

Estratégias para Aprimorar a Diversidade Categórica e Geográfica de Sistemas de Recomendação de POIs

Rodrigo Carvalho, Leonardo Rocha
Universidade Federal de São João del-Rei
(rodrigolcrocha)@ufsj.edu.br

Abstract

Currently, so-called Recommendation Systems (SRs) have been used to assist users in discovering relevant Points of Interest (POIs) on Location-Based Social Networks (LBSN), such as FourSquare and Yelp. Given the main challenges of data-sparse and geographic influence in this scenario, most of the work on POI recommendations has focused only on improving the effectiveness (i.e. accuracy) of the systems. However, there is a growing consensus that just effectiveness is not sufficient to assess the practical utility of these systems. In real scenarios, categorical and geographic diversities were identified as the main complementary dimensions for assessing user satisfaction and the usefulness of recommendations. The works in the literature are concentrated on only one of these concepts. In this work, we propose a new post-processing strategy, which combines these concepts in order to improve the user's interest in POIs. Our experimental results in the Yelp data sets show that our strategy can improve user satisfaction, considering different SRs and multiple diversification metrics. Our method is capable of improving diversity by up to 120 % without significant losses in terms of effectiveness.

Keywords

Diversity, Recommendation Systems, Points of interest

1 Introdução

Este artigo apresenta um resumo dos resultados obtidos na dissertação defendida no Programa de Pós-graduação em Ciência da Computação da Universidade Federal de São João del-rei em 12 de agosto de 2019. O trabalho foi desenvolvido pelo primeiro autor, num período de 24 meses, sob a orientação do último autor.

Atualmente, os Sistemas de Recomendação (SRs) têm sido usados para ajudar os usuários a descobrir Pontos de Interesse (POI) relevantes na redes Sociais baseadas em localização (*Location Based Social Network* - LBSN) [5], tais como Yelp e FourSquare. Grande parte dos SsR em POIs tem se concentrado na acurácia da tarefa de correspondência entre POIs e usuários [6], mas há um consenso crescente de que outras dimensões são também importantes para se avaliar a utilidade da recomendação em cenários reais, tais como diversificação categórica e geográfica [6]. Entretanto, observa-se um dilema, pois acurácia e diversidade representam interesses distintos: POIs que aumentam a acurácia são mais populares entre os usuários, por outro, POIs que aumentam a diversidade podem ser menos convencionais. Além disso, no que se refere a diversidade, ou os trabalhos se concentram na diversificação categórica [10] e não são capazes de

prover uma diversificação geográfica, ou se concentram na diversificação geográfica [6], e não proveem uma diversificação categórica.

A partir do contexto acima, definimos nossa primeira hipótese: **(H1): É possível explorar as diversidades categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade.** Para validá-la, propomos um novo método de diversificação *geográfica e categórica* (*Geo-Cat*), que visa equilibrar ambos os conceitos na recomendação de POIs. Trata-se de um **método de pós-processamento** a ser aplicado sobre uma lista de um recomendador base, no intuito de maximizar a diversidade e manter a relevância dos POIs (i.e., acurácia). Este *trade-off* é modelado de forma semelhante à Relevância Máxima Marginal (MMR) [1]. Validamos nossa proposta combinando-a com diversos recomendadores base, considerando dados de duas cidades oriundos da coleção *Yelp*. As avaliações de SsR presentes na literatura limitam-se a avaliar métricas inerentes às técnicas implementadas, não validando simultaneamente requisitos de qualidade e satisfação dos usuários. Assim, definimos nossa segunda hipótese: **(H2): É possível avaliar SsR considerando múltiplas perspectivas de qualidade, tais como acurácia, diversidade categórica e diversidade geográfica.** Para validar a H2, adaptamos a Teoria da Utilidade Multiatributos [8], comum em Teoria de Jogos, materializando-a como uma métrica (MAUT) que visa representar a eficácia de um SR em função de diferentes dimensões de qualidade. Em nossos resultados demonstramos que o *Geo-Cat* é o método que melhor equilibra todos os fatores de precisão, diversidade e novidade avaliados. Conseguimos ganhos significativos sobre o conceito de diversidade, em alguns casos acima de 120%, sem alterações no nível de relevância.

