

UNIVERSIDADE FEDERAL DE ITAJUBÁ - UNIFEI  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
CIÊNCIA E TECNOLOGIA DA COMPUTAÇÃO

RUM++: Uma Abordagem para  
Classificação de Usuários Baseada em  
Mineração de Logs.

**Helton Franco de Sousa**

Itajubá, 27 de março de 2019

**UNIVERSIDADE FEDERAL DE ITAJUBÁ - UNIFEI  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
CIÊNCIA E TECNOLOGIA DA COMPUTAÇÃO**

**Helton Franco de Sousa**

**RUM++: Uma Abordagem para  
Classificação de Usuários Baseada em  
Mineração de Logs.**

Dissertação submetida ao Programa de Pós-Graduação em Ciência e Tecnologia da Computação como parte dos requisitos para obtenção do Título de Mestre em Ciência e Tecnologia da Computação.

**Área de Concentração: Sistemas de Computação**

**Orientador: Prof. Dr. Laércio Augusto Baldochi**

**27 de março de 2019**

**Itajubá**

# Agradecimentos

Em primeiro lugar, agradeço a Deus por iluminar e guiar esta caminhada com saúde e força para superar todas as dificuldades e desafios encontrados até o momento.

Agradeço a todos aos professores que fui aluno durante o curso de Graduação e Pós Graduação da UNIFEI, os quais foram muito importantes na minha vida acadêmica. Agradeço especialmente ao professor Dr. Laércio Augusto Baldochi Júnior pela orientação durante o trabalho de mestrado, compreensão, apoio e por não medir esforços durante esse período. Agradeço também o apoio do professor Dr. Leandro Guarino pelo seu empenho, opiniões e críticas que sempre motivaram a busca de um trabalho de excelência.

Agradeço aos meus pais, Prof. Msc. Ilton Rosa de Sousa e Ronita Ribeiro Franco, por ter dado o dom da vida, educação e sempre me apoiar nas decisões tomadas até chegar aqui. Agradeço minha noiva, Cecília Marcondes Mota e toda a sua família, por estarem ao meu lado sempre e por me apoiarem durante todo esse tempo. Agradeço a todos os meus familiares que de uma forma ou de outra contribuíram para que eu pudesse concluir mais essa etapa. Meu muito obrigado a todos os meus amigos que estão sempre me apoiando e prontos para ajudar no que for preciso. Enfim, a todos que direta ou indiretamente fizeram parte do caminho percorrido até aqui, o meu muito obrigado.

*"O impossível não é um fato: é uma opinião."  
(Mario Sergio Cortella)*

# Resumo

Atividades diárias, como compras e serviços bancários, estão agora disponíveis em quase todos os lugares, o que torna a vida moderna mais cômoda e conveniente. No entanto, nem todos podem usufruir desses benefícios. A baixa alfabetização digital ainda impede que muitos usuários aproveitem ao máximo os serviços on-line. Um grupo que geralmente apresenta dificuldades relacionadas ao acesso à web são os idosos. À medida que as pessoas envelhecem, o controle motor, a acuidade visual e a cognição diminuem, o que faz com que os usuários mais velhos se esforcem mais para realizar tarefas em aplicações Web. Portanto, é importante detectar usuários com dificuldades para apoiá-los, por exemplo, fornecendo interfaces de usuário amigáveis. Para enfrentar esse problema, este trabalho propõe uma abordagem capaz de identificar padrões de uso comumente encontrados entre os idosos, os quais permitem a identificação de usuários com dificuldades enquanto eles navegam em aplicações Web. Logs das interações dos usuários com as interfaces são coletados e trabalhados até se realizar a extração dos padrões de comportamento. Nos experimentos realizados, foi possível obter uma eficiência superior a 84% na classificação desses usuários, mostrando a efetividade da abordagem proposta na identificação dos usuários que apresentam dificuldades no uso de aplicações Web. Assim, usando esta abordagem, os desenvolvedores podem implementar adaptações ou até mesmo interfaces próprias a esses usuários.

**Palavras-chaves:** Log de Clientes, Mineração de Logs, Idade, Experiência Tecnológica, Classificação

# Abstract

Daily activities such as shopping and banking are now available from almost everywhere, which makes modern life more convenient. However, not everyone may benefit from this convenience. Low web literacy still prevents many users to take full advantage from online services. A group that usually presents issues related to web access is the elderly. As people grow older, motor control, visual acuity and cognition decreases, which makes aging users struggle to perform tasks in web applications. Therefore, it is important to detect struggling web users in order to support them, for instance, by providing friendly user interfaces. In order to tackle this problem, this work proposes an approach that is able to identify usage patterns commonly found among the elderly. Our approach allows the identification of struggling users while they browse web applications. Logs of the interactions of the users on the interfaces are collected and worked until the behavior patterns are extracted. In the experiments, it was possible to obtain an efficiency greater than 84 % in the classification of these users, showing the effectiveness of the proposed approach in the identification of users who have difficulties in the use of Web applications. Thus, by using our approach, developers may code adaptations to support these users.

**Key-words:** Client Logs, Log Mining, Age, Technological Experience, Classification

# Lista de ilustrações

Figura 1 – Diagrama ilustrativo da Metodologia. . . . .	16
Figura 2 – Arquitetura da abordagem RUM e tecnologias utilizadas. . . . .	24
Figura 3 – Arquitetura da abordagem RUM. Fonte: Lira et al.[1] . . . . .	27
Figura 4 – Arquitetura da extensão RUM++ e tecnologias usadas. . . . .	30
Figura 5 – Sistema Bancário simulado. . . . .	36
Figura 6 – Árvore de decisão criada pelo algoritmo na classificação dos usuários do Estudo de Caso 1 baseado na experiência. Legenda: <i>duration</i> - Tempo de Duração (valores em milissegundos), <i>exp</i> - Experiente, <i>nexp</i> - Pouco Experiente . . . . .	40
Figura 7 – Plataforma LevelLearn . . . . .	42
Figura 8 – Árvore de decisão criada pelo algoritmo na classificação dos usuários do Estudo de Caso 2 baseado na experiência. Legenda: <i>scroll</i> - Quantidade de Scroll (valor médio), <i>digitação</i> - Quantidade de teclas digitadas (valor médio), <i>exp</i> - Experiente, <i>nexp</i> - Pouco Experiente . . . . .	45

# Lista de tabelas

Tabela 1 – Regra de ouro para interpretar o valor de um coeficiente de correlação [2] : . . . . .	33
Tabela 2 – Média de valores sobre os dados de perfil dos usuários do Estudo de Caso 1 . . . . .	37
Tabela 3 – Coeficientes de Correlação entre as características pessoais dos usuários do Estudo de Caso 1 e atributos coletados dos logs. . . . .	38
Tabela 4 – Média de valores sobre os dados de perfil dos usuários do Estudo de Caso 2 . . . . .	43
Tabela 5 – Coeficientes de Correlação entre as características pessoais dos usuário e atributos coletados dos logs. . . . .	44







# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
1.1	Objetivos	14
1.2	Metodologia	15
1.3	Organização da Dissertação	16
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>17</b>
2.1	Dificuldades de Interação com Dispositivos de Entrada	18
2.2	Diferenças de comportamento na Web	19
2.3	Discussão	20
<b>3</b>	<b>ABORDAGEM RUM</b>	<b>22</b>
3.1	Trabalhos Anteriores	22
3.2	Arquitetura da RUM	23
3.3	Coleta de Logs	25
3.4	Descoberta de Conhecimento	26
3.4.1	Etapas do KDD	26
3.5	Análise de Tarefas	28
3.6	Repositório de Conhecimento	28
3.7	Módulo de Serviços	28
<b>4</b>	<b>EXTENSÃO RUM++</b>	<b>29</b>
4.1	Arquitetura	29
4.1.1	Seleção e Pré-Processamento	29
4.1.2	Transformação	32
4.1.2.1	Atributos Indicados pela Literatura	32
4.1.2.2	Correlação	33
4.1.3	Mineração	33
4.1.4	Interpretação e Avaliação	34
<b>5</b>	<b>AVALIAÇÃO DA EXTENSÃO RUM++</b>	<b>35</b>
<b>5.1</b>	<b>Estudo de Caso 1</b>	<b>35</b>
5.1.1	Descrição do Estudo de Caso 1	35
5.1.2	Execução das Tarefas do Estudo de Caso 1	36
5.1.3	Preparação do Ambiente e Treinamento para as Tarefas do Estudo de Caso 1	37
5.1.4	Resultados do Estudo de Caso 1	37
5.1.4.1	Seleção e Pré-Processamento	38

5.1.4.2	Transformação . . . . .	38
5.1.4.3	Mineração . . . . .	39
5.1.4.4	Interpretação e Avaliação . . . . .	40
<b>5.2</b>	<b>Estudo de Caso 2 . . . . .</b>	<b>40</b>
5.2.1	Coeficiente de Experiência Tecnológica - CET . . . . .	41
5.2.2	Descrição e Execução do Estudo de Caso 2 . . . . .	42
5.2.3	Preparação do Ambiente e Treinamento para o Estudo de Caso 2 . . . . .	43
5.2.4	Resultados do Estudo de Caso 2 . . . . .	43
5.2.4.1	Seleção e Pré-Processamento . . . . .	43
5.2.4.2	Transformação . . . . .	44
5.2.4.3	Mineração . . . . .	44
5.2.4.4	Interpretação e Avaliação . . . . .	45
<b>5.3</b>	<b>Discussão sobre os Estudos de Caso . . . . .</b>	<b>46</b>
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>47</b>
<b>6.1</b>	<b>Trabalhos Futuros . . . . .</b>	<b>48</b>
	<b>APÊNDICES . . . . .</b>	<b>49</b>
	<b>APÊNDICE A – QUESTIONÁRIO UTILIZADO NO ESTUDO DE CASO 1 . . . . .</b>	<b>50</b>
	<b>APÊNDICE B – QUESTIONÁRIO UTILIZADO NO ESTUDO DE CASO 2 . . . . .</b>	<b>51</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>52</b>

# 1 Introdução

Com a evolução da tecnologia e da Internet, é cada dia mais comum que serviços tradicionais sejam oferecidos através de plataformas Web, nas quais é possível, por exemplo, realizar serviços bancários, governamentais, compras e buscas em geral, de modo a facilitar a vida cotidiana. Apesar do crescente aumento no uso da Web e de seus benefícios evidentes, parte significativa da população mais velha ainda sofre com a resistência ao uso dessa tecnologia, o que gera exclusão digital nessa parcela da população [3]. Os principais motivos que explicam a dificuldade que as pessoas mais velhas têm na utilização de aplicações Web estão associados com as mudanças nas capacidades sensoriais e cognitivas relacionadas com o aumento da idade. Esses distúrbios têm sido associados com a baixa velocidade de processamento das informações, maior tempo de reação, falhas na memorização e uma atenção deficiente [4].

Estudos mais recentes revelam que o avanço na idade nem sempre é um fator determinante para explicar a dificuldade de interação com aplicações Web [5, 6]. Experimentos demonstram que pessoas mais jovens, porém com baixa experiência tecnológica, podem apresentar dificuldades de interação até mais severas que usuários idosos com experiência tecnológica. Pessoas que envelheceram em contato com a tecnologia tendem a apresentar menor dificuldade do que outras que não tiveram esse contato [7]. Dessa forma, o avanço da idade, juntamente com a baixa experiência tecnológica, são fatores que limitam o bom desempenho no uso de aplicações Web.

Uma forma de contribuir para reduzir a dificuldade enfrentada por essas pessoas é o desenvolvimento de interfaces que utilizam diretrizes especialmente destinadas a esses usuários [8], tais como: utilizar alvos de interação maiores como links, botões e ícones; evitar barras de rolagem; não utilizar alvos que exigem duplos cliques; promover menos opções de escolha para o usuário; evitar o uso de fontes pequenas, entre outras.

Entretanto, o uso dessas diretrizes na Web nem sempre é possível. Aplicações Web, de modo geral, são desenvolvidas para atender a um grande número de usuários, de diferentes faixas etárias e experiências tecnológicas diversas. Um estudo desenvolvido por [Mbipom e Harper](#)[9] afirma que a acessibilidade está correlacionada com páginas Web consideradas esteticamente satisfatórias apenas em websites que apresentam um projeto de interface limpo, com poucos elementos de interação, o que não ocorre, por exemplo, nas aplicações de e-commerce, as quais geralmente apresentam dezenas de elementos de interação com o objetivo de atrair a atenção do usuário.

Uma abordagem que se mostra mais promissora no sentido de atender as necessidades de diferentes tipos de usuários é o uso de interfaces multicamadas [10, 11]. Essas

interfaces são projetadas para oferecer suporte ao aprendizado, oferecendo camadas de interface simplificadas para usuários novatos ou pessoas com dificuldades nas interações, e camadas mais complexas para usuários com habilidades de interação mais desenvolvidas.

Uma questão relevante para viabilizar o uso de interfaces multi-camada em aplicações web diz respeito à identificação transparente do usuário, de tal forma que a camada de interface adequada lhe seja provida. Para tanto, é necessário monitorar o usuário em busca de padrões de uso que o caracterizam.

Em trabalho anterior foi proposta a abordagem RUM [12] – Real-time Usage Mining – que permite analisar os logs dos usuários durante a navegação, com o objetivo de detectar padrões de comportamento. A abordagem RUM explora técnicas de mineração de uso da web – *Web Usage Mining* (WUM) – para identificar os padrões procurados. No sentido de compreender mais a fundo os detalhes de cada interação, a abordagem utiliza logs de cliente, coletados através do navegador do usuário. Esses logs são mais vantajosos que os logs de servidor, uma vez que fornecem maiores detalhes a respeito das interações dos usuários, reportando ações como clicks, teclas pressionadas, movimento do mouse, entre outras [13].

A abordagem RUM provou ser eficaz para analisar o comportamento do usuário em aplicações Web. Entretanto, para alcançar bons resultados, a abordagem RUM requer que um especialista da aplicação defina padrões de uso e associe esses padrões a perfis de usuário. Este trabalho propõe uma extensão na abordagem RUM de forma a permitir a detecção automática de padrões de uso associados a um perfil específico: usuário idosos e/ou com baixa experiência tecnológica.

## 1.1 Objetivos

O objetivo geral deste trabalho é permitir a detecção de usuários idosos e/ou com baixa experiência tecnológica baseado na análise de seus logs. Para tanto, é necessário minerar os logs em busca de padrões de comportamento característicos desses usuários. Diferentemente do que foi feito em [12], onde um especialista na aplicação web determina quais padrões de uso devem ser associados a cada perfil, o desafio neste trabalho é encontrar os padrões que caracterizam o perfil de usuário desejado.

