

Geographic-categorical diversification in POI recommendations*

Rodrigo Carvalho¹, Nicollas Silva², Luiz Chaves¹, Adriano C. M. Pereira², Leonardo Rocha¹

¹Universidade Federal de São João del-Rei, ²Universidade Federal de Minas Gerais
(rodrigo,luizfufs,lcrocha)@ufsj.edu.br;(ncsilvaa,adrianoc)@dcc.ufmg.br

ABSTRACT

Nowadays, Recommender Systems (RSs) have been used to help users to discover relevant Points Of Interest (POI) in Location Based Social Network (LBSN), such as Yelp and FourSquare. Due to the main challenges of data sparsity and the geographic influence in this scenario, most of works about POI recommendations has only focused on improving the system's accuracy. However, there is a consensus that just it is not enough to assess the practical effectiveness. In real scenarios, categorical and geographic diversities have been identified as key dimensions of recommendation utility. The few existing works are concentrated on just one of these concepts, singly. In this work, we propose a novel post-processing strategy to merge these concepts in order to improve the user interest in POIs. Our experimental results in Yelp datasets show that our strategy can improve users' satisfaction, considering different RS and multiple diversification metrics. Our method is able to improve the diversity up to 120% without significant accuracy losses.

CCS CONCEPTS

• Information systems → Recommender Systems.

KEYWORDS

POI Recommendation, Diversity, Geographic Diversity

ACM Reference Format:

Rodrigo Carvalho¹, Nicollas Silva², Luiz Chaves¹, Adriano C. M. Pereira², Leonardo Rocha¹. 2019. Geographic-categorical diversification in POI recommendations. In *Proceedings of Brazilian Symposium on Multimedia and the Web (WebMedia '19)*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3323503.3349554>

1 INTRODUÇÃO

Nos últimos anos, redes sociais baseadas na localização do usuário (*Location Based Social Network* - LBSN) têm se tornado cada vez mais populares na Web [7]. Um bom exemplo disso é o *Facebook* que vem aprimorando suas aplicações, permitindo que os usuários realizem *check-ins* em determinados lugares e/ou utilizem de sua localização para divulgar produtos no *marketplace* [20]. O registro de todas essas atividades representa um conjunto de *feedbacks* implícitos sobre

as experiências dos usuários, criando a oportunidade de associar outros serviços a LBSN. Um desses serviços é a utilização de Sistemas de Recomendação (SsR) para que possam apresentar aos usuários sugestões de pontos de interesse (i.e. *Points of Interest* - POI) a serem visitados, tais como museus, teatros, restaurantes, dentre outros.

Grande parte dos SsR em POIs tem se concentrado na acurácia da tarefa de correspondência entre POIs e usuários [7, 8]. Apesar disso, há um consenso crescente de que apenas a acurácia não é suficiente para avaliar a eficácia prática das recomendações. Em particular, a diversidade está sendo identificada como uma dimensão chave da utilidade da recomendação em cenários reais, tais como diversificação categórica e geográfica [8]. O objetivo de maximizar apenas a acurácia leva a modelos superespecializados nas preferências dos usuários, recomendando apenas itens semelhantes aos anteriormente visitados [11]. No cenário de POI, essa superespecialização pode ser ainda pior pela questão geográfica, ou seja, as recomendações ficam restritas a POIs que estejam em sub-regiões relacionadas a rotina do usuário (e.g., trabalho, casa, etc.) [4]. Encontramos na literatura alguns trabalhos focados na diversificação categórica de SsR em geral [5, 18, 19] mas que, no entanto, não são capazes de prover uma diversificação geográfica. Isso resulta na recomendação de POIs categoricamente "diversos" concentrados em uma pequena porção das sub-regiões em que o usuário alvo é mais ativo. No que se refere à diversificação geográfica, encontramos apenas um trabalho [8], mas que não é capaz de prover uma diversificação categórica.

Esse trabalho propõe 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), conhecida no campo de Recuperação de Informação (RI). Temos um parâmetro que controla o nível de relevância e diversidade em cada lista de recomendação. Enquanto a relevância é medida com base no *score* atribuído pelo recomendador base, a diversidade é uma combinação linear de dois fatores: a diversificação geográfica e a diversificação categórica. Mensuramos a diversificação geográfica com base no método de *Geographic-Diversification* (Geo-Div) [8], estado-da-arte nesse conceito e a diversificação categórica com base na combinação do método de diversificação do ILD [19] com o conceito de cobertura de categorias.

No intuito de consolidar nosso método, construímos um projeto experimental que avalia a aplicação de nossa proposta sobre diversos recomendadores base, tais como *Most Popular*, *User-kNN*, *WRFM*, e *USG*. Nessas análises, consideramos as duas principais cidades da base de dados da *Yelp* (i.e., *Las Vegas* e *Phoenix*). Nossos resultados mostram que conseguimos ganhos significativos sobre o conceito de diversidade, em alguns casos acima de 120%, sem alterações no nível

*This work was partially funded by the Brazilian National Institute of Science and Technology for the Web - INWeb, MASWeb, CAPES, CNPq, Finep, and Fapemig.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebMedia '19, October 29-November 1, 2019, Rio de Janeiro, Brazil

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6763-9/19/10...\$15.00

<https://doi.org/10.1145/3323503.3349554>

de acurácia (i.e. precisão e revocação). Além disso, comparamos o *Geo-Cat* com os métodos de diversificação estado-da-arte, que são específicos para determinados conceitos, como *intra-list distance* (ILD) [19], cobertura de categorias [5] e novidade dos POIs [18]. Comparamos todas essas abordagens utilizando a métrica da utilidade de múltiplos atributos (MAUT) [12, 16], comum em Teoria de Jogos. Em nossos resultados demonstramos que o *Geo-Cat* é o método que melhor equilibra todos os fatores de precisão, diversidade e novidade avaliados, destacando-se frente aos principais métodos existentes na literatura.

