

## **AVALIAÇÃO DE SOFTWARES ROTEIRIZADORES DE VEÍCULOS**

**Karin Cristina Siqueira Ramos**

Universidade do Vale do Itajaí

Centro de Educação Superior São José

**Mirian Buss Gonçalves**

Universidade Federal de Santa Catarina

Programa de Pós-Graduação em Engenharia de Produção

### **RESUMO**

Este trabalho apresenta uma metodologia para avaliação de softwares interativos que tem como base a ergonomia e as normas para qualidade de software. O diferencial introduzido na metodologia e que a difere de outros procedimentos tradicionais de análise ergonômica ou da qualidade, é a incorporação de aspectos comportamentais na avaliação. Essas informações podem ser obtidas a partir do uso de técnicas de preferência declarada. O ferramental desenvolvido foi testado e foi implementado para softwares para roteirização de veículos. Inicialmente são identificados os atributos essenciais para esse tipo de produto; na sequência foram realizadas entrevistas com especialistas e empresas do ramo de transporte. A partir dos resultados obtidos foi possível identificar o grau de importância de cada um dos atributos considerados, tornando possível a quantificação da conformidade do software em valores percentuais. Um software específico para este fim foi avaliado e os resultados obtidos mostraram a eficiência e a viabilidade da metodologia.

### **ABSTRACT**

This paper presents a methodology for evaluation of interactive software which is based on the ergonomics and norms of quality software. The differential aspect introduced into the methodology proposed here that makes it different from others traditional ergonomic analysis or quality procedures is the incorporation of behavior aspects in the evaluation. These information can be obtained from the use of the stated preference techniques. The developed tool was tested in vehicles routing softwares. Primarily, the essential attributes for this kind of product were identified; following interviews with specialists and companies of transportation were carried out. From the results obtained, it was possible to identify the degree of importance of each one of the considered attributes making even possible a quantification of the software conformity. Experimental results have been obtained by evaluating a vehicle routing software, showing the efficiency and feasibility of the proposed methodology.

### **1. INTRODUÇÃO**

Nas últimas décadas, a sociedade vem sofrendo profundas transformações, principalmente no que diz respeito à revolução tecnológica, que veio incrementar atividades, implementar outras até então impossíveis, e, inclusive, substituir a atuação de humanos em outras dessas. A rápida expansão das tecnologias, principalmente de comunicação e informação, tem levado a um aumento considerável do número de softwares disponíveis no mercado. Todavia, existem softwares que são colocados à disposição do público sem a preocupação de uma avaliação, seja do ponto de vista ergonômico ou mesmo funcional, o que pode levar, por exemplo, a erros na entrada de dados, problemas de processamento das informações, além de uso inadequado de cores e fontes. Problemas desse tipo comprometem a qualidade do produto, podendo, inclusive, torná-lo obsoleto.

A seleção de um software interativo implica obrigatoriamente em uma definição prévia dos critérios específicos para avaliar sua qualidade e adequabilidade (Cybis, 1994). Para tanto é necessário conhecer melhor o usuário em potencial do produto e qual tarefa o mesmo deseja implementar com o uso desse. Além disso, faz-se necessária a adoção de um método de avaliação capaz de considerar tanto as características funcionais quanto à “amigabilidade” da interface (Ray, 1994; Shneiderman, 1998).

A ergonomia, mais especificamente a ergonomia de software, oferece bases teóricas e metodológicas para enfrentar o problema da avaliação e seleção de softwares, dado que estuda as dificuldades relativas à relação existente entre o homem e a máquina, usando sempre o equilíbrio entre conforto, segurança e eficiência do usuário no uso de produtos informatizados (Hix e Schulman, 1991; Scapin e Bastien, 1997). Da mesma forma, existem atualmente diversas normas (ISO e NBR, por exemplo) que podem indicar diferentes critérios a serem considerados no processo de avaliação da qualidade de software. Elas em geral consideram importantes questões de programação, design e documentação, analisando tais critérios tanto quantitativa como qualitativamente.

Considerando que um mesmo software pode ser manipulado por diferentes pessoas, é importante conhecer as habilidades físicas, cognitivas e perceptivas desses usuários, fazendo considerações, inclusive, quanto a diferenças individuais e à diversidade cultural. Além disso, é importante entender que diferentes pessoas têm diferentes objetivos a alcançar com o uso de um software, portanto, cada uma delas terá uma avaliação particular dessa experiência. Assim sendo, é extremamente importante, na seleção de um produto informatizado, levar em consideração a opinião dos futuros usuários do mesmo (Vicente, 1999). Um dos métodos que tem se mostrado bastante eficaz na obtenção de informações relevantes dos usuários é a adoção de técnicas de preferência declarada (Ortuzar e Willumsen, 1994; Louviere *et al*, 2000). A partir das respostas obtidas, é possível determinar uma função utilidade que representa a atratividade do bem ou serviço em termos de seus atributos.