No sentido de alcançar o objetivo geral, os seguintes objetivos específicos foram definidos:

- i Levantar as dificuldades de interação de usuários idosos e/ou com baixa experiência tecnológica relatados na literatura;
- ii Identificar atributos nos logs de cliente capazes de expressar as dificuldades levantadas em (i) e investigar suas potencialidades;

- iii Utilizando os atributos relevantes obtidos em (ii), definir padrões que caracterizam usuários idosos e/ou com baixa experiência tecnológica;
- iv Desenvolver um mecanismo capaz de detectar os padrões definidos nos logs dos usuários;
- v Validar os padrões definidos com logs de aplicações web reais;

## 1.2 Metodologia

Para alcançar o primeiro objetivo específico, foi realizada uma revisão bibliográfica nas áreas de acessibilidade e envelhecimento, a qual apontou dificuldades de interação por parte dos usuários mais velhos com dispositivos de entrada. Foram investigados também outros trabalhos que encontraram diferenças na navegação e na execução de tarefas por parte de usuários mais velhos.

Com base nas informações levantadas na revisão bibliográfica, foi possível identificar atributos relevantes entre os mais de vinte presentes nos logs de cliente. Outros atributos não associados aos artigos revisados foram também investigados para verificar suas potencialidades na detecção de padrões capazes de indicar limitações nas interações com aplicações web. Dessa forma, o objetivo específico (ii) foi alcançado.

Para alcançar o terceiro objetivo específico, foi necessário analisar os logs de aplicações reais com o intuito de verificar como os atributos escolhidos variam entre os diferentes grupos de usuários considerados. Um estudo de caso com uma aplicação que simula as funcionalidades de um home banking foi realizado para essa finalidade. O estudo de caso permitiu analisar o comportamento dos atributos escolhidos nos logs dos usuários de cada grupo. Essa análise permitiu definir padrões capazes de caracterizar as interações de usuários idosos e/ou com baixa experiência tecnológica.

Para alcançar o quarto objetivo específico foi necessário estender o módulo de mineração de dados da abordagem RUM, implementando as funcionalidades necessárias para detectar os padrões definidos nos logs dos usuários.

Finalmente, para alcançar o quinto objetivo específico, foi realizado um estudo de caso com logs reais de uma aplicação de gamificação aplicada ao ensino. Os resultados desse estudo de caso mostraram que os padrões definidos foram eficientes na detecção dos usuários alvo. Para facilitar a compreensão sobre os passos da metodologia, a Figura 1 descreve suas etapas.

Além da extensão no módulo de mineração, este trabalho também realizou otimizações na abordagem RUM, visando melhorar seu desempenho. Uma otimização relevante foi a reimplementação do mecanismo de pré-processamento e transformação dos dados,

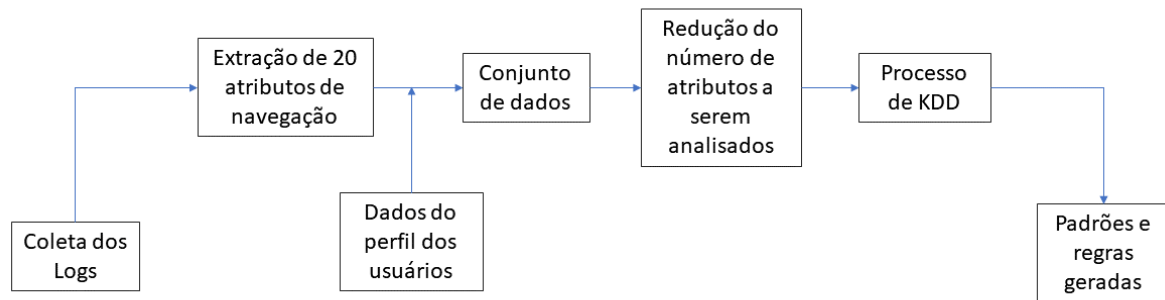


Figura 1 – Diagrama ilustrativo da Metodologia.

que passou a utilizar uma pilha ELK, solução tecnológica que utiliza ferramentas modernas que permitem processar um grande volume de logs de maneira rápida e eficiente. Toda essa extensão da abordagem recebeu o nome de RUM++.

### 1.3 Organização da Dissertação

Esta dissertação está estruturada em 5 capítulos. O Capítulo 2 apresenta uma revisão bibliográfica sobre o comportamento dos usuários mais velhos e/ou com baixa experiência tecnológica na Web. O Capítulo 3 apresenta a abordagem RUM, detalhando o seu funcionamento e sua estrutura e em seguida, no Capítulo 4, é apresentado a extensão RUM++, mostrando suas etapas e implementações realizadas. Neste capítulo é apresentado o passo a passo de todo o processo de extração de conhecimento sobre os logs, desde a etapa de coleta de logs até a extração dos atributos relevantes relacionados ao grupo de usuários. No Capítulo 5 são apresentados os estudos de casos realizados, seus resultados e discussões. Por fim, no Capítulo 6 são realizadas as considerações finais sobre a dissertação.



## 2 Revisão Bibliográfica

O impacto do envelhecimento dos usuários no uso de aplicações web tem sido reportado desde os anos 80 [14, 15, 16]. Desde então, um número vasto de trabalhos vem se esforçando para compreender e prover melhorias com o intuito de ajudar o grupo de adultos mais velhos e usuários idosos a utilizarem a web com mais facilidade e confiança. Parte desses trabalhos da literatura se dedicam a entender como as características dos usuários (idade, escolaridade, fatores cognitivos, questões culturais, habilidades motoras, sensoriais, experiência tecnológica anterior, habilidade e confiança com internet, entre outros fatores) influenciam na execução de tarefas na web [6, 17]. Já outros trabalhos se dedicam a entender os padrões, opiniões e expectativas dos usuários mais velhos em relação ao uso da web, definindo diretrizes e recomendações de design para interfaces [18, 8].

A idade, por si só, não é a única responsável por essas dificuldades apresentadas. Apesar da diminuição dos poderes cognitivos dos adultos mais velhos, muitos deles carregam uma grande experiência no uso de tecnologias web, principalmente aqueles que usam a tecnologia no trabalho. Essa experiência pode ajudar a compensar as deficiências cognitivas relacionadas à idade, como, por exemplo, desenvolvendo heurísticas que orientam suas atividades na web, mesmo sem sua consciência [19].

As dificuldades no uso de tecnologias enfrentadas por esses usuários muitas vezes os impedem de ter livre acesso aos serviços disponíveis na Web, o que os torna dependentes da ajuda de outras pessoas. O estudo conduzido por [Lara et al.](#)[3] revelou que muitos adultos mais velhos têm uma forte dependência do apoio de familiares ou amigos para ajudá-los a usar sites e serviços.

Das diversas definições de faixa etária para idosos e adultos mais velhos, neste trabalho, foi utilizado a definição sugerida pela *Web Accessibility Initiative* (WAI) em *World Wide Web Consortium* [20], segundo o qual jovens são classificados como aquele com faixa etária de 18 a 39 anos, adultos mais velhos com faixa etária entre 40 e 59 anos e idosos aqueles com faixa etária superior e igual a 60 anos. Já em relação à experiência tecnológica não existe uma regra, medida ou fórmula concreta que meça essa característica do usuário. Geralmente, a comunidade acadêmica utiliza o tempo e/ou a frequência de uso de computadores/smartphones como métrica de experiência [21, 5, 17].

Utilizar apenas a frequência de uso como métrica de experiência não é uma boa opção, como observado pelo trabalho de [Chadwick-Dias, McNulty e Tullis](#)[22]. Eles concluíram que apenas a frequência não foi o suficiente para medir a experiência do usuário, sendo necessário utilizar mais parâmetros. Nos seus estudos subsequentes [23], foram uti-

lizados mais atributos - tempo, frequência e a variedade de atividades realizadas na web - divididos em uma escala de 0 a 12, onde cada atributo contribuiu com 0 a 4 pontos.

Em relação às dificuldades enfrentadas pelo público alvo, o trabalho de [Czaja e Sharit\[24\]](#) demonstrou que adultos mais velhos com experiência tecnológica anterior foram mais eficazes ao realizar uma tarefa em relação àqueles com baixa experiência. Observando essas duas características, a pesquisa de [Crabb e Hanson\[5\]](#) descobriu que a idade tem baixa influência na navegação do usuário ao executar tarefas na Web. Portanto, considerar essas duas características (idade e experiência) é importante para encontrar os usuários que apresentam dificuldades de interação com as interfaces. Assim, a partir de agora, esse grupo de usuários será referenciado como público alvo.

Devido aos distúrbios causados pela idade, estudos apontam que mudanças nas características cognitivas dos usuários afetam o uso de dispositivos de entrada de computador como mouse e teclado [25, 26, 27, 28]. Além disso, a visão prejudicada também se torna outro fator que afeta o uso de computadores e da Internet, apresentando perda de foco, dificuldade no uso de telas pequenas e baixa compreensão de mensagens de erro [29, 30, 3].

Desta forma, a seguir são apresentados, trabalhos que investigam dificuldades de interação com dispositivos de entrada (mouse, teclado ou touchscreen) e trabalhos que encontraram diferenças de comportamento entre esses grupos na Web. A partir da análise desses trabalhos, foi possível investigar se os atributos contidos nos logs de cliente continham padrões que demonstravam essas dificuldades.

## 2.1 Dificuldades de Interação com Dispositivos de Entrada

Vários trabalhos vem se esforçando para entender o impacto do avanço da idade e a sua relação na utilização do mouse. O autor [Chaparro et al.\[31\]](#) realizou um experimento com 20 pessoas envolvendo jovens e adultos mais velhos, que executaram tarefas de clique e clique-arrasta. Em geral, os usuários mais velhos apresentaram movimentos menos variáveis e foram mais lentos que os mais jovens para alcançar alvos distantes. De forma geral, os jovens apresentaram desempenho melhor em relação aos idosos tanto para o mouse quanto para o *trackball*, uma espécie de mouse que permite navegar através de movimentos de rolagem.

Essa diferença também foi percebida por [Carvalho, Bessa e Magalhaes\[32\]](#), os quais relataram que os usuários mais velhos costumam gastar mais tempo que os mais jovens ao realizar tarefas envolvendo computadores. Uma peculiaridade percebida pelo autor foi que adultos mais velhos geralmente perdiam o local onde estava o cursor do mouse na tela. Para tentar encontrar o mouse novamente eles geralmente procuravam pelo cursor na tela ou movimentavam o mouse de forma desordenada até encontrá-lo [33].

O uso de *double-clicks* também foi reportado como atividade de difícil execução. Os resultados do trabalho dos autores [Smith, Sharit e Czaja](#)[25], reforçam as evidências de que os declínios relacionados ao envelhecimento impactam o uso do mouse. Os usuários mais velhos cometeram mais erros de cliques duplos, pois, segundo os autores, devido à dificuldade de manter o mouse parado e ao mesmo tempo apertar rapidamente o botão.

O uso massivo de dispositivos móveis nos últimos tempos fez com que grande parte dos usuários da web se aproximasse ainda mais de tecnologias *touchscreen*, melhorando a interação dos usuários mais velhos, segundo [Findlater et al.](#)[34]. Seus experimentos mostraram que os usuários idosos são significativamente mais lentos do que os mais jovens em geral, mas a tecnologia *touchscreen* reduziu essa diferença em relação ao *desktop*. Tamanho da tela e sua orientação também podem contribuir em um melhor desempenho na execução de tarefas em ambientes *touchscreen* [35] e ajudar a reduzir o número de erros cometidos em tarefas, como, por exemplo, envolvendo eventos de arraste [36].

Apesar dessas melhorias de desempenho na interação com *touchscreen*, quando se trata de tarefas envolvendo inserção de dados, os usuários ainda obtêm melhores resultados usando o teclado tradicional, segundo [Carvalho et. al.](#) [37]. No seu experimento envolvendo 60 pessoas (20 crianças, 20 jovens e 20 adultos mais velhos), ele avaliou a eficiência (tempo para completar a tarefa) e efetividade (número de erros) desses 3 grupos usando diferentes modalidades de inserção de dados. Os resultados mostraram que os usuários mais velhos cometeram menos erros quando utilizaram o teclado em vez da tecnologia *touchscreen* e foram 55% mais lentos em relação aos jovens quando realizaram as tarefas usando teclado.

## 2.2 Diferenças de comportamento na Web

As diferenças de comportamento entre usuários mais velhos e mais jovens podem ser percebidas até com o uso de logs de servidores. O trabalho de [Fairweather](#)[6] mostra que usuários mais velhos tem uma tendência a seguir caminhos diferentes e visitar zonas particulares de sites durante a navegação web.

Diferenças também foram encontradas nas estratégias utilizadas nos motores de pesquisa. O trabalho de [Sanchiz, Chevalier e Amadiou](#)[38] mostrou que os adultos mais velhos fizeram o uso de mais palavras-chave na sua pesquisa e ainda passaram mais tempo nas páginas do mecanismo de busca do que os jovens. Já no trabalho de [Czaja e Sharit](#)[24] foi analisada a relação entre o uso anterior da tecnologia por parte do usuário em relação ao desempenho para executar uma tarefa. Adultos mais velhos com experiência tecnológica anterior foram mais eficazes (cometeram uma quantidade menor de erros) em relação aqueles com baixa experiência tecnológica anterior.

O comportamento entre os jovens e adultos mais velhos são diferentes dependendo da complexidade da tarefa realizada, como mostra o autor [Sharit et al.](#)[39]. Em seu estudo

envolvendo 10 (dez) jovens com idades entre 18 e 39 anos e 40 (quarenta) usuários com idades avançadas (20 usuários entre 60 e 70 anos e 20 usuários entre 71 e 85 anos), o autor percebeu diferenças significativas no desempenho (tempo) de tarefas complexas entre os jovens e os adultos mais velhos, enquanto nas tarefas simples, essas diferenças de desempenho não foram percebidas.

Resultados semelhantes também foram encontrados pelos autores [Chevalier, Domes e Marquié](#)[40]. Neste trabalho, jovens e adultos mais velhos foram submetidos à tarefas específicas envolvendo a atividade de pesquisa na Internet, onde o objetivo era encontrar respostas para 3 (três) questões fáceis, 3 (três) questões difíceis e 3 (três) questões impossíveis (sem resposta). Os resultados mostraram que os participantes mais velhos eram menos precisos e usavam estratégias pouco eficientes em comparação com os participantes mais jovens, como por exemplo, uso menor de palavras chaves e tempo maior de pesquisa gastos no Google do que no conteúdo dos sites.

Já [Liao et al.](#)[41] realizou observações sobre a navegação de usuários com idade entre 19 a 87 anos em dois diferentes sites. O autor percebeu que usuários mais velhos procuravam todas as possíveis opções de navegação antes de realizar uma ação, enquanto os jovens costumavam ser mais diretos na sua escolha, optando pela primeira opção de navegação que encontravam. Para ambos os sites que o experimento aconteceu, os idosos navegaram em uma quantidade menor de páginas e tiveram um menor índice de páginas revisitadas.