Assim, destacamos como 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 métrica de utilidade de múltiplos atributos para que o contexto de recomendação para que SsR possam ser avaliados sob diferentes perspectivas em conjunto; e (3) uma vasta avaliação experimental dos principais trabalhos de diversificação existentes na literatura.

2 REFERENCIAL TEÓRICO

Conforme mencionado, além da precisão, os conceitos relacionados à diversidade têm sido considerados fundamentais para melhorar a atratividade e a acurácia de SsR [9, 11]. Dessa forma, apresentamos nessa seção os principais trabalhos que discutem esse conceito.

2.1 Diversidade em Sistemas de Recomendação

O conceito de diversificação foi introduzido em SR a partir de trabalhos relacionados na área de RI. Uma frequente abordagem utilizada é a *Maximal Marginal Relevance* (MMR), proposta em [2], que consiste em uma combinação linear de relevância e diversidade - Equação 1. O MMR é utilizado para encontrar a funcionalidade f de um documento d para uma *query* q , baseado nos D documentos disponíveis no sistema. Definindo o $\lambda = 0$, o MMR se comporta como um modelo de predição tradicional, não alterando os documentos recuperados. Por outro lado, aplicando o $\lambda = 1$, faz com que o MMR foque no processo de diversificação e altere a ordem dos documentos recuperados. O maior desafio consiste em definir os valores de λ e as melhores métricas para medir a diversidade dos documentos.

$$f(q, d, D) = (1 - \lambda) \cdot rel(q, d) - \lambda \cdot \max_{d_j \in D} sim(d, d_j) \quad (1)$$

A maioria das abordagens minimizam a redundância dos documentos aplicando o método MMR como uma etapa de pós-processamento das ferramentas de busca [2, 17, 22]. A adaptação dessa proposta para SsR consiste em reordenar uma lista de recomendação com n itens, cuja ordem original é definida por um recomendador, com base no objetivo de diversificação e os top- k itens reordenados, para $k \ll n$, são apresentados para o usuário. Os métodos de reordenação são, principalmente, destinados a maximizar: (1) a diversificação categórica do item; (2) a distância entre os itens da lista (*intra-list distance*); ou (3) a proporcionalidade das categorias dos itens.

No primeiro aspecto, Dang V. & Croft W. [5] propõem o *framework Binomial*, que minimiza a redundância dos tópicos retornados para uma consulta. Basicamente, o Binomial é baseado em uma métrica de diversidade definida como o produto de dois componentes: cobertura de categorias e a não redundância. A cobertura é definida pelo produto dos gêneros não cobertos e suas probabilidades. Por sua

vez, o conceito de não redundância é definido pelo nível de *tolerância* restante do usuário para cada gênero já coberto. A função que combina linearmente a relevância e a diversidade binomial é similar a MMR, sendo que a diversidade é representada pelo quanto o item pode aumentar a diversificação da lista existente:

$$div(i; R) = BinomDiv(R \cup \{i\}) - BinomDiv(R) \quad (2)$$

De acordo com o segundo aspecto, destacamos o *framework Intra-List Distance - LD*, proposto por Vargas et al. [19]. Nesse, a diversidade é baseada na distância entre os itens, medida pela correlação de *Pearson*. Especialmente, em alguns casos, é possível medir a distância baseada nos gêneros dos itens.

$$ILD = div(R) = \frac{2}{|R|(|R| - 1)} \sum_{i_k \in R, l < k} d(i_k, i_l) \quad (3)$$

Por sua vez, Vargas et al. [18] propõe um novo **Proportional Framework - PM2** que explora a diversidade baseada na proporcionalidade da relevância de cada documento d por cada tópico t . Basicamente, a fim de reduzir a redundância dos documentos, o PM2 também considera a proporcionalidade de um tópico t a outro não relacionado a d . A Equação 4 representa esta proporção, semelhante ao método MMR. $P(d|t)$ representa a probabilidade da relevância de um documento d para um tópico t , o termo $qt[i]$ representa o número de interações de um tópico t em uma *query* q . No domínio de recomendação, estes tópicos podem ser representados como categorias ou gêneros dos itens disponíveis.

$$f(q, d) = \lambda \cdot qt[i^*] \cdot P(d|t_{i^*}) + (1 - \lambda) \cdot \sum_{i \neq i^*} qt[i] \cdot P(d|t_i) \quad (4)$$

Contudo, embora alcancem resultados satisfatórios em cenários tradicionais, esses métodos não são diretamente aplicáveis no contexto de diversificação geográfica, fundamental no cenário de POIs. Alguns trabalhos até tentaram adaptar um grupo específico de métodos de reordenação que se baseiam na comparação par a par da distância do item [1, 2, 23]. Basicamente, essas propostas trocavam o conceito de “distância” para a distância geográfica entre dois POIs. Entretanto, esses métodos simplesmente maximizavam a distância entre POIs, recomendando POIs localizados longe uns dos outros. Tal abordagem não possui aplicabilidade prática, uma vez que, os usuários preferem POIs próximos à sua localização atual.

