

# Uma Metodologia para a Caracterização de Comportamento Evolutivo em Sistemas de Recomendação\*

Alan Cardoso<sup>1</sup>, Walkiria Resende<sup>1</sup>, Daniel Rocha<sup>1</sup>,  
Fernando Mourão<sup>2</sup>, Leonardo Rocha<sup>1</sup>, Wagner Meira Jr.<sup>2</sup>

<sup>1</sup> Computer Science Department - Federal University of São João Del Rei (UFSJ)  
Pc. Dr. Augusto Chagas Viegas 17 - DCOMP - 36300-088 São João Del Rei, Brazil

<sup>2</sup>Computer Science Department - Federal University of Minas Gerais (UFMG)  
Av. Antônio Carlos 6627 - ICEX - 31270-010 Belo Horizonte, Brazil

{alanc,walkiria,danielrocha,lcrocha}@ufsj.edu.br, {fhmourao,meira}@dcc.ufmg.br

**Resumo.** *Sistemas de Recomendação (SRs) tornaram-se ferramentas de crescente relevância para variadas aplicações comerciais na Web. Apesar de muitos esforços, SRs ainda requerem melhorias para tornar a recomendação mais eficaz e aplicável a vários cenários reais. Estudos recentes apontam a evolução temporal como uma forma primordial de melhorar SRs sem, entretanto, entender em detalhes como essa evolução se manifesta. Assim, propomos uma metodologia de caracterização evolutiva de usuários e aplicações a fim de prover um maior entendimento sobre a dinâmica temporal em SRs. A aplicação da metodologia em um cenário real mostrou-se útil inclusive para auxiliar na escolha de SRs mais aderentes a cada cenário.*

**Abstract.** *Recommender Systems (RSs) have become increasingly important tools for various commercial applications on the Web. Despite numerous efforts, RSs still need improvements to make recommendation more effective and applicable to many real scenarios. Recent studies point out the temporal evolution as a primordial manner for improving RSs without, however, understand in detail how this evolution occurs. Thus, we propose a methodology for evolute characterization of users and applications in order to provide a better understanding of this temporal dynamic in RSs. Applying the methodology in a real scenario has proved to be useful even to help in the choice of RSs adherents of each scenario.*

## 1. Introdução

O grande volume de dados disponível na WEB gerou nos últimos anos um cenário desafiador para variadas aplicações. Usuários possuem mais opções que efetivamente podem manipular [Adomavicius and Tuzhilin 2005]. Diversas aplicações comerciais, tais como Amazon, Last.Fm, dentre outras, provêem uma coleção de itens com milhares de produtos distintos. Embora a disponibilidade de um amplo conjunto de opções tenha sido um cenário desejado no passado, atualmente representa um desafio primordial. Tal mudança deve-se ao fato deste grande volume de opções

---

\*Este trabalho é parcialmente financiado por CNPq, CAPES, FINEP, FAPEMIG e INWeb.

estar “sufocando” os usuários, tornando a simples escolha de um objeto de interesse uma tarefa difícil. Neste contexto, os chamados Sistemas de Recomendação (SRs), que permitem filtrar essa quantidade de informação mostrando apenas o que pode ser de interesse do usuário, assumiram uma posição de crescente importância.

Diversas estratégias para recomendar produtos, informações e serviços aos usuários foram propostas recentemente [Adomavicius and Tuzhilin 2001, Burke 2002, Abbasse and Mirrokni 2007]. A principal ideia de SRs consiste em estimar potenciais itens de interesse aos usuários, baseado em um conhecimento prévio de seu comportamento, bem como características relevantes de itens. Embora a ideia seja simples, sua implementação apresenta diversos desafios computacionais que vão desde como modelar o comportamento dos usuários, a como utilizar as informações dessa modelagem para prover a recomendação. Por exemplo, o comportamento dos usuários pode ser representado por qualquer subconjunto de itens que ele tenha consumido, ou mesmo por itens ainda não consumidos mas que possam ser relevantes para o sistema ou para o usuário, dada uma métrica de interesse. Embora haja inúmeras propostas, SRs atuais ainda requerem melhorias para abordar tais desafios e tornar a recomendação mais eficaz e aplicável a uma gama mais ampla de cenários, tais como recomendação de viagens, serviços financeiros, dentre outros.

Para abordar tais desafios, algumas dimensões de análise vem sendo incorporadas a SRs. Particular atenção vem sendo dada à dimensão temporal [Koren 2009]. Isso porque o comportamento dos usuários muda ao longo do tempo (e.g., seus interesses por produtos mudam), bem como as aplicações (e.g., novos produtos são incorporados, forma de interação evolui, etc.). Como consequência, a modelagem utilizada deve ser continuamente atualizada, a fim de refletir tais mudanças naturais. Nesse sentido, uma atividade indispensável consiste em entender e mensurar a variabilidade associada ao comportamento dos usuários e aplicações, bem como a interação entre eles. Apesar da relevância deste entendimento, não encontramos na literatura trabalhos que objetivem analisar como essa evolução temporal, a qual denominamos comportamento evolutivo, se manifesta em ambientes de recomendação. Algumas questões podem prover informações importantes para SRs, tais como: *Com que frequência usuários costumam consumir um mesmo item?* ou *De quanto em quanto tempo um item já consumido é novamente requisitado?*

Assim, apresentamos nesse trabalho uma metodologia de caracterização evolutiva de usuários e aplicações, a qual é dividida em três etapas principais que representam uma visão hierárquica dos domínios de recomendação, com o objetivo de mensurar um conjunto, não fechado, de características que variam ao longo do tempo e que possam afetar a qualidade de SRs. Tais informações irão prover subsídios para que novas técnicas de SRs sejam propostas, bem como as tradicionais sejam adaptadas. Para validar a metodologia, elegemos o *Last.fm*, uma das maiores comunidades virtuais de música no mundo. Os resultados mostraram que, dentre outras coisas, o *Last.fm* é preponderantemente composto por atividades de novos usuários, o que acarreta um baixo volume de informações acerca de cada usuário. Além disso, trata-se de um ambiente em que grande parte dos itens são consumidos mais de uma vez pelos usuários. Tais observações permitiram auxiliar na identificação de técnicas de recomendação mais adequadas ao cenário avaliado.