Este trabalho tem por objetivo principal apresentar uma metodologia de avaliação de softwares interativos tanto a partir de critérios ergonômicos e de qualidade amplamente divulgados e reconhecidos quanto a partir das preferências dos usuários, obtidas através das técnicas de preferência declarada. Isso porque, no processo de avaliação de softwares, seja para adoção ou para o desenvolvimento de novos produtos, percebe-se a importância não apenas de questões de cunho quantificável, mas também aquelas de cunho qualitativo, que despertam o interesse de quem efetivamente usa o produto com frequência. Em especial, a metodologia foi implementada para a avaliação de softwares de roteirização de veículos.

## **2. METODOLOGIA DE AVALIAÇÃO DE SOFTWARE**

A metodologia proposta tem uma estrutura hierárquica composta por dois módulos. No Módulo 1 são avaliadas as características básicas do software, enquanto que no Módulo 2 são avaliadas as características específicas do mesmo. Cada um destes módulos será detalhado na sequência.

### **2.1. Módulo 1: Critérios Básicos**

Como características básicas consideram-se critérios ergonômicos e de qualidade que devem ser observados para qualquer tipo de software interativo. Para essa metodologia as características a serem analisadas baseiam-se em modelos de avaliação ergonômica que seguem a corrente francófona de ergonomia (Scapin e Bastien, 1997; Barthet, 1988; Valentin *et al*, 1993), bem como em reconhecidas normas de qualidade de software, basicamente as normas NBR 13596 (ABNT, 1996) e ISO 9241 (1992).

Este módulo consta de um *checklist* (lista de verificação) composto por perguntas que devem ser respondidas por um avaliador que, necessariamente, tenha tido algum contato direto com o

objeto de avaliação. Ele é composto por 47 perguntas divididas em 8 classes. Cada classe está associada a um dos critérios considerados relevantes para qualquer tipo de software interativo, quais sejam:

- Dispositivos de Apresentação e Linguagem de Interação, onde são avaliados o vocabulário, a legibilidade, a clareza e a homogeneidade das informações e comandos;
- Controle do Usuário, onde se avalia qual o nível de controle que o usuário tem sobre as ações e o sequenciamento das mesmas;
- Adaptabilidade, onde é avaliado se o software é passível de adaptação de acordo com as especificidades, níveis de experiência ou preferências do usuário, tanto no que diz respeito a sua interface quanto ao diálogo;
- Gestão de Erros, onde são avaliadas a forma de proteção e correção de erros, bem como a qualidade das mensagens de erro;
- Carga de Trabalho, onde se avalia basicamente a brevidade de ações e a carga de memorização;
- Usabilidade, onde são avaliados critérios que identificam se o software é fácil de ser operado;
- Eficiência, esta avaliada em relação ao tempo e a quantidade de recursos utilizados; e
- Funcionalidade, onde se avalia se de fato o software realiza o que propõe e se essa realização ocorre de maneira adequada.

Para a quantificação das questões presentes no Módulo 1, o avaliador é convidado a associar a cada questão um peso  $p_i$ , que representará o grau de importância daquela questão na avaliação. Deste modo, para a metodologia proposta, existem três graus de importância: questões muito importantes, que recebem peso  $p_i = 1,5$ ; questões importantes, que recebem peso  $p_i = 1$ ; e questões onde o critério não se aplica, as quais recebem peso  $p_i = 0$ . Cada questão tem apenas três alternativas de resposta e a cada uma delas é associado um valor numérico  $a_i$ , sendo que: a resposta sim tem associado a ela o valor  $a_i = 1$ ; a resposta parcialmente tem associada a ela o valor  $a_i = 0,5$ ; e a resposta não tem associada a ela o valor  $a_i = 0$ .

Sendo assim, para o critério  $j$ , com um conjunto  $k$  de perguntas a serem respondidas, calcula-se um valor percentual, denotado  $VC_j$ , que representa a nota que o critério  $j$  alcançou segundo as respostas e os pesos dados pelo avaliador às perguntas feitas. Essa nota é calculada então conforme a expressão:

$$VC_j = \frac{\sum_{i=1}^k p_i a_i}{\sum_{i=1}^k p_i} \cdot 100 \quad (1)$$

Tendo feito isso para cada um dos 8 critérios considerados, devem ser verificados então os seguintes requisitos:

- Requisito 1)  $VC_j \geq 30 \quad \forall j = 1, 2, \dots, n$
- Requisito 2)  $\overline{VC_j} \geq 55$ , onde  $\overline{VC_j}$  é a média aritmética de todos os critérios  $VC_j$ .

O primeiro requisito visa garantir que o software tenha pelo menos 30% de conformidade com cada um dos critérios individualmente. Isso porque, caso ocorra o contrário, o software pode acabar não sendo útil para realizar a tarefa pretendida. O segundo requisito visa garantir que a média aritmética de todos os critérios deve ser de, no mínimo, 55%. Isso quer dizer que, para satisfazer esse requisito, o software deve ter mais de 55% de conformidade com os critérios básicos considerados.

Caso o software que está sendo analisado satisfaça a estes requisitos, então a avaliação passa para o próximo estágio, o Módulo 2.