2.2 Diversidade em Recomendação de POIs

Além dos conceitos de diversidade mencionados anteriormente, igualmente importantes para recomendação de POIs, um recente trabalho aponta para a necessidade da diversificação geográfica. Han J. and Yamana H. [8] 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. Em particular, os autores argumentam que recomendar POIs próximos à localização atual de um usuário - uma forma popular de aplicação para dispositivos móveis - é menos útil. Ao invés disso, defendem que recomendar POIs, assumindo que os interesses dos usuários são distribuídos em várias sub-regiões de um distrito alcançam melhores resultados. Portanto, Han J. and Yamana H. [8] propõem um novo método de diversificação geográfica, denominado **Geo-Div**, que seleciona um conjunto de k itens proporcionalmente ao nível de atividade dos usuários em cada uma das sub-regiões de uma cidade. Este método é

baseado no MMR e define a diversidade do item de acordo com a proporcionalidade da recomendação, demonstrado nas Equações 5 e 6.

$$f(i, R) = rel(i)^{1-\lambda} + div_g(i, R)^\lambda \quad (5)$$

$$div_g(i, R) = \max(0, PR_g(\{i\} \cup R) - PR_g(R)) \quad (6)$$

3 ESTRATÉGIA DE DIVERSIFICAÇÃO

Como mencionado, as principais estratégias voltadas para diversificação consistem em estratégias de pós-processamento que ampliam o objetivo dos recomendadores. Tradicionalmente, os métodos de recomendação focam o seu treinamento em métricas de acurácia. Já as estratégias de diversidade visam adicionar outras métricas nesse processo de treino. Contudo, apesar das estratégias de *Binom*, *ILD* e *PM2* serem muito eficientes nos cenários clássicos de recomendação, elas não são projetadas para o domínio de recomendação de POIs. Tais estratégias não consideram o impacto do fator geográfico que define os pontos de interesse e, portanto, não são as mais adequadas para POIs. Por outro lado, a estratégia de *Geo-Div*, estado-da-arte em diversificação geográfica, não considera os aspectos intrínsecos dos POIs, como suas categorias ou mesmo perfil dos usuários que os visitam. Dessa forma, nossa estratégia é projetada para considerar ambos os conceitos de diversificação: categórica e geográfica.

3.1 Diversificação Categórica

Nesse trabalho, visamos alcançar a diversificação categórica por meio dos fatores que estão intrinsecamente relacionados aos POIs. Especificamente, estamos nos referimos à sua categoria (e.g., restaurante, cinema, museu, entre outros) e ao seu público alvo (i.e., os perfis dos usuários que os frequentam). Nossa proposta, neste caso, é combinar esses dois fatores em apenas uma métrica que visa otimizar a diversificação categórica dos recomendadores. Para o primeiro fator consideramos a métrica de *Genre Coverage*, proposta por *Puthiya et al.* [13], enquanto para o segundo fator, consideramos a métrica de *Intra-List Distance* (ILD), proposta por *Vargas et al.* [19]. Nossa proposta é agregá-las por meio de uma combinação linear, dando pesos iguais a ambas, como mostra a Equação 7. Representamos o quanto um item i influencia a lista de recomendação R sob os pontos de vista do *Genre Coverage* e do *ILD*. Como o *ILD* é calculado sobre toda a lista de recomendação, precisamos dividi-lo pelo tamanho da lista para utilizar o valor de cada item presente.

$$div_{cat}(i, R) = GenreCov(\{i\} \cup R) + \frac{ILD(\{i\} \cup R)}{|R|} \quad (7)$$

A métrica de *Genre Coverage* [13] visa diversificar as categorias de POIs recomendados. Basicamente, essa métrica representa a média de categorias relevantes presentes na lista de recomendação para o usuário. Valores altos de *Genre Coverage* indicam que os POIs recomendados cobrem um número maior de categorias relevantes. Nesse trabalho, consideramos como categorias relevantes aquelas que pertencem a POIs avaliados com uma nota superior a média do usuário alvo. Por outro lado, a métrica de *Intra-List Distance* (ILD) [19] visa identificar um conjunto de POIs capaz de cobrir o máximo possível de perfis dos usuários. O *ILD* mensura a dissimilaridade dos itens da lista baseado nos usuários que frequentaram os POIs. Nesse caso, os POIs são modelados como vetores de usuários, de forma que um POI é considerado mais similar a outros frequentados pelos mesmos usuários. Por essa razão, aumentar os valores

de *ILD* indica que os POIs recomendados estão relacionados a mais usuários distintos. Na prática, ambas as métricas são essenciais pois é possível que os POIs recomendados pertençam a várias categorias relevantes, tendo altos valores de *Genre Coverage*, mas sejam de categorias similares. Contudo, com a utilização de *ILD*, o modelo consegue capturar essas diferenças entre categorias na hierarquia.

