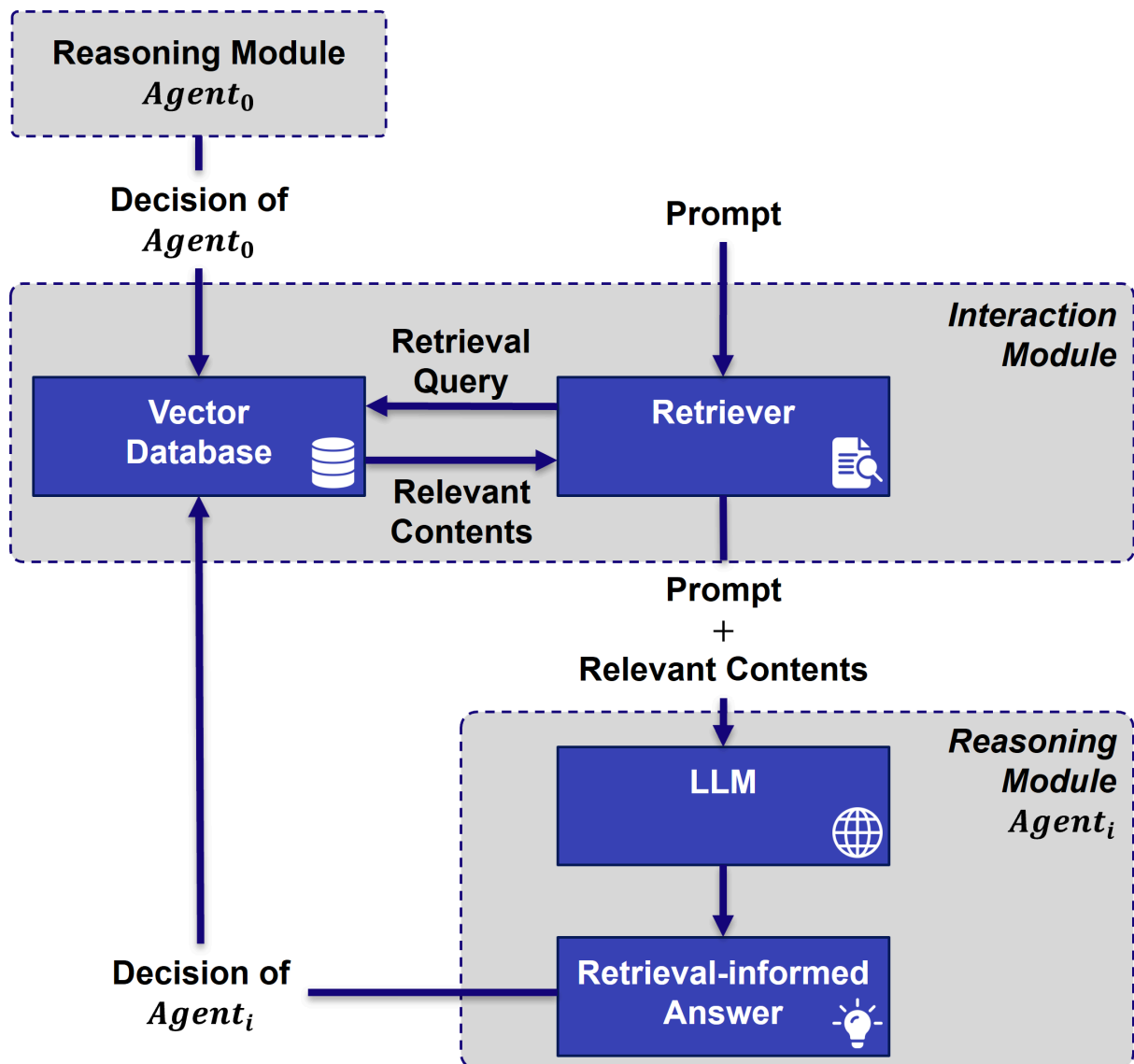
- Characterization Phase : 실제 사용자의 게시글을 이용해 프롬프트 기반으로 성격(personality)과 관심사(interests)를 추출해 에이전트 초기화.
- Simulation Phase : 에이전트들이 실제 소셜미디어 스타일의 환경에서 활동하며 결정(게시/리쉐어/비활동)을 내리도록 구성됨. Reasoning Module과 Interaction Module로 구성됨.
- Interaction Module은 RAG(Retrieval-Augmented Generation) 방식으로 에이전트들이 이전 행동을 기반으로 다른 에이전트의 게시글을 추천받거나 노출됨.

