

Recomendação de Música baseada em Filtros Colaborativos utilizando Informação Temporal

Ricardo Dias, Manuel J. Fonseca
Departamento de Engenharia Informática, INESC-ID/IST-UTL
Rua Alves Redol, 9, 100-029, Lisboa
ricardo.dias@ist.utl.pt, mjf@inesc-id.pt

Resumo

Diversos estudos realizados nos últimos anos revelaram que os hábitos de audição de música seguem determinados padrões ao longo do tempo, tais como a preferência por ouvir determinados artistas ou géneros musicais em certos períodos do dia e em determinados dias da semana. Porém, apenas recentemente têm surgido algoritmos de recomendação que tiram partido deste conhecimento para melhorar as recomendações. Uma das abordagens mais estudada e utilizada em produtos comerciais é a de Filtros Colaborativos, que se foca nas músicas ouvidas por utilizadores com gostos semelhantes, mas que deixa de fora o contexto temporal em que estas foram ouvidas. Neste trabalho, apresentamos uma análise da utilização de informação temporal em técnicas de filtros colaborativos baseadas em sessões. De forma a medir a importância do contexto temporal na recomendação, comparámos dois algoritmos: um algoritmo base, sem recurso a informação temporal; e um algoritmo que tira partido de características temporais das sessões. Estas características são usadas pelo algoritmo para inferir similaridade entre as sessões. Para avaliar os dois algoritmos em termos de precisão e ordenação das recomendações, medimos o HitRatio e Mean Reciprocal Rank. Os resultados evidenciam que a inclusão de características temporais neste tipo de algoritmos é benéfica para melhorar a precisão, mas não a ordenação das listas de recomendações geradas.

Palavras-Chave

Algoritmo, Recomendação, Hábitos de Audição, Filtros Colaborativos, Contexto Temporal

1 Introdução

Devido à expansão da internet e ao crescente aparecimento de serviços online para ouvir música, os utilizadores passaram a ter acesso a uma quantidade enorme de músicas para ouvir. No entanto, com o acesso directo a milhões de músicas, a tarefa de escolher que músicas ouvir torna-se penosa e difícil. Neste sentido, e de forma a aliviar o fardo de escolher individualmente as músicas a ouvir, os investigadores têm-se preocupado cada vez mais com o desenvolvimento de mecanismos automáticos de recomendação com base nas preferências dos utilizadores.

Tradicionalmente, os sistemas de recomendação podem ser classificados em cinco categorias: demográficos, filtros colaborativos (CF), baseados no contexto, baseados no conteúdo, ou métodos híbridos. Embora os mais estudados e com maior sucesso sejam os de filtros colaborativos, recentemente, as técnicas baseadas no contexto têm ganho importância, à medida que os investigadores identificam que determinados factores como o tempo, as actividades realizadas, os locais ou as emoções que sentimos, influenciam as nossas escolhas musicais. Estes estudos revelam também a preferência por determinadas músicas, artistas e géneros em certos momentos do dia e em dias específicos da semana [Baltrunas09, Herrera10], sugerindo

que técnicas que recorram ao contexto temporal sejam bem sucedidas.

No entanto, as técnicas de filtros colaborativos não exploram características temporais [Lee06], nem o facto de os utilizadores darem preferência a determinadas músicas para ouvir em sequência [Hansen09]. Um dos trabalhos mais recentes é o algoritmo desenvolvido por Park [Park11], onde o autor apresenta um algoritmo baseado em filtros colaborativos, em que utiliza sessões na recomendação para capturar sequência e repetição no processo de audição de músicas. Também Hu [Hu11] apresentou um novo algoritmo que explora a probabilidade uma música ser tocada num determinado período do dia. Porém, nenhuma destas técnicas utiliza informação temporal extraída a partir das sessões no processo de recomendação.

Neste artigo exploramos a utilização de informação temporal em técnicas de filtros colaborativos baseadas em sessões. Partindo de um algoritmo base de filtros colaborativos, propomos um conjunto de características temporais extraídas das sessões, nomeadamente, o período do dia, dia da semana, dia do mês e mês, de forma a inferir semelhança entre sessões e assim melhorar a recomendação. Para determinar a semelhança entre

sessões usando as características extraídas, utilizámos um algoritmo de aprendizagem não supervisionada, de forma a agrupar sessões semelhantes de acordo com as características temporais.