3.2 Diversificação Geográfica

De maneira complementar, visamos maximizar a diversificação geográfica dos POIs recomendados considerando o nível de atividade dos usuários em determinadas sub-regiões. Especificamente, dividimos uma área (e.g., cidade) em sub-regiões de 0.5km^2 e seguimos a direção do trabalho de *Han J. e Yamana H.* [8] que define a métrica de diversificação geográfica baseada na proporcionalidade PR_g . Fundamentalmente, a métrica de PR_g é o complemento da métrica de desproporcionalidade DP_g normalizada, como mostra a Equação 8 proposta por *Dang V. and Croft W.* [5]. 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 9. 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.

$$PR_g(P) = 1 - DP_g(P)/(|T_u| + 1/2) \quad (8)$$

$$DP_g(P) = \sum_{l \in T_u} f_l \cdot (v_l - s_l)^2 + n_{NR}^2/2 \quad (9)$$

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 [8]. Todas essas variáveis compõem a Equação 9, que mensura a desproporcionalidade de cada POI l do conjunto de treino do usuário.

Dessa forma, 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 10 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.

$$div_{geo}(i, R) = PR(\{i\} \cup R) - PR(R) \quad (10)$$

3.3 Diversificação Geográfica-Catégorica

Com as Equações 7 e 10 definidas, o *Geo-Cat* consiste então em uma combinação linear desses conceitos supracitados, como mostra a Equação 11. A primeira parte dessa equação refere-se a diversificação geográfica, mais especificamente à proporcionalidade geográfica de distribuição dos POIs proposta por Han J. & Yamana H. [8]. A segunda parte corresponde a diversificação catégorica dos POIs, em que utilizamos as duas principais métricas existentes na literatura: *ILD* e *Genre Coverage*. Para equilibrar os fatores geográficos e catégoricos, adicionamos um parâmetro β a ser ajustado com algum valor real entre 0 e 1. Assim, a configuração deste parâmetro permite modificar o comportamento da diversidade do nosso método *Geo-Cat*. Valores próximos de 0 resultam em recomendações mais diversas catégoricamente, enquanto valores próximos de 1 resultam em recomendações com maior diversificação geográfica.

$$div_{geo-cat}(i, R) = \beta \cdot div_{geo}(i, R) + (1 - \beta) \cdot div_{cat}(i, R) \quad (11)$$

Visando equilibrar os conceitos de diversidade e acurácia, propomos um algoritmo de otimização para maximizar a diversificação geográfica-catégorica de uma lista de POIs. Nossa função objetivo, apresentada na equação 12, corresponde matematicamente ao método de diversificação do *Geo-Cat* acrescido da relevância do POI para o usuário. Essa relevância, representada na variável $rel(i)$, representa o *score* atribuído por qualquer recomendador base ao POI_i . Nosso intuito é manter a qualidade da predição realizada, ao mesmo tempo que maximizamos o fator de diversificação proposto. Esse *trade-off* é equilibrado pelo parâmetro λ , um valor real entre 0 e 1. Quando λ é próximo de 1 são produzidas listas de recomendação que atribuem mais peso para a diversidade. Por outro lado, se λ é próximo de 0 são geradas listas de recomendações cujos pesos são maiores para a nota de relevância do POI, criando recomendações menos diversas e mais parecidas com as que foram geradas pelo recomendador base. Essa proposta é semelhante a utilizada em [8] para melhorar a diversificação geográfica das recomendações de POIs. Contudo, no *Geo-Cat* reconstruímos a segunda parte da equação com o objetivo de considerar simultaneamente a diversificação catégorica e geográfica.

$$f_o(i, R) = rel(i)^{1-\lambda} + div_{gc}(i, R)^\lambda \quad (12)$$

$$div_{gc}(i, R) = \max(0, div_{geo-cat}(\{i\} \cup R)) \quad (13)$$

Para maximizar o fator de diversidade definimos a Equação 13 e propomos uma otimização gulosa para selecionar POIs que potencializam o método *Geo-Cat*. Esse processo é representado no Algoritmo 1. Inicialmente, obtém-se uma lista temporária oriunda de qualquer recomendador base (linha 5). Na primeira iteração, o *Geo-Cat* calcula a função objetivo para cada um dos n POIs da lista temporária para encontrar o POI com o maior valor dentre todos. Em seguida, esse POI é incluído na lista de recomendações (como primeiro POI) e removido da lista temporária. Na segunda iteração, o algoritmo calcula a função objetivo (linha 11) acrescentando cada um dos POIs remanescentes na lista com o primeiro POI. Com isso,

o algoritmo detecta qual é o POI que junto com os que já foram adicionados anteriormente na lista de recomendações alcança o maior resultado na função objetivo. Esse processo é repetido por k vezes, a fim de montar uma lista com os top- k POIs que maximizam essa função objetivo.

Algorithm 1 GEO-CAT DIVERSIFICATION (u, R', k)

```

1:  $R \leftarrow \emptyset$ 
2: for  $i$  in  $k$  do
3:    $POI^* \leftarrow \max_{p \in R'}(rel(p)^{1-\lambda} + div_{gc}(p, R)^\lambda)$ 
4:    $R \leftarrow POI^*$ 
5:    $R' \leftarrow R' \setminus POI^*$ 
6: end for
7: return  $R$ 
```

4 CONFIGURAÇÃO EXPERIMENTAL

No intuito de avaliar nossa proposta, definimos um projeto experimental a ser aplicado em dados de cidades reais. Para isso, selecionamos os principais recomendadores base, bem como as técnicas estado-da-arte para diversificação como *baselines*. Para analisar o desempenho, selecionamos diversas métricas da literatura e adaptamos uma métrica utilizada na Teoria dos Jogos para avaliar um método sobre múltiplos atributos, como acurácia, novidade e diversidade.

