

That's my jam! Uma Análise Temporal sobre a Evolução das Preferências dos Usuários em uma Rede Social de Músicas

Fabiola S. F. Pereira¹, Cláudio D. G. Linhares¹, Jean R. Ponciano¹,
João Gama², Sandra de Amo¹, Gina M. B. Oliveira¹

¹Faculdade de Computação – Universidade Federal de Uberlândia, Brasil

²INESC TEC – Universidade do Porto, Portugal

{fabiola.pereira, claudiodgl, jean, deamo, gina}@ufu.br, jgama@fep.up.pt

Abstract. *User preferences are fairly dynamic, since users tend to exploit a wide range of information and modify their tastes accordingly over time. Existing models that capture the evolution of preferences in the music domain are very restricted and do not take into account social influence. This work proposes an analysis on the evolution of musical preferences of the users of a social music network, using temporal networks of similarity. We found that few users significantly vary their preferences. The tendency is similar artists and users maintaining their similarities over time.*

Resumo. *As preferências dos usuários são muito dinâmicas, uma vez que eles tendem a explorar uma vasta gama de informações e modificar seus gostos ao longo do tempo. Modelos existentes que capturam a evolução de preferências no domínio de músicas são muito restritos e não levam em conta a influência social. Este trabalho propõe uma análise sobre a evolução das preferências musicais dos usuários de uma rede social de músicas, utilizando redes temporais de similaridade. Como resultado, foi detectado que poucos usuários variam significativamente suas preferências. A tendência é que artistas e usuários semelhantes mantenham suas similaridades ao longo do tempo.*

1. Introdução

O que direciona a evolução das preferências dos usuários? Entender como essas preferências evoluem é uma das tarefas de personalização mais importantes nos contextos de recuperação da informação e sistemas de recomendação. Quanto mais se conhece sobre o usuário, melhor será a qualidade dos itens recomendados a ele. De fato, a observação de como um usuário modifica seus gostos e costumes ao longo do tempo tem sido explorada em diferentes domínios, como filmes [Siddiqui et al. 2014], leitura de notícias [Pereira et al. 2016] e visões políticas [Macropol et al. 2013]. No domínio da música o cenário não é diferente. Estudos na área de psicologia indicam que o gosto musical das pessoas muda e evolui ao longo do tempo [Bonneville-Roussy et al. 2013].

Nesse contexto de evolução de preferências musicais, os trabalhos em geral propõem observações sobre as variações dos gêneros musicais e artistas com base em atributos como popularidade e periodicidade de audição [Kapoor et al. 2013]. Uma outra linha de pesquisa busca por modelos de aprendizado da evolução das preferências musicais levando em conta a sequência de músicas ouvidas em *playlists* [Moore et al. 2013]. Nenhum desses estudos, entretanto, levam em consideração a influência social que o usuário

sofre durante o processo de evolução das suas preferências musicais. De acordo com a psicologia, influência social é quando o comportamento de uma pessoa faz com que outra pessoa mude de opinião ou execute uma ação que, de outro modo, não executaria [Michener 2005]. As redes sociais são instrumentos propícios para a ocorrência de tal fenômeno, pois elas naturalmente facilitam a construção de relações entre pessoas que compartilham os mesmos interesses. Nelas, usuários conectados trocam informações sobre novas tendências, preferências de consumo e opiniões entre seus pares. Em especial, analisar a evolução dessas redes torna-se interessante para a compreensão temporal de tais relações sociais [Aggarwal and Subbian 2014].

Tendo como motivação o cenário descrito, neste trabalho é investigada a evolução do gosto musical dos usuários da rede social de músicas *This Is My Jam*¹ (TIMJ), utilizando técnicas de análise de redes sociais temporais [Holme 2014]. A proposta é modelar duas redes temporais de similaridade – uma entre artistas e outra entre usuários, e analisá-las separadamente. As redes temporais de similaridade são aquelas em que os nós estão ligados com base em alguma semelhança entre eles em um determinado momento. Similaridade entre artistas significa que dois artistas são semelhantes se há uma grande intersecção entre seus públicos. No contexto de usuários, dois usuários similares são aqueles que compartilham do mesmo gosto musical. Entender como esse grau de semelhança se altera ao longo do tempo nessas redes leva à compreensão da evolução das preferências dos indivíduos.

O objetivo principal do trabalho é verificar se as preferências musicais dos usuários realmente mudaram ao longo do tempo. Para tanto, as seguintes perguntas de pesquisa serão respondidas: (1) artistas semelhantes em relação ao seu público mantêm a semelhança ao longo do tempo? (2) Usuários com mesmo gosto musical evoluem de maneira similar compartilhando das mesmas preferências? A originalidade deste trabalho está na proposta do uso de redes temporais de similaridade para análise da evolução das preferências dos usuários sobre dados no domínio de música.