2. Challenges

- 실제 사용자 행동을 잘 반영하는 에이전트 설계의 어려움 : 성격, 관심사, 게시·리쉐어 패턴 등을 단순 규칙으로 정의하기 힘들.

- 에이전트 간 상호작용, 노출(rec) 전략, 추천 노출 영향 등이 복잡하며 정형화된 매개변수로 포착하기 어려움.
- 대형 언어모델을 활용할 경우 프롬프트 설계, 컨텍스트 창(context window) 용량, 외부 기억 저장(vector database) 구축 등 기술적·리소스적 이슈 존재
- 시뮬레이션의 현실성(fidelity)을 확보하기 위해 실제 사용자 데이터와 비교 평가가 필요하고, 추천 전략이 사회적 영향(동질성, 극화) 등에 미치는 영향도 고려해야 함.

3. Method



- **3.1 Characterization Phase:**

- 실제 사용자의 트윗 데이터를 입력으로 사용. 프롬프트 기반으로 LLM에게 성격 및 관심사 기능을 묻고 응답 결과를 에이전트의 초기 상태로 설정함.
- 이 방식으로 성격(personality traits) 및 관심사(interests)를 설정하면, 에이전트가 다양하고 현실감 있는 사용자처럼 행동 가능함.
- **3.2 Simulation Phase:**
 - 두 모듈 존재:
 - Reasoning Module : 각 에이전트가 매 반복(iteration)마다 게시(original), 리쉐어(reshare), 비활동(idle) 중 하나를 선택. LLM이 프롬프트를 통해 해당 선택과 이유(reason), 콘텐츠(content)를 생성함.
 - Interaction Module : 에이전트의 행동과 콘텐츠를 벡터 데이터베이스에 저장. 이후 RAG를 통해 다른 에이전트의 게시물이나 추천 콘텐츠를 검색해 각 에이전트의 다음 행동에 영향을 줌. 이는 노출 전략(recommendation strategy)이 시뮬레이션에 반영됨.
- 추천 전략(노출 방식) 설정 :
 - Preference-based recommendation : 에이전트의 관심사에 맞춘 콘텐츠 노출
 - Random recommendation : 관심사와 상관없이 랜덤한 콘텐츠 노출
 이러한 전략이 커뮤니티 형성 및 에코챔버 발생 여부에 어떤 영향을 미치는지 탐구함.

4. Experiments

Table 1: Top-10 Keywords in Real Case and Simulation

Real Case	Simulation
realdonaldtrump	trump
trump	president
president	biden
biden	administration
joe Biden	freedom
people	maga
america	actions
covid	change
time	covid
maga	state

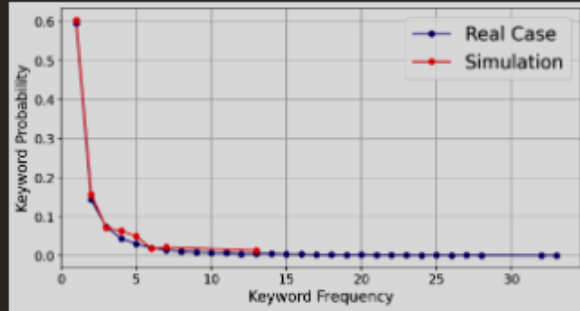
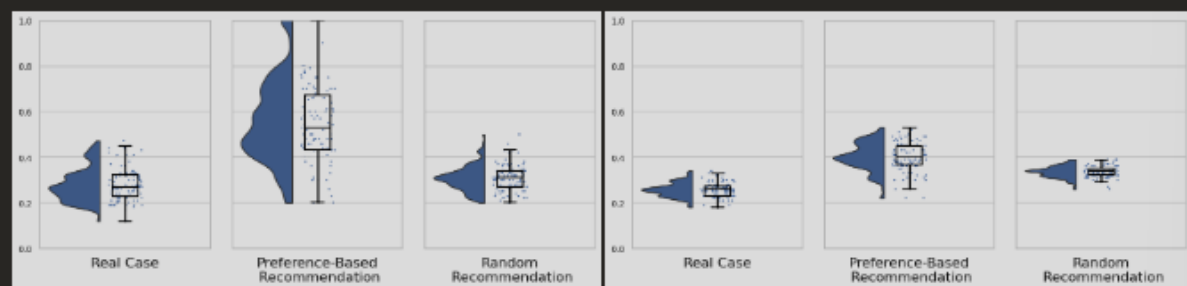


