

自由記述嗜好を考慮した多人数レシピ推薦における嗜好統合手法の比較分析

東 颯人[†] 宮森 恒[†]

[†] KyotoSangyoUniversity 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †{g2354323,miya}@cc.kyoto-su.ac.jp

あらまし 本稿では、自由記述で与えられる多様な曖昧な嗜好を前提とした多人数レシピ推薦の問題に取り組む。従来のレシピ推薦研究は単一ユーザを主対象としており、多人数の嗜好衝突を体系的に扱う研究は限られている。また、グループ推薦に関する既存手法の多くは、定型的に表現された嗜好入力を前提としており、自由記述嗜好を考慮した場合の有効性は十分に検証されていない。そこで、本稿では、自由記述で与えられる嗜好に対する多人数レシピ推薦において、嗜好統合手法が推薦品質および満足度の公平性に与える影響を定量的に分析する。具体的には、自由記述嗜好を正規化する前処理として、埋め込み表現、辞書ベース手法、大規模言語モデルを用いた補完手法を比較し、それらが多人数嗜好統合に与える影響を調査する。さらに、平均満足度最大化、最小満足度最大化、および満足度分散を考慮した統合手法を対象に、グループ人数や嗜好衝突度を変化させた設定のもとで性能評価を行う。

キーワード 多人数レシピ推薦, 自由記述, 嗜好統合, 嗜好正規化, 嗜好衝突度, 公平性

1 はじめに

近年、家族や友人、同僚など、複数人で食事をする機会は日常的に存在している。一方で、レシピサイトやSNSの普及により利用可能なレシピの選択肢は爆発的に増加しており、「何を食べるか」という意思決定は必ずしも容易ではない。特に複数人で食事を決める場合、各人の嗜好や制約が異なるため、意思決定の負担はさらに大きくなる。このような背景から、レシピ推薦システムに対する需要は高まっている。

従来のレシピ推薦研究の多くは、単一ユーザを対象とした推薦を主眼としており、個人の嗜好に基づいて最適なレシピを提示する手法が数多く提案されてきた [1], [2]。一方で、複数人の嗜好を同時に考慮する多人数レシピ推薦に関する研究は限定的であり、特に嗜好が衝突する状況を体系的に扱った研究は十分とは言えない。また、グループ推薦に関する既存研究の多くは、あらかじめ定型的に整理された嗜好入力を前提としており、自然言語による自由記述嗜好を入力とした場合の有効性については十分に検証されていない。

しかし、実際の利用場面では、「さっぱりしたものが食べたい」「辛い料理は苦手」「子供向けがよい」といったように、嗜好は自由記述で曖昧かつ主観的に表現されることが多い。このような自由記述嗜好には、表記揺れ、未知語、文脈依存の意味、曖昧な評価表現などが含まれ、嗜好を数値化・ベクトル化する際の大きな課題となる [3]。さらに、多人数の嗜好を単純に平均化するような統合手法では、一部のユーザの満足度が著しく低下し、推薦結果の公平性が損なわれる可能性がある。

そこで本研究では、自由記述で与えられる嗜好を前提とした多人数レシピ推薦に着目し、嗜好の正規化手法および嗜好統合手法の違いが、推薦品質および満足度の公平性に与える影響を定量的に分析することを目的とする。具体的には、自由記述嗜好

を数値表現へ変換する前処理として、埋め込み表現のみを用いる手法、辞書ベースによる正規化手法、大規模言語モデルを用いた文脈補完手法を比較する。さらに、平均満足度最大化、最小満足度最大化、満足度分散を考慮した嗜好統合手法を対象に、グループ人数や嗜好衝突度を変化させた条件下で評価実験を行う。

本研究の貢献は以下の三点にまとめられる。第一に、自由記述嗜好を対象とした多人数レシピ推薦において、異なる嗜好正規化手法が推薦結果に与える影響を体系的に比較・分析する点である。第二に、複数の嗜好統合手法を公平性の観点から評価し、満足度最大化と個人間の納得感のトレードオフを明らかにする点である。第三に、嗜好衝突度やグループ人数といった実利用を想定した条件変化が推薦性能に及ぼす影響を実験的に示す点である。