4.1 Base de Dados

Selecionamos os dados disponíveis na base do *Yelp Challenge*¹ relativos às cidades de Las Vegas e Phoenix nos Estados Unidos. Para ambas, realizamos o pré-processamento dos dados, recomendado pelos trabalhos da literatura, para selecionar apenas: (1) os POIs que foram visitados por cinco ou mais usuários; e (2) os usuários que realizaram vinte ou mais *check-ins* na cidade alvo. A quantidade de *check-ins*, usuários e POIs remanescentes são apresentadas no gráfico da Figura 1. Desses dados, 70% dos primeiros *check-ins* de cada usuário foram separados para o treino e os 30% restantes para testes.

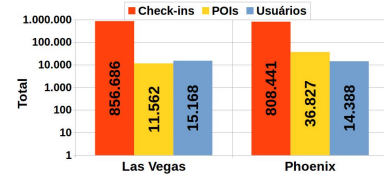


Figure 1: Usuários e POIs das cidades selecionadas.

Além disso, na base de dados do *Yelp* cada POI é categorizado hierarquicamente em uma árvore de categorias de muitos níveis. Por essa razão, assim como foi feito no trabalho de Han & Yamana [8], descartaram-se as categorias que estão abaixo do segundo nível da árvore para os experimentos. Esta filtragem é essencial porque as categorias são utilizadas para avaliar quanto um POI é similar a outro. Dessa forma, mede-se a distância entre as categorias de um POI i em relação às categorias de outro POI j usando a Equação 14 proposta por Castillo et al. [3], onde cat_i representa a categoria do POI i , cat_j indica a categoria do POI j e $sp(cat_i, cat_j)$ é a distância do caminho mais curto entre as categorias i e j na árvore. Ressalta-se ainda que, quando o i ou o j tem mais de uma categoria, é considerado o menor valor da distância entre eles [8].

¹Disponível em <https://www.yelp.com/dataset/challenge>

$$dissim_{cat}(i, j) = 1 - \frac{1}{(1 + sp(cat_i, cat_j))} \quad (14)$$

4.2 Recomendadores Seleccionados

Avaliamos o impacto da aplicação do *Geo-Cat* sob duas perspectivas. Em primeiro lugar, avaliamos a aplicação da nossa estratégia de pós-processamento sobre os recomendadores base. Para tal, seleccionamos quatro dos principais recomendadores clássicos, e um dos principais algoritmos utilizados no cenário de POIs. Os métodos tradicionais foram executados pela biblioteca do *MyMediaLite* com os parâmetros *default*. O método específico para POI foi executado com base no código disponibilizado por [8].

- **Most Popular:** recomenda os POIs mais populares entre aqueles que ainda não foram frequentados pelo usuário alvo.
- **User-kNN:** recomenda os POIs mais relevantes para os usuários similares ao usuário alvo, sob um raio de vizinhança.
- **WRMF:** recomenda os POIs mais bem avaliados com base nos fatores latentes dos usuários e dos POIs, resultantes da aplicação de um método de fatorização de matriz.
- **Item Attribute kNN:** recomenda os POIs mais similares aos POIs avaliados pelo usuário alvo, considerando os *atributos* (e.g., categorias) dos itens como base para a similaridade.
- **USG:** recomenda os POIs tal qual o *User-kNN*, porém calcula a similaridade dos usuários pelas informações geográficas e sociais.

Em segundo lugar, comparamos os resultados alcançados pelo *Geo-Cat* com as técnicas de pós-processamento que são estado-da-arte em diversificação. Seleccionamos os algoritmos **PM2** [5], **LD** [18], **Binom** [18] e **Geo-Div** [8], detalhados na Seção 2. Nesse caso, consideramos o **USG** como recomendador base para os algoritmos de diversificação por ser específico para o cenário de POIs, uma vez que utiliza informações sociais e geográficas para calcular a similaridade dos usuários [21].

4.3 Métricas de Avaliação

A fim de contrastar a nossa proposta com as estratégias existentes na literatura, seleccionamos diversas métricas relacionadas à acurácia (*Precision* e *Recall*), Diversidade e Novidade. Com o intuito de avaliar a eficácia da nossa proposta em distintos tamanhos de listas avaliamos as recomendações geradas com 5, 10 e 20 POIs para todas essas métricas.

- **Precision (Prec):** representa a probabilidade de um POI recomendado ser relevante para o usuário. Medimos a relevância verificando se o POI está ou não no conjunto teste desse usuário [14].
- **Recall (Rec):** representa a probabilidade que um POI relevante tem de ser recomendado para o usuário [14]. Utilizamos a mesma noção de relevância do *precision*, mencionada acima.
- **EPC:** avalia a novidade da lista de recomendação de acordo com a popularidade dos POIs. Quanto mais alto o valor de EPC menos populares são os POIs presentes na lista de recomendações [19].
- **ILD:** avalia a dissimilaridade entre as categorias dos POIs da lista de recomendação, tal qual mencionado anteriormente [18].
- **Genre Coverage (GC):** representa a média de categorias relevantes recomendadas para o usuário [13]. Novamente, a relevância está relacionada as categorias dos POIs presentes no teste.