## **2.2. Módulo 2: Critérios Específicos**

Neste módulo são avaliadas as características específicas do software, isto é, as peculiaridades de cada tipo de software interativo. Tais especificidades serão determinadas a partir de experimentos usando técnicas de preferência declarada. Para analisar tais características o trabalho é bem mais minucioso, sendo necessário identificar quais os critérios relevantes a serem analisados e seus respectivos pesos. Somente após isso é que a efetiva avaliação poderá ser feita, tendo como resposta final um índice de conformidade do software em relação aos critérios considerados.

Este módulo de avaliação foi estruturado em duas fases distintas porém, hierarquicamente dependentes: a Fase 1 – Experimento de Preferência Declarada e a Fase 2 – Avaliação do Software Específico.

A Fase 1 deve ser feita por especialistas no desenvolvimento e realização de experimentos de preferência declarada. Esta equipe, para atingir os objetivos pretendidos na metodologia proposta, deve seguir uma lista de procedimentos que consiste nos seguintes passos:

- Definir uma lista preliminar de atributos;
- Dividir os atributos, se necessário, em áreas de interesse;
- Se não houver divisão em áreas, desenhar um único experimento de preferência declarada;
- Se houver mais de uma área de interesse, para cada uma dessas, desenhar um experimento de preferência declarada, que dará origem a uma função utilidade local. Além disso, desenhar um outro experimento que identifique a importância relativa de cada uma dessas áreas, originando a função utilidade geral;
- Identificar um valor numérico para ser possível quantificar cada nível de cada atributo qualitativo;
- Aplicar o(s) experimento(s) de preferência declarada, entrevistando uma amostra significativa e representativa do público-alvo do software;
- Encontrar os coeficientes da(s) função(ões) utilidade ;
- Identificar o valor máximo da função utilidade geral, o que equivalerá a 100% de conformidade.

De posse desses valores, passa-se à Fase 2, sob responsabilidade do avaliador. Nesta segunda fase da avaliação, o avaliador está de posse dos pesos dos atributos (coeficientes das funções utilidade) e conseqüentemente das funções utilidade locais e a geral, o valor máximo desta última e a valoração dos níveis de cada atributo. Parte-se então para a análise particular de um software específico.

Neste momento surge a necessidade de se identificar a existência de algum atributo que seja expressamente indispensável para o tratamento do problema a partir do software, para aquele usuário específico. Caso não haja um atributo deste tipo, segue-se com a avaliação normalmente. Caso exista algum atributo deste tipo, então há a necessidade de se fazer uma análise mais detalhada do problema, avaliando-se, inclusive, a conveniência ou não de continuar a avaliação.

Sendo assim, o que se está analisando neste momento são as características de um software em particular segundo as expectativas dos usuários em potencial, essas identificadas a partir dos atributos levantados na fase anterior. Para cada atributo considerado nas funções utilidade, deve-se atribuir o valor numérico correspondente ao nível em que o software avaliado se encaixa. Tendo sido feito isso em todas as funções utilidade das diferentes áreas de interesse, resta calcular, para aquele software em especial, a função utilidade geral. De posse desse valor, através de uma regra de três, calcula-se o percentual de conformidade desse software.

### 3. IMPLEMENTAÇÃO DA METODOLOGIA – ROTEIRIZADORES DE VEÍCULOS

Como o módulo 1 da metodologia é fixo, para personalizá-la para a avaliação de um software específico, basta que a equipe de avaliação concentre seus esforços no módulo 2 da metodologia, mais especificamente no levantamento e quantificação dos atributos.

Para o caso específico da avaliação de softwares para roteirização de veículos, foi necessária inicialmente uma análise geral da estrutura dos softwares roteirizadores, identificando os requisitos básicos necessários segundo diversos autores (Golden *et al*, 1986; Assad, 1988; Ronen, 1988; Bodin, 1990; Melo e Ferreira, 2001; Hall, 2002). De posse destes requisitos, foi aplicado um questionário preliminar com pesquisadores da área para identificar, dentro da totalidade dos atributos elencados, aqueles que pudessem ser considerados os mais relevantes. Feita esta análise, restaram dezesseis atributos, cada um desses dividido em dois níveis. Porém, com esse elevado número de atributos divididos em níveis, a quantidade de combinações possíveis entre estes tornaria praticamente impossível a aplicação de um experimento de preferência declarada com um único conjunto. Sendo assim, o mais viável, em termos de planejamento, foi a divisão desses atributos em conjuntos (Louviere e Timmermans, 1990), levando em consideração o fato de que os atributos deveriam ser ortogonais, isto é, ter o mínimo de correlação entre si (Souza, 1999). Foram usados três conjuntos que são detalhados a seguir.