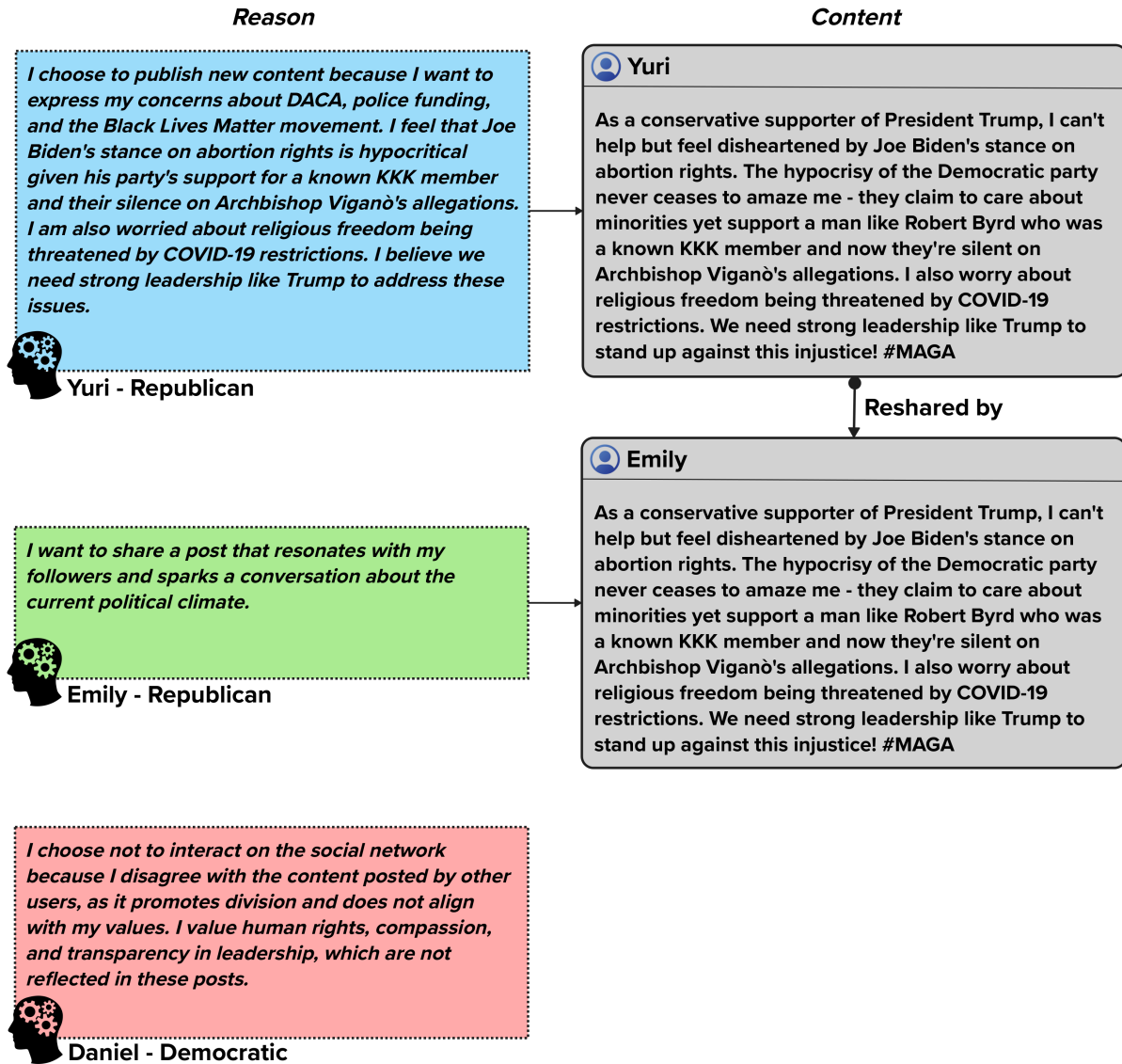
Figure 3: The distributions of keywords usage in real Twitter discussions and the simulation.

The x-axis represents the number of occurrences of a keyword, while the y-axis represents the probability of encountering a keyword with a specific frequency.