## 2. Trabalhos Relacionados

Sistemas de Recomendação (SRs) atualmente desempenham um papel importante em sistemas de comércio eletrônico, auxiliando usuários na descoberta de seus itens e serviços favoritos. Dessa forma, diversos estudos propõem novas estratégias de recomendar produtos, informações e serviços aos usuários em variados domínios [Burke 2002]. Entretanto, desafios diversos tornam a eficácia e aplicabilidade das técnicas atuais insuficientes para diversos cenários [Adomavicius and Tuzhilin 2005]. Alguns desses desafios vêm sendo estudados extensivamente, e métricas que permitem identificá-los e mensurá-los em domínios reais é alvo de constante investigação.

Tais desafios começam pela modelagem dos usuários. Como cada usuário pode ser modelado através de um subconjunto distinto de objetos (e.g., apenas por objetos consumidos por ele, ou por objetos considerados relevantes para o domínio), identificar a melhor modelagem é uma tarefa complexa. Apesar disso, a maior parte dos trabalhos em SRs tratam tal modelagem de maneira simplória, sem se preocupar em identificar características relevantes acerca do comportamento dos usuários. Por exemplo, métricas que quantificam a diversidade de consumo de cada usuário podem prover informações úteis sobre o tamanho do conjunto de objetos adequado para modelar usuários. Um segundo desafio refere-se à esparsidade para modelagem de dados em SRs, estabelecida pela própria natureza das aplicações comerciais. Como o número de objetos distintos nesses domínios é geralmente enorme, usuários estão aptos a consumirem apenas uma pequena parcela dos itens existentes. Além disso, há uma grande concentração dos usuários em torno de poucos objetos seguida de uma procura descendente em torno dos demais, um fenômeno conhecido como *Cauda Pesada* [Anderson 2006], acentuando a esparsidade dos dados. Neste contexto, mensurar o surgimento de novos usuários e itens ao longo do tempo em domínios de recomendação permite identificar melhor o impacto da esparsidade em SRs. Alguns trabalhos, inclusive, propõem técnicas específicas para abordar este problema em SRs [Wu and Li 2008].

Apesar do grande volume de opções existentes, a própria diversidade na recomendação é um desafio. Embora os domínios onde atuam os SRs apresentem uma grande diversidade de itens, as recomendações, em geral, são pouco diversificadas. Trabalhos recentes vêm tentando mensurar a diversidade no consumo dos usuários, bem como prover recomendações mais diversificadas. Em [Zhang and Hurley 2008], por exemplo, os autores tratam a diversidade em recomendação como um problema de otimização. Outro desafio recorrente em SRs é denominado *Cold Start* e refere-se à dificuldade em realizar recomendações sobre itens novos ou para usuários novos, visto que há pouca informação no sistema sobre tais itens e usuários [Schein et al. 2002]. De fato, um dos grandes desafios consiste em prover recomendações precisas quando pouco se sabe sobre um usuário [Adomavicius and Tuzhilin 2005].

Mais recentemente, um novo desafio vem sendo analisado em SRs: a evolução temporal dos dados [Koren 2009]. Tradicionalmente, SRS baseiam-se na premissa que o comportamento passado dos usuários se repete no futuro. Entretanto, tal premissa nem sempre é verdade, visto que dados mudam com o tempo. Por exemplo, novos objetos surgem e opiniões sobre os mesmos objetos variam ao longo do tempo. Assim, a análise destes dados necessita encontrar um balanceamento entre penalizar

efeitos temporais que têm baixo impacto sobre comportamentos futuros, enquanto captura tendências que refletem padrões recorrentes inerentes aos dados.

Os esforços sobre evolução temporal em SRs podem ser classificados em dois grupos. O primeiro compreende trabalhos cujo foco é avaliar a qualidade da recomendação ao longo do tempo. Em [Lathia et al. 2009], o impacto da dinâmica temporal sobre as recomendações é avaliado. Já em [Zhang and Hurley 2008], os autores avaliam como a diversidade da recomendação é afetada ao longo do tempo. No segundo grupo temos trabalhos que propõem novos modelos de recomendação que consideram a evolução temporal. Em [Koren 2009], os autores argumentam que propor modelos de recomendação que levem em consideração o tempo tende a ser mais eficaz que propor modelos complexos. Dessa forma, variações sobre os perfis dos usuários ao longo do tempo vêm sendo incorporados aos SRs [Stern et al. 2009].

Nosso trabalho se difere dos demais por analisar como a dinâmica temporal se manifesta em cenários de recomendação, além de avaliar como algumas métricas relacionadas aos desafios mencionados se comportam ao longo do tempo. Apesar dos vários esforços, não encontramos na literatura trabalhos que objetivem caracterizar e entender a evolução temporal em SRs. Acreditamos que tal entendimento é relevante não somente para propor técnicas que abordem a dinâmica temporal de maneira apropriada, mas também para prover um melhor entendimento sobre diversos desafios de recomendação. Quantificar cada um dos desafios mencionados, bem como a forma como tais desafios evoluem com o tempo, permite identificar quais devem ser priorizados, e conseqüentemente, quais técnicas melhor se adequam a cada domínio.

### 3. Metodologia

Nesta seção apresentamos nossa metodologia de análise do comportamento evolutivo em ambientes de recomendação. De forma a caracterizar dimensões distintas de cada domínio de análise, dividimos tal metodologia em três etapas principais, a saber: **Análise de Contexto do Sistema**, **Análise de Interação** e **Análise de Perfil de Usuários**. Cada etapa possui um conjunto específico, não fechado, de métricas capazes de capturar aspectos pertinentes do domínio, que variam ao longo do tempo e que podem afetar a qualidade da recomendação. A escolha dessas métricas está baseada na correlação das mesmas com os principais desafios em SRs atualmente, tal como apontado por estudos prévios descritos na seção 2. Além disso, novas métricas poderão ser incorporadas a medida que outros aspectos importantes forem identificados.