2. Trabalhos Correlatos

Evolução das preferências. A maioria dos trabalhos que tratam sobre a evolução das preferências do usuário têm como objetivo a recomendação de itens com base na variação dos perfis de cada usuário [Felício et al. 2016]. No domínio de músicas, o trabalho [Moore et al. 2013] desenvolveu um modelo probabilístico que incorpora o tempo para explorar como as preferências musicais de uma população evoluem. A ideia é representar usuários e músicas em um espaço Euclidiano, no qual suas posições mudam ao longo do tempo, demarcando a trajetória da evolução. Em [Summers and Popp 2015] os autores buscam por padrões de músicas ouvidas de acordo com diferentes contextos. O objetivo é identificar uma relação por exemplo, entre músicas natalinas e a época de Natal, com foco na predição da próxima música a ser tocada em uma *playlist*. São trabalhos de motivação similar a esta proposta, porém não utilizam técnicas de análise de redes temporais.

Análise de Redes Temporais. De acordo com [Aggarwal and Subbian 2014] existem duas maneiras de analisar uma rede que evolui ao longo do tempo: através de métodos de manutenção e através de métodos analíticos. No primeiro, é desejável manter os resultados de um processo de mineração de dados continuamente ao longo do tempo. Por

¹<https://www.thisismyjam.com/>

exemplo, em tarefas como predição de *links* e detecção de comunidades (clusterização) deseja-se sempre manter os modelos atualizados e com boa acurácia à medida que a rede evolui. No segundo, a ideia é quantificar e entender as mudanças que ocorreram na estrutura da rede. Tais modelos estão focados em analisar a mudança, ao invés de apenas se ajustarem a ela. Este trabalho adéqua-se a essa segunda maneira de analisar uma rede em evolução – por meio do entendimento das mudanças.

Redes de Similaridade. Modelar a similaridade entre pessoas na forma de rede não é uma tarefa nova, tendo sido aplicada nos domínios de percepção visual em pinturas [Felício et al. 2016] e co-autoria de artigos científicos [Silva et al. 2015]. Em [Cano 2004] são analisadas redes de similaridade entre artistas no domínio da música, porém a evolução dessas redes não é levada em conta.

3. Modelagem das Redes

A análise foi conduzida sobre a rede social de músicas *This Is My Jam* (TIMJ). Nela, os usuários podem compartilhar suas músicas favoritas com seus seguidores. Apenas uma música pode ser compartilhada por vez - o *jam* atual, que dura no máximo uma semana no *status* dos usuários. Além disso, os usuários podem curtir o *jam* de outros usuários. A base de dados da TIMJ foi liberada por [Jansson et al. 2015] e contém 219.940 artistas, 132.299 usuários e 2.095.441 *jams*, referente ao período de 26/08/2011 a 26/09/2015. As análises apresentadas neste artigo levam em consideração períodos trimestrais, totalizando 17 instantes de tempo. Foram geradas duas redes de similaridade, descritas a seguir.

3.1. Rede de Similaridade entre Artistas (*art-art*)

Nessa rede os nós são os artistas e as arestas ocorrem entre artistas que interpretam as músicas compartilhadas como *jam* por um usuário em comum em determinado intervalo de tempo. O peso das arestas é definido pelo número de usuários em comum. Formalmente, tem-se: $\mathcal{A} = (V, E)$, onde \mathcal{A} é uma rede não-dirigida, ponderada, sendo V o conjunto de artistas, e cada aresta $e = (u, v, t) \in E$ para $u, v \in V$, indica que u e v são intérpretes de *jams* compartilhados no instante t por um mesmo usuário. A função peso é definida por $w : E \rightarrow \mathbb{N}$, sendo $w(e)$ o número de usuários em comum entre u e v em t . Foi definido empiricamente um limiar de similaridade $\alpha = 20$ que atua como um filtro sobre a rede, onde $E = \{e | w(e) > \alpha\}$.

3.2. Rede de Similaridade entre Usuários (*usr-usr*)

Na rede não-dirigida ponderada *usr-usr* os nós são os usuários e as arestas indicam a similaridade entre eles com base nos artistas em comum que foram compartilhados como *jam* em um determinado instante de tempo. O peso das ligações representa a força da similaridade (*tie strength* [Brandao and Moro 2017]) entre os nós, calculado em função da quantidade de artistas em comum que dois usuários compartilharam ao mesmo tempo. Formalmente, tem-se $\mathcal{U} = (V, E)$, onde \mathcal{U} é um grafo não-dirigido, ponderado, sendo V o conjunto de usuários, e cada aresta $e = (u, v, t) \in E$ indica que u e v são usuários que compartilharam o mesmo artista como *jam* no tempo t . A função peso é definida por $w : E \rightarrow \mathbb{N}$, sendo $w(e)$ o número de artistas compartilhados em comum entre u e v em t . Também foi definido um limiar de similaridade $\alpha = 100$ para filtro sobre a rede, onde $E = \{e | w(e) > \alpha\}$.

A Tabela 1 resume as características das redes implícitas analisadas neste trabalho, extraídas da base de dados TIMJ. Na Tabela 2 estão as propriedades que garantem que as redes geradas são redes do mundo real [Zafarani et al. 2014], com distribuição de graus seguindo a lei de potência, alto coeficiente de clusterização e baixo comprimento médio de caminho entre os nós.