As principais contribuições desse trabalho: (1) o *Geo-Cat*, um novo método de diversificação geográfica-categórica capaz de prover recomendações que equilibram conceitos distintos como precisão, diversidade e novidade; (2) a adaptação da Teoria da Utilidade Multiatributos para SsR; e (3) uma vasta avaliação experimental dos principais trabalhos de diversificação. Os resultados desse trabalho resultaram na publicação de dois artigos [2, 3], além de uma submissão ao periódico *Expert Systems with Applications*.

2 Referencial Teórico

Diversificação foi introduzido em SR a partir de trabalhos relacionados na área de RI, sendo a mais utilizada a *Maximal Marginal Relevance* (MMR) [1], que propõe uma combinação linear de relevância e diversidade. A adaptação dessa proposta para SsR consiste em reordenar uma lista de recomendação com R itens, cuja ordem original, definida por um SR base, é revista com o objetivo de diversificação. Os $top-k$ itens reordenados, para $k \ll R$, são apresentados para o usuário. Conforme podemos ver na Equação 1, o MMR é utilizado para encontrar a funcionalidade f de um item i dentre os R itens recomendados por um SR base. Para $\lambda = 0$, o MMR retorna a lista idêntica ao SR base e para $\lambda = 1$ o MMR foca totalmente no

In: II Concurso de Teses e Dissertações (CTD 2020), São Luís, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2020.

© 2020 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

processo de diversificação. O maior desafio consiste em definir os valores de λ e as melhores métricas para medir a diversidade dos itens.

$$f(i, R) = rel(i)^{1-\lambda} + div(i, R)^\lambda \quad (1)$$

Na linha diversificação categórica, destacamos o trabalho [4], no qual os autores propõem o *framework Binomial*, que minimiza a redundância das categorias dos itens retornados por um determinado SR base. diversidade é representada pelo quanto o item pode aumentar a diversificação de categorias da lista existente. Uma abordagem complementar é apresentada em [11], o **Intra-List Distance - ILD**, no qual a diversidade é baseada na distância entre os itens, medida pela correlação de *Pearson*. Por fim, em [10] propõe-se o **Proportional Framework - PM2** que explora a diversidade baseada na proporcionalidade da relevância de cada item i por cada tópico t . Basicamente, o PM2 também considera a proporcionalidade de um tópico t a outro não relacionado a i . (categorias ou gêneros).

Um trabalho recente aponta para a necessidade da diversificação geográfica é apresentado em [6]. Os autores assumem que a rotina de visita dos usuários em seus distritos de maior atividade (e.g., casa, trabalho, etc) faz com que os principais métodos recomendem apenas POIs localizados nessa pequena porção de regiões. Os autores defendem que recomendar POIs, assumindo que os interesses dos usuários são distribuídos em várias sub-regiões de um distrito alcançam melhores resultados. Assim, propõem o **Geo-Div**, uma estratégia que seleciona um conjunto de k itens proporcionalmente ao nível de atividade dos usuários em cada uma das sub-regiões de uma cidade.

3 Solução Proposta - Geo-Cat

Visando abordar **H1**, nossa proposta consiste no **Geo-Cat**, um método de pós-processamento que possa ser aplicado em conjunto com qualquer recomendador base e que seja capaz de prover listas de recomendação que sejam diversificadas categoricamente e geograficamente, simultaneamente. De forma muito simplificada, o Geo-Cat consiste em alterar a segunda parte da Equação 1 de acordo com a Equação 2:

$$div_{geo-cat}(i, R) = \beta \cdot div_{geo}(i, R) + (1 - \beta) \cdot div_{cat}(i, R) \quad (2)$$

onde β é o parâmetro para equilibrar os fatores geográficos e categóricos.

No que se refere à diversificação categórica (div_{cat}), nossa proposta é combinar a métrica de *Genre Coverage*, proposta por *Puthiya et al.* [9], e a métrica de *Intra-List Distance* (ILD), proposta por *Vargas et al.* [11], agregando-as por meio de uma combinação linear, dando pesos iguais a ambas, como mostra a Equação 3. Representamos o quanto um item i influencia a lista de recomendação R sob os pontos de vista do *Genre Coverage* e do *ILD*. A métrica de *Genre Coverage* [9] visa diversificar as categorias de POIs recomendados. Por outro lado, a métrica de *Intra-List Distance* (ILD) [11] visa identificar um conjunto de POIs capaz de cobrir o máximo possível de perfis dos usuários.

