

01

Adoção de Técnicas Fuzzy para Análise de Usabilidade de *Software*

Alisson Marques da Silva¹Dainae Kellen Piassi²Michel Pires da Silva³Thiago Magela Rodrigues Dias⁴

Resumo

O constante desenvolvimento tecnológico exige dos desenvolvedores de aplicações o aprimoramento de suas técnicas para a criação de soluções. Neste contexto, a usabilidade de *software*, caracteriza-se como uma das maiores preocupações em todo o processo de desenvolvimento. O intuito é prover, ao usuário final, diversos benefícios, tais como: facilidade de uso, agilidade e minimização de erros. Como auxílio para se obter um processo que possibilite avaliar a usabilidade de uma determinada aplicação, este trabalho define uma técnica baseada em métodos fuzzy, cujo objetivo é validar a usabilidade do *software* por meio da avaliação de conjunto de dados providos de pesquisas realizadas junto aos usuários do *software* selecionado para análise. Como resultado, espera-se agilizar o desenvolvimento e minimizar a quantidade de erros inseridos no produto ao longo de sua elaboração.

Palavras-Chave: Engenharia de *Software*, Usabilidade, Lógica Fuzzy.

Abstract

The constant technological development requires the application developers to improve its techniques for creating solutions. In this context, *software* usability, is characterized as a major concern throughout the development process. The aim is to provide to the end user several benefits such as: ease of use, speed and error minimization. As an aid to obtain a process which enables it to evaluate the usability of a particular application, this paper defines a technique based on fuzzy methods, whose purpose is to validate the usability of *software* through the provided data set of evaluation research conducted with users of the *software* selects for analysis. As a result, it is expected to expedite the development and minimize the amount of errors inserted in the product throughout its development.

Keywords: *Software* Engineering, Usability, Fuzzy Logic.

¹CEFET-MG – Campus Divinópolis; alissonmarques@gmail.com.

²IFMG – Campus Bambuí; daiane.piassi@gmail.com.

³CEFET-MG – Campus Divinópolis; michel.silva@gmail.com.

⁴CEFET-MG – Campus Divinópolis; thiago@div.cefetmg.br.

1. Introdução

A quantidade de aplicações atualmente disponíveis faz com que o nível de exigência dos usuários aumente de forma considerável a cada dia. Isso obriga o setor a aprimorar suas técnicas e métodos de produção de soluções. Prover melhorias na agilidade, usabilidade e redução de riscos e erros são alguns dos fatores que podem garantir uma melhor aceitação do produto. Em razão disso, o aprimoramento de técnicas para avaliar a qualidade ao longo do desenvolvimento de software se torna, cada vez mais, frequente e necessária. Nesse contexto, a identificação de problemas de usabilidade se torna um fator crítico para qualquer solução proposta.

Segundo [15], a usabilidade é um fator que pode prover como resultado facilidades de uso, eficiência, eficácia e satisfação. Sua avaliação e melhoria vai além desses quesitos, abrangendo também o quão eficiente e agradável o produto se apresentará. Segundo [8] e [9] as preocupações com esses fatores podem prover um produto final caracterizado por ser fácil de usar, rápido de aprender, de manutenção branda, eficiente e eficaz em suas tarefas principais e com um alto grau de satisfação para com seus *stackholders*.

Na literatura pode-se encontrar trabalhos cujo foco é propor métricas e técnicas para se garantir uma boa usabilidade para o usuário final. Esses, por sua vez, utilizam padrões, como, a norma ISO/IEC para definir as diretrizes de avaliação e garantia de qualidade. Uma das normas aplicáveis para esse fim é a ISO/IEC 9126. Nela, se define usabilidade como fator de qualidade necessária para garantir produtividade e segurança em um contexto de uso específico [1]. Uma forma de atingir tais objetivos está no uso de testes de usabilidade.

Testes de usabilidade se caracterizam como métodos formais de avaliação, centrados nos usuários e utilizam diferentes técnicas para avaliação, como: questionários aplicados [16, 17], *checklists* [10, 2], modelos formais [6, 4], base de conhecimento [7], ensaios de interação ou sistemas de monitoramento [13]. Objetivam minimizar os prejuízos e visam alcançar: redução de chamadas de suporte, redução do custo do treinamento, aumento da aceitação relacionada à fidelidade do usuário ao produto e diferencial do produto em relação ao concorrente.

Neste trabalho objetiva-se aprimorar a avaliação do nível de usabilidade e automatizar os testes com a

aplicação de técnica de inteligência computacional, mais especificamente Lógica *Fuzzy*. Para isso, será utilizado uma metodologia baseada em coleta de requisitos e Lógica *Fuzzy* para análise de usabilidade de software. O uso de questionários aplicados é uma técnica prospectiva que envolve a opinião do usuário e a avaliação entre ele e o sistema. Esses são utilizados para coletar as opiniões e requisitos primários, fazendo ao final uma avaliação tanto qualitativa quanto quantitativa do produto avaliado. Já a Lógica *Fuzzy* será empregada na avaliação e classificação dos requisitos do sistema, uma vez que ela permite inferir conclusões e obter respostas de informações vagas, ambíguas e imprecisas que podem estar presentes nos dados coletados pelo processo de aplicação dos questionários.