Para avaliar o desempenho do algoritmo desenvolvido, comparámo-lo com o algoritmo base sem recurso a informação temporal. A avaliação experimental consistiu em prever a próxima música para ser tocada na sessão corrente, usando informação das músicas já tocadas anteriormente. Os resultados mostram que a inclusão de informação temporal no processo de recomendação melhora significativamente a precisão, mas não a ordenação das listas de músicas recomendadas, sugerindo que mais esforço deve ser dedicado a esta última componente.

2 Filtros Colaborativos baseados em Sessões

Tipicamente, as técnicas de filtros colaborativos não utilizam informação temporal para recomendar músicas aos utilizadores. Porém, Park [Park11] desenvolveu uma técnica de filtros colaborativos (*SSCF*) que captura implicitamente o contexto temporal, recorrendo a sessões, conjuntos de músicas ouvidas em sequência num determinado período temporal. Nesta técnica, Park mantém toda a estrutura inerente às técnicas de filtros colaborativos, alterando apenas a informação utilizada para gerar as recomendações (sessões em detrimento de utilizadores).

Um ponto importante desta técnica é a fórmula de predição do *rating* da próxima música a ser tocada (ver Equação 1). O *rating* previsto para uma música nova i na sessão activa s é determinado de acordo com a mesma fórmula que seria usada se esta fosse uma técnica de filtros colaborativos baseada em utilizadores [Celma10] e não em sessões.

$$\hat{r}_{s,i} = \bar{r}_s + \frac{\sum_{v \in S(s)^k} \text{sim}(s,v)(r_{v,i} - \bar{r}_v)}{\sum_{v \in S(s)^k} \text{sim}(s,v)} \quad (1)$$

$S(s)^k$ representa o número de k sessões vizinhas da sessão activa s , \bar{r}_s e \bar{r}_v representam o *rating* médio para as sessões s e v em comparação. $r_{v,i}$ representa a frequência com que a música i foi tocada na sessão v , e $\text{sim}(s,v)$ a similaridade entre as sessões s e v .

3 Utilização de Informação Temporal

Tal como referido anteriormente, a técnica *SSCF*, apesar de tirar partido da organização das músicas ouvidas em sessões, não utiliza características temporais sobre as sessões, passíveis de modelar os hábitos de audição do ponto de vista temporal. Nesse sentido, para caracterizar temporalmente as sessões, definimos um vector de características contendo quatro propriedades. Embora estas características tenham sido escolhidas a partir de alguns trabalhos relacionados [Lee06, Baltrunas09, Herrera10, Hu11], nenhuma delas foi utilizada anteriormente para determinar similaridade entre sessões. Apresentamos em seguida uma breve descrição das características.

- **Período do Dia:** captura a preferência de os uti-

lizadores ouvirem determinadas músicas (artistas e géneros), em determinados períodos do dia (por exemplo, de manhã, à tarde, ou à noite)

- **Dia da Semana:** modela a preferência de os utilizadores ouvirem músicas diferentes em dias diferentes da semana, por exemplo, durante um dia de trabalho ou ao fim de semana
- **Dia do Mês:** captura as preferências de ouvir músicas em dias diferentes do mês, possivelmente associados a diferentes fases do mês
- **Mês:** captura o facto dos utilizadores ouvirem algumas músicas em determinados meses, possivelmente associadas as estações do ano ou a períodos de trabalho / lazer

Com base na hipótese de que estas características são úteis para modelar as preferências ao longo do tempo, decidimos recorrer a um mecanismo de aprendizagem não supervisionado para agrupar sessões semelhantes e assim definir uma métrica para comparar sessões. Através do uso desta técnica conseguimos agrupar sessões semelhantes de acordo com as características temporais e por conseguinte modelar os hábitos de audição ao longo do tempo.

Neste sentido, treinámos um Modelo de Mistura Gaussiano (*GMM*), usando o algoritmo *Expectation Maximization algorithm* (*EM*), com os parâmetros padrão. Como resultado, cada sessão tem associado um vector de distribuições de probabilidades, com tantos valores como o número de grupos, ao invés de apenas um grupo associado.