O Conjunto 1, denominado Quanto à Capacidade Algorítmica do Roteirizador, foi composto pelos seguintes atributos:

- Janelas de tempo: horário pré-determinado por um cliente para o recebimento de mercadorias.
- Tempo de carga e descarga: o tempo que se irá levar para carregar e/ou descarregar o veículo.
- Considera a velocidade: um software pode adotar uma velocidade fixa (por exemplo, a velocidade média) ou variável (usando diferentes velocidades, dependendo da rodovia ou do horário).
- Horário de início e fim da viagem: a política de jornada de trabalho da empresa pode ser um fator importante a se considerar. Caso não haja nenhuma restrição na empresa de horário para início ou fim de uma viagem, então esse fator não é relevante.

- Roteiros por veículo: número de roteiros que o software pode alocar a cada veículo por dia.
- Restrições de circulação de veículos: horários onde é proibida a circulação de veículos de grande porte, rodízio de veículos de acordo com o número da placa e vias de mão única são exemplos de restrições de circulação de veículos.

**Tabela 1:** Níveis dos atributos do Conjunto 1

Atributo	Valoração	Níveis	Variável
Janelas de Tempo	0	Rígidas	$X_{11}$
	1	Flexíveis	
Tempo de carga e descarga	0	Não considera	$X_{12}$
	1	Considera	
Considera a velocidade	0	Fixa	$X_{13}$
	1	Variável	
Horário de início e fim da viagem	0	Não considera	$X_{14}$
	1	Considera	
Roteiros por veículo	0	Um único	$X_{15}$
	1	Múltiplos	
Restrições de circulação de veículos	0	Não considera	$X_{16}$
	1	Considera	

A especificação de cada nível destes atributos, bem como sua valoração e a variável associada são mostrados na tabela 1. Como dito anteriormente, cada atributo foi dividido em 2 níveis: o nível associado ao valor zero, representando a forma mais básica do atributo ou mesmo a inexistência dele no software; e o nível associado ao valor 1, representando a forma mais completa do atributo, ou tão somente a sua existência ou consideração. Este padrão irá se repetir também para os demais conjuntos constantes no experimento.

O Conjunto 2, intitulado Quanto às Restrições Consideradas, foi composto pelos seguintes atributos:

- Número de bases: por base entende-se o ponto de saída e/ou chegada dos veículos, podendo esta ser única ou múltiplas.
- Tipo de veículo: alguns softwares só trabalham com roteirização considerando apenas um único tipo de veículo, enquanto que outros consideram vários tipos de veículo.
- Duração máxima de um roteiro: em algumas situações esta pode ser uma variável relevante, principalmente no que diz respeito às leis trabalhistas ou à política de carga horária da empresa.
- Capacidade do veículo: neste caso considera-se como informação importante o limite máximo possível de embarque em termos de peso e/ou volume.
- Tamanho do veículo por cliente: alguns clientes têm restrições quanto ao tamanho do veículo que fará a entrega ou coleta de produtos (principalmente pelo fato que certos veículos podem facilitar o processo de embarque/desembarque).

A especificação de cada nível dos atributos deste conjunto, os valores associados a esses e as variáveis representativas na função utilidade são mostrados na tabela 2.

**Tabela 2:** Níveis dos atributos do Conjunto 2

Atributo	Valoração	Níveis	Variável
Número de bases (origem/destino)	0	Uma única	

	1	Múltiplas	$X_{21}$
Tipo de veículo	0	Único tipo	$X_{22}$
	1	Vários tipos	
Duração máxima de um roteiro	0	Não considera	$X_{23}$
	1	Considera	
Capacidade do veículo (ton/m <sup>3</sup> )	0	Não considera	$X_{24}$
	1	Considera	
Tamanho do veículo por cliente	0	Não considera	$X_{25}$
	1	Considera	

O Conjunto 3, nomeado Quanto à Interface e o Tratamento de Dados Reais, tem os seguintes atributos:

- Roteirização em tempo real: possibilidade de inserção de pontos de coleta ou entrega na rota que um veículo está fazendo, isso em tempo real.
- Mudança manual da solução: também conhecido como *drag and drop*, é a capacidade ou não de mudar a solução final dada pelo software, caso isso seja preciso (por exemplo, se for para privilegiar as exigências de algum cliente preferencial).
- Recursos gráficos: um software pode apresentar o roteiro apenas através de uma lista sequencial, através de pontos ligados por linhas ou pode permitir a visualização da rota através de mapas e figuras explicativas.
- Cálculo real de distância e tempo: um software pode ou não ser capaz de considerar informações reais de tráfego, considerando, por exemplo, engarrafamentos e acidentes para a previsão do horário de chegada de uma mercadoria para um certo cliente.
- Interação com outros softwares: o software pode ou não ter alguma interação com outros programas computacionais, como rastreadores, WMS, GPS etc.

A especificação de cada nível destes atributos, bem como sua valoração e a variável associada são mostradas na tabela 3.