Tabela 1. Descrição das redes implícitas geradas a partir da rede social de músicas *This Is My Jam*.

Rede	# nós	# arestas	Semântica	Limiar de Similaridade (α)
<i>art-art</i>	1670	3588	similaridade entre artistas	20
<i>usr-usr</i>	4388	9567	similaridade entre usuários	100

Tabela 2. Descrição das propriedades que garantem que as redes implícitas de similaridade geradas a partir da rede social de músicas são redes do mundo real.

Propriedade	<i>art-art</i>	<i>usr-usr</i>
distribuição dos graus	lei de potência	lei de potência
coeficiente de clusterização local médio	0,548	0,743
comprimento médio do caminho	4,57	3,849

3.3. Relação entre Redes de Similaridade e Preferências do Usuário

A ideia de investigar as preferências do usuário utilizando similaridade baseia-se no fato de que uma mudança na similaridade é um indício de que os usuários estão mudando seus gostos musicais. Do ponto de vista dos artistas, os nós densamente conectados indicam que aqueles artistas são ouvidos pelos mesmos usuários e, portanto, são similares em relação ao seu público. Com o passar do tempo, os artistas podem deixar de ser, tornarem-se ou manterem-se similares. Quanto mais mudanças nessa similaridade, mais os usuários estão mudando suas preferências. De fato, a rede *art-art* provê uma visão da evolução das preferências do usuário de uma maneira global. Os usuários podem, por exemplo, mudar suas preferências com base no lançamento de novos artistas ou músicas ou com base em acontecimentos como falecimento ou indicação para prêmios.

Em relação à similaridade entre usuários, a intuição da rede *usr-usr* é que os usuários estão conectados com base na semelhança entre suas preferências musicais. Uma mudança estrutural da rede pode indicar que os usuários estão deixando de ter gostos parecidos e tornando-se similares a outros. É uma visão do usuário em relação aos demais membros da comunidade musical, portanto uma evolução de maneira local. Uma pessoa pode mudar suas preferências com base em influências recebidas do meio externo no seu dia-a-dia, deixando de ser similar aos seus pares até então.

4. Metodologia

A análise temporal das redes de similaridade foi baseada em *snapshots* de cada instante de tempo (trimestre). Logo, não é uma abordagem incremental, pois cada métrica foi obtida baseada em apenas um instante de tempo, esquecendo o passado.

4.1. Métricas de Rede

Foram definidas as métricas ponte, influência e versatilidade para análise local dos nós das redes. Cada nó (artista ou usuário) tem o seu comportamento em relação a essas métricas. Também foi analisada a maneira como os nós evoluem em relação às suas vizinhanças ou comunidades. Tais definições são descritas a seguir.

1. **Comunidades:** foi utilizada a métrica *modularity* [Newman 2006] para detecção de comunidades em redes. Ela indica o quanto a estrutura de comunidade encontrada foi criada aleatoriamente. Assim, deseja-se dividir a rede em partições de modo que se a probabilidade de dois nós se conectarem aleatoriamente for baixa, eles devem estar na mesma partição e vice-versa [Zafarani et al. 2014]. No contexto de similaridade, tem-se que dentro de cada partição encontrada pela métrica estão os nós mais similares entre si.
2. **Ponte:** os “nós ponte” são aqueles que ligam duas ou mais comunidades, mantendo a rede conexa. A métrica utilizada foi a centralidade *betweenness* [Brandes 2001]. Essa medida indica o quão importante o nó é para conectar outros nós. Em redes de similaridade, um nó com alto valor de ponte é aquele que concentra características dos nós das comunidades que ele conecta.
3. **Influência:** um nó influente é aquele importante para os nós da sua vizinhança. A métrica *PageRank* [Brin and Page 1998] foi utilizada para detectar a influência de um nó nas redes de similaridade. Um nó com alto grau de influência na rede de similaridade tem interferência sobre seus pares, que por sua vez também têm alto grau de influência e assim por diante. Na similaridade entre artistas, por exemplo, um artista influencia demais artistas parecidos com ele.
4. **Versatilidade:** a versatilidade indica o quão próximo um nó está dos demais nós da rede. A métrica de centralidade *closeness* [Brandes 2001] foi utilizada. Quanto maior o *closeness*, mais similar é o nó em relação aos demais nós da rede, representando portanto um “nó eclético”.

4.2. Análise e Avaliação da Evolução das Redes

Uma primeira análise da evolução foi conduzida baseada em *insights* obtidos por meio da visualização das redes. De fato, técnicas de visualização de redes temporais têm sido propostas [Beck et al. 2017] e auxiliam na compreensão do problema em análise. O *layout* temporal gerado pela ferramenta DyNetVis [Linhares et al. 2017] foi utilizado, bem como as visualizações estruturais geradas pela ferramenta Gephi [Bastian et al. 2009].