Cabe ressaltar que, embora as etapas sejam independentes e possam ser aplicadas de forma separada, elas representam uma visão hierárquica dos ambientes de recomendação. Inicialmente, objetivamos, na Análise de Contexto do Sistema, avaliar o ambiente de recomendação de forma a entender como é definida ao longo do tempo a oferta de itens, bem como as regras de negócio estabelecidas pelo domínio de recomendação. Em seguida, avaliamos na Análise de Interação como os usuários interagem com o sistema no decorrer de seu tempo de vida no sistema. Por fim, caracterizamos como os usuários do domínio se comportam quanto ao consumo dos itens disponibilizados, bem como este comportamento se altera ao longo do tempo. Nas seções subseqüentes descreveremos em detalhes os objetivos e principais métri-

cas relacionadas a cada uma das etapas mencionadas de nossa metodologia.

### 3.1. Análise de Contexto do Sistema

O objetivo dessa primeira etapa da metodologia é entender o ambiente avaliado. Identificar características inatas aos objetos e a forma de interação entre objetos e usuários, definidas pelo próprio ambiente, representam aspectos pertinentes para esta análise. Por exemplo, o “consumo” de músicas pode se diferir essencialmente do “consumo” de vídeos. Isso porque, em geral, assume-se que usuários ouvem uma mesma música repetidas vezes com mais frequência que assistem a um mesmo vídeo. Outro aspecto pertinente, seria a distribuição de popularidade dos itens em cada domínio. Enquanto em alguns domínios, tais como de música e vídeos, itens populares são ordens de grandeza mais consumidos que itens pouco populares, em outros, como recomendação de restaurantes, essa diferença não é tão proeminente. Dessa forma, consideramos que características como estas definem um conjunto de “parâmetros” gerais para a recomendação em cada domínio, por serem capazes de informar aos sistemas de recomendação aspectos que os permitem se adaptar a cada domínio. Por exemplo, saber que um mesmo usuário consome frequentemente repetidas vezes um mesmo item pode significar ao recomendador que ele pode fazer uma mesma recomendação, para o mesmo usuário, mais de uma vez.

De forma a realizar essa análise, apresentamos abaixo uma lista de métricas que julgamos relevantes. A escolha de tais métricas foi feita através de um levantamento sistemático de informações que podem ser diretamente utilizadas por sistemas de recomendação. É importante mencionar que, tal como nas demais etapas, não estamos interessados em listar um conjunto fechado de todos os “parâmetros” existentes e possíveis de serem mensurados.

Métrica	Descrição
Distribuição do tempo de vida no sistema (AC-1)	Para cada usuário, determina seu tempo de vida medido em termos de unidades temporais distintas, desde seu cadastro no sistema até o momento da análise. Com isso, verificamos se usuários de um determinado domínio tendem a se manterem no sistema por muito tempo.
Distribuição do número de itens consumidos (AC-2)	Para cada usuário, mensura o número total de itens consumidos durante todo o seu tempo de vida até o momento da análise. Essa distribuição mostra se, em geral, usuários consomem muitos itens.
Distribuição de popularidade dos itens (AC-3)	Para cada item distinto, determina o número de usuários diferentes que o consumiram pelo menos uma vez durante todo o período analisado. Com isso, avaliamos a probabilidade de um item se tornar popular no sistema.
Taxa de surgimento de usuários e itens distintos consumidos (AC-4)	Para cada momento de análise, calculamos o número de itens e usuários distintos presentes no sistema. Dessa forma, é possível identificar a tendência de diversificação do ambiente em termos de usuários e produtos.
Taxa de repetição no consumo de itens (AC-5)	Para cada momento de análise, dividimos o número total de itens consumidos pelo número de itens distintos consumidos por todos os usuários do sistema. Essa informação é particularmente útil ao recomendador pois mensura sua “liberdade” em oferecer repetidas vezes, ao mesmo usuário, um mesmo item.

Tabela 1. Métricas para Análise de Contexto (AC) do Sistema

### 3.2. Análise de Interação

Uma vez identificadas as características do domínio de recomendação, nosso próximo passo é entender como os usuários utilizam e interagem com o sistema. Aspectos tais como com que frequência consomem itens, o intervalo entre consumo, assiduidade de uso do sistema, dentre outras, são informações importantes acerca de como os sistemas são utilizados. A utilidade de tais informações para os SRs pode ser ilustrada pelo uso da informação de intervalo entre consumos ao longo do tempo, tal como definido na tabela 2. Usuário com um intervalo menor entre consumos consecutivos pode impor um exigência maior por diversidade sobre suas recomendações, visto que

Métrica	Descrição
Tempo de uso do sistema (AI-1)	Definida como a quantidade de unidades temporais do tempo de vida de um usuário em que ele efetivamente consome ou acessa itens no sistema.
Intervalo entre acessos ao longo do tempo (AI-2)	O intervalo entre acessos é dado pelo intervalo de tempo médio entre acessos consecutivos de um mesmo usuário. Uma análise destes intervalos ao longo do tempo mostra se usuários em um sistema estão deixando de usar o sistema ou não.
Frequência de consumo por tempo de vida (AI-3)	Calcula o número médio de itens consumidos pelos usuários por cada momento distinto de seus tempos de vida. Esta informação pode ser relevante por avaliar se o sistema possui uma tendência ascendente de consumo. Uma questão relevante intimamente relacionada com esta análise é se usuários "mais velhos" no sistema possuem um consumo mais elevado ou não.
Distribuição da taxa de consumo por acesso (AI-4)	Definida pela divisão entre o número de acessos que um item obtém pelo número de vezes que foi consumido (em cenários onde esta distinção é válida).

**Tabela 2. Métricas para Análise de Interação (AI)**

o consumo de itens muito similares em um curto período de tempo pode incomodar os usuários.

As principais métricas definidas para esta etapa são descritas na tabela 2. Tal conjunto de métricas pode ser ampliado de forma a capturar outras informações relevantes para os recomendadores. Por exemplo, métricas baseadas em *log* de acesso, que determinam o caminho de navegação dos usuário no sistema, ou horário de acesso, dentre outros, são informações potencialmente relevantes.