**Tabela 3:** Níveis dos atributos do Conjunto 3

Atributo	Valoração	Níveis	Variável
Roteirização em tempo real	0	Impossível	$X_{31}$
	1	Possível	
Mudança manual da solução	0	Impossível	$X_{32}$
	1	Possível	
Recursos gráficos	0	Escassos	$X_{33}$
	1	Mapas/figuras	
Cálculo real de distância e tempo	0	Impossível	$X_{34}$
	1	Possível	
Interação com outros softwares	0	Nenhuma	$X_{35}$
	1	Existente	

Assim, para cada área de interesse considerada, pôde-se estabelecer uma função utilidade local, como segue:

$$F(C1) = \alpha_{11}X_{11} + \alpha_{12}X_{12} + \alpha_{13}X_{13} + \alpha_{14}X_{14} + \alpha_{15}X_{15} + \alpha_{16}X_{16} \quad (2)$$

$$F(C2) = \alpha_{21}X_{21} + \alpha_{22}X_{22} + \alpha_{23}X_{23} + \alpha_{24}X_{24} + \alpha_{25}X_{25} \quad (3)$$

$$F(C3) = \alpha_{31}X_{31} + \alpha_{32}X_{32} + \alpha_{33}X_{33} + \alpha_{34}X_{34} + \alpha_{35}X_{35} \quad (4)$$

onde  $X_{ij}$  são os valores que serão associados a cada um dos dezesseis atributos e  $\alpha_{ij}$  são os coeficientes ponderadores, que devem ser calibrados segundo o procedimento estatístico da máxima verossimilhança.

Para cada um dos conjuntos, foi necessário aplicar um experimento de preferência declarada. Para a efetiva aplicação das entrevistas aos usuários em potencial, foram adotados, para cada um dos três conjuntos, ensaios fatoriais fracionados baseados nos arranjos ortogonais de Taguchi (Souza, 1999). Além disso, devido ao elevado número de alternativas e atributos envolvidos, foi feita uma distribuição das alternativas em blocos, seguindo a técnica de blocos incompletos balanceados (Yates, 1970). Desta forma, todos os três conjuntos ficaram com um número reduzido de alternativas (8) divididos em 7 blocos com 4 alternativas cada. A apresentação das alternativas aos entrevistados para a efetivação da pesquisa foi feita através de cartões contendo ilustrações e textos para auxiliar na compreensão das situações.

Houve ainda a necessidade de aplicar mais um experimento que de alguma forma identificasse a importância relativa de cada um dos conjuntos. Deste modo, foi importante também criar um quarto conjunto, o qual foi nomeado de “Conjunto 4 – Quanto ao software como um todo”, que tem essa função integradora. A tabela 4 apresenta os três atributos associados a esse conjunto, bem como seus níveis, respectiva valoração e variáveis associadas. Assim sendo, a partir do grau de relevância de cada área segundo o entrevistado, puderam ser determinados os coeficientes ( $\beta_i$ ) que representam o peso relativo entre cada conjunto. Assim, obteve-se a função utilidade geral, que no caso representa uma única função utilidade que identifica a importância de cada um dos dezesseis atributos considerados relevantes na avaliação do software em questão, tendo esta então a seguinte expressão:

$$FU = \beta_1 F(C_1) + \beta_2 F(C_2) + \beta_3 F(C_3) \quad (5)$$

**Tabela 4:** Níveis dos atributos do Conjunto 4

Atributo	Valoração	Níveis	Variável
Capacidade algorítmica	0	Pequena	$F(C_1)$
	1	Grande	
Tratamento das restrições	0	Inserir poucas restrições	$F(C_2)$
	1	Inserir várias restrições	
Interface e tratamento de dados reais	0	Pouco satisfatório	$F(C_3)$
	1	Satisfatório	

Neste caso, para a realização das entrevistas, apesar de existirem somente 3 atributos, optou-se também por trabalhar com um delineamento mais simplificado. Ao invés de considerar o fatorial completo com as 8 alternativas, decidiu-se trabalhar com o fracionado, considerando assim apenas 4 alternativas (Kocur *et al*; 1982). Essa decisão foi tomada pensando-se em não alongar ainda mais o experimento, já que o entrevistado, nessa altura dos trabalhos, já havia feito 21 escolhas (uma escolha para cada um dos 7 blocos de cada um dos 3 conjuntos).

As entrevistas foram realizadas nos meses de janeiro e fevereiro de 2004, sendo entrevistados, face a face ou eletronicamente, tanto pesquisadores quanto empresas do ramo de logística e transporte. Anualmente a Revista Tecnológica (2003a, 2003b) realiza um levantamento dos operadores logísticos em atuação no Brasil, tendo sido relacionados nesse ano um total de aproximadamente 140 empresas. A pesquisa foi enviada para cada uma delas, tendo recebido



um retorno de aproximadamente 10% das mesmas. Em relação aos pesquisadores, foram selecionados 20 pesquisadores brasileiros que tem desenvolvido estudos sobre roteirizadores de veículos, tendo destes um retorno de aproximadamente 60%.