Assim, uma vez que as sessões são descritas como vectores de distribuições de probabilidades, podemos inferir similaridade entre eles comparando os vectores. Com este método, sessões com probabilidades maiores para os mesmos grupos, serão mais similares do que as que têm valores mais baixos. Para estimar a similaridade entre os diferentes vectores de distribuições de probabilidades recorreremos posteriormente à divergência de *Kullback-Leibler*.

Assim, esta medida de semelhança pode ser utilizada na Equação 1, pois permite estimar quão similares duas sessões são, e influenciar quais as próximas músicas que devem ser ouvidas a seguir na sessão activa.

4 Avaliação Experimental

Para avaliar a importância da utilização de informação temporal no algoritmo de filtros colaborativos baseado em sessões (*TSSCF*), comparámo-lo com o algoritmo base *SSCF* [Park11], recorrendo às métricas padrão *HitRatio* (*HR*) e *Mean Reciprocal Rank* (*MRR*), indicadores de precisão e posição das recomendações

4.1 Dataset e Geração de Sessões

Como dados de entrada para avaliar os algoritmos utilizámos o *Last.fm dataset* [Celma10]. Este consiste num registo das músicas ouvidas por 992 utilizadores, com 19.150.868 entradas, respectivas a 176.948 artistas únicos

e 1.500.661 músicas únicas. Partindo destes dados, procedemos à geração de sessões. Embora existam diversas técnicas para gerar sessões [Pabarskaite07], decidimos utilizar a abordagem do *intervalo temporal*, dado que o dataset utilizado apenas disponibiliza a marca temporal em que a música foi ouvida sem informação adicional sobre como ou quanto tempo esta foi ouvida. Baseado no trabalho relacionado utilizámos diferentes valores como limite para gerar sessões: 10, 15, 20, 25 e 30 minutos. Utilizando estes valores obtivemos datasets com mais de 1.000.000 sessões, com uma média de 11 a 18 músicas por sessão.

4.2 Procedimento

De forma a estimar a capacidade de conseguir prever qual a próxima música a ser tocada numa sessão activa, escolhemos aleatoriamente 1.000 sessões do dataset para serem utilizadas como interrogações (sessões com menos de duas músicas foram descartadas). Após este passo removemos a última música de todas as sessões, de forma a poder ser usada para avaliação dos algoritmos. Usando as interrogações registámos o *top-100* das recomendações (n), para posteriormente podermos medir a presença e posição das músicas a recomendar (últimas músicas das sessões). Repetimos esta experiência 20 vezes e em cada iteração escolhemos aleatoriamente outro conjunto de 1.000 interrogações, de forma a efectuarmos comparações estatísticas.

4.3 Parâmetros

De acordo com a Equação 1, o número de vizinhos (k) é o parâmetro essencial para as técnicas de filtros colaborativos, e no algoritmo proposto, essencial para validar a métrica de similaridade entre sessões. Para avaliar os dois algoritmos adoptámos os valores descritos em [Park11].

O outro parâmetro necessário para o algoritmo proposto é o número de grupos usado para treinar o GMM. Para encontrar o valor mais adequado para os diferentes datasets de sessões utilizámos validação cruzada. 8 foi escolhido como o melhor valor para os 5 datasets.

4.4 Resultados

No geral, os resultados obtidos para o HitRatio (ver Figura 1-cima), evidenciam que a utilização de informação temporal melhora a precisão das recomendações em cerca de 3%, o correspondente a uma melhoria de cerca de 35%. No entanto, apesar de existirem mais acertos nas músicas a tocar, em relação ao MRR, o algoritmo base apresenta resultados duas vezes mais elevados que o algoritmo proposto. Para validar estes resultados realizámos dois testes Kruskal-Wallis. Tanto as diferenças nos valores de HitRatio ($H(2) = 1560.115, P < 0.01$) como de MRR foram consideradas significativas ($H(2) = 1064.659, P < 0.01$).