Alvos pequenos contidos nos sites podem atrapalhar o desempenho desse grupo na sua interação com a Web. [Hwang, Hollinworth e Williams](#)[42] perceberam que aumentar os alvos (elementos clicáveis) dos sites foi capaz de reduzir o tempo de execução de uma tarefa em até 14% e as taxas de erro em até 50%.

## 2.3 Discussão

Fica claro que as questões de interação relacionadas à idade são agravadas quando o usuário possui baixa experiência tecnológica. Portanto, tanto a idade quanto a experiência tecnológica estão associadas às dificuldades de interação com aplicações web.

De modo geral, os trabalhos [31, 32, 25, 33] apontam dificuldades com a utilização do mouse por parte dos adultos mais velhos envolvendo cliques duplos e no uso de *scroll*. Essa conclusão nos levou a investigar atributos dos logs relacionados à essas ações, como, por exemplo, quantidade total de cliques duplos e velocidade média de *scroll*.

Carvalho et. al. [37] apontou lentidão na interação dos dispositivos de entrada e uma quantidade maior de erros cometidos pelos usuários mais velhos. Isso nos levou a investigar atributos relacionados à digitação como, por exemplo, velocidade média de

digitação e quantidade total de erros de preenchimento.

A dificuldade visual e de interação com alvos pequenos [42] fez com que fosse investigado também o uso de *Zoom* nas páginas e a quantidade total de cliques, já que o usuário pode tentar acertar o alvo várias vezes. Por fim, o número total de páginas visitadas também foi analisado, já que os idosos costumavam navegar em um número menor de páginas [41].

Além dos vários atributos fornecidos pela abordagem RUM, novas implementações foram necessárias para coletar os atributos levantados nesta seção.

## 3 Abordagem RUM

Definida como uma abordagem para análise do comportamento do usuário de aplicações Web em tempo real, a RUM [12] explora o uso de mineração de logs de cliente com o objetivo de encontrar padrões de navegação. A RUM coleta, armazena e processa logs de cliente, extraíndo padrões de comportamento e informações sobre a usabilidade da interface. Assim, esta abordagem encapsula o conhecimento e o processamento necessários à análise do comportamento do usuário.

Através da RUM, as aplicações Web podem consumir padrões de comportamento do usuário e informações sobre possíveis problemas de usabilidade em tempo real e, através de ações pré-programadas pelo desenvolvedor, adaptações podem auxiliar o usuário de acordo com suas necessidades e preferências pessoais. Ao contrário das abordagens da literatura, a RUM atribui a responsabilidade da adaptação ao especialista da aplicação Web, fornecendo informações relevantes sobre o comportamento do usuário em tempo real, e minimizando para o desenvolvedor o esforço de implementação necessário para construir uma aplicação Web adaptativa.

### 3.1 Trabalhos Anteriores

A Web apresenta um certo padrão estrutural no qual sites são compostos por uma coleção de páginas que, por sua vez, são compostas de vários elementos como hiperlinks, tabelas, formulários, etc., que geralmente são agrupados por elementos de estrutura HTML. Ao explorar a fundo esse padrão e considerando que os elementos das interfaces são geralmente compartilhados entre várias páginas, [Vasconcelos e Baldochi\[43\]](#) propuseram o COP, um modelo de interface que visa facilitar a padronização destes elementos.

Os principais conceitos no COP são: Contêiner, Objeto e Página. Um objeto é definido como qualquer elemento da página com o qual o usuário possa interagir, como hiperlinks, campos de texto, imagens, botões, etc. Um contêiner é qualquer elemento de página que contenha um ou mais objetos. Finalmente, a página é uma interface que contém um ou mais contêineres.

Além de explorar o fato de que os contêineres e objetos podem aparecer em várias páginas, o modelo COP também explora semelhanças de objetos e contêineres em uma única página. Em qualquer página, um objeto pode ser único (usando seu id) ou similar a outros objetos em termos de formatação (por exemplo, borda, tipo de fonte, cor, etc.) e/ou em termos de conteúdo (textos e imagens). O mesmo se aplica aos contêineres: um contêiner pode ser identificado de uma forma única ou pode ser classificado como

semelhante a outros contêineres, mas apenas em termos de formatação.

O modelo COP foi a base para o desenvolvimento do USABILICS e consequentemente da abordagem RUM. USABILICS é um sistema de avaliação de usabilidade remota orientado por tarefas que avalia a execução de tarefas calculando a similaridade entre a sequência de eventos produzidos pelos usuários e aqueles previamente capturados pelos avaliadores. Usando o USABILICS, os avaliadores podem se beneficiar do modelo COP para definir tarefas genéricas, economizando tempo e esforço para avaliar as tarefas. A abordagem fornecida pela USABILICS é composta por quatro atividades principais: definição de tarefa, registro, análise de tarefas e recomendações.

Uma restrição da ferramenta USABILICS é a incapacidade de definir e analisar tarefas em dispositivos móveis, como smartphones e tablets. Para resolver essa limitação, foi desenvolvido uma ferramenta chamada MOBILICS [44], que estende os principais módulos do USABILICS para suportar a avaliação da usabilidade de aplicativos da web para dispositivos móveis. As ferramentas USABILICS e MOBILICS executam a avaliação de usabilidade após os usuários concluírem suas interações. No entanto, detectar as dificuldades dos usuários durante a navegação é essencial para ajudá-los a atingir as metas desejadas no aplicativo da Web.

## 3.2 Arquitetura da RUM

A RUM foi projetada para coletar, armazenar, processar e analisar logs de cliente gerados em quaisquer aplicações Web, sejam estáticas ou dinâmicas, e fornecer informações sobre o comportamento do usuário em tempo real. Para isso, a arquitetura da RUM está estruturada em cinco módulos, destacados na Figura 2, que são:

- **Coleta de logs:** coleta e armazena as ações dos usuários detectadas na aplicação Web;
- **Análise de tarefas:** realiza a avaliação remota e automática de usabilidade em tempo real baseada em tarefas previamente definidas pelo especialista da aplicação Web;
- **KDD automatizado:** do termo em inglês *Knowledge Discovery in Databases* (descoberta de conhecimento em bases de dados). Detecta padrões de comportamento a partir do histórico de navegação dos usuários da aplicação Web;
- **Repositório de conhecimento:** armazena e gerencia os padrões de comportamento detectados, considerando parâmetros fornecidos pelo especialista da aplicação Web.

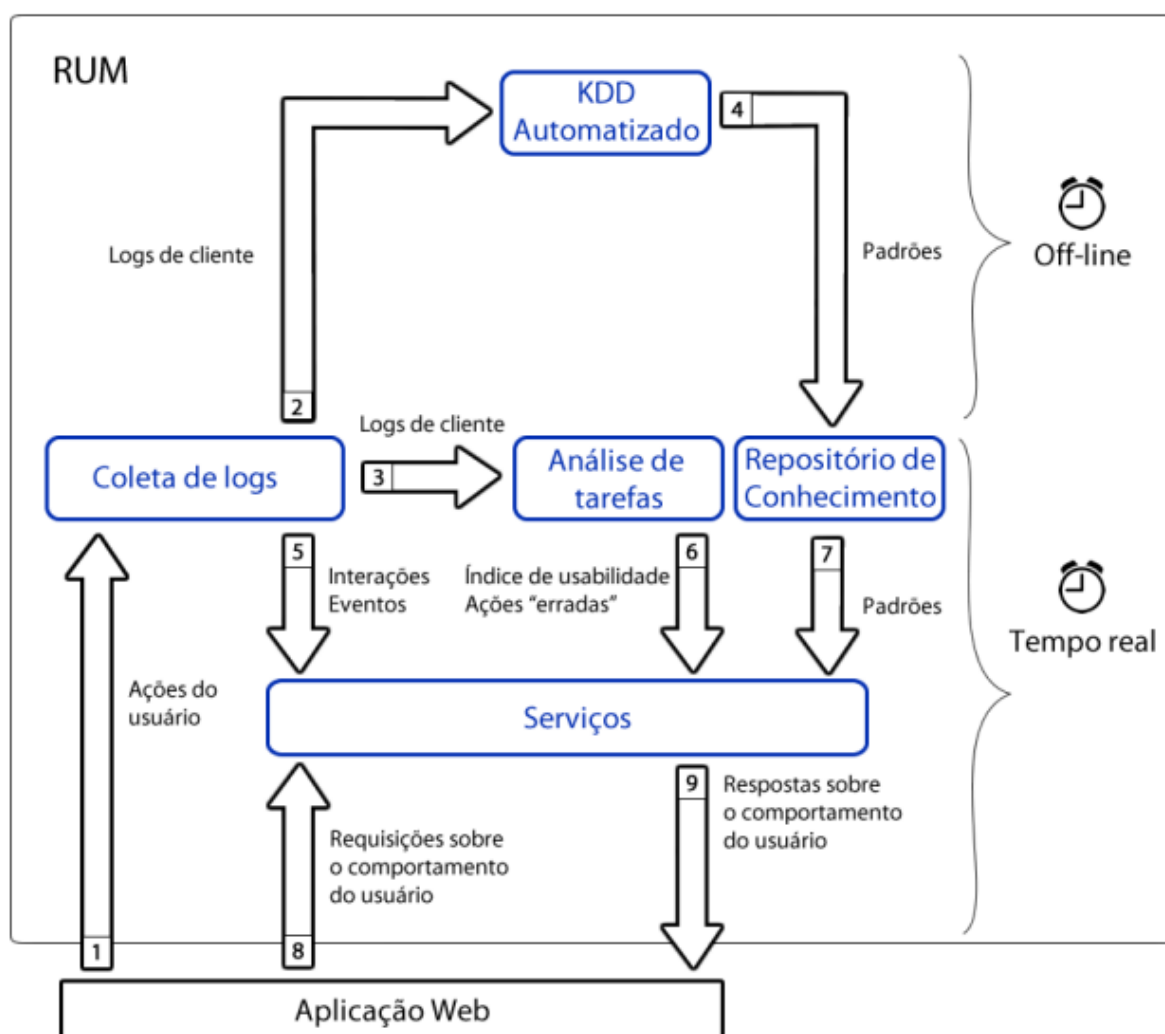


Figura 2 – Arquitetura da abordagem RUM e tecnologias utilizadas.

- **Serviços:** recebe requisições da aplicação Web e fornece informações do comportamento do usuário em tempo real.

A Figura 2 mostra o funcionamento da abordagem RUM onde as setas indicam o fluxo de dados entre os módulos e a numeração indica a seqüência desse fluxo. Inicialmente, o módulo de coleta captura as ações realizadas pelos usuários nas aplicações web (fluxo 1). Em seqüência, as ações detectadas são transformadas e armazenadas em logs para serem analisados pelos módulos de Análise de Tarefas e de KDD Automatizado (fluxo 2 e 3).

O módulo de análise de tarefas realiza a avaliação remota e automática de usabilidade da aplicação Web em tempo real, usando tarefas previamente definidas pelo especialista da aplicação. Periodicamente, como indicado pelo fluxo 2, o módulo de descoberta automatizada de conhecimento consome os logs gerados pelo módulo de coleta



e executa o processo de Descoberta de Conhecimento em Bases de Dados (da sigla em inglês, KDD), detectando padrões de comportamento a partir do histórico dos usuários.

Os padrões de comportamento detectados alimentam o módulo de Repositório de Conhecimento (fluxo 4), que define a relevância de cada padrão, adequando o repositório às mudanças do comportamento dos usuários com o passar do tempo. Em tempo real, através do módulo de Serviços, a aplicação Web requisita informações sobre o comportamento do usuário (fluxo 8) para disparar adaptações pré-programadas na interface. As respostas (fluxo 9) das requisições são: sequências das últimas ações do usuário (fluxo 5), o resultado da avaliação de usabilidade em tempo real (fluxo 6), e os padrões de comportamento executados pelo usuário ativo (fluxo 7).

### 3.3 Coleta de Logs

No módulo de Coleta de Logs, a abordagem RUM faz uso do RUB Toolkit [45], que contém um mecanismo de coleta baseado em log de clientes, desenvolvido em JavaScript. Após a realização de um cadastro de domínio dentro da RUM, é disponibilizado um trecho de código HTML para que seja inserido no cabeçalho da página ou de todo o site. Esse trecho fará a inserção da biblioteca *jUsabilics.js* na aplicação Web, o que dará início à coleta.

A estrutura gerada para identificar o elemento clicado segue o modelo de interface COP [46], que facilita a identificação e comparação dos componentes em que o usuário interagiu. Além do elemento manipulado pelo usuário, diversos atributos são coletados pelo RUB Toolkit:

- *timestamp*: Sequência de caracteres codificados que identificam o exato momento em que determinado evento ocorreu.
- *event*: Tipo de evento realizado (click, mouseover, focusin, load, entre outros) [47].
- *currentURL* e *previousURL*: Link do URL acessado no momento e o anterior.
- *objectGeneratedId* e *objectParentId*: Objeto HTML interagido e o seu respectivo pai, usado pelo modelo COP.
- *browserHeight* e *browserWidth*: Altura e largura da tela do navegador.

Quando ocorre a interação do usuário com a página Web, a biblioteca *jUsabilics.js* captura todas essas informações e as concatena em um único trecho de texto, usando um carácter especial como separador (&), por exemplo:

```
timestamp=1235678&event=click &currentURL=...&previousURL=... &objectGeneratedId=...&objectParentId=... &browserHeight=...&browserWidth=....
```

O RUB Toolkit permite que toda a coleta seja feita de forma automática, não sendo necessário fazer anotações e realizar o experimento ao mesmo tempo, evitando a sobrecarga de tarefas ao realizar o experimento.

O uso de logs de cliente é fundamental na abordagem para que seja possível mensurar de forma precisa e não intrusiva todo o comportamento do usuário, evitando erros de medição e subjetividade. Ao mesmo tempo, possibilita que essa abordagem seja dinâmica, podendo ser replicada e executada em todo e qualquer site, independente do dispositivo utilizado, não sendo necessária a instalação de softwares ou aplicativos.

## 3.4 Descoberta de Conhecimento

KDD (*Knowledge Discovery in Databases*) é um processo genérico para encontrar informação em grandes conjuntos de dados. [Frawley, Piatetsky-Shapiro e Matheus\[48\]](#) definiram como “o processo de extração não-trivial de informações dos dados, informações que estão implicitamente presentes nos dados, previamente desconhecidas e potencialmente úteis para o usuário”.

Para alcançar o propósito da abordagem RUM, o processo de KDD explora os logs de cliente e extrai padrões relevantes sobre o comportamento dos usuários. O objetivo do processo de KDD da RUM nos logs de cliente é caracterizar o usuário. A motivação para isso é que cada aplicação Web possui diferentes usuários com interesses e objetivos específicos. Para isso, os logs de cliente são explorados a fim de detectar padrões sequenciais e regras de associação a partir de atributos que caracterizam o comportamento dos usuários.