### 3.3. Análise de Perfil de Usuários

Por fim, nossa metodologia preocupa-se em entender o comportamento dos usuários no sistema com relação aos produtos consumidos, definido como o perfil do usuário. Trata-se, de fato, de uma análise quantitativa sobre o consumo de itens. Tal análise baseia-se em duas principais dimensões: a diversidade de itens consumidos por cada usuário, bem como a temporização de suas ações. Entende-se por temporização a mensuração dos intervalos temporais entre ações de um mesmo usuário no sistema, de forma a compreender padrões comportamentais recorrentes ao longo do tempo. O entendimento sobre esse perfil é essencial para orientar a recomendação de forma mais individualizada. Por exemplo, saber que usuários, ou mesmo um usuário específico, possuem uma diversidade média de  $X$  itens consumidos por semana informa ao recomendador que não deve prover mais de  $X$  itens distintos por semana a cada usuário. A seguir, definimos as principais métricas utilizadas para ambas dimensões da Análise de Perfil de Usuários.

É importante mencionar que na definição de algumas métricas mencionadas, tanto para a análise de Diversidade de Consumo quanto para a de Temporização das ações, utilizamos um conjunto ordenado de itens para cada usuário que é baseado em relevância relativa. Podemos definir a relevância relativa entre itens considerando aspectos distintos, tais como frequência de consumo, similaridade entre itens, probabilidade de alcance em uma rede de "navegação" entre itens, dentre outros. Dessa forma, itens mais relevantes ficam nas primeiras posições do conjunto, enquanto itens pouco relevantes se mantêm nas últimas posições. Além disso, as análises propostas podem ser realizadas considerando-se granularidades temporais distintas (e.g., semanas, meses, semestres, dentre outras), bem como tamanhos de conjuntos distintos. Com isso, é possível definir uma avaliação mais ampla, capaz de contrastar o comportamento evolutivo dos usuários em granularidades temporais distintas, bem como para conjuntos de itens de tamanhos variados.

**Diversidade de consumo:** objetiva avaliar como os usuários se comportam em termos da diversidade dos objetos consumidos. Além disso, variações dessa diversi-

dade ao longo do tempo, bem como sua tendência evolutiva, são foco de investigação nesta etapa da análise.

Métrica	Descrição
Distribuição de diversidade (AD-1)	Para cada usuário, determina o número de itens distintos consumidos durante todo seu tempo de vida no sistema. Com isso é possível saber o perfil de consumo dos usuários com relação a diversificação de itens.
Diversidade média por tempo de vida (AD-2)	Para cada "idade" do usuário no sistema, mensurada através da unidade temporal de análise, verifica-se quantos itens distintos em média os usuários consomem. Assim, verificamos a tendência de diversificação dos usuários no decorrer de seu tempo de vida.
Diversidade de itens em um conjunto ordenado de tamanho $N$ (AD-3)	Para cada momento $X$ de análise, mensuramos a porcentagem de interseção entre os itens do conjunto ordenado de tamanho $N$ de $X$ com os conjuntos de cada período $Y$ posterior a $X$ .
Variabilidade da Relevância dos Itens (AD-4)	Determina a relevância média dos itens consumidos por cada usuário em cada momento distinto a partir da primeira vez que ele consome cada item.

**Tabela 3. Métricas para Análise da Diversidade (AD) de consumo**

**Temporização das ações:** objetiva compreender padrões comportamentais recorrentes ao longo do tempo, tais como o período de tempo no qual os usuários consomem os itens e o tempo necessário para um item deixar de ser consumido ou voltar a ser consumido.

Métrica	Descrição
Distribuição do Período de Estabilidade (AT-1)	Determina o período contínuo de tempo no qual itens se mantêm como os mais relevantes para cada usuário. Uma distribuição destes valores mostra se, em geral, itens se mantêm relevantes por um longo período em um domínio.
Probabilidade de Retorno (AT-2)	Determina a probabilidade de um item mais relevante no momento $X$ ser consumido pelo menos uma vez em cada momento $Y$ posterior a $X$ . Esta medida está relacionada com possibilidade de um item voltar a ser relevante no futuro, dado que é relevante no presente.
Tempo de Saturação (AT-3)	Determina o período médio de tempo necessário para um item deixar de ser consumido, a partir do momento que ele foi consumido pela primeira vez por cada usuário.
Momento de Retorno (AT-4)	Determina o período de tempo médio que um item que deixou de ser consumido pela primeira vez por um usuário voltar a ser consumido.

**Tabela 4. Métricas para Análise da Temporização (AT) das ações**

Na seção seguinte, discutiremos mais a fundo as métricas apresentadas, avaliando-as sobre dados provenientes do *Last.fm*. Apresentamos também as conclusões que podem ser obtidas a partir da aplicação da metodologia e como tais conclusões podem auxiliar na identificação de técnicas de SRs mais apropriadas.

## 4. Estudo de Caso

### 4.1. Descrição da Base de Dados

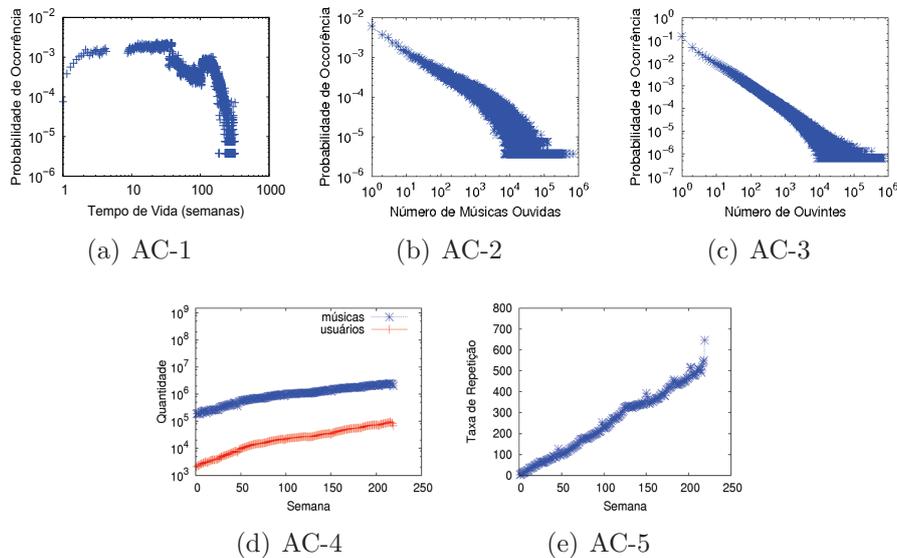
De forma a entender melhor as análises realizadas, apresentamos o conjunto de dados utilizado em nossas análises, provenientes do sistema *Last.Fm*<sup>1</sup>. O *Last.fm* é um dos maiores sistema de rádio e comunidade musical *online* do mundo, criado em 2002 no Reino Unido. Até o momento de aquisição dos dados utilizados, estimava-se que o *Last.fm* possuía mais de 30 milhões de usuários ativos e mais de 7 milhões de faixas musicais distintas em seu banco de dados.