- **PRg:** avalia a distribuição dos POIs recomendados nas subáreas. Valores altos indicam que os POIs estão geograficamente distribuídos nas subáreas onde o usuário é mais ativo. Valores baixos indicam que os POIs recomendados estão concentrados em poucas subáreas ou espalhados aleatoriamente [5, 8].

4.4 Teoria da Utilidade Multiatributo (MAUT)

Diferente da literatura, esse trabalho propõe avaliar 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. Por essa razão, adaptamos a Métrica da Utilidade Multiatributo (MAUT), utilizada no domínio de Teoria dos Jogos [12] para manipular problemas em que os resultados são dependentes de dois ou mais atributos [16]. Em nosso trabalho, a MAUT visa representar a eficácia de um método em função das métricas de qualidade [10]. Nesse caso, 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 [15].

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 15, 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 [16]. Assim, após obter os resultados do modelo aditivo da Equação 15 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 por meio dessa adaptação da MAUT conseguimos avaliar a qualidade dos recomendadores considerando diversas perspectivas simultaneamente.

$$U_j = \sum_{i=1}^{|M|} U_{i,j} \cdot w_i \quad (15)$$

5 RESULTADOS & DISCUSSÕES

Primeiramente, implementamos os algoritmos e métricas descritos, respectivamente, nas seções 4.2 e 4.3 em um ambiente de experimentos próprio e integrado à biblioteca *MyMediaLite* [6]. Dessa forma, realizamos um estudo preliminar para configurar os parâmetros do nosso método de recomendação *Geo-Cat*. Em seguida, avaliamos nossa proposta em duas partes, por meio de: (1) um comparativo entre os resultados alcançados pelos recomendadores base e pelo *Geo-Cat* após realizar o pós-processamento; e (2) outro comparativo entre resultados alcançados pelo *Geo-Cat* com os dos algoritmos de diversificação estado-da-arte. Por fim, agregamos os resultados por meio da métrica MAUT no intuito de identificar a estratégia capaz de equilibrar os principais requisitos de qualidade.

Algoritmo	Las Vegas						Phoenix					
	Prec@5	Rec@5	EPC@5	ILD@5	PRg@5	GC@5	Prec@5	Rec@5	EPC@5	ILD@5	PRg@5	GC@5
ItemAttributeKNN	0,0146	0,0060	0,8740	0,1835	0,1129	0,2548	0,0064	0,0021	0,9255	0,1852	0,0241	0,2573
Geo-cat (ItemAttributeKNN)	0,0228 ▲	0,0089 ▲	0,8625 ▼	0,4846 ▲	0,3878 ▲	0,4869 ▲	0,0121 ▲	0,0039 ▲	0,9214 ▼	0,4735 ▲	0,1710 ▲	0,5340 ▲
MostPopular	0,0228	0,0090	0,1062	0,7242	0,1109	0,2170	0,0141	0,0046	0,3041	0,4682	0,0108	0,3405
Geo-cat (MostPopular)	0,0391 ▲	0,0148 ▲	0,4255 ▲	0,7071 ▼	0,3354 ▲	0,4720 ▲	0,0188 ▲	0,0064 ▲	0,5495 ▲	0,5404 ▲	0,1445 ▲	0,5706 ▲
UserKNN	0,0501	0,0200	0,5276	0,6115	0,2227	0,2776	0,0386	0,0140	0,6907	0,4762	0,1180	0,3486
Geo-cat (UserKNN)	0,0506 ●	0,0206 ●	0,6241 ▲	0,6790 ▲	0,3921 ▲	0,4784 ▲	0,0367 ●	0,0135 ●	0,7732 ▲	0,5757 ▲	0,2862 ▲	0,5327 ▲
USG	0,0548	0,0237	0,5805	0,6070	0,2503	0,2740	0,0460	0,0182	0,7407	0,4695	0,1537	0,3533
Geo-cat (USG)	0,0497 ▼	0,0214 ●	0,6543 ▲	0,6868 ▲	0,4216 ▲	0,5033 ▲	0,0380 ▼	0,0153 ▼	0,7989 ▲	0,5926 ▲	0,3261 ▲	0,5491 ▲
WRMF	0,0520	0,0204	0,4714	0,6013	0,2106	0,2807	0,0319	0,0111	0,5934	0,4734	0,0792	0,3521
Geo-cat (WRMF)	0,0521 ●	0,0207 ●	0,5885 ▲	0,6741 ▲	0,3844 ▲	0,4854 ▲	0,0331 ●	0,0117 ●	0,7093 ▲	0,5560 ▲	0,2551 ▲	0,5433 ▲
Algoritmo	Prec@10						Rec@10					
	Prec@10	Rec@10	EPC@10	ILD@10	PRg@10	GC@10	Prec@10	Rec@10	EPC@10	ILD@10	PRg@10	GC@10
ItemAttributeKNN	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 (ItemAttributeKNN)	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 (MostPopular)	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 (UserKNN)	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 (USG)	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 (WRMF)	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					
	Prec@20	Rec@20	EPC@20	ILD@20	PRg@20	GC@20	Prec@20	Rec@20	EPC@20	ILD@20	PRg@20	GC@20
ItemAttributeKNN	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 (ItemAttributeKNN)	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 (MostPopular)	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 (UserKNN)	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 (USG)	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 (WRMF)	0,0407 ●	0,0635 ●	0,6085 ▲	0,6569 ▲	0,5104 ▲	0,7253 ▲	0,0258 ▲	0,0357 ▲	0,7176 ▲	0,5554 ▲	0,3515 ▲	0,7412 ▲

Table 1: Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas.