Segundo [Fayyad, Piatetsky-Shapiro e Smyth\[49\]](#), o processo de KDD é composto por cinco passos: Seleção, Pré-Processamento, Transformação, Mineração de Dados e Interpretação e Avaliação, que são mostrados na Figura 3.

### 3.4.1 Etapas do KDD

O processo de KDD é iterativo e os passos anteriores podem ser retomados quando necessário. A seguir, há uma breve descrição do que ocorre em cada passo do processo:

#### 1. Seleção

- **Compreensão do domínio da aplicação:** inclui conhecimento relevante e os objetivos da aplicação;
- **Criação de conjunto de dados para descoberta:** inclui selecionar um conjunto de dados ou focar em um subconjunto dos dados disponíveis em que a descoberta será realizada;

#### 2. Pré-Processamento

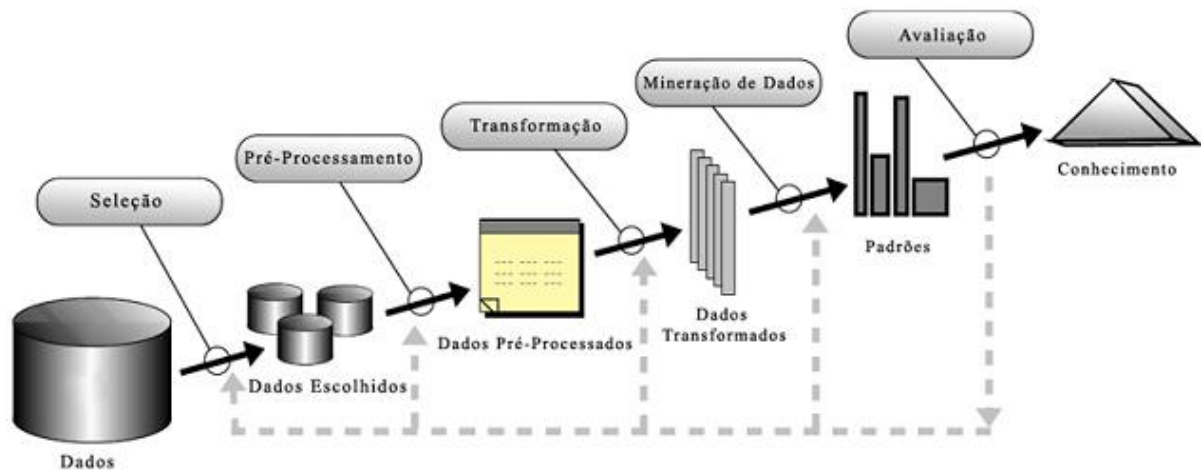


Figura 3 – Arquitetura da abordagem RUM. Fonte: [Lira et al.\[1\]](#)

- **Limpeza e pré-processamento dos dados:** inclui tarefas básicas como remover ruídos ou desvios, se necessário;

### 3. Transformação

- **Redução e reprojção:** inclui encontrar características úteis para representar os dados, dependendo do objetivo da tarefa, e usar método de redução ou transformação de dimensionalidade para reduzir o número efetivo de variáveis que devem ser analisadas;

### 4. Mineração de Dados

- **Escolha da técnica de mineração de dados:** inclui definir a técnica de mineração adequada aos dados, tais como: agrupamento, classificação, regressão, descoberta, regras de associação, entre outras.
- **Escolha dos algoritmos de mineração e de seus parâmetros:** inclui selecionar um algoritmo referente à técnica escolhida e definir seus parâmetros;
- **Mineração de dados:** inclui a busca pelos padrões no conjunto de dados;

### 5. Interpretação e Avaliação

- **Interpretação:** inclui interpretar os padrões detectados e, se necessário, retornar a alguns dos passos anteriores;
- **Consolidação e avaliação:** inclui incorporar os resultados em um sistema ou simplesmente documentá-los.

### 3.5 Análise de Tarefas

O módulo de análise de tarefas da RUM tem como objetivo avaliar a usabilidade de aplicações Web em interfaces para computadores e dispositivos móveis. Na abordagem RUM, é possível realizar essa avaliação em tempo real com o propósito de prover à aplicação Web informações sobre possíveis dificuldades que o usuário enfrenta enquanto navega, de forma que possam ser disparadas adaptações na interface que o auxiliem na interação.

### 3.6 Repositório de Conhecimento

Este módulo é responsável por armazenar de forma estruturada os padrões extraídos do processo KDD automatizado e gerenciar o conhecimento sobre o comportamento do usuário, com a participação do especialista da aplicação Web. Todo padrão armazenado no repositório é priorizado por sua frequência, que é um contador de sua ocorrência. Através de sua frequência, o repositório aprende o comportamento dos usuários e, com o passar do tempo, descarta padrões que deixaram de ser executados.

Os padrões associados pelo especialista a perfis de usuários ou a ações para adaptação são armazenados em um banco de dados específico. Esse banco de dados é consumido pelo módulo de Serviços para comparar, em tempo real, as ações do usuário ativo com os padrões classificados pelo especialista.

### 3.7 Módulo de Serviços

Este módulo é responsável por exportar as informações para as aplicações Web. De modo geral, o RUM fornece dois serviços principais para o desenvolvedor do aplicativo: (i) análise de tarefas e (ii) detecção de padrões de uso. O primeiro explora trabalhos anteriores sobre avaliação de usabilidade [46]. O módulo Análise de Tarefas avalia a execução de tarefas calculando a similaridade entre a sequência de eventos produzidos pelos usuários e aqueles previamente definidos pelo especialista da aplicação. Esse serviço é relevante para detectar usuários que estão com dificuldades para executar tarefas.

A detecção de padrões de uso, por outro lado, depende de um processo de KDD para descobrir padrões de logs. O processo de KDD implementado na RUM encontra padrões sequenciais, ou seja, sequências de ações executadas pelos usuários. A abordagem, em seguida, depende de um especialista da aplicação para corresponder padrões sequenciais a perfis de usuário. Além disso, o especialista pode associar uma determinada ação a ser acionada quando um usuário realiza uma ação associada ao seu perfil. É importante notar que essas ações são implementadas pelo desenvolvedor da aplicação. Assim, o papel do RUM é informar a aplicação web sobre a ocorrência de uma ação.

## 4 Extensão RUM++

O processo original de KDD implementado na RUM não consegue detectar efetivamente a diferença de desempenho entre dois usuários que executam uma mesma tarefa de maneira geral, já que seu mecanismo de detecção de padrão considera apenas a diferença na sequência de ações executadas pelo usuário.

Para entender como os usuários executam uma tarefa, é fundamental analisar outras informações contidas nos logs, como o tempo de execução de cada tarefa, o movimento do mouse, a quantidade de cliques, a rolagem, entre outras. Para classificar os usuários de acordo com os perfis de desempenho, foi realizada extensões no módulo de KDD da abordagem RUM. A extensão da abordagem resultante foi renomeada como RUM ++.

A principal contribuição deste trabalho está dentro da etapa de KDD da abordagem RUM, que tem como finalidade, destacar quais atributos contidos nos logs estão relacionados às dificuldades apresentadas pelos usuários mais velhos e/ou com baixa experiência tecnológica.

Portanto, este capítulo apresenta as extensões realizadas na abordagem RUM a fim de permitir a detecção dos usuários alvos deste trabalho.

### 4.1 Arquitetura

Dos módulos disponíveis na abordagem RUM, todos ainda são necessários para utilizar a extensão RUM++. Pequenas alterações foram realizadas na etapa de Coleta de Logs para capturar variáveis que identificam o autor das ações. A principal mudança aconteceu no módulo de Descoberta de Conhecimento, onde cada etapa de KDD (Seleção e Pré-Processamento, Transformação, Mineração e Interpretação e Avaliação) recebeu novas implementações. A Figura 4 mostra a arquitetura da extensão RUM++, expondo as tecnologias utilizadas em seus módulos, os quais são apresentados à seguir.

#### 4.1.1 Seleção e Pré-Processamento

Inicialmente são removidos os ruídos e informações irrelevantes contidos nos arquivos de logs (requisições de arquivos CSS, JavaScript e imagens). Além dos logs, os dados do perfil de cada usuário também são selecionados, como idade, tempo de uso de computadores e/ou frequência de uso. Após essa etapa, os dados que resumem a navegação são pré-processados e depois são atrelados com os dados de perfil de cada usuário (idade e experiência tecnológica).

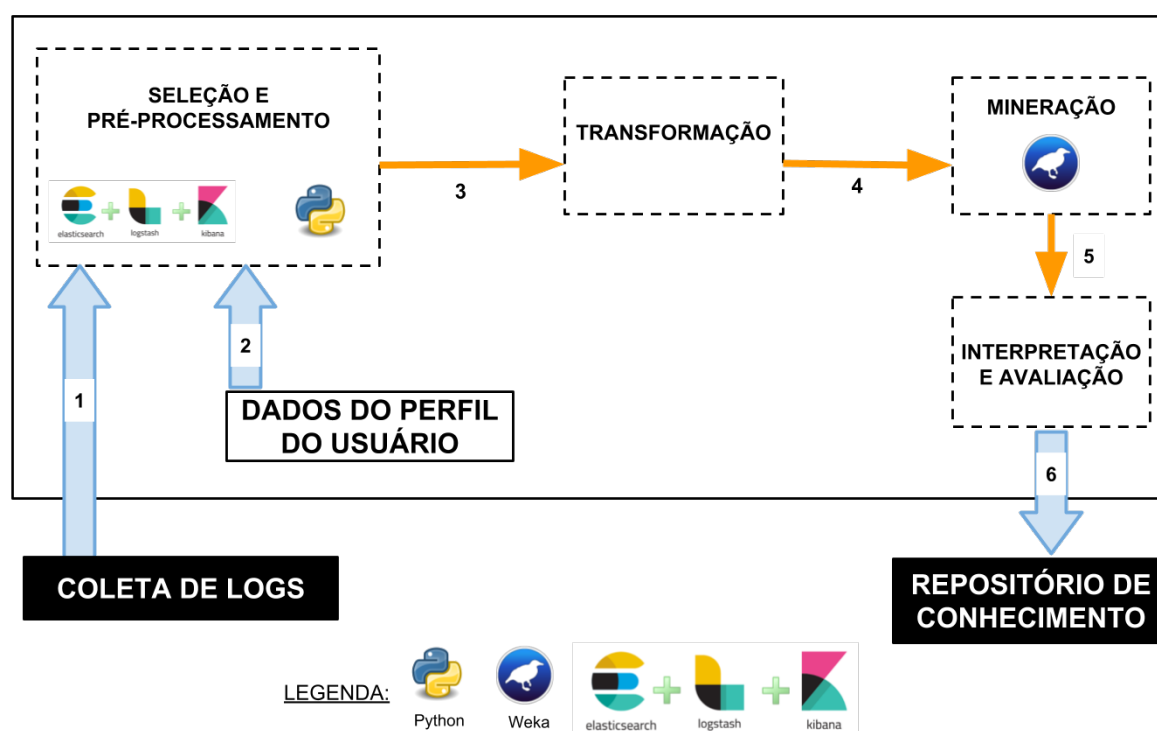


Figura 4 – Arquitetura da extensão RUM++ e tecnologias usadas.

Toda a etapa de Seleção e Pré-processamento é realizada pelo Logstash, uma ferramenta de gerenciamento de logs disponível para uso na Web. Essa ferramenta se torna mais útil quando integrada com outros dois sistemas, ElasticSearch (motor de busca) e Kibana (ferramenta para visualização de dados), formando assim a pilha ELK.

A ferramenta Logstash foi configurada para ler o arquivo de log de 5 em 5 segundos, e ao identificar uma nova linha inserida, pré-processa as informações em atributos resumidos de navegação. Após pré-processados, esses logs são indexados no próprio Logstash e disponibilizados para pesquisa através do ElasticSearch que fornece uma API (*Application Programming Interface*) onde requisições de busca podem ser executadas para recuperar os dados de navegação. Com o Kibana, é possível visualizar os logs e realizar pesquisas na própria interface. Ao final da etapa de pré-processamento, um total de 20 atributos de navegação estão disponíveis para uso, são eles:

1. Quantidade de *clicks*: Durante a navegação, é contabilizado a quantidade de eventos de click;
2. Quantidade de *double clicks*: Durante a navegação, é contabilizado a quantidade de eventos de clicks duplos;
3. Elemento da página que foi clicado: Identifica o tipo de elemento clicado, como por exemplo, botão, link, ícone, texto ou qualquer elemento HTML presente na página;

4. Uso de comandos de atalho: Identifica se o usuário utilizou comandos como selecionar todo o texto, copiar, colar e recortar ;
5. Quantidade de zoom utilizado nas páginas: É contabilizado a quantidade de vezes que foi realizado o aumento ou decréscimo do tamanho das telas durante a navegação;
6. Caminho da aplicação web percorrido pelo usuário: Em que página o usuário começou a navegação, por onde ele passou até onde ele parou;
7. Quantidade de teclas digitadas: Contabiliza a quantidade de dígitos executados durante a navegação e a quantidade em cada um dos campos preenchidos;
8. Sequência de eventos realizadas pelo usuário: Identifica a sequência de páginas e eventos realizados;
9. Quantidade de erros de preenchimento de formulário: Contabiliza a quantidade de erros de preenchimento devido a dados inválidos e campos não preenchidos;
10. Mudança de orientação de tela: Detecta a mudança de orientação no ambiente mobile;
11. Identificar regiões onde usuário movimenta com o mouse: É contabilizado a quantidade de vezes que o mouse passou sobre os elementos da página;
12. Quantidade total de eventos realizados: É contabilizado e somado todos os possíveis eventos disparados pelos usuários como os eventos de mouse, teclado e tela;
13. Duração total da navegação dentro da aplicação web: Diferença de tempo entre o início e fim da última interação do usuário com o sistema;
14. Tempo gasto em cada página: Diferença de tempo entre o fim do carregamento da página até a saída para outra página ou fechamento do browser;
15. Velocidade de rolagem: Contabiliza o tempo médio da velocidade que o usuário realizou o scroll nas páginas acessadas;
16. Quantidade de scroll: Quantidade de eventos do tipo scroll realizados;
17. Velocidade de digitação: Contabiliza o tempo médio entre teclas digitadas durante o preenchimento de campos de entrada;
18. Número de páginas visitadas: Contabiliza a quantidade total de páginas visitadas durante a navegação;

19. Quantidade de eventos do tipo *focus*: O evento *focus* ocorre quando um campo de texto ganha foco do cursor. Este evento é geralmente utilizado em campos do tipo *input*, *select* e *links*;
20. Quantidade de visitas por página: Quantidade de vezes em que cada página recebeu o acesso durante uma seção de navegação;

### 4.1.2 Transformação

Nesta etapa, o objetivo é reduzir a dimensionalidade das variáveis que devem ser analisadas. Segundo [Guyon e Elisseeff\[50\]](#), essa redução pode contribuir para uma maior facilidade na visualização, redução do custo computacional e melhor desempenho por parte dos classificadores. O processo de Transformação passa por duas etapas de redução: a primeira faz uso dos trabalhos da literatura discutidos no Capítulo 2 para escolher os atributos que podem indicar dificuldades de interação. Levantados esses atributos, a segunda etapa faz uso do artifício de Correlação, quando são destacados os que mais estão correlacionados com a idade e/ou experiência tecnológica dos usuários.