本論文の構成は以下の通りである。第2章では関連研究について述べる。第3章では、自由記述嗜好を考慮した多人数レシピ推薦の提案手法について説明する。第4章では、実験設定および評価指標について述べる。第5章では、評価実験の結果を示し、正規化手法および嗜好統合手法の違いが推薦性能に与える影響について考察する。最後に、第6章で本研究をまとめ、今後の課題について述べる。

2 関連研究

2.1 単語・文書のベクトル表現

自然言語処理分野では、Word2Vec や fastText などの分散表現を用いて単語の意味を捉える手法が提案されてきた [1]。レシピ推薦においても、レシピ内の単語をベクトル化することで、従来のキーワード検索では不可能だった「材料名は異なるが意味的に近いレシピ」の抽出が可能となっている (矢野ら, 2013)。近年では、OpenAI の text-embedding-3-small に代表される高

次元の埋め込みモデルが登場し、自由記述文全体のコンテキストを高精度に捉えることが可能となった（川寄ら, 2024）。一方で、これらの埋め込み表現は意味的類似性の判定には優れるものの、「特定食材の除外」といった明示的な制約条件や、ユーザーごとの嗜好の強弱を直接的に制御しにくいという課題がある。

2.2 食材嗜好に関する研究

ユーザーの食嗜好を扱う研究は、アンケートや過去の調理履歴に基づく定量化が主流であった。例えば、食材単位の「好き・嫌い」を数値化し、それに基づいてレシピをフィルタリングする手法がある（高畑ら, 2011）。しかし、実際の利用シーンでは「こってりしたもの」や「子供が喜ぶもの」といった抽象的かつ自由な形式で嗜好が表現されることが多い。これに対し、大規模言語モデル（LLM）を用いて自由回答を特定のレシピカテゴリや類似単語へ変換・正規化する試みが行われているが、多種多様な表現をどの程度の解像度で集計すべきかについては議論の余地がある。

2.3 多人数レシピ推薦・グループ推薦

複数ユーザーを対象とするグループ推薦では、個人の満足度の平均値を最大化する手法や、不満を持つユーザーを最小限に抑える手法などの嗜好統合アルゴリズムが提案されてきた [4], [5]。しかし、先行研究の多くは「ユーザーの嗜好がすでに数値化されている」ことを前提としており、自由記述から得られた曖昧な嗜好情報をどのように統合し、推薦結果へ反映させるべきかの検討は不十分である。特に人数が増えるほど嗜好の把握と「すり合わせ」の難易度は増大するため（川寄ら, 2024）、自由記述の正規化プロセスと嗜好統合アルゴリズムの組み合わせが推薦精度およびユーザーの納得感に与える影響を分析することは、実用上極めて重要である。

3 提案手法

本研究では、自由記述で与えられる多人数の嗜好を対象としたレシピ推薦において、嗜好の正規化・嗜好統合・推薦・評価を一貫したパイプラインとして設計し、正規化手法と統合手法の組み合わせが推薦品質および公平性に与える影響を分析する手法を提案する。

3.1 手法全体の流れ

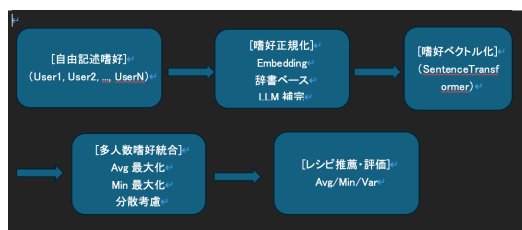


図1 自由記述嗜好を考慮した多人数レシピ推薦の全体構成

本研究では、自由記述で与えられる多人数の嗜好に対して、

嗜好正規化手法および多人数嗜好統合手法を組み合わせ、推薦品質および満足度の公平性への影響を比較分析する。

提案手法は以下の5段階から構成される。

1. 自由記述嗜好の入力
2. 自由記述嗜好の正規化
3. ユーザ嗜好のベクトル化
4. 多人数嗜好の統合
5. レシピ推薦および評価

各処理段階を明確に分離することで、どの要素が推薦結果に影響を与えているかを定量的に比較可能とする。

3.2 自由記述嗜好の正規化

自由記述嗜好は、表記揺れ、曖昧表現、暗黙的な制約を含むため、そのままでは推薦に利用しにくい。そこで本研究では、以下の3手法を用いて正規化を行い、比較する。

3.2.1 埋め込み単独手法

各ユーザーの自由記述嗜好文全体を文埋め込みモデルに直接入力し、意味空間上のベクトルとして表現する。文脈的意味を包括的に捉えられる一方で、調理制約や忌避食材といったハード制約を明示的に扱いにくいという課題を持つ。本研究では本手法をベースラインとする。