5.1 Configurações de Parâmetros

Na nossa proposta, o comportamento do *Geo-Cat* pode ser modificado pelos parâmetros λ , para configurar de forma geral o peso do fator diversidade, e β , para guiar a diversidade para o fator geográfico e/ou categórico. Portanto, para validar os parâmetros ideais para nossa abordagem, variamos os valores de λ e β entre 0 e 1 com intervalos de 0,25 a cada teste, resultando em 21 combinações possíveis (quando $\lambda = 0$ o valor de β não é considerado).

Dessa forma, é apresentado no gráfico da Figura 2 o comportamento do *Geo-Cat*, considerando as recomendações de 10 POIs, sob todas as métricas, em diferentes configurações de parâmetros. Primeiramente, observa-se pouca variação em *precision* e *recall*. Os valores dessas métricas tendem a cair quando $\lambda = 1$ uma vez que o *Geo-Cat*, nesse caso, considera que apenas a diversidade é importante. Observa-se também que aumentar o valor do parâmetro λ (primeiro valor apresentado no eixo X) faz com que as métricas de diversidade alcancem melhores resultados. Por fim, percebe-se que o parâmetro β (segundo valor apresentado no eixo X) funciona perfeitamente como um *threshold* entre as diversificações geográfica e categórica. Logo, conforme o valor de β aumenta, observa-se que o valor de PR_g também aumenta, pois estamos atribuindo mais peso para a diversificação geográfica.

Para definir os melhores parâmetros para nossa estratégia, utilizamos a métrica MAUT considerando todas as métricas de acurácia, novidade e diversidade ao mesmo tempo. Também consideramos que todas as métricas de avaliação são igualmente importantes e possuem pesos iguais. De acordo com os resultados obtidos, a combinação $\lambda = 0,75$ e $\beta = 0,75$ apresenta os melhores resultados em média para todas as cidades considerando recomendações de 5, 10 e 20 POIs. Da mesma forma, executamos os mesmos testes de parâmetros para os algoritmos *baselines* de diversificação e pelas mesmas razões selecionamos a melhor configuração para cada um: **Binom** ($\lambda = 0,75$); **Geo-div** ($\lambda = 0,5$); **LD** ($\lambda = 0,25$); **PM2** ($\lambda = 1,0$).

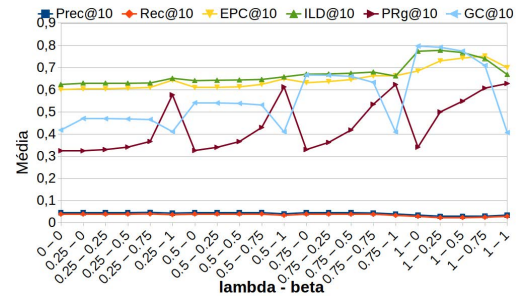


Figure 2: Resultados do Geo-Cat com diferentes parâmetros.

5.2 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 métricas 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,



Figure 3: Gráficos de ganhos e perdas percentuais dos algoritmos de diversificação.

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 *WMRF* 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 existentes na literatura, mostrando ser uma potencial estratégia a ser aplicada em domínios reais.

5.3 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. Disponibilizamos uma tabela detalhando dos resultados por meio do endereço¹. Novamente, comparamos os resultados utilizando o teste de *Wilcoxon* com 95% de

confiança. Por uma questão de limitação de espaço, resumimos os resultados por meio dos gráficos apresentados na Figura 3, que apresentam diferenças percentuais de cada estratégia aplicada sobre o *USG*. Para cada estratégia, calculamos o quanto ela ganhou ou perdeu dos resultados do SR base (*USG*) e plotamos essa variação.

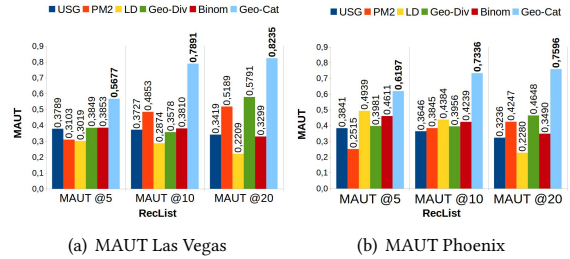


Figure 4: Valores de MAUT

Primeiramente, analisando os resultados referentes às métricas de acurácia, temos que todas as estratégias de pós-processamento apresentaram pequenas variações negativas, na maioria das vezes sem significância estatística, demonstrando que grande parte dessas estratégias são capazes de gerar recomendações diversificadas sem afetar a acurácia das mesmas. O *Geo-Cat* se destaca por ser a estratégia que apresenta as menores diferenças quando comparado às demais estratégias. Do ponto de vista das métricas de diversidade e novidade (i.e. *EPC*, *ILD*, *PR_g* e *GC*), é possível observar que nenhuma das estratégias de pós-processamento consideradas apresenta os melhores resultados sobre todas as métricas consideradas. Por exemplo, o *LD* é método que apresenta os melhores resultados quando observamos a métrica de *ILD*. Esse método utiliza a similaridade entre os POIs para aumentar a diversidade das recomendações,