#### 4.1.2.1 Atributos Indicados pela Literatura

Relacionado ao movimento do mouse, foram escolhidos a velocidade média de scroll, quantidade de duplo cliques e o número total de eventos realizados, já que os trabalhos de [Chaparro et al.](#); [Carvalho, Bessa e Magalhaes](#); [Smith, Sharit e Czaja](#); [Hollinworth e Hwang\[31, 32, 25, 33\]](#) apontam dificuldades relacionadas a essas atividades.

Já em relação ao uso de teclado, foi investigado a velocidade média de digitação dos usuários, quantidade de eventos do tipo *focus* (quantas vezes o usuário selecionou o campo de entrada), tempo total para realizar uma tarefa, número total de teclas digitadas e a quantidade de erros de preenchimento, pois o trabalho de [Carvalho et. al. \[37\]](#) destaca estas atividades.

A dificuldade de interação com alvos pequenos citada por [Hwang, Hollinworth e Williams\[42\]](#) pode ocasionar o uso de mais cliques, tentando acertar o alvo, ou o uso de outros artifícios como o de aumentar a tela (Zoom). Portanto, optou-se por investigar a quantidade de zoom e cliques realizados. [Liao et al.\[41\]](#) reportou que os adultos mais velhos costumam visitar um número maior de páginas, então, foi selecionado também o atributo quantidade total de páginas visitadas para ser analisado.

Como resultado, um grupo de 12 atributos foram selecionados como candidatos para mineração - 1, 2, 5, 7, 9, 12, 13, 15, 16, 17, 18 e 19. Para certificar a relevância desses atributos, foi utilizado uma técnica de correlação para garantir que a seleção dos atributos estivesse correta e contribuir possivelmente na redução da dimensionalidade da análise.



Tabela 1 – Regra de ouro para interpretar o valor de um coeficiente de correlação [2] :

Coeficiente da Correlação	Interpretação
0.8 a 1 (-0.8 a -1)	Muito Forte
0.60 a 0.79 (-0.60 a -0.79)	Forte
0.4 a 0.59 (-0.4 a -0.59)	Moderada
0.2 a 0.39 (-0.2 a -0.39)	Fraca
0 a 0.19 (0 a -0.19)	Muito Fraca

#### 4.1.2.2 Correlação

O objetivo nessa etapa foi investigar se seria possível reduzir ainda mais a dimensão da análise, escolhendo os atributos mais relacionados com a idade e/ou experiência tecnológica. Portanto, foi aplicado a Correlação de Pearson com o objetivo de identificar quais atributos, dos 12 disponíveis, são mais relevantes em relação à idade e experiência tecnológica dos usuários. O resultado da correlação é comparado com a Tabela 1, a qual são destacados como atributos relevantes aqueles que apresentam coeficientes de correlação do tipo Moderado, Forte e Muito Forte.

Dos diversos métodos de correlação existentes, a escolha na utilização da Correlação de Pearson se ao deu fato de outros trabalhos relacionados na área já utilizarem essa técnica [23, 51, 4, 6] e que segundo o autor David [52], a Correlação de Pearson é uma boa escolha para casos onde o número de amostras são maiores que 25, já que as amostras dos estudos de caso eram de 32 e 44 para o estudo de caso 1 e 2 respectivamente.

É importante ressaltar que o resultado final da etapa de correlação pode variar. Devido à diversidade de aplicações web existentes e ações executadas dentro deles, pode ocorrer casos onde certos atributos não sejam excitados ou até mesmo explorados o suficiente, deixando assim, de carregar as diferenças de comportamento entre os grupos de usuários analisados.

Ao final desta etapa, o conjunto de dados é transformado em um formato de arquivo compatível com a entrada do algoritmo que será utilizado na mineração.

#### 4.1.3 Mineração

O arquivo gerado no processo de Transformação serve como entrada de dados para o software WEKA [53], o qual deve realizar o processo de classificação. Neste trabalho, foi utilizado o algoritmo *RepTree* (*Reduces Error Pruning Tree Classifier*) para classificação pelo fato de ser um algoritmo de aprendizado de árvore de decisão rápida [54], além da possibilidade de utilizá-lo com atributos numéricos. Ao final da execução, o algoritmo *RepTree* constrói uma árvore de decisão usando informação de ganho e variância e poda a árvore usando poda de erro reduzido, de forma a simplificá-la, removendo seções da árvore que tem pouco poder para classificar as instâncias.

Das opções disponíveis de treinamento do classificador, foi escolhido a de *Cross-validation* [55] com valor padrão (Número de amostras - 1). Essa técnica permite avaliar a capacidade de generalização do modelo juntamente com sua base de dados. O resultado da classificação é repassado para etapa seguinte de Interpretação e Avaliação.

#### 4.1.4 Interpretação e Avaliação

Apesar da etapa de mineração oferecer recursos poderosos, ainda existe um enorme desafio em relação à precisão e à consistência nos resultados gerados pelas técnicas de mineração. O objetivo dessa etapa é interpretar e analisar os resultados da classificação e assim decidir se são relevantes.

Ao final da etapa de mineração, o algoritmo de classificação expõe seu grau de eficiência em uma escala de 0 a 100%. É papel do especialista da aplicação Web analisar esse resultado baseado no experimento, e caso o resultado seja eficaz, as regras geradas na etapa de mineração são inseridas ou atualizadas no repositório de conhecimento.

Vale ressaltar que todo o processo apresentado, da Seleção até a Interpretação dos Resultados, é interativo e adaptativo, ou seja, pode ser executado várias vezes e a cada vez, o sistema é capaz de criar novas regras aprimoradas para serem atualizadas no repositório de conhecimento.

## 5 Avaliação da Extensão RUM++

Neste trabalho, foram realizados dois estudos de caso com o intuito de avaliar a extensão RUM++. O primeiro (Estudo de Caso 1) foi realizado em um sistema de *home banking* simulado e o segundo (Estudo de Caso 2), em um sistema real de gamificação utilizado por alunos de uma faculdade. No início deste trabalho, foram realizadas tentativas de executar os experimentos envolvendo também o grupo de idosos (mais de 60 anos), porém houve uma grande dificuldade de recrutamento desses usuários. Além disso, a baixa alfabetização tecnológica desse grupo dificulta a realização das tarefas.

### 5.1 Estudo de Caso 1

Para o Estudo de Caso 1, um total de 32 participantes foram recrutados e separados em dois grupos de acordo com a faixa etária: jovem de 18 a 39 anos (grupo 1A) e de adultos mais velhos, de 40 a 59 anos (grupo 1B). 14 usuários foram recrutados para o grupo 1A, e 18 usuários, para o grupo 1B. É importante ressaltar que nenhum usuário havia tido contato prévio com o sistema web utilizado no estudo de Caso.

Antes de realizar as tarefas, os participantes responderam um questionário (Apêndice A) com o objetivo de coletar informações sobre o perfil e o conhecimento prévio sobre suas experiências tecnológicas.

#### 5.1.1 Descrição do Estudo de Caso 1

Neste estudo de caso, foi utilizado uma aplicação Web baseada em serviços de *internet banking*. Para o desenvolvimento da aplicação, foi feito um levantamento prévio nos serviços de *home banking* do Brasil e sobre a disposição das informações nas suas interfaces, principais tarefas disponíveis aos usuários e o layout, buscando criar um ambiente mais próximo da realidade.

Realizado todo esse levantamento, replicou-se essas características no sistema bancário simulado, como mostra a Figura 5. Das diversas tarefas disponíveis, foram exploradas as seguintes neste estudo:

- Realizar transferência entre contas: Ação utilizada para mover uma quantia de valor de uma conta origem para uma conta destino. Em relação à origem, o usuário deve selecionar o tipo de conta da qual será retirado o dinheiro, se é conta corrente ou conta poupança. Com relação ao destino, o usuário deve informar o valor, número da agência e número da conta.

Figura 5 – Sistema Bancário simulado.

- Pagar boleto bancário: Ação comumente utilizada no Brasil, na qual o usuário insere uma sequência de caracteres que carrega informações de pagamento, como data de vencimento, número do banco e valor a ser pago. Ao inserir essa sequência de caracteres, o sistema automaticamente verifica sua veracidade e processa o pagamento.
- Visualizar extrato bancário: Essa funcionalidade relata o histórico de transações da conta do usuário, sendo necessário informar uma data de início e uma data final para realizar o extrato das transações. É necessário também selecionar qual tipo de conta ele deseja consultar, se é conta corrente ou conta poupança.

Cada usuário participante foi cadastrado no sistema bancário e poderia acessá-lo através de login e senha. Para permitir que os usuários cumprissem as tarefas propostas, cada usuário iniciou sua conta bancária com saldo de \$10.000,00, tanto em conta corrente quanto em conta poupança.

### 5.1.2 Execução das Tarefas do Estudo de Caso 1

Para executar as tarefas do Estudo de Caso 1 por completo, o usuário deveria realizar 7 tarefas. Usuários que não conseguiram realizar todas as tarefas, ou pularam alguma delas, foram retirados da avaliação, para que não houvesse interferência na análise e correlação dos resultados. As seguintes tarefas foram definidas:

1. Fazer uma transferência de R\$ 1.675,89 da conta corrente do usuário para a conta-corrente do cliente de agência: 49083 e conta corrente: 107396;
2. Fazer uma transferência de R\$ 831,10 da conta poupança do usuário para a conta-corrente do cliente de agência: 49083 e conta corrente: 107396;
3. Fazer pagamento do boleto de número: 68876 24201 99836 00000008261 144277 6 39264019452719 para a data de 17/05/2017 no valor de R\$ 742,21;

4. Fazer uma transferência de R\$ 3.905,63 da conta conta corrente do usuário para a conta poupança do cliente de agência: 49083 e conta corrente: 107396;
5. Fazer uma transferência de R\$ 3.311,10 da conta poupança do usuário para a conta poupança do cliente de agência: 49083 e conta corrente: 107396;
6. Visualizar o extrato de conta corrente entre os dias 01/05/2017 e 17/05/2017 e, em seguida, salvá-lo em PDF;
7. Visualizar o extrato de conta poupança entre os dias 21/05/2017 e 05/06/2017 e, em seguida, salvá-lo em PDF;

### 5.1.3 Preparação do Ambiente e Treinamento para as Tarefas do Estudo de Caso 1

Para reduzir os ruídos, todos os usuários usaram um computador contendo mouse e teclado padrão, assim, todos realizaram as tarefas no mesmo ambiente computacional. Antes de executar as tarefas, os usuários foram instruídos sobre o sistema web a ser utilizado e as tarefas a serem realizadas, para se familiarizarem com a aplicação.

Após essa etapa de treinamento e retenção de conhecimento, os usuários foram autorizados a realizar a iniciar. Ao final da última tarefa, os participantes foram instruídos a fechar o navegador para que eventos extras não interferissem nas análises futuras.

### 5.1.4 Resultados do Estudo de Caso 1

No primeiro estudo de Caso, 32 participantes realizaram as tarefas, sendo que 14 eram usuários jovens e 18 adultos. Quando a experiência tecnológica é analisada, considerou-se os usuários que tem entre 0 a 4 anos de uso com computadores como usuários pouco experientes. Aqueles que tinham mais de 5 anos de uso, foram considerados usuários experientes. Dos 32 participantes, 22 foram considerados usuários experientes e 10 com pouca experiência.

Em relação ao público do estudo de caso, a média de idade entre os jovens que realizaram a tarefa foi de 21,6 anos. A média de idade entre os adultos mais velhos foi de 51,7 anos. O atributo tempo de experiência também foi analisado. O grupo de jovens obteve uma média de 9,6 anos de experiência com computadores e os adultos, 5,7 anos de experiência, como mostra Tabela 2.

Tabela 2 – Média de valores sobre os dados de perfil dos usuários do Estudo de Caso 1

	Jovens	Adultos
Idade (anos)	21,6	51,7
Tempo de Experiência (anos)	9,6	5,7

#### 5.1.4.1 Seleção e Pré-Processamento

Do total de 32 participantes, 26 conseguiram completar a tarefa com sucesso. Apenas 4 adultos não conseguiram completar as atividades porque falharam ao encontrar o caminho correto para realizar uma tarefa ou fecharam o navegador antes do fim, devido a tentativas exaustivas de preencher corretamente os campos. As tarefas de 1 jovem também foram descartadas pois o participante demorou um longo tempo para fechar o navegador após realizar as tarefas, gerando ações extras gravadas nos logs que inviabilizaram a análise.

Após realizar essa seleção, os dados de navegação dos usuários que completaram a tarefa com sucesso foram submetidos à etapa de Pré-Processamento, a partir do qual foram obtidos os 20 atributos de navegação. Os atributos de navegação foram cruzados com os dados do perfil dos usuários e avançou-se para a etapa de Transformação.

#### 5.1.4.2 Transformação

Utilizando o conhecimento extraído da literatura, realizou-se a redução de dimensionalidade que resultou em 12 atributos a serem analisados. O método da correlação foi então aplicado para tentar reduzir ainda mais a dimensionalidade da análise e entender quais atributos poderiam ser mais relevantes em relação à idade e à experiência tecnológica. O resultado está disposto na Tabela 3.

Tabela 3 – Coeficientes de Correlação entre as características pessoais dos usuários do Estudo de Caso 1 e atributos coletados dos logs.

ID	Descrição	Idade	Experiência
1	Quantidade de <i>Click</i>	0,019	-0,219
2	Quantidade de <i>Double Click</i>	0,372	0,074
5	Quantidade de Zoom	0,0	0,0
7	Quantidade de teclas digitadas	0,679	-0,569
9	Erros de Preenchimento	0,0	0,0
12	Quantidade Total de Eventos	0,383	-0,527
13	Tempo de Duração	0,607	-0,804
15	Velocidade Média entre <i>Scroll</i>	0,274	-0,546
16	Quantidade de <i>Scroll</i> Realizados	0,379	-0,357
17	Velocidade Média de Digitação	0,506	-0,716
18	Quantidade de Páginas Visitadas	0,063	-0,110
19	Quantidade de <i>Focus</i>	-0,423	0,001

Para definir quais atributos foram mais relevantes, utilizou-se a Tabela 1 apresentada na Seção 4.1.2.1 que auxilia a interpretação dos valores do coeficiente de correlação. Foram considerados atributos relevantes aqueles nos quais os coeficientes de correlação estavam classificados como *Moderado*, *Forte* e *Muito Forte*.