$$div_{cat}(i, R) = GenreCov(\{i\} \cup R) + \frac{ILD(\{i\} \cup R)}{|R|} \quad (3)$$

De maneira complementar, visamos maximizar a diversificação geográfica dos POIs recomendados considerando o nível de atividade dos usuários em determinadas sub-regiões. Assim como em [6]

dividimos uma área (e.g., cidade) em sub-regiões de 0,5x0,5km, utilizando a métrica de diversificação geográfica baseada na proporcionalidade PR_g , que corresponde ao complemento da métrica de desproporcionalidade DP_g normalizada, como mostra a Equação 4 proposta em [4]. Essa desproporcionalidade DP_g mensura se o comportamento do usuário está ou não dividido proporcionalmente nas sub-regiões de uma cidade. Para isso, DP_g leva em conta cada POI l presente no histórico de visitas do usuário e a cobertura desse POI na região, conforme mostra a Equação 5. A cobertura c de um POI é calculada por um método iterativo onde c é inicializado com o resultado da divisão entre o número de POIs visitados pelo usuário e a quantidade de lugares recomendados. Em seguida, esse valor é alterado de acordo com os vizinhos de cada POI: (1) se o POI recomendado possui vizinhos, c é dividido pela quantidade de vizinhos do POI e distribuído para cada um deles, que armazenam a soma destes valores ao longo das iterações; ou (2) se o POI não possui vizinhos, o valor de c é acumulado (n_{NR}) para ser utilizado no cálculo da desproporcionalidade. Assim como nos trabalhos da literatura, definimos como vizinhos os POIs que foram consumidos pelo mesmo usuário e estão em um raio de 100 metros.

Além disso, esses valores de cobertura são utilizados para definir a variável binária f_l de cada POI. Neste contexto, f_l é igual a 1 quando o valor de cobertura acumulado pelo POI é menor que 1. Caso contrário o valor assumido é 0. Dessa forma, o modelo garante que a desproporcionalidade seja calculada apenas para POIs que receberam valores de cobertura bem distribuídos. POIs que receberam um valor de cobertura muito alto significa que estão muito próximos geograficamente dos POIs recomendados e devem ser penalizados com baixa proporcionalidade. Por sua vez, v_l representa o valor esperado de cobertura acumulada e s_l representa os valores de cobertura mensurados. Por essa razão, o valor de v_l observado sempre será 1 uma vez que quando o valor de cobertura c não é distribuído de forma uniforme entre os POIs do conjunto de treinamento, a desproporcionalidade é máxima [6]. Todas essas variáveis compõem a Equação 5, que mensura a desproporcionalidade de cada POI l do conjunto de treino do usuário. Assim, maximizar a diversificação geográfica significa sugerir POIs de distintas sub-regiões que já são frequentadas pelo usuário, aumentando assim sua proporcionalidade em todas as regiões. Para isso, propomos na Equação 6 maximizar o valor de PR . Para cada POI i a ser testado, calcula-se a diferença entre o novo PR , adicionando i , e o PR atual. Assim, espera-se recomendar um número de POIs por sub-região que seja proporcional ao comportamento do usuário no passado.

$$PR_g(P) = 1 - DP_g(P) / (|T_u| + 1/2) \quad (4)$$

$$DP_g(P) = \sum_{l \in T_u} f_l \cdot (v_l - s_l)^2 + n_{NR}^2 / 2 \quad (5)$$

$$div_{geo}(i, R) = PR(\{i\} \cup R) - PR(R) \quad (6)$$

4 Teoria da Utilidade Multiatributo (MAUT)

Para abordar a **H2**, propomos a avaliação de todas as métricas de qualidade em conjunto. Pretendemos analisar o quão eficaz os métodos de pós-processamento são para equilibrar todos os critérios de qualidade. Assim, adaptamos a Métrica da Utilidade Multiatributo (MAUT), utilizada no domínio de Teoria dos Jogos [8] para manipular problemas em que os resultados são dependentes de dois ou mais atributos. Em nosso trabalho, a MAUT visa representar a eficácia de um método em função das métricas de qualidade [7]. Nesse caso,