Nossa análise foi realizada sobre uma amostra de dados do *Last.Fm*. Estes dados foram coletados através de uma API fornecida pelo próprio *Last.fm*<sup>2</sup>. Tal API nos permite obter informações relativas às diversas entidades de dados, tais como artistas, álbuns, faixas, usuários, entre outros. Consideramos como relevantes para a nossa análise apenas as informações referentes aos usuários e faixas. Foram coletadas informações para um conjunto de 146.973 usuários distintos e um total de 1.515.258 faixas diferentes observados no *Last.fm* no intervalo de tempo de 12/02/2005 a 26/04/2009.

<sup>1</sup>Disponível em <http://www.last.fm/>

<sup>2</sup><http://www.last.fm/api>

## 4.2. Análise de Contexto do Sistema



**Figura 1. Métricas de Análise de Contexto**

Os gráficos da figura 1 apresentam as métricas descritas para a Análise de Contexto na seção anterior. Iniciamos nossa análise pela distribuição do tempo de vida do usuário no sistema apresentada na figura 1 (a). Podemos observar que grande parte dos usuários coletados possuem entre 10 e 50 semanas de vida no sistema. Períodos muito curtos (menos de 5 semanas) ou muito longo (mais de 150 semanas) possuem uma baixa probabilidade de ocorrência. Isso mostra que embora os usuários explorem o sistema por algum tempo, o *Last.fm* não é capaz de mantê-los assíduos por muito tempo.

Na figura 1 (b) apresentamos o número de músicas ouvidas por cada usuário. Podemos observar que a probabilidade de usuários ouvirem poucas músicas no sistema é muito alta. Além disso, é importante perceber que o gráfico obedece uma lei de potência, enfatizando a grande diferença de probabilidade entre se ouvir poucas e muitas músicas. Além de se manterem por pouco tempo no sistema, os usuários consomem poucos itens, fazendo com que os recomendadores para esse cenário tenham que tratar do problema de *Cold Start* [Schein et al. 2002].

Nossa próxima análise refere-se à distribuição de popularidade das músicas no sistema, tal como mostra a figura 1 (c). Grande parte das músicas em nosso conjunto de dados foram ouvidas por menos de 100 usuários distintos, e apenas um conjunto restrito delas foram ouvidas por muitos usuários. Dessa forma, o *Last.fm* representa um cenário em que pouquíssimos itens conseguem se tornar populares. Conseqüentemente, estratégias de recomendações simples baseadas em popularidade não seriam apropriadas para grande parte dos itens.

No gráfico da figura 1 (d) apresentamos o número de usuários e itens que surgem a cada momento no sistema. Como não temos o conjunto de usuários e músicas que surgem no *Last.fm* globalmente, o que este gráfico mostra é o número de usuários novos a cada semana e o número de músicas distintas que esses usuários, presentes em cada semana, ouviram. Como podemos observar, há um aumento significativo de usuários únicos analisados a cada semana no sistema, bem como um aumento de músicas distintas ouvidas. Tal crescimento na quantidade de faixas e usuários é

ligeiramente ascendente, gerando uma esparsidade crescente de informação para um SR, que, para ser eficaz, deverá ser hábil em lidar com esse problema.

Por fim, temos a análise da taxa de repetição apresentada na figura 1 (e). Podemos observar que o *Last.fm* não só constitui-se como um ambiente com alta taxa de repetição, como também por uma tendência crescente desta taxa. Ou seja, ao longo do tempo os usuários tendem a ouvir cada vez mais as mesmas músicas, apesar da quantidade de músicas novas aumentar ao longo do tempo. Essa observação apresenta duas implicações importantes para os recomendadores. A primeira é que itens já consumidos pelos usuários poderão ser recomendados por diversas vezes. A segunda é que os recomendadores deverão ser robustos quanto ao problema de superespecialização, evitando que as recomendações fiquem sempre ancoradas em itens já conhecidos pelos usuários.

Dessa forma, a análise de contexto nos sugere que o *Last.fm* representa um cenário desafiador para recomendadores, uma vez que, apesar da grande diversidade de usuários e itens, para grande parte dos usuários o histórico de utilização é pequeno, concentrado em poucos itens e com alta taxa de repetição no consumo.

### 4.3. Análise de Interação

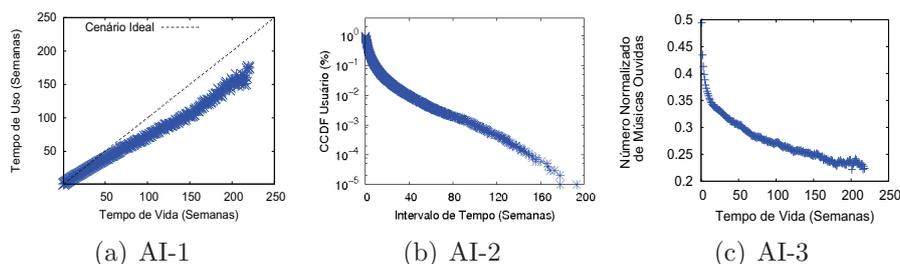


Figura 2. Métricas de Análise de Interação

Iniciamos a Análise de Interação no *Last.fm* observando a figura 2 (a), que apresenta a Taxa de Uso do sistema. Tanto o Tempo de Uso como o Tempo de Vida são medidos em relação ao número de semanas distintas. A curva denominada “Cenário Ideal” representa o cenário em que o tempo de uso é igual ao tempo de vida, ou seja, os usuários acessam o sistema em todas as semanas durante seu tempo de vida. É interessante observar que ao longo do tempo os usuários tendem a usar cada vez menos o sistema, ou seja, a medida que os usuários “envelhecem” no sistema o tempo de uso se distancia mais do tempo de vida. Neste caso, estratégias de recomendação mais incisivas são necessárias, a fim de manter os usuários no sistema.