De modo geral, as correlações envolvendo a experiência do usuário se mostraram mais *Fortes* comparadas com as da idade, para a maioria dos casos, como mostra a Tabela 3. Interessante notar que a quantidade total de eventos realizados apresentou uma correlação *Fraca* em relação à idade, mas uma correlação *Moderada* quando comparada a experiência. Essa correlação *Fraca* ocorreu devido tanto aos jovens quanto aos adultos, pois eles dispararam quantidades semelhantes de ações, reduzindo o valor da correlação.

Quando analisada a velocidade média entre Scroll, a correlação foi *Fraca* quando envolvia a idade e *Moderada* envolvendo a experiência tecnológica. Observou-se que diversos tipos de scroll foram realizados, usando scroll do mouse, tecla para baixo, botão *Page Down* e o botão *End*. Tanto o botão *End* quanto o botão *Page Down*, costumam ser mais rápidos para deslizar a página em relação ao scroll do mouse e a tecla para baixo. Esse tipo de técnica foi mais utilizada por pessoas com experiência tecnológica maior.

A alta correlação envolvendo a velocidade média de digitação se dá pelo fato que grande parte dos idosos se mostraram mais cautelosos na digitação, tendo a precaução de olhar o dígito a ser inserido antes de digitá-lo e também à queda nos fatores cognitivos que o avanço da idade acarreta. Os atributos quantidade de Zoom e erros de preenchimento apresentaram valores igual a zero, pois poucas ações desses dois tipos foram disparadas durante o as tarefas.

Interpretando os resultados do Estudo de Caso 1 contidos na Tabela 3, é possível afirmar que os atributos 7, 13, 17 e 19 são relevantes quando comparados com a Idade. Já os atributos 7, 12, 13, 15 e 17 são destacados como relevantes quando comparados com a experiência tecnológica.

#### 5.1.4.3 Mineração

Destacados os atributos mais significativos, é relevante saber se com esses atributos seria possível criar regras de classificação para distinguir os usuários em grupos baseado na Idade e na Experiência Tecnológica usando aprendizado de máquina.

Para esse objetivo foi utilizado o software WEKA [53], utilizando o classificador *REPTree* para entender se os atributos destacados como relevantes pela abordagem são suficientes para uma classificação eficiente, capaz de validar a abordagem proposta.

Na abordagem utilizando classificação, amostras de treinamento são utilizadas com o objetivo de fazer com que o classificador aprenda uma regra geral para realizar classificações futuras. Após várias entradas de treinamento, o classificador é capaz de responder se a entrada seguinte pertence a um grupo ou outro.

Das opções disponíveis de treinamento, foi escolhido a de *Cross-validation* com valor padrão 25 (Número de amostras - 1). Após aplicar o algoritmo, obteve-se 65,38% de eficiência na classificação envolvendo o grupo baseado na idade (jovens e adultos) e 84,61%

de eficiência na classificação envolvendo o grupo baseado na experiência tecnológica (baixa e alta experiência).

#### 5.1.4.4 Interpretação e Avaliação

Após realizar a etapa de mineração, é papel do especialista avaliar se o resultado obtido é satisfatório para o cenário que se encontra. Observando os resultados da etapa de mineração da seção anterior 5.1.4.3, utilizar o atributo Experiência seria mais eficiente do que utilizar o atributo Idade para esta classificação.

Observando a árvore de decisão criada pelo algoritmo *REPTree* na classificação baseado na Experiência do Usuário, mostrado na Figura 6, percebe-se que apenas o atributo *Tempo de Duração* seria necessário para realizar a classificação neste estudo de Caso.

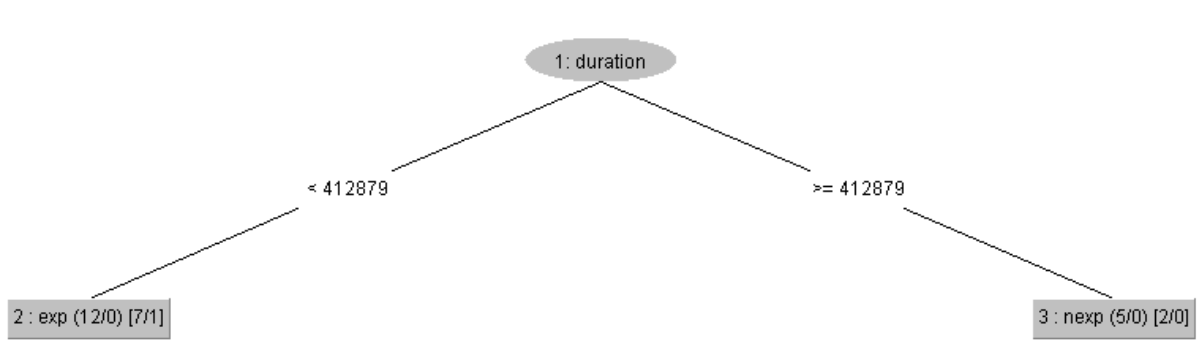


Figura 6 – Árvore de decisão criada pelo algoritmo na classificação dos usuários do Estudo de Caso 1 baseado na experiência. Legenda: *duration* - Tempo de Duração (valores em milissegundos), *exp* - Experiente, *nexp* - Pouco Experiente

Portanto, já que o grau de eficiência da classificação baseado na experiência é satisfatório (84,61%), as regras geradas pela etapa de mineração (árvore de decisão) podem ser incluídas no repositório de conhecimento da abordagem RUM. Essa inclusão é feita de forma automática após aprovação do especialista na extensão RUM++.

## 5.2 Estudo de Caso 2

No Estudo de Caso 2, um total de 44 participantes foram recrutados e separados em dois grupos de acordo com a faixa etária: jovens de 18 a 39 anos (grupo 2A) e adultos mais velhos, de 40 a 59 anos (grupo 2B). 16 usuários foram recrutados para o grupo 2A, e 28 usuários, para o grupo 2B. É importante ressaltar que nenhum usuário havia tido contato com o sistema web utilizado no estudo de caso.

Foram incluídas novas perguntas no questionário utilizado no Estudo de Caso 2 (Apêndice 2) em relação ao Estudo de Caso 1, para obter informações como frequência



de uso de computadores, frequência de uso de smartphones, quais tipos de sites o usuário costuma acessar pelo computador e quais tipos de sites costuma acessar pelo smartphone.

Com a experiência adquirida no Estudo de Caso 1 e com o avanço nas revisões bibliográficas, ficou claro que utilizar apenas o tempo de uso como métrica de experiência não seria uma boa opção [22]. No Estudo de Caso 1, foi utilizado apenas do tempo de uso como parâmetro de experiência, mas foi perceptível que utilizar somente este atributos não era suficiente para essa medida, principalmente em países em desenvolvimento como o Brasil, onde há uma educação tecnológica tardia da população em relação a outros países da América do Norte e Europa. Por isso, introduzimos o uso do CET - Coeficiente de Experiência Tecnológica - com o objetivo de quantificar a experiência tecnológica do usuário considerando o tempo e a frequência de uso de computadores.

### 5.2.1 Coeficiente de Experiência Tecnológica - CET

É evidente que quanto mais informações o usuário fornecer sobre sua experiência tecnológica para o avaliador, maior será a precisão dessa característica, mas ao mesmo tempo, menos transparente e automática se torna a detecção desses usuários na web. Por isso, empiricamente, introduzimos o uso do CET - Coeficiente de Experiência Tecnológica - que utiliza o tempo e a frequência de uso como parâmetros que ponderam o grau de experiência do usuário.

O tempo de uso (TU) com computadores foi mensurado através de uma escala de 5 pontos:

- 1 ponto: Menos de 1 ano de uso
- 2 pontos: Entre 2 e 3 anos de uso
- 3 pontos: Entre 4 a 6 anos de uso
- 4 pontos: Entre 7 a 10 anos de uso
- 5 pontos: Mais de 10 anos de uso

A frequência de uso (FU) com computadores foi mensurado através de uma escala de 7 pontos, onde 1 ponto corresponde ao uso de 1 vez por semana e 7 pontos corresponde ao uso de 7 vezes por semana. Ao final, o CET foi definido como uma média ponderada do tempo e da frequência:

$$CET = \frac{\left(\frac{TU}{5} + \frac{FU}{7}\right)}{2}$$

Quanto mais próximo de 0 (zero) for o valor do CET, menor é a experiência tecnológica do usuário e quanto mais próximo de 1 (um), maior é experiência tecnológica

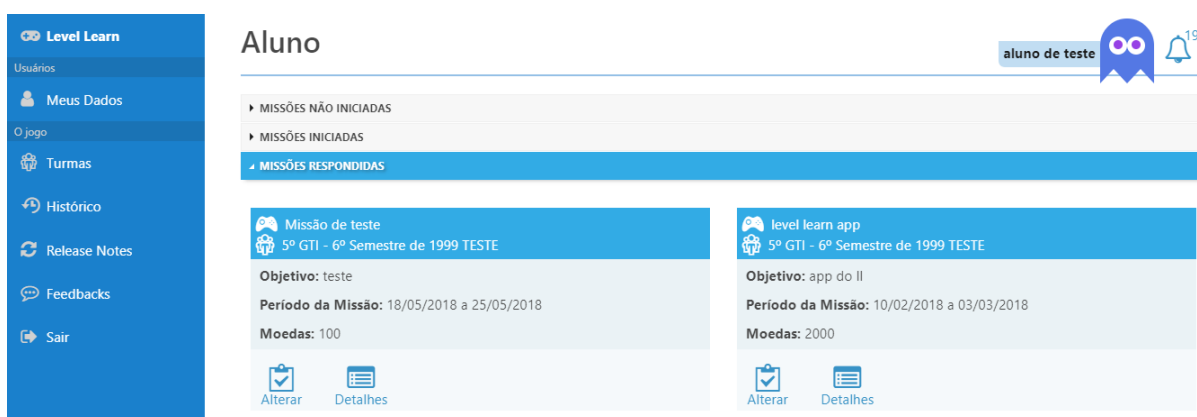


Figura 7 – Plataforma LevelLearn

do usuário. Fica a cargo do especialista da aplicação Web avaliada, ponderar em que faixas de coeficiente seus usuários alvos se encontram.

### 5.2.2 Descrição e Execução do Estudo de Caso 2

Neste estudo de caso foi proposta a execução de algumas tarefas na plataforma LevelLearn [56], um sistema de gamificação utilizado por alunos e professores da Faculdade de Tecnologia de Guaratinguetá (FATEC), que tem como objetivo garantir uma maior interação entre alunos e conteúdo escolar, através de desafios e recompensas. Para realizar o Estudo de Caso 2, o usuário deveria executar as seguintes tarefas:

1. Localizar a aba de MISSÕES NÃO INICIADAS e responder a missão TESTE disponível, clicando no botão RESPONDER;
2. Digitar o seguinte texto como resposta: *Belo Horizonte teve a temperatura mais baixa do ano ao amanhecer nesta segunda-feira (21), segundo o Instituto Nacional de Meteorologia (Inmet). Os termômetros registraram 8,7°C às 6h na Região da Pampulha. A sensação térmica foi de 7°C.*;
3. Após digitar o texto, clicar no botão RESPONDER;
4. Depois de respondido, clicar no botão VOLTAR;
5. Localizar a aba de MISSÕES RESPONDIDAS e alterar a missão TESTE clicando no botão ALTERAR;
6. Inserir no final do texto, os seguinte trecho: *A menor temperatura anterior havia sido registrada neste domingo (20) na Estação Meteorológica da Mata do Cercadinho, na Região Centro-Sul da capital mineira, onde fez 12,1°C por volta das 7h.*;
7. Após digitar o texto, clicar no botão RESPONDER;

8. Depois de respondido, clicar no botão VOLTAR;
9. Acessar a aba TURMAS e clicar no botão RANKING;
10. Navegar até final da lista e clicar no botão VOLTAR;
11. Clicar na botão HISTÓRICOS;
12. Clicar no botão SAIR;

### 5.2.3 Preparação do Ambiente e Treinamento para o Estudo de Caso 2

Para reduzir os ruídos, todos os usuários usaram um computador contendo mouse e teclado padrão. Assim, todos realizaram as tarefas no mesmo ambiente computacional.

Antes de executar as tarefas, os usuários foram instruídos sobre o sistema web a ser utilizado para se familiarizarem com a aplicação. Ao final da última tarefa, os participantes foram instruídos a fechar o navegador para que eventos extras não interferissem nas análises futuras.

### 5.2.4 Resultados do Estudo de Caso 2

No segundo estudo de caso, 44 participantes realizaram as tarefas, sendo que 16 eram usuários jovens e 28 adultos. Quando a experiência tecnológica é analisada, foi utilizado o conceito do CET, no qual os usuários que obtiveram um coeficiente menor que 0,5 foram considerados de baixa experiência tecnológica, e os que obtiveram valor maior ou igual a 0,5 foram considerados de alta experiência tecnológicas.

Em relação ao público do estudo de caso, a média de idade entre os jovens que realizaram a tarefa foi de 23,59 anos com um CET de 0,88. Já a média de idade entre os adultos mais velhos foi de 54,27 anos com um CET de 0,44, como mostra Tabela 4.

Tabela 4 – Média de valores sobre os dados de perfil dos usuários do Estudo de Caso 2

	Jovens	Adultos
Idade (anos)	23,59	54,27
Tempo de Experiência (anos)	8,21	5,93
Frequência de Uso (por semana)	6,62	2,4
CET	0.88	0.44

#### 5.2.4.1 Seleção e Pré-Processamento

Dos 44 participantes, todos conseguiram completar a tarefa com sucesso, não havendo descarte de nenhum usuário. Portanto, ao final desta etapa, foram obtidos 20 atributos de navegação extraídos dos logs. Cruzou-se os atributos de navegação com os dados de perfil dos usuários e avançou-se para a etapa de Transformação.

### 5.2.4.2 Transformação

Utilizando o conhecimento extraído da literatura, assim como no Estudo de Caso 1, realizou-se a redução de dimensionalidade que resultou em 12 atributos a serem analisados. Aplicou-se o método da correlação para tentar reduzir ainda mais a dimensionalidade da análise e entender quais atributos poderiam ser mais relevantes em relação à idade e à experiência tecnológica. O resultado está disposto na Tabela 5.

Tabela 5 – Coeficientes de Correlação entre as características pessoais dos usuário e atributos coletados dos logs.