¹<https://drive.google.com/file/d/1VWZcTjmcjo-fyYujNDA5FTnea6ij0MM9/view>

ou seja, recomenda POIs o mais diferente possível entre si. Apesar dessa estratégia melhorar o *ILD*, o mesmo não é capaz de apresentar melhorias expressivas em outras métricas, sobretudo em $PR_g@N$. Outro exemplo é o método Geo-Div que em sua proposta original visa a diversificação geográfica dos POIs. Apesar de conseguir apresentar os melhores resultados em *PGr*, para demais métricas apresenta um resultado bem inferior aos outros métodos. Esse é um fato importante de se destacar, uma vez que na maioria dos trabalhos apresentados na literatura, apenas um ou duas dessas métricas são consideradas nas avaliações. Contudo, observando os resultados referentes ao Geo-Cat, temos que seus resultados estão sempre entre os melhores obtidos para cada uma das métricas avaliadas, se apresentando como a melhor solução dentre todas as estratégias.

Para confirmar as conclusões obtidas no parágrafo anterior, novamente utilizamos a Métrica da Utilidade Multiatributo (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. 15). Conforme apresentado na Seção 4.4, o valor da MAUT está entre 0 e 1: quando mais próximo de 1, melhor é o comportamento de um determinado método frente às múltiplas métricas avaliadas. Os resultados referentes à essa análise estão apresentados nos gráficos da Figura 4. 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 & TRABALHOS FUTUROS

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, 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 [12]. 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. Outras duas questões relevantes podem ser discutidas em trabalhos futuros: (1) É possível mitigar o problema da esparsidade em recomendação de POIs e ainda assim obter bons índices de diversidade e novidade?; e (2) Qual o desempenho dos SsR de POIs para o problema de *cold*

start? Além disso, almejamos a realização de experimentos *on-line* com grupos de usuários distintos a fim de validar a real utilidade dos resultados apresentados neste artigo.

REFERENCES

- [1] Idir Benouaret and Dominique Lenne. 2016. A package recommendation framework for trip planning activities. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 203–206.
- [2] Jaime G Carbonell and Jade Goldstein. 1998. The Use of MMR and Diversity-Based Reranking for Reordering Documents and Producing Summaries. (1998).
- [3] L CASTILLO, E ARMENGOL, E ONAINDIA, L SEBASTIA, J GONZALEZBOTTICARIO, A RODRIGUEZ, S FERNANDEZ, J ARIAS, and D BORRAJO. 2008. samap: An user-oriented adaptive system for planning tourist visits. *Expert Systems with Applications* 34, 2 (feb 2008), 1318–1332. <https://doi.org/10.1016/j.eswa.2006.12.029>
- [4] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM.
- [5] 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.
- [6] Zeno Gantner, Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2011. MyMediaLite: A Free Recommender System Library. In *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys 2011)*.
- [7] Jean-Benoît Griesner, Tael Abdesslem, and Hubert Naacke. 2015. POI recommendation: towards fused matrix factorization with geographical and temporal influences. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 301–304.
- [8] Jungkyu Han and Hayato Yamana. 2017. Geographical Diversification in POI Recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems - RecSys '17*. ACM Press. <https://doi.org/10.1145/3109859.3109884>
- [9] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen, and John T Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)* 22, 1 (2004), 5–53.
- [10] 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>
- [11] Sean M. McNee, John Riedl, and Joseph A. Konstan. 2006. Being accurate is not enough. In *CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*. ACM Press. <https://doi.org/10.1145/1125451.1125659>
- [12] 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).
- [13] Shameem A Puthiya Parambath, Nicolas Usunier, and Yves Grandvalet. 2016. A coverage-based approach to recommendation diversity on similarity graph. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 15–22.
- [14] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer.
- [15] Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor (Eds.). 2011. *Recommender Systems Handbook*. Springer US. <https://doi.org/10.1007/978-0-387-85820-3>
- [16] Stuart Russell. 2004. *Inteligência Artificial (Tradução da Segunda Edição)*. Elsevier.
- [17] Rodrygo LT Santos, Craig Macdonald, and Iadh Ounis. 2010. Exploiting query reformulations for web search result diversification. In *Proceedings of the 19th international conference on World wide web*. ACM, 881–890.
- [18] Saúl Vargas, Linas Baltrunas, Alexandros Karatzoglou, and Pablo Castells. 2014. Coverage, redundancy and size-awareness in genre diversity for recommender systems. In *Proceedings of the 8th ACM Conference on Recommender systems*. ACM, 209–216.
- [19] Saúl Vargas and Pablo Castells. 2011. Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems - RecSys '11*. ACM. <https://doi.org/10.1145/2043932.2043955>
- [20] Shaojung Sharon Wang and Michael A Stefanone. 2013. Showing off? Human mobility and the interplay of traits, self-disclosure, and Facebook check-ins. *Social Science Computer Review* 31, 4 (2013), 437–457.
- [21] Mao Ye, Peifeng Yin, and Wang-Chien Lee. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*. ACM, 458–461.
- [22] Mi Zhang and Neil Hurley. 2008. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2008 ACM conference on Recommender systems*. ACM, 123–130.
- [23] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. 2005. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web - WWW '05*. ACM Press.