Algoritmo	Las Vegas						Phoenix					
	Prec@10	Rec@10	EPC@10	ILD@10	PRg@10	GC@10	Prec@10	Rec@10	EPC@10	ILD@10	PRg@10	GC@10
ItemNN	0,0138	0,0110	0,8784	0,2113	0,1640	0,3197	0,0056	0,0035	0,9275	0,2125	0,0415	0,3282
Geo-cat	0,0193 ▲	0,0150 ▲	0,8632 ▼	0,4333 ▲	0,4721 ▲	0,5873 ▲	0,0095 ▲	0,0058 ▲	0,9234 ▼	0,3990 ▲	0,1928 ▲	0,6076 ▲
MostPopular	0,0218	0,0167	0,1981	0,6984	0,1540	0,3564	0,0134	0,0085	0,3780	0,4776	0,0254	0,4474
Geo-cat	0,0331 ▲	0,0254 ▲	0,4264 ▲	0,7048 ▲	0,3951 ▲	0,5970 ▲	0,0166 ▲	0,0110 ▲	0,5316 ▲	0,5121 ▲	0,1713 ▲	0,6663 ▲
UserKNN	0,0428	0,0338	0,5561	0,6293	0,2893	0,4180	0,0340	0,0240	0,7129	0,5012	0,1771	0,4789
Geo-cat	0,0443 ●	0,0357 ▲	0,6300 ▲	0,6742 ▲	0,4852 ▲	0,6170 ▲	0,0338 ●	0,0246 ●	0,7698 ▲	0,5660 ▲	0,3664 ▲	0,6533 ▲
USG	0,0460	0,0392	0,6006	0,6236	0,3248	0,4181	0,0371	0,0287	0,7495	0,4946	0,2220	0,4836
Geo-cat	0,0446 ●	0,0383 ●	0,6620 ▲	0,6795 ▲	0,5342 ▲	0,6329 ▲	0,0356 ●	0,0278 ●	0,7927 ▲	0,5772 ▲	0,4301 ▲	0,6671 ▲
WRMF	0,0458	0,0356	0,5159	0,6134	0,2742	0,4142	0,0286	0,0197	0,6303	0,4908	0,1276	0,4727
Geo-cat	0,0468 ●	0,0366 ●	0,5956 ▲	0,6653 ▲	0,4761 ▲	0,6202 ▲	0,0293 ●	0,0206 ●	0,7058 ▲	0,5545 ▲	0,3232 ▲	0,6596 ▲
Algoritmo	Prec@20	Rec@20	EPC@20	ILD@20	PRg@20	GC@20	Prec@20	Rec@20	EPC@20	ILD@20	PRg@20	GC@20
ItemKNN	0,0122	0,0186	0,8783	0,2558	0,2185	0,4068	0,0052	0,0066	0,9275	0,2520	0,0657	0,4175
Geo-cat	0,0155 ▲	0,0239 ▲	0,8687 ▼	0,3816 ▲	0,4650 ▲	0,6473 ▲	0,0070 ▲	0,0086 ▲	0,9242 ▼	0,3415 ▲	0,1842 ▲	0,6412 ▲
MostPopular	0,0224	0,0320	0,3068	0,6770	0,2216	0,5068	0,0121	0,0149	0,4446	0,4181	0,0640	0,5630
Geo-cat	0,0271 ▲	0,0404 ▲	0,4375 ▲	0,7005 ▲	0,4006 ▲	0,7085 ▲	0,0136 ▲	0,0176 ▲	0,5420 ▲	0,4807 ▲	0,1812 ▲	0,7270 ▲
UserKNN	0,0375	0,0589	0,5902	0,6440	0,3561	0,5780	0,0286	0,0405	0,7354	0,5277	0,2419	0,6105
Geo-cat	0,0385 ●	0,0603 ●	0,6389 ▲	0,6695 ▲	0,5230 ▲	0,7254 ▲	0,0295 ●	0,0422 ●	0,7718 ▲	0,5637 ▲	0,4062 ▲	0,7398 ▲
USG	0,0385	0,0645	0,6244	0,6396	0,3995	0,5760	0,0303	0,0464	0,7628	0,5221	0,2971	0,6149
Geo-cat	0,0386 ●	0,0643 ●	0,6633 ▲	0,6736 ▲	0,5863 ▲	0,7388 ▲	0,0305 ●	0,0465 ●	0,7916 ▲	0,5678 ▲	0,4853 ▲	0,7524 ▲
WRMF	0,0406	0,0634	0,5611	0,6292	0,3412	0,5702	0,0246	0,0338	0,6669	0,5190	0,1885	0,6007
Geo-cat	0,0407 ●	0,0635 ●	0,6085 ▲	0,6569 ▲	0,5104 ▲	0,7253 ▲	0,0258 ▲	0,0357 ▲	0,7176 ▲	0,5554 ▲	0,3515 ▲	0,7412 ▲