As comunidades foram avaliadas com base em nós selecionados. Dado um nó $v \in V$, para cada instante de tempo t , sua comunidade C_t^v foi detectada. Depois, a persistência de tal comunidade foi obtida por meio do coeficiente de Jaccard:

$$com(v) = \frac{|C_1^v \cap \dots \cap C_n^v|}{|C_1^v \cup \dots \cup C_n^v|} \quad (1)$$

onde $n = 17$ corresponde à quantidade de instantes de tempo analisada. Quanto mais próximo de 1, maior é a persistência da comunidade de v .

Por fim, a avaliação da evolução de todos os nós das redes ocorreu em função da variância $var_c(v)$ de cada nó $v \in V$ em relação a uma métrica de centralidade local c durante n instantes de tempo, definida como:

$$var_c(v) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2)$$

onde n é 17 representando a quantidade de instantes de tempo e x é a distribuição da centralidade c do nó v .

5. Resultados Obtidos

As redes de similaridade entre artistas e usuários foram analisadas de forma que a evolução de cada uma leva à conclusão sobre a evolução das preferências dos usuários de maneira global e local, respectivamente.

5.1. Como foi a evolução dos artistas?

Uma visão estrutural, estática, da rede *art-art* é dada na Figura 1(a). É possível observar que não é uma rede densa e que é composta por diversas comunidades, em sua maioria, pequenas.

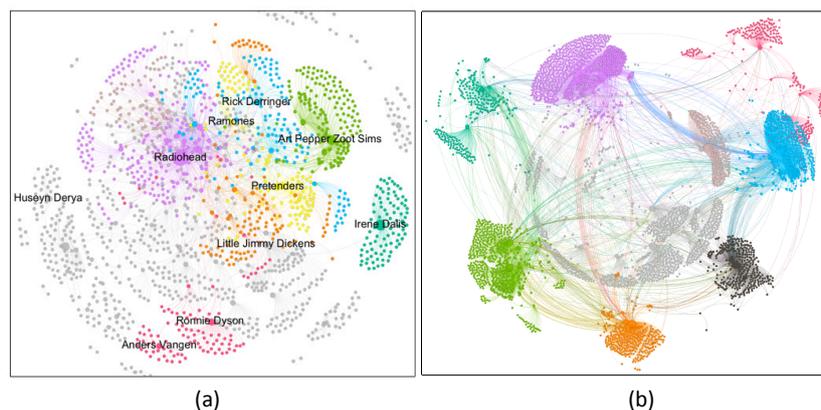


Figura 1. Visão geral das redes *art-art* (a) e *usr-usr* (b). As cores indicam a comunidade e o tamanho dos nós indica a influência. Alguns dos artistas mais influentes estão em destaque.

A visualização temporal da rede de similaridade entre artistas, mostrada na Figura 2(a), ilustra a intensidade de comunicação em relação aos 200 nós mais populares ao longo do tempo. No *layout* temporal, os nós são posicionados no eixo vertical e cada instante de tempo no eixo horizontal ilustra as arestas existentes entre os nós naquele momento [Linhares et al. 2017]. É uma perspectiva de atividade temporal na rede. Popularidade [Zafarani et al. 2014], no domínio dos artistas, indica aqueles que mais apareceram como *jam* nos perfis dos usuários considerando todo o período. São os artistas mais curtidos. Pode-se perceber uma interação intensa entre os trimestres 2 e 8 na rede de similaridade, ou seja, uma grande quantidade de artistas populares foram ouvidos pelos mesmos usuários.

Desconsiderando a evolução temporal, a Tabela 3 lista os 10 artistas que se destacam na rede estática, considerando todo o período de aproximadamente 4 anos. É uma visão geral da rede de similaridade *art-art*, que não provê informação sobre a evolução temporal de tais artistas. É interessante observar que os artistas *Radiohead*, *David Bowie*

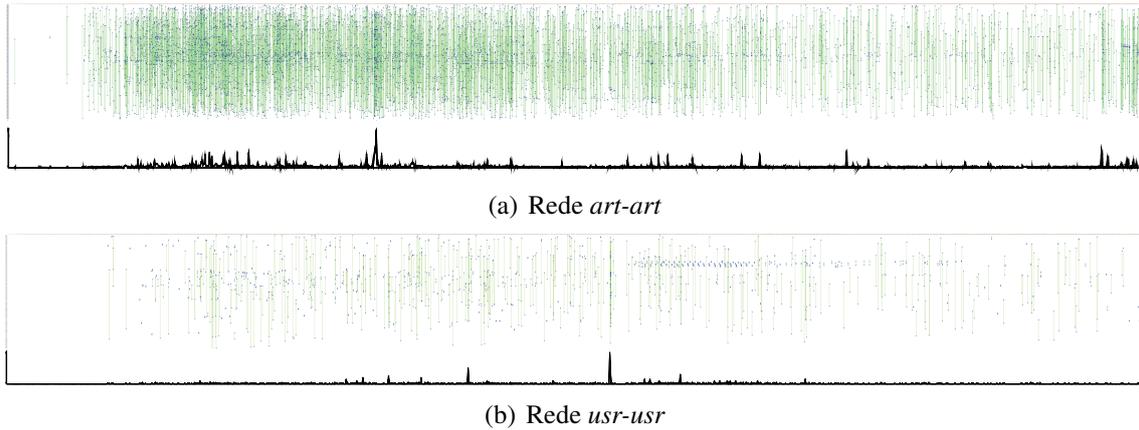


Figura 2. Visualização temporal das redes de similaridade considerando os 200 nós mais populares de cada domínio, durante aproximadamente 4 anos com granularidade diária. Visualização gerada pelo software DyNetVis [Linhares et al. 2017].

e *R.E.M.* são aqueles mais centrais em relação a todas as métricas calculadas. Em termos de preferências, a análise estática leva à conclusão de que são os artistas mais preferidos da comunidade de músicas TIMJ.