Nossa próxima análise é sobre o intervalo médio entre acessos no sistema. A figura 2 (b) apresenta a distribuição acumulada complementar desses intervalos (CCDF)<sup>3</sup>. Podemos observar que apenas 0,1% dos usuários possuem um intervalo maior que 10 semanas, ou seja, a grande maioria possui um intervalo médio entre acessos muito pequeno. Além disso, temos que a medida que o intervalo de tempo aumenta, decresce a porcentagem de usuários seguindo dois regimes lineares distintos: em um primeiro momento, até aproximadamente a décima semana, o decaimento é bem acentuado, a partir deste ponto o decaimento torna-se mais suave. Esse resultado, associado ao apresentado pela figura 1 (a), nos mostra que

<sup>3</sup>Do inglês *Complementary Cumulative Distribution Function*.

no *Last.fm* novos usuários acessam o sistema com frequência, entretanto tendem a abandonar o mesmo rapidamente ao invés de ficar ouvindo músicas de maneira esparsa temporalmente. Essa informação é importante para o recomendador, o qual poderá priorizar usuários que possuem um intervalo entre acessos pouco maior que 1 semana, pois estes potencialmente abandonarão o sistema.

A figura 2 (c) mostra o número médio de músicas ouvidas por usuários do sistema ao longo do tempo. Esta média foi calculada sobre os valores normalizados pelo número máximo de músicas ouvidas por cada usuário em uma única semana. É possível perceber um decaimento no número de músicas ouvidas ao longo do tempo, ou seja, em geral, os usuários tendem a usar o sistema com menos frequência. Novamente, a recomendação pode ser uma estratégia relevante para atenuar este decaimento na frequência de uso.

A partir das análises desta etapa conseguimos vislumbrar melhor o relacionamento dos usuários com o *Last.fm*, caracterizando-o como um sistema preponderantemente composto por atividades de novos usuários que apresentam uma tendência descendente de consumo. Tal comportamento pode estar relacionado a algum tipo de limitação do sistema, como recomendações de baixa qualidade ou falta de algumas funcionalidades. Algumas das métricas definidas na seção 3 não puderam ser avaliadas em virtude de não possuímos informações de acesso em nossos dados.

#### 4.4. Análise de Perfil de Usuários

Como descrito na seção 3, dividimos nossa análise de perfil sobre o *Last.fm* em duas etapas: Análise de Diversidade e Análise de Temporização. Nas subseções seguintes descreveremos os principais resultados encontrados em cada etapa. O entendimento deste perfil, juntamente com o entendimento do ambiente e das interações, obtidos anteriormente, compõem um conjunto de informações extremamente rico para SRs.

##### Diversidade de consumo

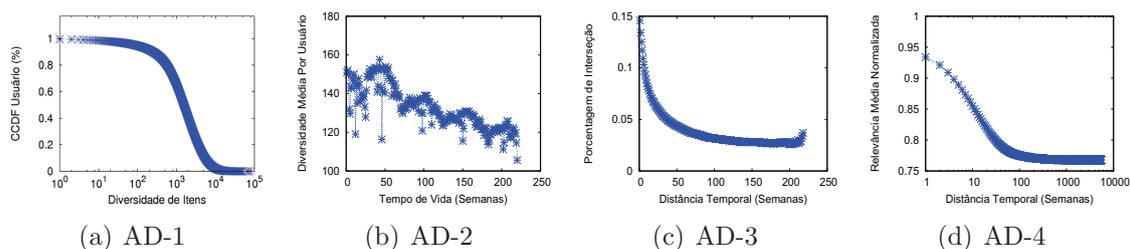


Figura 3. Métricas de Análise de Diversidade

Na figura 3 (a) apresentamos a distribuição de diversidade de itens consumidos pelos usuários durante todo o seu tempo de vida no sistema. Podemos observar que a grande maioria dos usuários (mais de 90%) escutam no máximo 200 músicas distintas, enquanto que poucos usuários possuem uma diversidade de músicas muito elevada (i.e., acima de 5.000). Isso mostra que os usuários no *Last.fm* possuem um comportamento de consumo focado em um pequeno número de itens distintos. Dessa forma, é possível traçar perfis de consumidores baseado em quão diversificados são seus históricos de consumo.

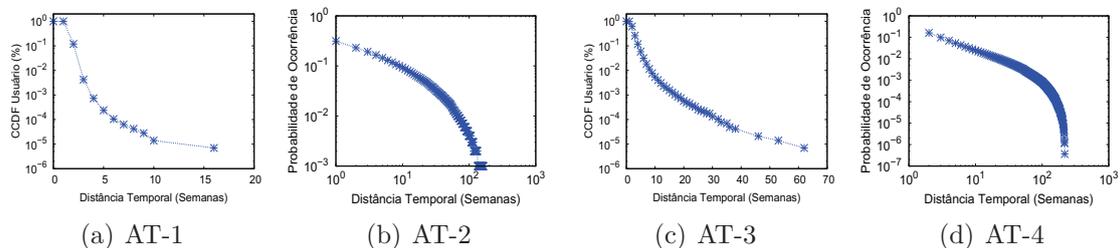
De forma complementar, na figura 3 (b) apresentamos a quantidade média de itens consumidos por usuário ao longo do tempo. Observamos que os usuários ouvem

em média 150 músicas distintas em suas primeiras semanas de vida. Entretanto, a medida que os usuários envelhecem no sistema, a diversidade de itens consumidos reduz significativamente. Esta informação pode ser relevante aos SRs, uma vez que os informa, a cada momento, o limite de diversidade de itens a serem recomendados.