Tabela 1: Geo-Cat aplicado sobre as listas geradas pelos SsR base (● empates estatísticos, ▲ ganhos estatísticos e ▼ perdas estatísticas.)

consideramos M como o conjunto de métricas (i.e. Prec, Rec, EPC, ILD, GC e PRg) e $m_{i,j}$ a utilidade da métrica i para o método j (i.e. PM2, LD, Binom, Geo-Div e Geo-Cat). Para todas as métricas, quanto maior o valor alcançado, melhor é o método. Logo, $m_{i,j}$ pode ser mapeada diretamente pelo valor obtido pelo método j na métrica i .

Assim, para cada métrica, calcula-se o *utility score* de um método por meio da normalização min-max dada por: $(U_{i,j} = (m_{i,j} - \min(m_i)) / (\max(m_i) - \min(m_i)))$. Essa normalização possibilita ranquear os métodos de uma determinada métrica do melhor ($U_{i,j} = 1$) para o pior ($U_{i,j} = 0$). A função de utilidade de um método U_j pode ser definida na forma de um modelo aditivo como é apresentado na Equação 7, na qual w_i representa o peso da métrica i e normalmente satisfaz a condição $\sum_{i=1}^{|M|} w_i = 1$. O peso padrão de cada atributo é $1/|M|$, e pode ter variações se, porventura, algumas das métricas forem consideradas mais importantes que as demais. Em nosso caso, assumimos que todas as métricas são independentes e igualmente importantes no processo de decisão. Assim, após obter os resultados do modelo aditivo da Equação $U_j = \sum_{i=1}^{|M|} U_{i,j} \cdot w_i$ para cada método de pós-processamento, consideramos como melhor aquele que obtém maior resultado. Essa é a segunda grande contribuição do nosso trabalho, uma vez que conseguimos avaliar a qualidade dos recomendadores considerando diversas perspectivas simultaneamente.

$$U_j = \sum_{i=1}^{|M|} U_{i,j} \cdot w_i \quad (7)$$

5 Resultados Experimentais

5.1 Configuração Experimental

5.1.1 - Base de Dados: Selecionamos os dados de 2 das 5 cidades que possuem maior número de POIs na base do *Yelp Dataset Challenge*: (1) Las Vegas (15.168 usuários e 11.562 POIs); e (2) Phoenix (14.388 usuários e 36.827 POIs).

5.1.2 - Algoritmos de Recomendação: Avaliamos a aplicação da nossa estratégia de pós-processamento sobre os recomendadores base **Most Popular**, **User-kNN**, **WRMF**, **Item Attribute kNN** e **USG**. Em seguida, comparamos os resultados alcançados pelo *Geo-Cat* com as técnicas de pós-processamento estado-da-arte em diversificação: **PM2**, **LD** e **Binom** utilizando como SR base o **USG**.

5.1.3 - Métricas de Avaliação: selecionamos as métricas clássicas de avaliação: **Precision (Prec)**, **Recall (Rec)**, **EPC**, **ILD**, **Genre Coverage (GC)** e **PRg**. Avaliamos as recomendações geradas com 10 e 20 POIs para todas essas métricas.

5.1.4 - Configuração de Parâmetros: Através de um experimentação em um conjunto de validação, encontramos os melhores parâmetros utilizando a MAUT com pesos iguais para todas as métricas. As melhores combinações foram *Geo-Cat* ($\lambda = 0,75$ e $\beta = 0,75$); **Binom** ($\lambda = 0,75$); **Geo-div** ($\lambda = 0,5$); **LD** ($\lambda = 0,25$) e **PM2** ($\lambda = 1,0$).

5.2 Resultados

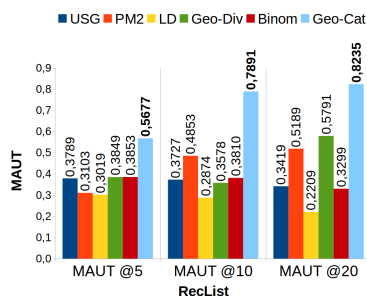
Na Tabela 1 apresentamos o comparativo entre os resultados dos SsR base sem e com o uso do *Geo-Cat* e na Figura 1 o comparativo entre resultados do *Geo-Cat* com os de outros algoritmos de diversificação, agregando-os por meio da MAUT.