A cada entrevistado inicialmente eram apresentados os motivos da entrevista. A seguir solicitava-se ao mesmo que preenchesse algumas perguntas de cunho social e profissional, para melhor identificação dos entrevistados no processo de segmentação da amostra, caso isso fosse relevante no futuro. Em seguida era iniciada a apresentação sequenciada dos 3 primeiros conjuntos. Para cada um destes os cartões eram apresentados seguindo a divisão em blocos delineada, onde cada bloco de cartões era identificado através de uma cor. Era solicitado então ao entrevistado que escolhesse dentro de cada um dos blocos a sua alternativa preferida. Cada escolha era devidamente anotada em formulário próprio. Por fim eram apresentados ao entrevistado os cartões relativos ao Conjunto 4, sendo solicitado nesse momento que o mesmo ordenasse os cartões segundo suas preferências e anotasse suas escolhas em local apropriado, no mesmo formulário.

Os resultados obtidos a partir das entrevistas e codificados de forma apropriada foram posteriormente inseridos em um programa computacional desenvolvido para ajustes de modelos de preferência declarada, no caso, o programa LMPC (Souza, 1999). Através deste, os parâmetros são calibrados pela máxima verossimilhança, usando o método de ajuste de Newton-Raphson. As funções utilidade objetos de estudo foram aquelas descritas nas equações de (2) a (5) apresentadas anteriormente.

**Tabela 5: Resultados da calibração dos parâmetros do Conjunto 1**

Atributo	Coefficiente	Erro	Teste t-student
Janelas de tempo	0,7370	0,2036	3,6202
Tempo carga/descarga	1,2096	0,2229	5,4268
Considera velocidade	0,9922	0,2337	4,2456
Horário início/fim	1,2360	0,2266	5,4544
Roteiros por veículo	0,9112	0,2278	4,0004
Restrições circulação	1,4139	0,2483	5,6942

Eficiência = 0,8571

LR  $(-2[F(0)-F(B)]) = 157,2912$

Número de Casos = 175

Rho = 0,3242

**Tabela 6: Resultados da calibração dos parâmetros do Conjunto 2**

Atributo	Coefficiente	Erro	Teste t-student
Número de bases	0,9164	0,2751	3,3312
Tipo de veículo	1,5842	0,2751	5,7588
Duração máxima roteiro	1,6223	0,3054	5,3125
Capacidade do veículo	1,4057	0,2343	5,9996
Tamanho veículo/cliente	0,5018	0,2343	2,1417

Eficiência = 0,8571

LR  $(-2[F(0)-F(B)]) = 208,3326$

Número de Casos = 175

Rho = 0,4294

As tabelas 5, 6 e 7 apresentam os ajustes obtidos para cada uma das 3 funções utilidade locais referentes aos conjuntos 1, 2 e 3 respectivamente, enquanto que na tabela 8 pode ser visto o ajuste encontrado para a função utilidade geral. Em cada uma dessas tabelas são apresentados também testes estatísticos importantes para a verificação da confiabilidade dos resultados obtidos, sendo os mais relevantes: teste t-student, likelihood e rho (falso rho quadrado).

**Tabela 7: Resultados da calibração dos parâmetros do Conjunto 3**

Atributo	Coefficiente	Erro	Teste t-student
Roteiriza tempo real	0,9344	0,2140	4,3672
Mudança manual	1,0546	0,2140	4,9287
Recursos gráficos	1,2689	0,2577	4,9240
Cálculo real dist/tempo	1,2164	0,2303	5,2828
Interação outro software	1,0562	0,2303	4,5869

Eficiência = 0,8571

Número de Casos = 175

LR  $[-2(F(0)-F(B))]= 168,6456$

Rho = 0,3476

A aplicação do teste *t-student* revela um grau de confiabilidade de, no mínimo, 98% para todos os 19 atributos analisados, já que em todos os conjuntos o número de casos, ou seja, o número de entrevistas, é maior do que 30. Na expressão  $-2[F(0) - F(B)]$ ,  $F(0)$  representa o valor da função log – verossimilhança (*log – likelihood*) para todos os coeficientes nulos e  $F(B)$  representa o valor otimizado da função e é determinado usando a distribuição  $\chi^2$ . Os valores obtidos com este teste, para cada um dos 4 conjuntos, indica que, em todos os casos, é possível rejeitar a hipótese da nulidade de todos os parâmetros simultaneamente. Por fim, o teste estatístico do *rho* ( $\rho^2$ ) apresenta um ótimo desempenho em todos os 4 conjuntos de análise, tendo valores entre 0,3 e 0,43, valores apontados como excelentes, de acordo com Ortuzar e Willumsen (1994).