Tabela 3. Top-10 artistas com maiores valores médios das métricas locais. Artistas que se repetiram em todas as métricas estão em destaque.

	Ponte	Influência	Versatilidade
1	Radiohead	Radiohead	Radiohead
2	Art Pepper n Zoot Sims	Irene Dalis	David Bowie
3	Ramones	David Bowie	The Beatles
4	Queen	R.E.M.	Queen
5	Pretenders	Little Jimmy Dickens	The Rolling Stones
6	Rick Derringer	Art Pepper n Zoot Sims	Bob Dylan
7	Pale Saints	Ramones	Nick Cave
8	David Bowie	Beastie Boys	R.E.M.
9	R.E.M.	Joe McPhee Quartet	The Kinks
10	Billie Holiday	Pretenders	Nina Simone

A Tabela 4 ilustra a evolução dos top-10 artistas com maior variância das métricas locais. A primeira conclusão é que as métricas são complementares e não necessariamente a variância em relação a uma, leva à variância em relação às demais. Os artistas que mais variaram em relação à versatilidade não variaram em relação à influência e ponte. Já entre ponte e influência houve uma sobreposição de 5 artistas dentre os 10 que mais variaram. A variância da centralidade para cada um dos artistas na Tabela 4 mostra que as preferências dos usuários que os escutaram oscilaram ao longo do tempo. Em determinados momentos, *Ramones* e *Pretenders*, por exemplo, foram muito influentes, em outros deixaram de ser.

As comunidades na rede de similaridade entre artistas representam os grupos de artistas que são ouvidos pelos mesmos usuários. Pelos *insights* visuais, dois padrões de evolução de comunidades foram detectados na rede. O primeiro corresponde às comunidades que existiram durante todo o tempo de observação (~ 4 anos), ou seja, uma quantidade representativa de nós que faziam parte de uma comunidade no primeiro instante de tempo permaneceu conectada durante os demais instantes. O segundo refere-

Tabela 4. Top-10 artistas com maior variância para cada uma das métricas locais. Artistas que se repetiram em pelo menos duas métricas estão em destaque.

	Ponte	Influência	Versatilidade
1	Art Pepper n Zoot Sims	Irene Dalis	Ronnie Dyson
2	Ramones	Little Jimmy Dickens	Dire Straits
3	Pretenders	Art Pepper n Zoot Sims	Kanye West
4	Radiohead	Radiohead	Roberta Flack
5	Rick Derringer	Beastie Boys	Aphex Twin
6	Pale Saints	Joe McPhee Quartet	Flower Travellin'band
7	Queen	Pretenders	Refused
8	Billie Holiday	Ricky Nelson	Self
9	The Damned	Ramones	Nicki Minaj
10	R.Seiliog	Rick Derringer	Rocket Ship

se às comunidades que surgiram e desapareceram ao longo da evolução. A Figura 3 ilustra a evolução de duas comunidades durante três trimestres. A primeira, cujo no base é *Radiohead*, possui o comportamento persistente durante a evolução da rede, com $com(Radiohead) = 0,76$. A segunda, com *Ricky Nelson* como artista base, apenas se tornou volumosa no instante $t = 15$, sendo $com(RickyNelson) = 0,13$.

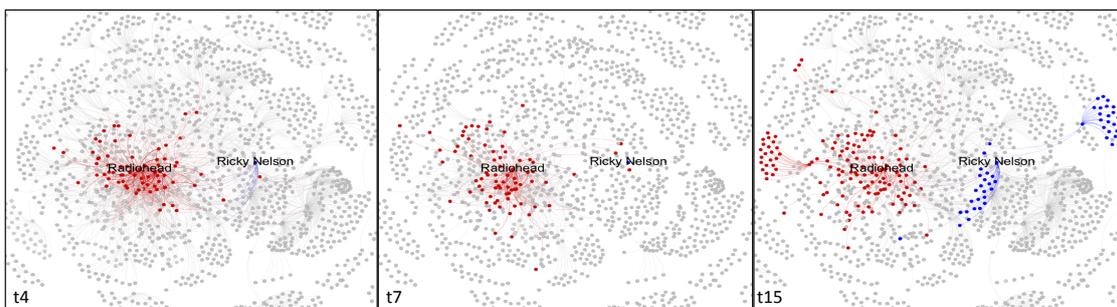


Figura 3. Evolução de duas comunidades a que pertencem *Radiohead* e *Ricky Nelson* em três instantes de tempo diferentes. *Radiohead* pertence a uma comunidade persistente. *Ricky Nelson* pertence a uma comunidade efêmera.