5.3 Avaliação do Geo-Cat nos SsR Tradicionais

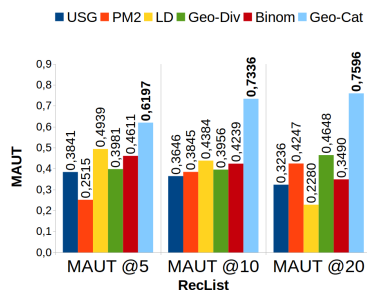
Nesta primeira análise, nosso objetivo é avaliar se a utilização de nossa estratégia de pós-processamento *Geo-Cat* é capaz de melhorar a qualidade das recomendações geradas pelos SsR tradicionais. Os resultados são apresentados na Tabela 1. Em ambas as cidades, para cada métrica e tamanho de lista de recomendação, comparamos os resultados do SR base com e sem o *Geo-Cat* utilizando o teste de *Wilcoxon* com 95% de confiança. De maneira geral, podemos observar que o *Geo-Cat* conseguiu melhorar o resultado de grande parte das métricas avaliadas para os SsR base utilizados e considerando todos os tamanhos N da lista de recomendação.

Primeiramente, considerando as métrica de acurácia (*Precision* e *Recall*), observamos que a utilização do *Geo-Cat* foi capaz de prover resultados estatisticamente melhores ou equivalentes àqueles obtidos pelos SsR base. Nesse caso, destacamos os resultados obtidos pelo método *Most Popular*. Tradicionalmente, esse recomendador apresenta apenas os POIs mais populares (i.e., mais frequentados). Com nosso método, conseguimos apresentar POIs distintos dos anteriores, aumentando a diversidade e novidade das recomendações, e ainda assim, satisfazer mais os interesses dos usuários, aumentando as métricas de *Precision* e *Recall*. Considerando as métricas de diversidade e novidade, observamos novamente que a utilização do

Geo-Cat foi capaz de prover resultados muito superiores aos obtidos pelos outros SsR. Em especial, nos métodos altamente focados em acurácia, como os recomendadores de *WRF* e *USG*, conseguimos melhorar a diversidade e novidade dos itens sem apresentar perdas significativas de acurácia. Por exemplo, no caso da cidade de Phoenix, considerando lista de 10 recomendações do *USG*, obtivemos uma melhora de 120% para a métrica de PR_g . Uma única exceção é observada para o algoritmo *Item Attribute-KNN* sobre a métrica de *EPC*, porém com uma diferença muito pequena (inferior à 1%). O *Item Attribute-KNN*, por si só, alcança bons resultados de novidade pois apresenta POIs aleatórios para o usuário, porém com uma acurácia muito baixa, inferior a todos os demais. Essa pequena diferença na diversidade e novidade é significativamente compensada na acurácia das recomendações, uma vez que o *Geo-Cat* é capaz de melhorar em até 56% a métrica *Precision*. Esses resultados mostram que nossa abordagem é eficaz para melhorar a maioria dos recomendadores, potencialmente útil de ser aplicada em domínios reais.



(a) MAUT Las Vegas



(b) MAUT Phoenix

Figura 1: Valores de MAUT

5.4 Avaliação do Geo-Cat com os Baselines

A segunda análise consiste em comparar o *Geo-Cat* com outras estratégias de diversificação de POIs existentes na literatura. Nessas análises, consideramos como SR base o *USG* por ser específico para o cenário de POIs. Os parâmetros considerados em todas as estratégias foram aqueles descritos na Seção 5.1. Para sumarizar todos os resultados encontrados, utilizamos a Métrica da Utilidade Multia-tributo (MAUT), originalmente utilizada em Teoria dos Jogos, com o objetivo de avaliar cada um dos métodos considerando simultaneamente as várias métricas analisadas. Assim como na definição dos parâmetros dos métodos utilizados (Seção 5.1), também consideramos que todas as métricas de avaliação são igualmente importantes e possuem pesos iguais ($w = 0, 1667$ na Equação. 7). Conforme apresentado na Seção 4.4, o valor da MAUT está entre 0 e 1: quando mais próximo de 1, melhor é comportamento de um determinado método

frente às múltiplas métricas avaliadas. Os resultados referentes à essa análise estão apresentados nos gráficos da Figura 1. Primeiramente, podemos observar novamente que há diferenças entre os métodos avaliados. Enquanto o *LD* apresenta bons resultados quando consideramos listas de recomendação de tamanho 5 para Phoenix, nessa mesma cidade o *LD* apresenta os piores resultados para uma lista de tamanho 20. Por outro lado, por essa análise é possível observar que o *Geo-Cat* é aquele que apresenta os melhores resultados quando consideramos todos as métricas simultaneamente.