**Tabela 8: Resultados da calibração dos parâmetros do Conjunto 4**

Atributo	Coefficiente	Erro	Teste t-student
Capacidade algorítmica	1,3313	0,4082	3,2616
Tratamento restrições	1,9316	0,4246	4,5496
Interface e dados reais	1,3958	0,4059	3,4388

Eficiência = 1,0000

Número de Casos = 75

LR  $[-2(F(0)-F(B))]= 68,3199$

Rho = 0,4299

Portanto, os coeficientes ajustados, quando inseridos na expressão (5), geram a função utilidade geral, dada por:

$$FU = 1,3313F(C_1) + 1,9316F(C_2) + 1,3958F(C_3) \quad (6)$$

sendo que

$$F(C_1) = 0,7370X_{11} + 1,2096X_{12} + 0,9922X_{13} + 1,2360X_{14} + 0,9112X_{15} + 1,4139X_{16} \quad (7)$$

$$F(C_2) = 0,9164X_{21} + 1,5842X_{22} + 1,6223X_{23} + 1,4057X_{24} + 0,5018X_{25} \quad (8)$$

$$F(C_3) = 0,9344X_{31} + 1,0546X_{32} + 1,2689X_{33} + 1,2164X_{34} + 1,0562X_{35} \quad (9)$$

A partir dos resultados obtidos é possível calcular o valor máximo da função utilidade geral, partindo-se do pressuposto que todas as utilidades locais alcancem individualmente seus valores máximos (isto acontecendo quando todos os atributos respectivos atingirem seus níveis máximos de pontuação); o valor máximo da função utilidade geral equivalerá então a 100% de conformidade com os atributos especificados. Sabendo que cada um dos atributos considerados pode assumir o valor máximo igual a 1, tem-se:  $FU_{\max} = 28,02113$ . Assim, quando um software específico for avaliado segundo esses atributos, basta calcular sua função utilidade geral. Tendo em mãos este valor, aplica-se uma regra de três para identificar o quão conforme o mesmo está em relação aos atributos considerados.

Para ilustrar a aplicação da metodologia proposta no presente trabalho, escolheu-se como objeto de avaliação um software de roteirização amplamente utilizado por empresas nacionais e internacionais. Relatos de algumas empresas de distribuição que fazem uso do software apontam que a partir da adoção do mesmo obtiveram reduções significativas – de 10 a 25 por cento – no tempo do ciclo do pedido e no número de veículos de sua frota, conseguindo com isso que seus clientes diretos considerassem suas entregas mais eficientes. A partir da avaliação feita através da metodologia proposta, o software que aqui, por motivos éticos, vamos chamar de “xyz” apresenta resultados bastante satisfatórios, tanto em relação aos aspectos ergonômicos e de qualidade (Módulo 1) quanto em relação aos critérios considerados importantes segundo usuários desse tipo de programa (Módulo 2).

Tanto a avaliação dos critérios básicos quanto à atribuição dos valores dos atributos em cada uma das funções utilidades locais (Módulo 1 e Módulo 2 da metodologia proposta, respectivamente) foram feitas a partir de pesquisas com um revendedor autorizado do software no Brasil, de questionamentos a algumas empresas usuárias e através do próprio conhecimento das autoras. Sendo assim, o início da avaliação do Módulo 1 se deu a partir da associação dos pesos e posteriormente das respostas para cada pergunta de cada um dos 8 critérios constantes. Tendo sido feitas os devidos cálculos, o software foi considerado 79% conforme (requisito 2), sendo que os critérios com menor adequabilidade (requisito 1) foram Gestão de Erros, com 57,1% e Carga de Trabalho, com 60%. Já nos critérios 7 e 8, respectivamente Eficiência e Funcionalidade, o software foi considerado 100% adequado.

Tendo satisfeito os dois requisitos relativos ao Módulo 1, passou-se então para a avaliação dos critérios específicos dos softwares para roteirização de veículos. Esta segunda parte da avaliação (Módulo 2) se deu mais facilmente, apenas necessitando verificar, para cada um dos 16 atributos  $X_{ij}$  em que nível o software avaliado se insere. Assim, foi possível calcular os valores das funções utilidades locais do software “xyz” :

$$F(C1) = 0,7370 \cdot 1 + 1,2096 \cdot 0 + 0,9922 \cdot 1 + 1,23605 \cdot 1 + 0,9112 \cdot 1 + 1,4139 \cdot 1 = 5,2903 \quad (10)$$

$$F(C2) = 0,9164 \cdot 1 + 1,5842 \cdot 1 + 1,6223 \cdot 1 + 1,4057 \cdot 1 + 0,5018 \cdot 0 = 5,5286 \quad (11)$$

$$F(C3) = 0,9344 \cdot 1 + 1,0546 \cdot 1 + 1,2689 \cdot 1 + 1,2164 \cdot 1 + 1,0562 \cdot 1 = 5,5305 \quad (12)$$

Colocando estes valores na equação (6), encontra-se o valor da função utilidade global igual a 25,44. Para finalizar a avaliação do software, bastou tomar como referência o valor máximo da função utilidade geral (28) e aplicar uma regra de três com o valor da função utilidade obtida para o software “xyz” (25,44). De posse deste resultado, pode-se concluir que o software avaliado segundo a metodologia proposta está 90,86% em conformidade com os atributos específicos considerados essenciais para um software para roteirização de veículos.