Analizamos também a interseção entre itens de maior relevância consumidos por usuários ao longo do tempo, figura 3 (c). Para isso, definimos a métrica de relevância simplesmente como o número de vezes que cada música foi ouvida em cada semana distinta, e utilizamos as 100 músicas mais relevantes para cada usuário em cada semana. Temos que a porcentagem de interseção é bem reduzida já entre as semanas adjacentes (i.e. menos de 15%). Além disso, essa interseção cai rapidamente ao longo do tempo, obedecendo um lei de potência, chegando à um valor próximo de zero em 54 semanas. Ou seja, no intervalo de um ano o hábito dos usuários muda quase que completamente. Com isso, podemos informar aos recomendadores a porcentagem de músicas atualmente consumidas por cada usuários que pode ser trocada a cada intervalo de tempo distinto (e.g., a cada semana, quinzena ou mês).

Por fim, apresentamos na figura 3 (d) a análise de variabilidade da relevância dos itens consumidos ao longo do tempo. Novamente, definimos a métrica de relevância como a frequência com que cada música foi ouvida em cada semana distinta. A frequência de cada par usuário/música foi normalizada pela maior frequência com que o usuário ouviu em uma única semana cada música. A partir da primeira vez que um usuário ouviu cada música, calculamos sua relevância média para os momentos subsequentes. Por fim, definimos um valor médio de relevância para todos os pares usuário/música observados em cada momento subsequente. Observamos que a relevância média das músicas tende a ser alta em momentos próximos ao momento em que elas foram ouvidas pela primeira vez. Nos momentos posteriores essa relevância reduz de forma acentuada, se estabilizando em valores muito baixos. Isso mostra que as músicas se mantêm interessantes aos usuários por um curto período de tempo, próximo ao momento em que eles as ouviram pela primeira vez. Dado que o *Last.fm* é um ambiente com repetição de consumo, esse decaimento de relevância nos sugere que as recomendações repetidas estão mais propensas a terem sucesso em momentos próximo à primeira vez que cada música é ouvida.

### Temporização das ações



**Figura 4. Métricas de Análise de Temporização**

Primeiramente, avaliamos a distribuição do período de estabilidade no *Last.fm*, figura 4 (a). Para essa análise consideramos as 5 músicas mais ouvidas de cada usuário em cada momento distinto ao longo do tempo. Podemos observar que, em média, quase todas as músicas permanecem por até duas semanas entre as 5 mais ouvidas. Entretanto, já na terceira semana (i.e., distância temporal 2) apenas 2,5% das músicas ainda se mantêm entre as favoritas. Isso mostra claramente que

a preferência dos usuários se mantém estável por pouco tempo, gerando uma certa dinamicidade no conjunto de músicas favoritas que deve ser capturada por SRs.

Nossa próxima avaliação é sobre a probabilidade de retorno das músicas. Assim, apresentamos na figura 4 (b) a probabilidade média de uma música favorita ser ouvida em cada distância temporal a partir da semana em que a mesma se tornou a favorita. Observamos que apenas um terço das músicas que se tornaram favorita são ouvidas novamente na semana seguinte. Além disso, apenas 23% delas são ouvidas novamente após duas semanas. Esse decaimento segue uma lei de potência até a centésima semana. Tais observações sugerem que não seria uma boa estratégia recomendar músicas favoritas após algumas semanas, pois os usuários podem ficar “saturados” rapidamente. Outro ponto interessante é que, a partir da centésima semana a probabilidade de um usuário ouvir sua música favorita é zero, o que mostra que músicas mais velhas podem tornar-se “esquecidas” pelos usuários.

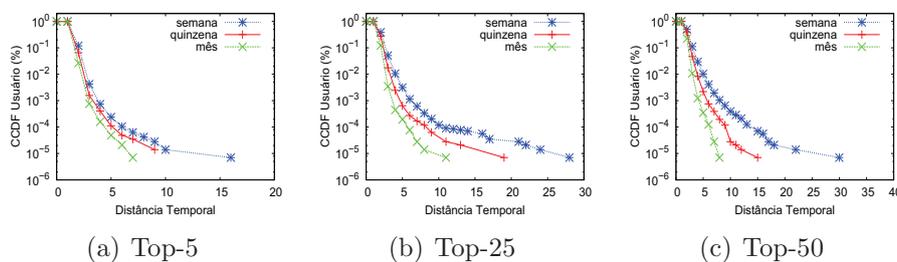
De forma complementar às análises anteriores, apresentamos na figura 4 (c) os resultados referentes à análise do tempo de saturação. Para essa análise mensuramos por quanto tempo, em média, um usuário consegue ouvir músicas que já estiveram, em algum momento, entre as 5 mais ouvidas. Podemos observar que a probabilidade de qualquer música ser ouvida por duas (98%) ou três (62%) semanas consecutivas, após se tornarem uma das 5 mais ouvidas é muito maior do que durante mais de quatro semanas (6% apenas). Isso mostra que, em geral, os usuários estão interessados em ouvir uma música por um curto período de tempo, apresentando posteriormente um interesse significativamente menor para estas músicas.

Também foi avaliado o momento de retorno médio das músicas favoritas de cada usuário. A figura 4 (d) apresenta a probabilidade média das músicas favoritas serem ouvidas novamente para cada distância temporal distinta a partir do momento em que o usuário deixou de ouvi-las pela primeira vez. Observamos que pouco mais de 15% dessas músicas são ouvidas uma semana depois da primeira semana que o usuário deixa de ouvi-las. As probabilidades exibem um comportamento descendente que também segue uma lei de potência para distâncias temporais até 150 semanas, aproximadamente. Isto mostra que, de fato, ao longo do tempo as músicas se tornam “esquecidas”. Dessa forma, temos duas importantes implicações para SRs. A primeira é que uma vez que o usuário deixa de ouvir uma música favorita, dificilmente ele vai retornar a ouvi-la em um futuro próximo. A segunda é que em um futuro distante tais músicas poderiam ser uma boa forma de diversificar as recomendação para os usuários.