3.2.2 辞書ベース手法

本研究では、辞書ベース手法として、形態素解析に基づくルールベースの嗜好正規化を行った。具体的には、MeCabを用いて自由記述嗜好文を形態素解析し、品詞情報に基づいて名詞、形容詞、および動詞のみを抽出する。これにより、料理嗜好に関係しやすい語を選択的に残し、助詞や助動詞などの文法的要素を除去した簡潔な表現へと正規化する。

本手法では、あらかじめ定義した品詞選択ルールを簡易的な辞書として用いることで、自由記述文を構造化された単語列へ変換する。正規化後のテキストは、多言語文埋め込みモデル（SentenceTransformer）を用いてベクトル表現へ変換し、推薦計算に利用する。

一方で、本手法は単語単位の抽出に基づくため、文脈的な意味関係や暗黙的な嗜好・制約を捉えることが困難であり、表現の柔軟性に限界があるという課題を有する。

3.2.3 LLMによる文脈補完手法

本研究では、自由記述嗜好に含まれる曖昧表現や暗黙的な制約を明示化するため、大規模言語モデル（LLM）を用いた文脈補完手法を導入する。具体的には、ユーザーの自由記述嗜好文に対して、料理推薦に適した「嗜好の正規化文」へ書き換える処理を行う。

LLM補完は以下の手順で実施した。まず、LLMに対して、「暗黙の嗜好や制約を明示すること」、「好みの味や調理法、制約を具体化すること」、「一文から二文の簡潔な説明文とすること」、「新たな嗜好を追加しないこと」といった条件を明示的に指示する。その上で、自由記述嗜好文を入力し、文脈を保ったまま構造化された嗜好表現へと変換する。

例えば、「出汁 旨味 和食 揚げ物 NG」という自由記述に対しては、出汁の効いた和食を好み、揚げ物を避けたいという嗜

好・制約が明示的に表現された文章へと補完される。このように、LLM を用いることで、辞書ベース手法では捉えにくい文脈の意図や暗黙的制約を正規化嗜好として反映することが可能となる。

3.3 ユーザ嗜好のベクトル化

正規化後の嗜好情報は、文埋め込みモデル (multilingual-e5-base) を用いてベクトル化する。埋め込み単独手法では原文を入力とし、辞書ベースおよび LLM 手法では正規化後のテキストを入力とすることで、正規化方法以外の条件を統一した比較を行う。

3.4 多人数嗜好の統合

複数ユーザの嗜好ベクトルを統合し、グループ全体を表す嗜好ベクトルを生成する。このような嗜好統合は、グループ推薦における基本的課題として広く研究されてきた [4], [5]。本研究では以下の統合手法を比較対象とする。

3.4.1 平均満足度最大化

ユーザ嗜好ベクトルの平均をグループ嗜好として用いる手法であり、全体として高い平均満足度を得やすい一方、一部ユーザの不満が無視される可能性がある。

3.4.2 最小満足度最大化

各ユーザの満足度のうち最小値を重視する手法であり、強い不満を回避できるが、推薦結果が無難なものに偏りやすい。

3.4.3 満足度分散を考慮した手法

平均満足度に加えてユーザ間の満足度分散を考慮し、満足度の偏りが大きい推薦を抑制する。これにより、公平性と推薦品質のバランスを図る。

3.5 レシピ推薦と評価

各レシピをテキスト表現として埋め込みベクトル化し、グループ嗜好ベクトルとのコサイン類似度に基づいてランキングを生成する。評価指標として、平均満足度、最小満足度、満足度分散を用い、推薦品質と公平性を総合的に評価する。

4 実験

4.1 研究課題

本研究では、自由記述嗜好を考慮した多人数レシピ推薦における、嗜好正規化手法および嗜好統合手法の有効性を検証するため、以下の研究課題 (RQ) を設定する。

- **RQ1**: 自由記述嗜好の正規化手法は、ユーザの潜在的な意図をどの程度正確にベクトル空間へ反映できるか。
- **RQ2**: 異なる嗜好統合手法は、グループ全体の満足度と個人間の公平性にどのような影響を与えるか。