#### 4. CONCLUSÕES

Resumidamente, é possível concluir que a metodologia de avaliação de softwares desenvolvida tem seus méritos principalmente devido às seguintes características: a linearidade, com a verificação de requisitos mínimos a serem atendidos na primeira etapa de avaliação (Módulo 1); a inserção da visão (leia-se preferências) dos especialistas e usuários na avaliação do software (Módulo 2); a quantificação, em termos percentuais, da conformidade do software; e a alta flexibilidade da metodologia, podendo esta ser adaptada à avaliação de diferentes tipos de softwares interativos.

Quanto ao caso específico da aplicação da metodologia na avaliação de softwares roteirizadores, vale ressaltar a importância que os resultados obtidos têm para a área de transportes. Diversos autores enfatizaram em seus trabalhos a necessidade de identificar quais os critérios que devem ser adotados na seleção de tais softwares, além da pouca discussão na literatura sobre tal questão. A metodologia proposta no presente trabalho, a partir de sua aplicação específica para esse tipo de software, trouxe essa contribuição significativa, não apenas elencando os critérios a serem considerados, mas, mais ainda, dando valores significantes do grau de importância de cada um deles para o processo de escolha.

## REFERÊNCIAS

- ABNT (1996) **NBR 13596 - Tecnologia de informação: Avaliação de produto de software: Características de qualidade e diretrizes para o seu uso**. Associação Brasileira de Normas Técnicas, Rio de Janeiro.
- ASSAD, A. A. (1988) Modeling and implamentation issues in vehicle routing. **Vehicle Routing: Methods and Studies**, p. 7-46.
- BARTHET, M. F. (1988) **Logiciels Interactifs et Ergonomie**. Dunod, Paris.
- BODIN, L. D. (1990) Twenty years of routing and scheduling. **Operations Research**, v. 38, n. 4, p. 571-579.
- CYBIS, W. (1994) **A identificação dos objetos de interfaces homem - computador e seus atributos ergonômicos**. Tese. (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina. Florianópolis, 1994.
- GOLDEN, B.; BODIN, L. e GOODWIN, T. (1986) Microcomputer-based vehicle routing and scheduling software. **Computers and Operations Research**, v. 13, n. 2/3, p. 277-285.
- HALL, R. (2002) **Vehicle Routing Software Survey**. Disponível em: [www.lionhrtpub.com/forms/surveys/Vehicle\\_Routing/vrss.html](http://www.lionhrtpub.com/forms/surveys/Vehicle_Routing/vrss.html). Acesso em: 07 dezembro 2002.
- HIX, D. e SCHULMAN, R. S. (1991) Human-Computer interface development tools: a methodology for their evaluation. **Communications of the ACM**, v. 34, n. 3, p. 74-87.
- ISO 9241-10 (1996) **Ergonomic requirements for Office work with visual display terminals (VDTs). Part 10: Dialogue principles**. International Organization for Standardization, Genève.
- KOCUR, G.; ADLER, W. e AUNET, B. (1982) **Guide to forecasting travel demand with direct utility assessment (Technical Report)**. U. S. Department of Transportation. UMTA-NH-11-0001-82-1.
- LOUVIERE, J. J. , HENSHER, D. e SWAIT, J. (2000) **Stated Choice Methods**.Cambridge University Press, Cambridge.
- LOUVIERE, J.; TIMMERMANS, H. **A review of recent advances in decompositional preference and choice models**. *Tijdschrift voor Econ. En Soc. Geografie*, v. 81, n. 3, p. 214-224, 1990.
- MELO, A. C. e FERREIRA FILHO, J. V. (2001) Sistemas de Roteirização e Programação de Veículos. *Pesquisa Operacional*, v. 21, n. 2, p. 223 – 232.
- ORTUZAR, J. e WILLUMSEN, L. (1994) **Modelling transportation**. 2. ed. John Wiley and Sons: England.
- RAY, E. (1994) **User Interface Design**. Prentice-Hall, New Jersey.
- REVISTA TECNOLÓGICA. São Paulo: Publicare, n. 91, junho 2003a.
- REVISTA TECNOLÓGICA. São Paulo: Publicare, n. 92, julho 2003b.
- RONEN, D. (1988) Perspectives on practical aspects of truck routing and scheduling. **European Journal of Operational Research**, n. 35, p. 137-145.
- SCAPIN, D. e BASTIEN, J. M. (1997) Ergonomic criteria for evaluating the ergonomic quality of interactive systems. **Behaviour & Information Technology**, v. 16, n. 4/5, p. 220 - 231.
- SHNEIDERMAN, B. (1998) **Designing the user interface**. 3. ed. Addison-Wesley, USA.
- SOUZA, O. A. de. (1999) **Delineamento experimental em ensaios fatoriais utilizados em preferência declarada**. Tese. (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina. Florianópolis.
- VALENTIN, A.; VALLERY, G. e LUCONGSANG, R. (1993) **L'évaluation ergonomique des logiciels: une démarche itérative de conception**. Anact, Montrouge.
- VICENTE, K. (1999) **Cognitive Work Analysis – Toward Safe, Productive and Healthy Computer-Based Work**. Lawrence Erlbaum Associates, USA.
- YATES, F. (1970) **Experimental design: selected papers of Frank Yates**. Charles Griffin, London.