Análises desta etapa nos permitem definir usuários do *Last.fm* como detentores de um consumo pouco diversificado e com grande dinamicidade, variando rapidamente o conjunto de músicas favoritas. Além disso, cada música é consumida mais frequentemente em momentos próximos à primeira vez que os usuários a ouviram. Observamos também que tais usuários ouvem uma mesma música por um período contínuo de tempo curto, e uma vez que a deixam de ouvir dificilmente voltam.

Por fim, é importante salientar que várias das métricas analisadas nessa seção podem ser reavaliadas considerando tanto granularidades temporais quanto tamanhos de conjunto ordenados distintos. Com isso, é possível contrastar o comportamento evolutivo considerando as diferentes perspectivas de análise. Nos gráficos da

figura 5, por exemplo, apresentamos uma análise da distribuição de período de estabilidade para variadas granularidades temporais, agrupadas por distintos tamanhos de conjuntos ordenados. Um aspecto interessante a observar é que a medida que aumentamos a granularidade temporal, mais acentuado é o decaimento das probabilidades, mostrando que em períodos de tempo maiores observa-se uma diversidade maior entre os itens consumidos pelos usuários. Podemos observar também que, a medida que aumentamos o tamanho do conjunto analisado, maiores serão as diferenças de probabilidade entre granularidades temporais maiores e menores. Isso mostra que conjuntos maiores de itens favoritos são mais estáveis em granularidades temporais menores. Dessa forma, seria possível definir estratégias de recomendações distintas combinando comportamento dos usuários de distintas granularidades temporais, bem como tamanhos de conjuntos ordenados, de forma a prover recomendações mais precisas.



**Figura 5. Análise da Métrica AT-1 para Distintos Top-N**

De maneira geral, as informações provenientes de cada uma das etapas de nossa metodologia podem auxiliar na escolha de técnicas de recomendações mais adequadas para cada cenário. Para o *Last.fm*, por exemplo, observamos que o problema de *Cold Start* é mais proeminente que o problema de esparsidade. Logo, técnicas como as propostas em [Schein et al. 2002] são preferíveis. Além disso, notamos que uma modelagem baseada apenas nas  $N$  músicas mais ouvidas por um usuário pode ser mais promissor que considerar todas as músicas ouvidas por ele em um dado período, pois a diversidade de músicas é baixa e há uma alta taxa de repetição no sistema. Considerando o desafio de evolução temporal, vimos que granularidades temporais menores são melhores tanto para a modelagem de comportamento dos usuários quanto para atualizar o “conhecimento” dos SRs, visto que usuários no *Last.fm* são altamente dinâmicos, modificando o conjunto de músicas ouvidas rapidamente. Verificamos também que a diversificação na recomendação é um problema maior que a precisão neste domínio pois usuários consomem frequentemente itens que já conhecem. Dessa forma, técnicas como as apresentadas em [Zhang and Hurley 2008] para diversificar a recomendação são especialmente pertinentes para o *Last.fm*.

Cabe salientar a necessidade de validação dessas observações e conclusões. Adotamos como estratégia de validação a implementação de uma nova técnica de SR, a qual incorpora diversas das observações levantadas, e posterior contrastação de resultados entre a técnica proposta e técnicas tradicionais. Tal estratégia está em desenvolvimento atualmente e sua aplicação sobre o *Last.fm* representa nosso próximo passo.

## 5. Conclusão e Trabalhos Futuros

Neste trabalho, apresentamos uma metodologia de caracterização evolutiva de usuários e aplicações a fim de fornecer subsídios para que novas técnicas de recomendação sejam propostas, bem como técnicas tradicionais sejam adaptadas. De forma a verificar a aplicabilidade de nossa metodologia, avaliamos como estudo de caso o sistema de músicas *Last.fm*. Mensurações de um conjunto não fechado de características pertinentes ao *Last.fm* nos permitiram entendê-lo melhor, bem como identificar técnicas de recomendação mais apropriadas para tal domínio. Dentre outras conclusões, verificamos que técnicas específicas para abordar os problemas de *Cold Start* e baixa diversidade nas recomendações são muito relevantes para o *Last.Fm*.

Como trabalhos futuros, temos, primeiramente, a validação das conclusões obtidas. Para tanto, definiremos uma nova técnica de SR, a qual incorporará diversas das observações levantadas para o *Last.fm*. Resultados desta nova técnica serão contrastados com técnicas tradicionais de forma a verificar a relevância dessas novas informações. Posteriormente, temos a análise de outras métricas, bem como a aplicação de nossa metodologia em outros cenários de recomendação.

## Referências

- Abbasse, Z. and Mirrokni, V. (2007). A recommender system based on local random walks and spectral methods. In *Proc. of the 9th WebKDD*, pages 102–108.
- Adomavicius, G. and Tuzhilin, A. (2001). Expert-driven validation of rule-based user models in personalization applications. *DMKD*, 5(1):33–58.
- Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749.
- Anderson, C. (2006). The long tail: How endless choice is creating unlimited demand.
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–470.
- Koren, Y. (2009). Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD*, pages 447–456. ACM New York, NY, USA.
- Lathia, N., Hailes, S., and Capra, L. (2009). Evaluating collaborative filtering over time. In *In ACM SIGIR Workshop on the Future of IR Evaluation*, Boston, USA.
- Schein, A. I., Popescul, A., Ungar, L. H., and Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. In *Proc. of the 25th SIGIR*, pages 253–260. ACM Press.
- Stern, D. H., Herbrich, R., and Graepel, T. (2009). Matchbox: large scale online bayesian recommendations. In *Proc. of the 18th WWW*, pages 111–120, New York, NY, USA. ACM.
- Wu, J. and Li, T. (2008). A modified fuzzy c-means algorithm for collaborative filtering. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*, pages 1–4. ACM.
- Zhang, M. and Hurley, N. (2008). Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the ACM RS*, pages 123–130. ACM.