4.2 実験設定

4.2.1 データセット

実験には、楽天レシピ API から取得した複数カテゴリ (和食・洋食・中華・その他) のレシピデータを用いた。

4.2.2 比較手法

自由記述嗜好の正規化手法として、以下の3手法を比較対象とする。

- 埋め込み単独手法
- 辞書ベース手法
- LLM による文脈補完手法

さらに、多人数嗜好統合手法として、平均満足度最大化、最小満足度最大化、および満足度分散を考慮した手法を用いる。

4.2.3 評価指標

推薦結果の評価には、以下の指標を用いる。

- **平均満足度**: グループ全体の推薦品質を表す指標。
- **最小満足度**: 最も満足度の低いユーザの満足度を表し、公平性を評価する指標。
- **満足度分散**: ユーザ間の満足度のばらつきを表す指標。

4.3 実験手順

まず、各ユーザの自由記述嗜好を、嗜好正規化手法ごとにベクトル化する。次に、異なる嗜好統合手法を用いてグループ嗜好ベクトルを生成し、レシピとの類似度に基づいて推薦を行う。最後に、各評価指標を算出し、手法間で比較分析を行う。

4.4 結果

表1に、衝突度および制約強度の異なる9条件における、正規化手法別の top1 推薦に対する満足度指標 (Avg, Min, Var) および推薦可能レシピ数を示す。

4.5 正規化手法ごとの推薦性能比較

まず、全条件を通じて平均満足度 (Avg) は 0.624~0.677 の範囲に分布しており、正規化手法間で極端な差は見られなかった。同様に、最小満足度 (Min) も 0.582~0.657 の範囲に収まり、全体としていずれの手法においても一定水準以上の推薦品質が確保されていることが確認された。一方で、条件ごとに最も高い満足度を示す正規化手法には違いが見られ、衝突度および制約の強さに応じた傾向が確認された。低衝突条件 (1~3) では、Embedding, 辞書ベース, LLM のいずれの手法においても Avg および Min に大きな差は見られなかった。例えば1【低×弱】では Embedding 手法が Avg=0.645 と最も高い値を示したものの、辞書ベース (0.633), LLM (0.624) との差は小さく、手法間の優劣は限定的であった。

衝突度が中程度の条件 (4~6) では、LLM 補完手法の優位性がより明確となった。特に4【中×弱】および5【中×中】では、LLM 手法が Avg および Min の両面で他手法を上回り、4では Avg=0.673, Min=0.657 と最も高い値を示した。一方で6【中×強】では、Avg は辞書ベース手法、Min は Embedding 手法が最大となり、制約が強い条件下では LLM 手法が必ずしも最良とはならないことが示された。

さらに衝突度が高い条件 (7~9) では、7【強×弱】において辞書ベースおよび LLM 手法が同程度の高い Avg を示したのに対し、8【強×中】および9【強×強】では LLM 補完手法が Avg および Min の両面で最も高い値を記録した。特に9では、LLM 手法が Avg=0.669 と他手法を上回っており、高衝突・高

表 1 条件別・正規化手法ごとの満足度指標 (Top1)

条件	衝突×制約	手法	Avg	Min	Var	推薦数
1	低×弱	Embedding	0.645	0.599	0.0009	5
		辞書	0.633	0.582	0.0012	5
		LLM	0.624	0.586	0.0004	5
2	低×中	Embedding	0.645	0.603	0.0011	5
		辞書	0.632	0.597	0.0009	5
		LLM	0.651	0.603	0.0013	5
3	低×強	Embedding	0.649	0.615	0.0005	5
		辞書	0.647	0.621	0.0004	5
		LLM	0.656	0.615	0.0008	5
4	中×弱	Embedding	0.659	0.624	0.0004	5
		辞書	0.652	0.619	0.0004	5
		LLM	0.673	0.657	0.0002	5
5	中×中	Embedding	0.658	0.616	0.0006	2
		辞書	0.654	0.612	0.0007	2
		LLM	0.677	0.639	0.0007	2
6	中×強	Embedding	0.659	0.623	0.0002	5
		辞書	0.669	0.642	0.0004	5
		LLM	0.648	0.616	0.0003	5
7	強×弱	Embedding	0.642	0.624	0.0004	5
		辞書	0.667	0.654	0.0001	5
		LLM	0.667	0.647	0.0002	5
8	強×中	Embedding	0.666	0.620	0.0009	2
		辞書	0.659	0.618	0.0008	2
		LLM	0.675	0.637	0.0007	2
9	強×強	Embedding	0.661	0.618	0.0007	5
		辞書	0.656	0.619	0.0006	5
		LLM	0.669	0.624	0.0008	5