ID	Descrição	Idade	Frequência	Tempo	CET
1	Quantidade de <i>Click</i>	-0,213	0,199	0,100	0,166
2	Quantidade de <i>Double Click</i>	0,054	0,046	-0,165	-0,063
5	Quantidade de Zoom	0,0	0,0	0,0	0,0
7	Quantidade de teclas digitadas	0,597	-0,756	-0,689	-0,784
9	Erros de Preenchimento	0,178	-0,104	-0,046	-0,084
12	Quantidade Total de Eventos	0,569	-0,703	-0,654	-0,736
13	Tempo de Duração	0,753	-0,855	-0,641	-0,815
15	Velocidade Média entre <i>Scroll</i>	-0,156	0,107	-0,066	0,026
16	Quantidade de <i>Scroll</i> Realizados	0,648	-0,616	-0,601	-0,659
17	Velocidade Média de Digitação	0,700	-0,750	-0,605	-0,737
18	Quantidade de Páginas Visitadas	0,127	-0,101	0,011	-0,063
19	Quantidade de <i>Focus</i>	-0,173	0,190	0,037	0,127

Continuou-se utilizando a Tabela 1 apresentada na Seção 4.1.2.1 que auxilia a interpretação dos valores do coeficiente de correlação.

Analisando o atributo 7 (Quantidade de teclas digitadas), foi possível perceber que os adultos mais velhos e usuários com baixa experiência tecnológica enfrentavam dificuldades ao digitar caracteres especiais como "o", "(" e ")" o que acarretava uma quantidade maior de teclas digitadas e eventos (atributo 12). Outra deficiência constatada foi a dificuldade de certos usuários alvos a encontrem em que trecho do texto haviam parado, gastando mais tempo para continuar a tarefa e contribuindo para o atributo 13 ganhar mais relevância na correlação.

Interpretando os resultados do Estudo de Caso 2 contido na Tabela 5, os atributos 7, 12, 13, 16 e 17 foram relevantes em todas as ocasiões (Idade, Frequência de Uso e CET). De forma geral, considerando os atributos mais relevantes, as correlações envolvendo a experiência tecnológica do usuário (CET) se mostraram mais expressivas comparadas com as da idade.

### 5.2.4.3 Mineração

Destacados os atributos mais relevantes, é intrigante saber se com esses atributos, seria possível criar regras de classificação para distinguir os usuários em grupos baseado

na Idade e na Experiência Tecnológica usando aprendizado de máquina. Foi utilizado novamente o software WEKA [53], utilizando o algoritmo de classificação *REPTree*. Das opções disponíveis de treinamento, foi escolhido a de *Cross-validation* com valor padrão 43 (Número de amostras - 1). Após aplicar o algoritmo, obteve-se 68,18% de eficiência na classificação envolvendo o grupo baseado na idade (jovens e adultos) e 84,09% de eficiência na classificação envolvendo o grupo baseado na experiência tecnológica (CET).

#### 5.2.4.4 Interpretação e Avaliação

Novamente, o especialista deve avaliar se o resultado obtido é satisfatório para o cenário que se encontra. Observando os resultados da etapa de mineração da Seção anterior 5.2.4.3, assim como ocorreu no Estudo de Caso 1, utilizar o atributo Experiência é mais eficiente do que utilizar o atributo Idade para a classificação.

Observando a árvore de decisão criada pelo algoritmo *REPTree* na classificação baseada na experiência tecnológica do usuário apresentada na Figura 8, percebe-se que apenas os atributos *Quantidade de Scroll* e *Quantidade de Teclas Digitadas* são necessários para realizar a classificação desse estudo de caso.

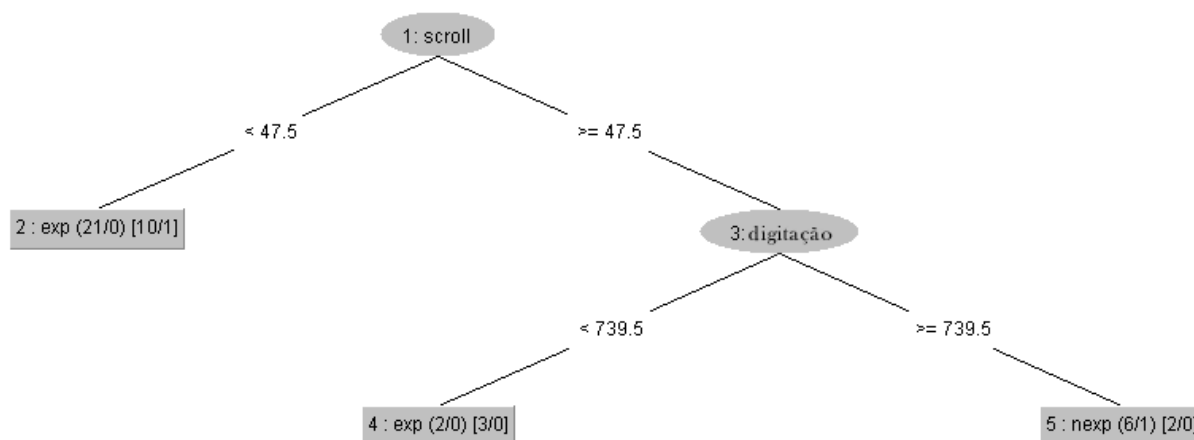


Figura 8 – Árvore de decisão criada pelo algoritmo na classificação dos usuários do Estudo de Caso 2 baseado na experiência. Legenda: *scroll* - Quantidade de Scroll (valor médio), *digitação* - Quantidade de teclas digitadas (valor médio), *exp* - Experiente, *nexp* - Pouco Experiente

Portanto, já que o grau de eficiência da classificação baseado na Experiência (CET) é satisfatório (84.09%), as regras geradas pela etapa de mineração (árvore de decisão) podem ser incluídas no repositório de conhecimento da abordagem RUM. Essa inclusão é feita de forma automática após aprovação do especialista na extensão RUM++.

### 5.3 Discussão sobre os Estudos de Caso

Depois de executado as tarefas do Estudo de Caso, é possível identificar com mais clareza as vantagens que a metodologia trouxe para o trabalho. São inúmeros os benefícios gerados desde a automação da coleta da iteração até a descoberta dos atributos que estão correlacionados com os usuários alvos.

Dos atributos analisados, 7 (sete) deles se mostraram relevantes na relação entre as dificuldades de uso de aplicações Web com os usuários alvos. Os atributos 7, 12, 13, 16 e 17 se mostraram relevantes nos dois Estudos de Caso, enquanto os atributos 15 e 19 foram relevantes apenas no Estudo de Caso 1 e o atributo 16 se mostrou relevante apenas no Estudo de Caso 2. Isso reforça ainda mais a afirmação de que os resultados obtidos na descoberta dos atributos relevantes podem variar de acordo com o site, público alvo, características dos usuários, entre outras propriedades que influenciam de forma direta e indireta nos resultados da abordagem.

Observando os resultados, é possível corroborar a afirmação dos autores [Crabb e Hanson](#)[5] de que a idade tem influência menor no comportamento dos usuários do que a experiência tecnológica. Tanto no Estudo de Caso 1 e 2, as diferenças de comportamento entre os usuários alvos estavam mais relacionados à experiência tecnológica do usuário do que a idade.

O alerta reportado pelos autores [Chadwick-Dias, McNulty e Tullis](#)[22] de que considerar apenas a frequência de uso de computadores não era suficiente para definir a experiência tecnológica, foi contornado com o uso do CET (Coeficiente de Experiência Tecnológica), que utilizou o tempo e a frequência de uso de computadores.

De forma geral, a extensão RUM++ se mostrou capaz de extrair os atributos contidos nos logs e destacar aqueles que tem relação com as dificuldades de comportamento na web apresentadas pelos usuários com idade avançada e/ou com baixa experiência tecnológica.

## 6 Conclusão

Os benefícios trazidos pelas aplicações Web acabam não atingindo todos os usuários. Indivíduos com idade avançada, por exemplo, geralmente apresentam dificuldades para realizar tarefas on-line. Com o passar dos anos, é recorrente a diminuição das habilidades motoras, sensoriais e cognitivas, o que dificulta o acesso desse grupo aos benefícios que a tecnologia oferece. Por outro lado, a população idosa é cada vez mais crescente. Os baixos índices de mortalidade infantil atrelados ao aumento da qualidade de vida em alguns países, faz com que a população idosa no mundo continue a crescer. Em 2015, 8,5% das pessoas em todo o mundo (617 milhões) apresentavam 65 anos ou mais. Segundo o relatório "An Aging World: 2015" [57], este percentual irá saltar para quase 17% da população mundial até 2050 (1,6 bilhão).

As dificuldades de navegação não estão somente relacionadas à idade, mas também com a experiência tecnológica [5, 22] dos usuários. Essa experiência pode contribuir de forma positiva para reduzir os esforços no uso das aplicações Web por parte dos adultos mais velhos e/ou com baixa experiência tecnológica.

Este trabalho propôs uma abordagem para apoiar os grupos de usuários com dificuldades de interação na Web, fornecendo uma maneira de detectá-los de forma transparente enquanto navegam. O trabalho estende a abordagem RUM [12], que permite analisar logs de usuários durante a navegação para detectar padrões de comportamento.

A abordagem RUM++ apresentada neste trabalho, estende o método tradicional de KDD [49] dentro da abordagem RUM. A principal contribuição da abordagem RUM++ está nas etapas de Pré-Processamento e Transformação dos dados, nas quais utiliza tanto o conhecimento da literatura quanto a técnica de correlação para reduzir a dimensionalidade da análise. Os dados transformados e fornecidos pelo algoritmo de mineração permitem obter bons resultados na classificação dos usuários.

Foram realizados dois estudos de caso para validar a extensão proposta. Ao final do primeiro estudo de caso ficou clara a necessidade de utilizar a frequência de uso como métrica de experiência tecnológica. Por isso, no segundo estudo de caso, foi introduzido o uso do CET (Coeficiente de Experiência Tecnológica), o qual utiliza o tempo e a frequência de uso como métricas de experiência tecnológica.

Ainda sobre a extensão proposta, é notório o êxito da sua utilização através dos estudos de caso, os quais demonstram como extrair os atributos relevantes contidos nos logs. O fato das ações serem gravadas sem interferir no modo como os usuários utilizam a aplicação, auxilia para que os participantes vivenciem o experimento como se fosse uma tarefa do dia-a-dia.

A utilização da abordagem deve ocorrer de forma controlada de maneira a evitar a obtenção de dados críticos e sensíveis, como senhas e números de cartão de crédito por exemplo. É necessário ficar atento para que esses dados não sejam capturados e que não haja violação da privacidade dos usuários .

Por fim, vale destacar que este trabalho alcançou seus objetivos, tendo as seguintes contribuições:

- i Identificação dos atributos nos logs de cliente que caracterizam o comportamento dos usuários idosos e/ou com baixa experiência tecnológica;
- ii Implementação de uma abordagem capaz de coletar o comportamento dos usuários de forma autêntica e transparente e, ao final, extrair atributos relevantes de forma prática e segura;
- iii Introdução ao uso do CET (Coeficiente de Experiência Tecnológica) que caracteriza a experiência tecnológica do usuário baseado no tempo e frequência de uso de computadores e smartphones;

## 6.1 Trabalhos Futuros

Algumas sugestões de trabalhos futuros são apresentadas a seguir. Essas sugestões visam contribuir com o desenvolvimento de novos trabalhos e incentivar pesquisas que auxiliem na compreensão do comportamento dos usuários em aplicações web:

- Investigação sobre a inclusão de mais características pessoais na composição do CET (Coeficiente de Experiência Tecnológica);
- Aplicação da extensão RUM++ em dispositivos móveis;
- Aprimoramento da extensão RUM++ para que possa ser realizada a classificação dos usuários em tempo real baseada nos grupos alvos indicados pelo especialista do sistema Web;
- Desenvolvimento de uma plataforma que permita o gerenciamento de todas as etapas da extensão RUM++, possibilitando a integração com as aplicações Web que queiram receber notificações de quando seus usuários alvo estiverem navegando na aplicação.



# Apêndices

# APÊNDICE A – Questionário Utilizado no Estudo de Caso 1

1. Qual seu nome?
2. Qual sua idade?
3. A quanto anos você usa COMPUTADOR para acessar a internet?
4. Você às vezes tem alguma dificuldades em usar o computador ou smartphone?
5. Caso você tenha dificuldade, a quem você pede ajuda?
6. Você usa alguma rede social ou aplicativo de comunicação?
7. Dos aplicativos e redes redes sociais, quais você usa e/ou participa?
8. Você usa a internet para realizar compras, transferências ou executar algum serviço bancário?
9. Caso você não use a internet para realizar compras, transferências e serviços, diga o motivo.

## APÊNDICE B – Questionário Utilizado no Estudo de Caso 2

1. Qual seu nome?
2. Qual sua idade?
3. Com que frequência você usa o COMPUTADOR na semana para acessar a internet?
4. Com que frequência você usa SMARTPHONE na semana para acessar a internet?
5. Caso você use, com que frequência você utiliza a ferramenta LevelLearn na semana?
6. A quanto tempo você usa COMPUTADOR para acessar a internet?
7. A quanto tempo você usa SMARTPHONE para acessar a internet?
8. Quais tipos de sites você costuma acessar pelo COMPUTADOR?
9. Quais tipos de sites você costuma acessar pelo SMARTPHONE?