A Figura 4 ilustra a evolução da influência de determinados artistas. É um *heatmap* no qual quanto mais escura a cor, maior a influência. É possível observar que a influência de artistas como *Radiohead* e *Arctic Monkeys* se manteve alta na maior parte do tempo. Entretanto, para os demais artistas, o comportamento detectado é que em geral o pico de influência dura de 2 a 3 trimestres. É uma observação complementar àquela obtida pela análise da evolução das comunidades.

Por fim, o comportamento da variância de todos os nós da rede para cada uma das métricas é mostrado na Figura 5. A média das variâncias tende a zero, ou seja, a maioria dos artistas manteve seus valores de centralidade ao longo do tempo, com baixos índices de variação. A distribuição da quantidade de artistas e suas respectivas variâncias é mostrada nos histogramas. Os top-10 artistas na Tabela 4 estão entre os *outliers* dos *boxplots*. A conclusão é que as preferências dos usuários não variaram em relação à grande maioria dos artistas. Artistas que eram ouvidos por determinados usuários em comum no início, permaneceram sendo ouvidos pelos mesmos usuários em comum.

5.2. Como foi a evolução dos usuários?

A rede de similaridade entre usuários possui características de evolução diferentes da rede de similaridade entre artistas. É uma rede mais densa, porém com menos comunidades,

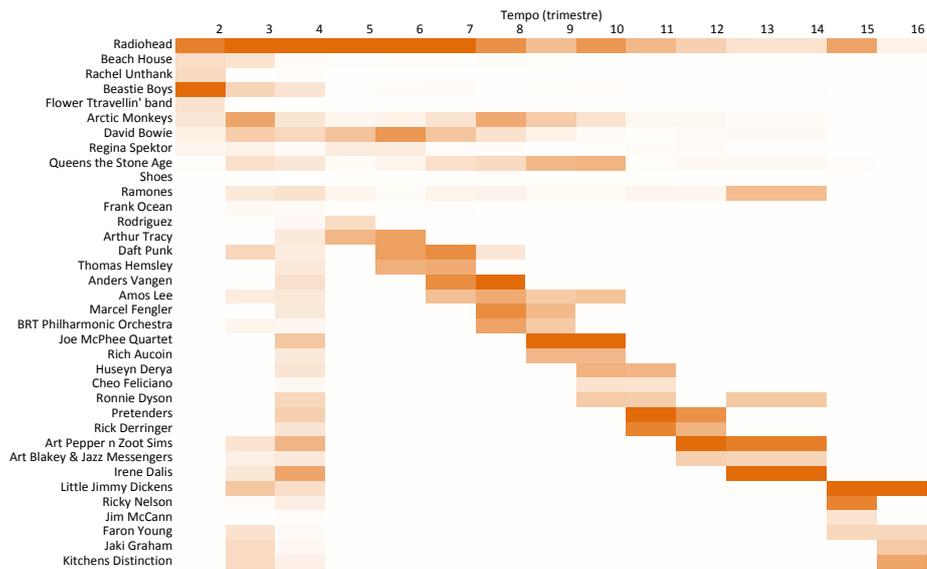


Figura 4. Heatmap baseado na influência dos artistas. Quanto mais escuro, maior a influência. Os artistas da figura são aqueles que em pelo menos um instante de tempo estiveram dentre os 5 mais influentes.

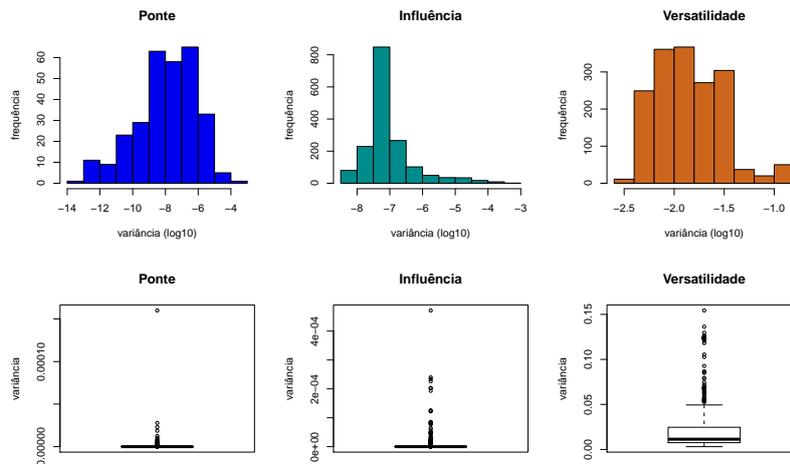


Figura 5. Comportamento da variância dos artistas durante o período observado em relação às métricas de ponte, influência e versatilidade.

sendo estas mais bem definidas de acordo com os *insights* de visualização (Figura 1(b)).

No contexto de usuários, popularidade foi definida em função do número de seguidores que os usuários possuem. Quanto mais seguidores, mais popular é aquele usuário. Observando a Figura 2(b), a rede *usr-usr* é esparsa quanto à interação entre os usuários populares, que, conseqüentemente, não são similares quanto às suas preferências musicais. A conclusão é que usuários populares formam comunidades de usuários similares a eles e que essas comunidades em geral não se misturam.