表 2 条件別に最も高い Avg を示した正規化手法

条件	衝突×制約	最良手法 (Avg)
1	低×弱	Embedding
2	低×中	LLM
3	低×強	LLM
4	中×弱	LLM
5	中×中	LLM
6	中×強	辞書
7	強×弱	辞書/LLM
8	強×中	LLM
9	強×強	LLM

表 3 衝突度別における統合手法間の順位一致傾向 (5 人条件)

衝突度	平均最大化	最小最大化	分散考慮
低 (1-3)	上位 1-2 位一致	上位 1-2 位一致	変動大 (特に 3)
中 (4-6)	1 位は概ね一致	1 位は概ね一致	全順位変動多い
高 (7-9)	一致限定的	一致限定的	ほぼ全条件で変動

制約条件における有効性が示唆される。

満足度分散 (Var) は、全条件において 0.0001~0.0013 と非常に小さく、いずれの正規化手法においても極端な不満が発生していないことが分かる。ただし、低衝突条件 1 では辞書ベース手法が Var=0.0012 と比較的大きな値を示した一方、LLM 補完手法は 4【中×弱】において Var=0.0002 と最も小さく、満足度のばらつきを抑える効果が確認された。

また、推薦数については、制約が中程度かつ衝突度が中～高の条件 (5, 8) において、いずれの正規化手法でも推薦レシピ数が 2 件に制限される現象が確認された。これは、嗜好制約の組合せによって候補集合自体が大きく縮小される状況を反映していると考えられる。

4.6 嗜好統合手法別ランキングの安定性

平均満足度最大化、最小満足度最大化、分散考慮の 3 つの統合手法について、ランキング変動を比較した。衝突度が低い条件 (1~3) では、平均満足度最大化および最小満足度最大化において、上位 1~2 位の順位が多くの手法で一致しており、ランキングは比較的安定していた。一方、分散考慮手法では、特に 3 (低×強) において順位変動が大きく、公平性を重視する統合が推薦順位に強く影響することが示唆された。衝突度が中低度の条件 (4~6) では、いずれの正規化手法においても、上位 1 位は大きくの場合一致したが、2 位以降の順位は大きく変動した。特に分散考慮手法では、全順位が異なるケースが多く見られた。衝突度が高い条件 (7~9) では、平均満足度最大化および最小満足度最大化においても順位の一貫性は限定的であり、分散考慮手法ではほぼすべての条件で異なるランキングが生成された。

4.7 グループ人数およびユーザ評価との関係

グループ人数を変更した実験では、人数を減らすにつれて Avg および Min が上昇し、満足度分散 (Var) は減少する傾向が確認された。特に 4 人グループの場合に最も高い満足度が得られた。また、5 人グループにおいて推薦数が 2 件に制限されていた条件では、人数を減らすことで推薦数が 4 件に増加することが確認された。ユーザ評価との一致については、9 (衝突度：高×制約：強) においてのみ、LLM 手法の平均満足度最大化によるランキングがユーザ評価と一致した。他の条件では、一致するランキングは確認されなかった。

5 考察

本実験から、自由記述嗜好を前提とした多人数レシピ推薦において、正規化手法および嗜好統合手法が推薦結果に与える影響が明らかとなった。このような推薦品質と公平性のトレード