Table 2: Results of Political Leaning Analysis and Interaction Patterns

	Political Leaning Analysis				Interaction Patterns		
	Consistent Users	Changed Users	Spearman Coefficient	P-Value	Original Publications	Non Interactions	Reshares
Pref-Based	88%	12%	0.494	<0.05	38.9%	10.5%	50.6%
Random 1	86%	14%	0.498	<0.05	90.0%	1.6%	8.4%
Random 2	88%	12%	0.490	<0.05	89.1%	2.5%	8.4%
Random 3	89%	11%	0.509	<0.05	90.1%	1.9%	8.0%





	Homophily Analysis		Controversy Metrics		
	Modularity	Homophily?	RWC	BCC	GMCK
Pref-Based	0.375	Yes, -10% inter-cluster edges (p-value <0.05)	0.692	0.423	0.334
Random 1	0.416	No, +4% inter-cluster edges	0.423	0.189	0.071
Random 2	0.416	No, +6% inter-cluster edges	0.541	0.230	0.268
Random 3	0.428	No, +4% inter-cluster edges	0.431	0.362	-0.009

- Intel i7-11800H, 16 GB RAM, NVIDIA RTX 3060 Laptop GPU. 약 15시간/시뮬레이션(10 iteration 기준) 수행.
- 오픈소스 LLM으로 Dolphin 2.1 Mistral 7B 사용 (비상업용 모델)

- 트위터 2020 미국 대선 관련 데이터: 2020년 6월 ~ 12월, 12 백만 트윗, 1.1 백만 유저.
- 봇으로 판단된 계정은 제거함. 사용자 중 100명을 선정해 정치 성향(annotation) 부여: 공화당 73명, 민주당 27명. 이들을 에이전트로 초기화함.
- RQ1: 키워드 사용(top 10), 관심사(정치 성향), 콘텐츠 유사성(semantic similarity) 분석.
- RQ2: 에이전트 간 리쉐어 활동(sharing activity), 동질성(homophily), 에코챔버(echo chamber) 분석.

5. Results

- RQ1 결과:
 - 키워드 사용 빈도 분포에서 실제 사용자와 에이전트 간 유사성 확인됨.
 - 정치성향 상관관계 : preference-based 추천과 랜덤 추천 모두에서 에이전트의 정치성향이 원사용자와 강한 양의 상관관계 나타냄(Spearman ≈ 0.49 , $p < 0.05$)
 - 콘텐츠 유사성(Self-Similarity)에서 에이전트가 실제 사용자보다 주제 집중도가 높게 나타남(권고: preference-based 설정 시) → 에이전트가 반복적으로 유사 콘텐츠를 생산함.
- RQ2 결과:
 - 리쉐어 활동 : preference-based 추천에서는 리쉐어 비율이 높았음($\approx 50.6\%$) vs 랜덤 추천($\approx 8\%$) → 추천 전략이 활동 형태에 큰 영향.)
 - 동질성 분석 : preference-based 추천 시 에이전트 간 이질적(inter-cluster) 연결이 약화되며 동질적 집단 형성됨(homophily 강화) $\sim 10\%$ 적음.
 - 에코챔버 분석 : preference-based 추천 시 논쟁성(controversy) 지표가 더 높게 나타남 → 노출 전략이 에코챔버 강화 요인으로 작용

6. Insight

- 이 연구는 LLM-기반 제너레이티브 에이전트를 ABM 프레임워크에 통합함으로써 소셜 미디어 시뮬레이션의 현실성(fidelity)을 높일 수 있음을 보여줌.
- 특히 shared embedding 수준이 아니라 개별 사용자성향과 관심사를 반영해 에이전트를 초기화한 점이 중요함.

- 규칙 기반 에이전트 설계 대비 유연성 및 표현력 향상
- 실제 사용자 데이터와 비교 평가하여 검증됨
- 사용된 LLM이 여전히 규모나 맥락 처리 능력 면에서 최신 모델 대비 제약이 있을 수 있음
- 시뮬레이션이 10 iteration 등 상대적으로 짧은 시간/단계로 진행됨 → 장기 변화나 동적 환경 변화 반영이 제한되어 있을 수 있음
- 추천 전략이 단순화된 형태(pseudo recommendation)로 구현됨 → 실제 추천 시스템 복잡도를 반영하기 어려움