A evolução consiste em, a cada trimestre, uma comunidade se destaca com grande volume de interações entre os nós pertencentes a ela, com um nó central sempre em destaque quanto às três métricas investigadas. Tal comportamento pode ser observado

pelo *heatmap* na Figura 6, que explora a evolução da métrica de influência para alguns usuários. Os usuários destacados são aqueles que em pelo menos um trimestre estiveram dentre os 5 mais influentes. Diferentemente da rede de artistas, em todos os trimestres, cada um dos 5 usuários mais influentes pertencem a comunidades distintas. A conclusão é que as comunidades nessa rede são efêmeras, ou seja, não existe um grupo de usuários similares que perdura durante todo o período de observação. Além disso, os usuários não estiveram ativos o tempo todo na comunidade. Os picos de interação mostram que a participação na rede social, compartilhando artistas favoritos, é sazonal.

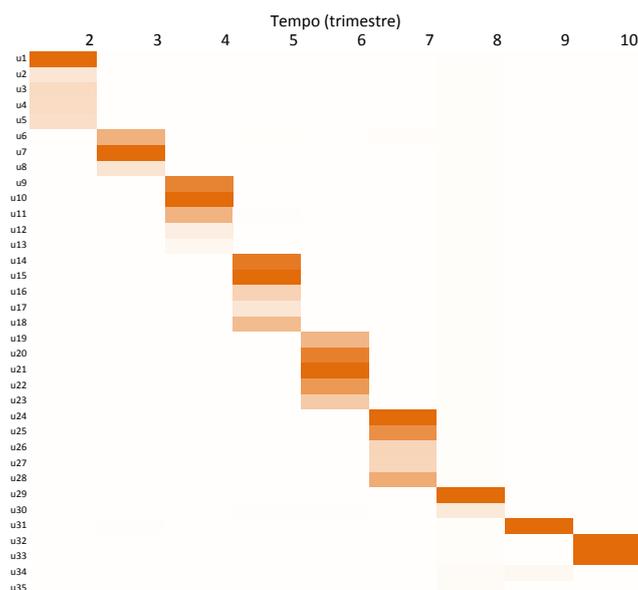


Figura 6. *Heatmap* baseado na influência dos usuários. Quanto mais escuro, maior a influência. Os usuários da figura são aqueles que em pelo menos um trimestre estiveram dentre os 5 mais influentes. Para todos os trimestres, os 5 mais influentes pertencem a comunidades diferentes.

O comportamento da variância entre as centralidades de toda a rede *usr-usr* é mostrado na Figura 7. Assim como na rede de artistas, na rede de usuários, a variância da maioria dos nós não foi significativa, tendendo a zero. Quando os usuários estiveram ativos na comunidade, eles se mantiveram similares. Ou seja, os mesmos usuários escutaram os mesmos artistas em comum durante o período de observação. A conclusão é que não houve variação nas preferências da maioria dos usuários da rede, uma vez que mantiveram suas similaridades.

5.3. Discussão

É importante ressaltar que os resultados obtidos possuem as seguintes limitações: (i) granularidade de tempo escolhida (trimestre). A variação da granularidade pode levar a conclusões diferentes. (ii) O limiar de similaridade (α) escolhido empiricamente. Quanto menor o limiar, mais conexões existirão na rede e vice-versa. Foi conduzida uma variação empírica de 10 valores diferentes para cada rede, na qual todas as redes geradas continuaram com propriedades do mundo real [Zafarani et al. 2014]. Como as análises não investigaram variações menores, focando apenas no comportamento das massas, o limiar não restringe as conclusões obtidas. (iii) Existência de muitos dados nulos, representando falta de atividade dos usuários. É característica da rede TIMJ a existência de usuários

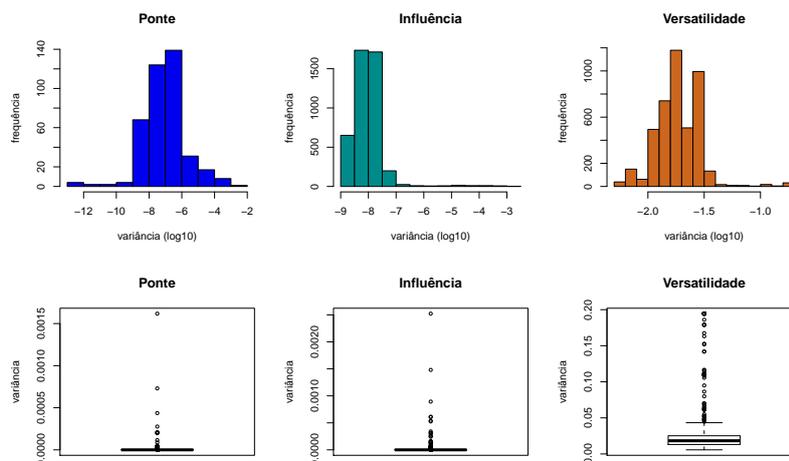


Figura 7. Comportamento da variância dos usuários durante o período observado em relação às métricas de ponte, influência e versatilidade.

sazonais, que não registraram atividades na rede durante todo o período de observação. Tais usuários podem ter mudado suas preferências musicais, porém deixaram de utilizar a TIMJ.