# Referências

- 1 LIRA, K. et al. Utilizando mineração de dados e sistemas multiagentes na análise da evasão em educação a distância por meio do perfil dos alunos. In: . [S.l.: s.n.], 2016. 7, 27
- 2 EVANS, J. *Straightforward Statistics for the Behavioral Sciences*. Brooks/Cole Publishing Company, 1996. ISBN 9780534231002. Disponível em: <<https://books.google.com.br/books?id=8Ca2AAAAIAAJ>>. 8, 33
- 3 LARA, S. M. et al. A study on the acceptance of website interaction aids by older adults. *Univers. Access Inf. Soc.*, Springer-Verlag, Berlin, Heidelberg, v. 15, n. 3, p. 445–460, ago. 2016. ISSN 1615-5289. Disponível em: <<http://dx.doi.org/10.1007/s10209-015-0419-y>>. 13, 17, 18
- 4 LABERGE, J. C.; SCIALFA, C. T. Predictors of web navigation performance in a life span sample of adults. *Human Factors*, v. 47, n. 2, p. 289–302, 2005. PMID: 16170939. Disponível em: <<http://dx.doi.org/10.1518/0018720054679470>>. 13, 33
- 5 CRABB, M.; HANSON, V. L. Age, technology usage, and cognitive characteristics in relation to perceived disorientation and reported website ease of use. In: *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*. New York, NY, USA: ACM, 2014. (ASSETS '14), p. 193–200. ISBN 978-1-4503-2720-6. Disponível em: <<http://doi.acm.org/10.1145/2661334.2661356>>. 13, 17, 18, 46, 47
- 6 FAIRWEATHER, P. G. How older and younger adults differ in their approach to problem solving on a complex website. In: *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, NY, USA: ACM, 2008. (Assets '08), p. 67–72. ISBN 978-1-59593-976-0. Disponível em: <<http://doi.acm.org/10.1145/1414471.1414485>>. 13, 17, 19, 33
- 7 O'BRIEN, M. A.; ROGERS, W. A.; FISK, A. D. Understanding age and technology experience differences in use of prior knowledge for everyday technology interactions. *ACM Trans. Access. Comput.*, ACM, New York, NY, USA, v. 4, n. 2, p. 9:1–9:27, mar. 2012. ISSN 1936-7228. 13
- 8 ZAPHIRIS, P.; GHIAWADWALA, M.; MUGHAL, S. Age-centered research-based web design guidelines. In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2005. (CHI EA '05), p. 1897–1900. ISBN 1-59593-002-7. Disponível em: <<http://doi.acm.org/10.1145/1056808.1057050>>. 13, 17
- 9 MBIPOM, G.; HARPER, S. The interplay between web aesthetics and accessibility. In: *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, NY, USA: ACM, 2011. (ASSETS '11), p. 147–154. ISBN 978-1-4503-0920-2. Disponível em: <<http://doi.acm.org/10.1145/2049536.2049564>>. 13
- 10 SHNEIDERMAN, B. Promoting universal usability with multi-layer interface design. In: *Proceedings of the 2003 Conference on Universal Usability*. New York, NY, USA: ACM, 2003. (CUU '03), p. 1–8. ISBN 1-58113-701-X. Disponível em: <<http://doi.acm.org/10.1145/957205.957206>>. 13

- 11 LEUNG, R. et al. Multi-layered interfaces to improve older adults's initial learnability of mobile applications. *ACM Trans. Access. Comput.*, ACM, New York, NY, USA, v. 3, n. 1, p. 1:1–1:30, set. 2010. ISSN 1936-7228. Disponível em: <<http://doi.acm.org/10.1145/1838562.1838563>>. 13
- 12 VASCONCELOS, L. G.; BALDOCHI, L. A.; SANTOS, R. D. C. Rum: An approach to support web applications adaptation during user browsing. In: GERVASI, O. et al. (Ed.). *Computational Science and Its Applications – ICCSA 2018*. Cham: Springer International Publishing, 2018. p. 76–91. ISBN 978-3-319-95165-2. 14, 22, 47
- 13 ATTERER, R.; WNUK, M.; SCHMIDT, A. Knowing the user's every move: User activity tracking for website usability evaluation and implicit interaction. In: *Proceedings of the 15th International Conference on World Wide Web*. New York, NY, USA: ACM, 2006. (WWW '06), p. 203–212. ISBN 1-59593-323-9. Disponível em: <<http://doi.acm.org/10.1145/1135777.1135811>>. 14
- 14 CZAJA, S. J. et al. Age related differences in learning to use a text-editing system. *Behaviour & Information Technology*, Taylor & Francis, v. 8, n. 4, p. 309–319, 1989. 17
- 15 ELIAS, P. K. et al. Acquisition of word-processing skills by younger, middle-age, and older adults. *Psychology and Aging*, American Psychological Association, v. 2, n. 4, p. 340, 1987. 17
- 16 ZANDRI, E.; CHARNESS, N. Training older and younger adults to use software. *Educational Gerontology: An International Quarterly*, Taylor & Francis, v. 15, n. 6, p. 615–631, 1989. 17
- 17 PRIEST, L.; NAYAK, L.; STUART-HAMILTON, I. Website task performance by older adults. *Behav. Inf. Technol.*, Taylor & Francis, Inc., Bristol, PA, USA, v. 26, n. 3, p. 189–195, maio 2007. ISSN 0144-929X. Disponível em: <<http://dx.doi.org/10.1080/01449290500402668>>. 17
- 18 SALES, M. B. de; CYBIS, W. de A. Development of a checklist for the evaluation of the web accessibility for the aged users. In: *Proceedings of the Latin American Conference on Human-computer Interaction*. New York, NY, USA: ACM, 2003. (CLIHC '03), p. 125–133. ISBN 978-1-4503-4324-4. Disponível em: <<http://doi.acm.org/10.1145/944519.944533>>. 17
- 19 ROGERS, B. L. Measuring online experience: It's about more than time! v. 2003. 17
- 20 ARCH, A. *Web Accessibility for Older Users: A Literature Review*. [S.l.]. Disponível em: <<http://www.w3.org/TR/2008/WD-wai-age-literature-20080514/>>. 17
- 21 ZHANG, B.; RAU, P.-L. P.; SALVENDY, G. Design and evaluation of smart home user interface: Effects of age, tasks and intelligence level. *Behav. Inf. Technol.*, Taylor & Francis, Inc., Bristol, PA, USA, v. 28, n. 3, p. 239–249, maio 2009. ISSN 0144-929X. Disponível em: <<http://dx.doi.org/10.1080/01449290701573978>>. 17
- 22 CHADWICK-DIAS, A.; MCNULTY, M.; TULLIS, T. Web usability and age: How design changes can improve performance. In: *Proceedings of the 2003 Conference on Universal Usability*. New York, NY, USA: ACM, 2003. (CUU '03), p. 30–37. ISBN 1-58113-701-X. Disponível em: <<http://doi.acm.org/10.1145/957205.957212>>. 17, 41, 46, 47

- 23 CHADWICK-DIAS, A.; TEDESCO, D.; TULLIS, T. Older adults and web usability: Is web experience the same as web expertise? In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2004. (CHI EA '04), p. 1391–1394. ISBN 1-58113-703-6. Disponível em: <<http://doi.acm.org/10.1145/985921.986072>>. 17, 33
- 24 CZAJA, S. J.; SHARIT, J. Age differences in the performance of computer-based work. *Psychology and aging*, American Psychological Association, v. 8, n. 1, p. 59, 1993. 18, 19
- 25 SMITH, M. W.; SHARIT, J.; CZAJA, S. J. Aging, motor control, and the performance of computer mouse tasks. *Human factors*, SAGE Publications Sage CA: Los Angeles, CA, v. 41, n. 3, p. 389–396, 1999. 18, 19, 20, 32
- 26 CHAPARRO, A. et al. Range of motion of the wrist: implications for designing computer input devices for the elderly. *Disability and rehabilitation*, Taylor & Francis, v. 22, n. 13-14, p. 633–637, 2000. 18
- 27 CHOU, J.-R.; HSIAO, S.-W. A usability study on human-computer interface for middle-aged learners. *Comput. Hum. Behav.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 23, n. 4, p. 2040–2063, jul. 2007. ISSN 0747-5632. Disponível em: <<http://dx.doi.org/10.1016/j.chb.2006.02.011>>. 18
- 28 KANG, N. E.; YOON, W. C. Age and experience-related user behavior differences in the use of complicated electronic devices. *Int. J. Hum.-Comput. Stud.*, Academic Press, Inc., Duluth, MN, USA, v. 66, n. 6, p. 425–437, jun. 2008. ISSN 1071-5819. Disponível em: <<http://dx.doi.org/10.1016/j.ijhcs.2007.12.003>>. 18
- 29 JAKOBI, P. Using the world wide web as a teaching tool: Analyzing images of aging and the visual needs of an aging society. *Educational Gerontology*, Taylor & Francis, v. 25, n. 6, p. 581–593, 1999. 18
- 30 BLAKE, M. Internet access for older people. In: MCB UP LTD. *Aslib Proceedings*. [S.l.], 1998. v. 50, n. 10, p. 308–315. 18
- 31 CHAPARRO, A. et al. The impact of age on computer input device use:: Psychophysical and physiological measures. *International Journal of Industrial Ergonomics*, Elsevier, v. 24, n. 5, p. 503–513, 1999. 18, 20, 32
- 32 CARVALHO, D.; BESSA, M.; MAGALHAES, L. Different interaction paradigms for different user groups: An evaluation regarding content selection. In: *Proceedings of the XV International Conference on Human Computer Interaction*. New York, NY, USA: ACM, 2014. (Interacci&#243;n '14), p. 40:1–40:6. ISBN 978-1-4503-2880-7. Disponível em: <<http://doi.acm.org/10.1145/2662253.2662293>>. 18, 20, 32
- 33 HOLLINWORTH, N.; HWANG, F. Cursor relocation techniques to help older adults find 'lost' cursors. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2011. (CHI '11), p. 863–866. ISBN 978-1-4503-0228-9. Disponível em: <<http://doi.acm.org/10.1145/1978942.1979068>>. 18, 20, 32

- 34 FINDLATER, L. et al. Age-related differences in performance with touchscreens compared to traditional mouse input. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2013. (CHI '13), p. 343–346. ISBN 978-1-4503-1899-0. Disponível em: <<http://doi.acm.org/10.1145/2470654.2470703>>. 19
- 35 PEDERSEN, E. W.; HORNBÆK, K. An experimental comparison of touch interaction on vertical and horizontal surfaces. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. New York, NY, USA: ACM, 2012. (NordiCHI '12), p. 370–379. ISBN 978-1-4503-1482-4. Disponível em: <<http://doi.acm.org/10.1145/2399016.2399074>>. 19
- 36 UMEMURO, H. Lowering elderly japanese users' resistance towards computers by using touchscreen technology. *Univers. Access Inf. Soc.*, Springer-Verlag, Berlin, Heidelberg, v. 3, n. 3-4, p. 276–288, out. 2004. ISSN 1615-5289. Disponível em: <<http://dx.doi.org/10.1007/s10209-004-0098-6>>. 19
- 37 CARVALHO, D.; BESSA, M.; MAGALHAES CARRAPATOSO, E. Age group differences in performance using diverse input modalities: Insertion task evaluation. In: *Proceedings of the XVII International Conference on Human Computer Interaction*. New York, NY, USA: ACM, 2016. (Interaccion 16), p. 12:1–12:8. ISBN 978-1-4503-4119-6. Disponível em: <<http://doi.acm.org/10.1145/2998626.2998664>>. 19, 20, 32
- 38 SANCHIZ, M.; CHEVALIER, A.; AMADIEU, F. How do older and young adults start searching for information? impact of age, domain knowledge and problem complexity on the different steps of information searching. *Comput. Hum. Behav.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 72, n. C, p. 67–78, jul. 2017. ISSN 0747-5632. Disponível em: <<https://doi.org/10.1016/j.chb.2017.02.038>>. 19
- 39 SHARIT, J. et al. Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. *ACM Trans. Comput.-Hum. Interact.*, ACM, New York, NY, USA, v. 15, n. 1, p. 3:1–3:25, maio 2008. ISSN 1073-0516. Disponível em: <<http://doi.acm.org/10.1145/1352782.1352785>>. 19
- 40 CHEVALIER, A.; DOMMES, A.; MARQUIÉ, J.-C. Strategy and accuracy during information search on the web. *Comput. Hum. Behav.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 53, n. C, p. 305–315, dez. 2015. ISSN 0747-5632. Disponível em: <<http://dx.doi.org/10.1016/j.chb.2015.07.017>>. 20
- 41 LIAO, C. et al. A comparison of website usage between young adults and the elderly. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, v. 44, n. 24, p. 4–101–4–101, 2000. Disponível em: <<https://doi.org/10.1177/154193120004402432>>. 20, 21, 32
- 42 HWANG, F.; HOLLINWORTH, N.; WILLIAMS, N. Effects of target expansion on selection performance in older computer users. *ACM Trans. Access. Comput.*, ACM, New York, NY, USA, v. 5, n. 1, p. 1:1–1:26, set. 2013. ISSN 1936-7228. Disponível em: <<http://doi.acm.org/10.1145/2514848>>. 20, 21, 32
- 43 VASCONCELOS, L.; BALDOCHI, L. Usabilics: remote usability evaluation and metrics based on task analysis. In: *Proceedings of the 10th Brazilian Symposium on Human*

*Factors in Computer Systems & 5th Latin American Conference on Human-Computer Interaction*. [S.l.: s.n.], 2011. v. 303. 22

44 GONCALVES, L. F. et al. Supporting adaptation of web applications to the mobile environment with automated usability evaluation. In: *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2016. (SAC '16), p. 787–794. ISBN 978-1-4503-3739-7. Disponível em: <<http://doi.acm.org/10.1145/2851613.2851863>>. 23

45 VASCONCELOS, L. G.; SANTOS, R. D. C.; BALDOCHI, L. A. Exploiting client logs to support the construction of adaptive e-commerce applications. In: *2016 IEEE 13th International Conference on e-Business Engineering (ICEBE)*. [S.l.: s.n.], 2016. p. 164–169. 25

46 VASCONCELOS, L. G. de; BALDOCHI JR., L. A. Towards an automatic evaluation of web applications. In: *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2012. (SAC '12), p. 709–716. ISBN 978-1-4503-0857-1. Disponível em: <<http://doi.acm.org/10.1145/2245276.2245410>>. 25, 28

47 HTML Event Attributes. <[https://www.w3schools.com/jsref/dom\\_obj\\_event.asp](https://www.w3schools.com/jsref/dom_obj_event.asp)>. Accessed: 2017-08-14. 25

48 FRAWLEY, W. J.; PIATETSKY-SHAPIRO, G.; MATHEUS, C. J. Knowledge discovery in databases: An overview. *AI Mag.*, American Association for Artificial Intelligence, Menlo Park, CA, USA, v. 13, n. 3, p. 57–70, set. 1992. ISSN 0738-4602. Disponível em: <<http://dl.acm.org/citation.cfm?id=140629.140633>>. 26

49 FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. The kdd process for extracting useful knowledge from volumes of data. *Communications of the ACM*, ACM, v. 39, n. 11, p. 27–34, 1996. 26, 47

50 GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. *Journal of machine learning research*, v. 3, n. Mar, p. 1157–1182, 2003. 32

51 KARANAM, S. et al. Modeling and predicting information search behavior. In: *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics*. New York, NY, USA: ACM, 2015. (WIMS '15), p. 7:1–7:12. ISBN 978-1-4503-3293-4. Disponível em: <<http://doi.acm.org/10.1145/2797115.2797123>>. 33

52 DAVID. Tables of the ordinates and probability integral of the distribution of the correlation in small samples. Issued by the Biometrika Office, University College, and printed at the University Press, Cambridge, England, 1938. 33

53 HALL, M. et al. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, n. 1, p. 10–18, nov. 2009. ISSN 1931-0145. Disponível em: <<http://doi.acm.org/10.1145/1656274.1656278>>. 33, 39, 45

54 WITTEN, I. H.; FRANK, E.; HALL, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. 3. ed. Amsterdam: Morgan Kaufmann, 2011. (Morgan Kaufmann Series in Data Management Systems). ISBN 978-0-12-374856-0. Disponível em: <<http://www.sciencedirect.com/science/book/9780123748560>>. 33



- 55 KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1-55860-363-8. Disponível em: <<http://dl.acm.org/citation.cfm?id=1643031.1643047>>. 34
- 56 GUARINO, L. *Plataforma LevelLearn*. 2018. <<http://levellearn.com.br/>>. [Online; Acessado dia 22-Julho-2018]. 42
- 57 HE, W. et al. *An Aging World : 2015*. [s.n.], 2016. (International population reports). Disponível em: <<https://books.google.com.br/books?id=C2auDAEACAAJ>>. 47