オフは、従来のグループ推薦研究においても指摘されてきた課題である [4], [5]。まず、正規化手法については、全体的な満足度指標の絶対値に大きな差は見られなかったものの、衝突度および制約が高まる条件下では、LLM 手法が平均満足度および最小満足度の両面で優位となる傾向が確認された。これは、LLM による補完が、自由記述嗜好に含まれる曖昧な表現や潜在的制約をより適切に解釈できたためと考えられる。一方で、6 (中×強) のように制約が厳しい条件では、LLM 手法の性能が必ずしも最良とはならなかった。これは、制約が過度に強い場合、LLM による意味拡張がかえって制約解釈の揺らぎを生み、満足度の低下につながった可能性を示唆している。嗜好統合手法に関しては、平均満足度最大化は多くの条件で安定した数値を示した一方、ランキングの差異が小さく、衝突状況を十分に反映できていない可能性がある。最小満足度最大化および分散考慮手法は、公平性を重視する観点では有効であるが、特に衝突度が高い条件ではランキングの不安定性が顕著であった。さらに、ユーザ評価との一致が限定的であった点から、数値指標としての満足度と実際の人間の評価との間には乖離が存在することが示された。特に高衝突・高制約条件においてのみ LLM 手法が一致したことは、人間の判断が暗黙的に高度な意味統合を行っている可能性を示唆している。以上より、自由記述嗜好を扱う多人数レシピ推薦においては、LLM を用いた正規化が有効となる場面は多いものの、制約の強さや衝突度に応じて手法を使い分ける必要があることが明らかとなった。

6 ま と め

本研究では、自由記述で与えられる多様かつ曖昧な嗜好を前提とした多人数レシピ推薦問題に対して、嗜好正規化手法および多人数嗜好統合手法の違いが推薦品質および満足度の公平性に与える影響を定量的に分析した。

まず、自由記述嗜好の正規化手法として、埋め込み単独手法、辞書ベース手法、および LLM による文脈補完手法を比較した。その結果、全条件を通じて平均満足度および最小満足度の絶対値に大きな差は見られなかったものの、衝突度および制約強度が高い条件においては、LLM 補完手法が平均満足度および最小満足度の両面で他手法を上回る傾向が確認された。特に高衝突・高制約条件では、LLM 補完手法がグループ内の満足度のばらつきを抑制し、公平性の高い推薦を実現できることが示唆された。

一方で、制約が強い条件の一部では、LLM 補完手法が必ずしも最良の結果を示さないケースも確認され、意味補完による解釈の揺らぎが満足度低下につながる可能性が示された。この結果は、自由記述嗜好に対する高度な意味解釈が有効である一方で、制約条件の厳密な取り扱いには慎重な設計が必要であることを示している。

次に、多人数嗜好統合手法の比較から、平均満足度最大化は全体として安定した推薦品質を示す一方、特定ユーザの不満が反映されにくい傾向が確認された。最小満足度最大化および満足度分散を考慮した手法は、公平性の観点では有効であるもの

の、衝突度が高い条件では推薦ランキングの変動が大きくなることが明らかとなった。これらの結果から、多人数推薦においては、推薦品質と公平性の間に明確なトレードオフが存在することが示された。

以上より、自由記述嗜好を扱う多人数レシピ推薦においては、嗜好正規化と嗜好統合を分離して設計し、衝突度や制約強度に応じて LLM による文脈補完と公平性を考慮した嗜好統合手法を適切に組み合わせることが、推薦品質と利用者の納得感を両立する上で有効であると結論付ける。

今後の課題として、評価対象となるユーザ数およびレシピ数の拡大、実利用環境における主観評価の導入、および LLM による嗜好補完過程の説明可能性向上が挙げられる。

7 謝 辞

本研究の一部は科研費 23K11342 の助成を受けたものである。

文 献

- [1] 矢野 達也, 林 豊洋, 大橋 健, “単語のベクトル表現を用いたレシピ推薦システム,” 人工知能学会研究資料, SIG-Challenge-053-4(3/24),2019.
- [2] 高畑 麻理, 上田真由美, 中島 伸介 “食材に対する好き嫌いを考慮した料理レシピ推薦手法の提案,” DEIM Forum 2011 E3-5
- [3] 川崎翔太, 飯塚雅文, 石川玲奈, 三村明農, 森田航平, 吉村魁人, 淵田孝康 “多数の参加者の好みを考慮したレシピ推薦システムの提案,” 2024 年度電気・情報関係学会九州支部連合大会, 09-1P-08, 2024. DOI: 10.11527/jceek.2024.0_184.
- [4] J. Masthoff, “Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers,” User Modeling and User-Adapted Interaction, 2004.
- [5] S. Amer-Yahia et al., “Group Recommendation: Semantics and Efficiency,” Proc. VLDB, 2009.
- [6] R. Burke, “Hybrid Recommender Systems: Survey and Experiments,” User Modeling and User-Adapted Interaction, 2002.