6 Conclusão & Discussões

Neste trabalho propomos o *Geo-Cat*, um novo método de pós-processamento que visa diversificar as recomendações no cenário de POIs, considerando conjuntamente aspectos categóricos e geográficos. Por meio de uma avaliação experimental, considerando bases de dados tradicionais e diversos SsR base, nossos resultados apontam que o *Geo-Cat* é capaz de potencializar a diversidade de todos os recomendadores utilizados. Além disso, quando comparamos o *Geo-Cat* frente aos principais modelos de pós-processamento de diversificação (Tabela 1), vemos que nosso método é capaz de superá-los em algumas das principais métricas de avaliação. Sobretudo, é possível observar que o *Geo-Cat* apresenta os melhores resultados ao avaliarmos todas as métricas simultaneamente, adaptando a métrica de multiatributos *MAUT* [8] (Figura 1). Em especial, a proposta de adaptação da *MAUT* representa um avanço significativo nas avaliações de SsR, uma vez que os mesmos podem ser avaliados sob múltiplas perspectivas, evitando avaliações enviesadas por alguma métrica específica. Como trabalhos futuros, pretendemos considerar outros algoritmos mais elaborados para se otimizar a reordenação da lista de POIs de um SR base, tais como meta-heurísticas bioinspiradas.

Referências

- [1] Jaime G Carbonell and Jade Goldstein. 1998. The Use of MMR and Diversity-Based Reranking for Reordering Documents and Producing Summaries. (1998).
- [2] Rodrigo Carvalho, Nicollas Silva, Luiz Chaves, Adriano C. M. Pereira, and Leonardo C. da Rocha. 2019. Geographic-categorical diversification in POI recommendations. In *WebMedia*.
- [3] Luiz Chaves, Nicollas Silva, Rodrigo Carvalho, Adriano C. M. Pereira, and Leonardo C. da Rocha. 2019. Exploiting the user activity-level to improve the models' accuracy in point-of-interest recommender systems. In *WebMedia*.
- [4] Van Dang and W Bruce Croft. 2012. Diversity by proportionality: an election-based approach to search result diversification. In *Proceedings of the international ACM SIGIR conference on Research and development in information retrieval*. ACM.
- [5] Jean-Benoît Griesner, Talel Abdesslem, and Hubert Naacke. 2015. POI recommendation: towards fused matrix factorization with geographical and temporal influences. In *Proceedings of the 9th ACM RecSys*. ACM, 301–304.
- [6] Jungkyu Han and Hayato Yamana. 2017. Geographical Diversification in POI Recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems - RecSys '17*. <https://doi.org/10.1145/3109859.3109884>
- [7] Alessio Ishizaka and Philippe Nemery. 2013. *Multi-criteria Decision Analysis: Methods and Software*. Wiley. <https://www.amazon.com/Multi-criteria-Decision-Analysis-Methods-Software/dp/1119974070?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xml2&camp=2025&creative=165953&creativeASIN=1119974070>
- [8] Silvia Beatriz Neiva and Luiz Flavio Autran Monteiro Gomes. 2009. A aplicação da teoria da utilidade multiatributo à escolha de um software de e-procurement. *Revista Tecnologia* 28, 2 (2009).
- [9] Shameem A Puthiya Parambath, Nicolas Usunier, and Yves Grandvalet. 2016. A coverage-based approach to recommendation diversity on similarity graph. In *RecSys*. 15–22.
- [10] Saül Vargas, Linas Baltrunas, Alexandros Karatzoglou, and Pablo Castells. 2014. Coverage, redundancy and size-awareness in genre diversity for recommender systems. In *RecSys*.
- [11] Saül Vargas and Pablo Castells. 2011. Rank and relevance in novelty and diversity metrics for recommender systems. In *RecSys*. <https://doi.org/10.1145/2043932.2043955>