6. Conclusão

Este trabalho é uma contribuição no campo de ciência de dados aplicada: uma análise temporal sobre uma base de dados utilizando técnicas de mineração e análise de redes sociais. Foi conduzida uma análise sobre a evolução das preferências musicais dos usuários de uma rede social de músicas. A proposta foi modelar o domínio através de redes temporais de similaridade. Foi detectado que a maioria dos artistas e usuários semelhantes mantiveram suas similaridades ao longo do tempo, levando à conclusão de que as preferências musicais da maioria dos usuários não mudaram significativamente. Os usuários são ecléticos em relação à diversidade de artistas que apreciam, porém tal diversidade é mantida ao longo do tempo. Como trabalho futuro, pode-se utilizar o modelo de similaridade proposto aplicado à recomendação de músicas. Além disso, usar outras métricas de rede e atributos do domínio como gêneros musicais são extensões naturais deste trabalho.

Agradecimentos

Este trabalho é apoiado pela agências brasileiras CNPq, CAPES e Fapemig. Os autores agradecem à Microsoft Azure pelo financiamento à pesquisa (*research sponsorship* 65c28dfb-a346-455b-a644-c847ff5ac284).

Referências

- Aggarwal, C. and Subbian, K. (2014). Evolutionary network analysis: a survey. *ACM Computing Surveys*, 47(1):10–36.
- Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks.
- Beck, F., Burch, M., Diehl, S., and Weiskopf, D. (2017). A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum*, 36(1):133–159.
- Bonneville-Roussy, A., Rentfrow, P., Xu, M., and Potter, J. (2013). Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood. *J Pers Soc Psychol.*, 4(105):703–717.

- Brandao, M. and Moro, M. (2017). Tie strength analysis: New metrics and open problems. *VI Brazilian Workshop on Social Network Analysis and Mining*, pages 682–687.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25:163–177.
- Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. In *World Wide Web*, pages 107–117.
- Cano, P. (2004). The emergence of complex network patterns in music artist networks. In *Int. Symposium on Music Information Retrieval*, pages 466–469.
- Felício, C. Z., de Almeida, C. M. M., Alves, G., Pereira, F. S. F., Paixão, K. V. R., and de Amo, S. (2016). *29th Canadian Conference on Artificial Intelligence*, chapter Visual Perception Similarities to Improve the Quality of User Cold Start Recommendations.
- Holme, P. (2014). Analyzing temporal networks in social media. *Proceedings of the IEEE*, 102(12):1922–1933.
- Jansson, A., Raffel, C., and Weyde, T. (2015). This is my jam – data dump. *16th Int. Society for Music Information Retrieval Conference*.
- Kapoor, K., Srivastava, N., Srivastava, J., and Schrater, P. (2013). Measuring spontaneous devaluations in user preferences. In *ACM SIGKDD KDD*, pages 1061–1069.
- Linhares, C. D. G., Travençolo, B. A. N., Paiva, J. G. S., and Rocha, L. E. C. (2017). Dynetvis: A system for visualization of dynamic networks. In *SAC*, pages 187–194.
- Macropol, K., Bogdanov, P., Singh, A. K., Petzold, L., and Yan, X. (2013). I act, therefore i judge: Network sentiment dynamics based on user activity change. In *Int. Conf. on Advances in Social Networks Analysis and Mining*, ASONAM '13, pages 396–402.
- Michener, H. A. (2005). *Psicologia Social*. Cengage Learning.
- Moore, J., Chen, S., Turnbull, D., and Joachims, T. (2013). Taste over time: the temporal dynamics of user preferences. In *Int. Society for Music Information Retrieval Conf.*
- Newman, M. E. (2006). Modularity and community structure in networks. *Proc Natl Acad Sci U S A*, 103(23):8577–8582.
- Pereira, F. S. F., de Amo, S., and Gama, J. (2016). On using temporal networks to analyze user preferences dynamics. In *Int. Conf. Discovery Science*, 408–423.
- Siddiqui, Z. F., Tiakas, E., Symeonidis, P., Spiliopoulou, M., and Manolopoulos, Y. (2014). xstreams: Recommending items to users with time-evolving preferences. In *Int. Conf. on Web Intelligence, Mining and Semantics*, pages 22:1–22:12.
- Silva, V., Sampaio, F., and Oliveira, J. (2015). Temporal analysis of co-authorship networks: A study on the interactions of authors in the brazilian journal of computing in education. *Brazilian Journal of Computers in Education (RBIE)*, 23(2).
- Summers, C. and Popp, P. (2015). Temporal music context identification with user listening data. *16th International Society for Music Information Retrieval Conference*.
- Zafarani, R., Abbasi, M. A., and Liu, H. (2014). *Social Media Mining: An Introduction*. Cambridge University Press, New York, NY, USA.