De forma a avaliar o impacto do mecanismo de geração de sessões, testámos os dois algoritmos utilizando os 5 datasets gerados. A Figura 2 evidencia as principais diferenças na utilização dos diferentes limites. Os resultados clarificam o que já tinha referido anteriormente no caso geral, onde a inclusão de informação temporal para

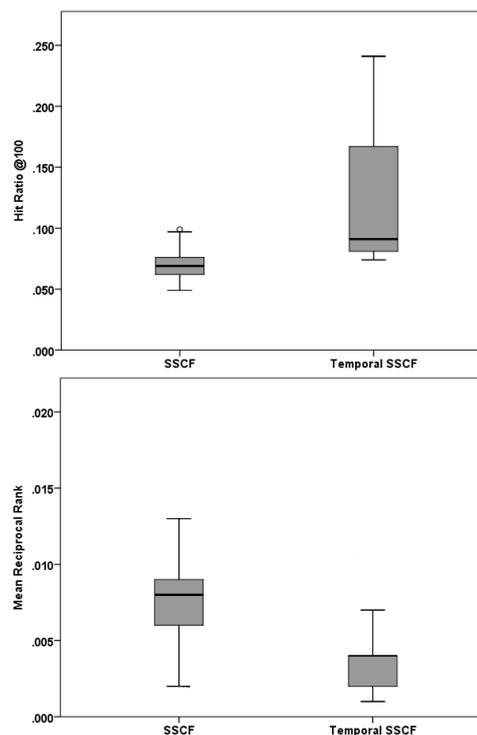


Figura 1. Resultados globais de HitRatio (em cima) e Mean Reciprocal Rank (em baixo).

a recomendação melhora os valores de precisão (HitRatio), mas no entanto não tem o mesmo efeito na posição das músicas desejadas nas listas de recomendações. De realçar também que o algoritmo TSSCF mantém constantes os valores de HitRatio para os diferentes datasets, e que os valores de MRR são melhores para os datasets gerados com limites menores ou maiores. Para posteriormente podermos comparar a utilização de diferentes conjuntos de vizinhos, aplicámos testes de Mann-Whitney de forma a determinar qual o dataset com que se obtiveram os melhores resultados. Não encontramos diferenças significativas entre usar $th = 15$ ou $th = 20$ em termos de HitRatio ($z = -0.293, p > 0.770$), no entanto, as diferenças de MRR entre estes dois datasets são significativas ($z = -2.356, p < 0.05$). Com base neste resultado, as comparações seguintes são realizadas apenas usando o dataset gerado com o limite $th = 20$.

A Figura 3 compara o impacto da utilização de diferentes números de vizinhos (k) nos valores de HitRatio e MRR para os dois algoritmos, no dataset gerado com limite $th = 20$. Os gráficos evidenciam que valores pequenos de vizinhança levam a melhores valores tanto de HitRatio como de MRR. De facto, com $k = 30$ os dois algoritmos apresentam os melhores resultados. Incrementos no número de vizinhos pioram os resultados em ambos os algoritmos, no entanto este decaimento é maior no algoritmo TSSCF, tal como a figura mostra.

Não obstante o facto do algoritmo SSCF ultrapassar clara-

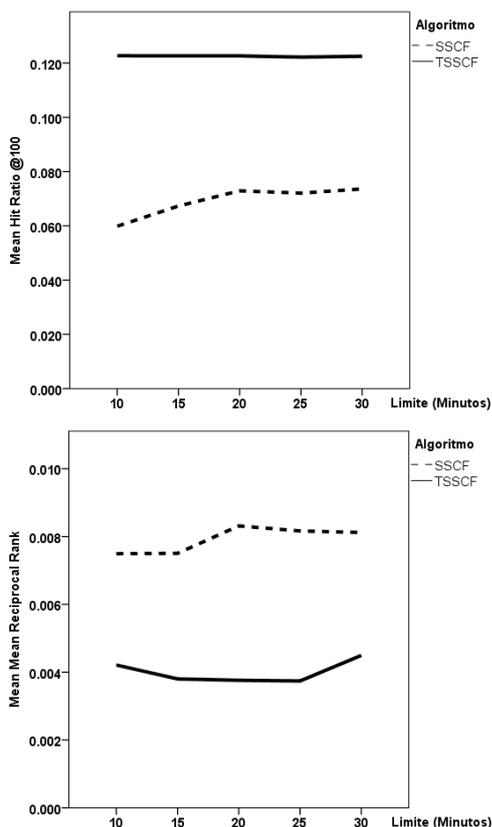


Figura 2. Impacto nas métricas HitRatio (em cima) e Mean Reciprocal Rank (em baixo) do limite utilizado para gerar sessões

mente o algoritmo TSSCF em termos de MRR, é importante salientar que estes valores são muito pequenos (menos de 0.01), o que evidencia a dificuldade em melhorar o posicionamento das músicas pretendidas.

5 Conclusões e Trabalho Futuro

Neste artigo explorámos a utilização de informação temporal para melhorar a recomendação de música em técnicas de filtros colaborativos baseadas em sessões. Para esse fim, propusemos um conjunto de características temporais extraídas das sessões, de forma a determinar semelhança entre elas. Os resultados da avaliação conduzida mostram que a utilização deste tipo de características melhora a precisão na recomendação, mas não a posição das recomendações. Como trabalho futuro, pretendemos avaliar o desempenho do algoritmo desenvolvido com recurso a testes com utilizadores. O nosso objectivo principal é validar se o algoritmo desenvolvido pode ser utilizado em tempo real para recomendar músicas aos utilizadores com base nos seus hábitos passados mais recentes.

6 Agradecimentos

Este trabalho foi financiado por fundos nacionais através da FCT –Fundação para a Ciência e a Tecnologia, no projecto PEst-OE/EEI/LA0021/2013. Ricardo Dias foi supor-

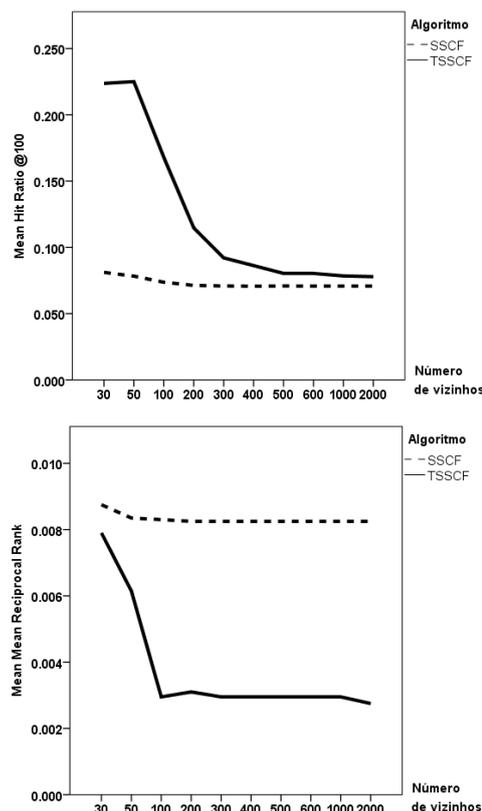


Figura 3. Impacto do número de vizinhos nos valores de HitRatio (em cima) e Mean Percentile Rank (em baixo) para $th = 20$.

tado pela FCT, bolsa SFRH/BD/70939/2010.

7. REFERÊNCIAS

- [Baltrunas09] Linas Baltrunas, Xavier Amatriain, and Via Augusta. Towards Time-Dependant Recommendation based on Implicit Feedback. In *CARS, RecSys*, 2009.
- [Celma10] Oscar Celma. *Music Recommendation and Discovery*. Springer, 1st edition, 2010.
- [Hansen09] D.L. Hansen and J. Golbeck. Mixing it up: recommending collections of items. In *CHI*, 2009.
- [Herrera10] Perfecto Herrera. Rocking around the clock eight days a week: an exploration of temporal patterns of music listening. In *WOMRAD*, pages 7–10, 2010.
- [Hu11] Yajie Hu and M. Ogihara. Nexttone Player: A Music Recommendation System Based on User Behavior. In *ISMIR*, 2011.
- [Lee06] J Lee. Music for my mood: A music recommendation system based on context reasoning. *Smart Sensing and Context*, 2006.
- [Pabarskaite07] Zidrina Pabarskaite and Aistis Raudys. A process of knowledge discovery from web log data: Systematization and critical review. *JGIS*, 28:79–104, 2007.
- [Park11] Sung Eun Park, Sangkeun Lee, and Sang-goo Lee. Session-Based Collaborative Filtering for Predicting the Next Song. *First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering*, 2011.