2. Revisão da Literatura

A definição de usabilidade foi introduzida por Miller em 1971 por meio de métricas relacionadas à facilidade de uso. Miller identificou vários critérios para mensurar a facilidade de uso, sendo alguns desses: tempo de aprendizado, número de erros e tolerância a falhas [12].

A primeira norma criada a partir dos conceitos apresentados por Miller foi a ISO/IEC 9126 que se refere à qualidade de software. Com uma abordagem orientada ao produto e ao usuário, considera a usabilidade um conjunto de atributos de software relacionados ao esforço necessário para seu uso e para o julgamento individual por determinados usuários [8]. Foi por meio dessa norma que o termo usabilidade ultrapassou os limites do ambiente acadêmico para fazer parte do vocabulário de outras áreas de conhecimento, como tecnologia da informação e interação homem máquina.

Em [16] foi definido um conjunto de métricas para avaliação de usabilidade, baseado em uma revisão da literatura realizada na busca de uma base de dados científica brasileira para seleção dos principais pesquisadores sobre o assunto usabilidade. Após selecionados os principais autores foi realizada uma análise da similaridade entre os trabalhos propostos por estes autores e realizada a unificação e construção de um conjunto de seis métricas para avaliação de usabilidade de software. O modelo proposto é composto pelas seguintes métricas:

- 1) **Facilidade de Aprender:** É a capacidade do software possibilitar ao usuário aprender a manuseá-lo.
- 2) **Facilidade de Relembrar:** Avalia as funcionalidades do sistema para que sejam fáceis de lembrar, mesmo após o usuário ficar certo período de tempo sem utilizá-lo.
- 3) **Controle de Erros:** É a capacidade do software possibilitar ao usuário operá-lo e controlá-lo.
- 4) **Eficiência:** É a capacidade do software prover o esperado pelo usuário.
- 5) **Eficácia:** Avalia como as tarefas foram concluídas e com que frequência elas produzem erros.
- 6) **Satisfação:** É a capacidade do software em atrair o usuário, ser agradável.

A Tabela 1 detalha as métricas acima apresentadas e seus respectivos objetivos.

Métrica	Objetivos
Facilidade de Aprender	Facilidade de navegação; Realização da tarefa com relação à clareza das mensagens; Número de possibilidades para realizar a mesma tarefa; Número de itens para a realização de todas as funcionalidades.
Facilidade de Relembrar	Consistência entre as interfaces; Relembrar a execução de tarefa após período sem usar o sistema; Capacidade que o sistema tem de guiar, dar informações e dicas; Ganho de produtividade com relação à reutilização de tarefa.
Controle de Erros	Clareza nas mensagens de erro; Retrabalho com relação à quantidade de erros provocados; Quantidade de erros.
Eficiência	Controle de ações; Produtividade do sistema; Velocidade na realização das tarefas.
Eficácia	Utilização de passos desnecessários; Quantidade de tempo para utilizar determinada tarefa; Eficiência e adaptação das tarefas.
Satisfação	Sentimento do usuário em relação à interação do sistema; Clareza das mensagens de ajuda; Sentimento em relação à clareza das mensagens e recuperação de erros.

Tabela 1. Métricas e construtos.

Além do exposto, a Lógica *Fuzzy* surge como

outro importante tópico para o desenvolvimento deste trabalho. A Teoria de Conjuntos *Fuzzy* foi concebida por L.A. Zadeh com o objetivo de fornecer um ferramental matemático para o tratamento de informações de caráter impreciso ou vago. A Lógica *Fuzzy*, baseada nessa teoria, foi inicialmente construída a partir dos conceitos já estabelecidos de lógica clássica; operadores foram definidos à semelhança dos tradicionalmente utilizados e outros foram introduzidos ao longo do tempo, muitas vezes por necessidades de caráter eminentemente prático.

Os conjuntos *fuzzy* constituem uma ponte no caminho de aproximar o raciocínio humano ao da lógica executada pela máquina. Na prática, as variáveis nebulosas caracterizam-se por pertencer simultaneamente a vários conjuntos numa determinada proporção entre eles. Esta proporção é designada por grau de pertinência. Este conceito de dualidade estabelece que um determinado estado, pode e deve coexistir com o seu oposto, e deste modo faz da Lógica Nebulosa, uma lógica natural, até mesmo inevitável.

Um conjunto clássico A é uma relação de elementos x existentes em um universo de discurso X ($A \subset X$). Cada elemento x do universo de discurso pode pertencer ao conjunto A ($x \in A$) ou não ($x \notin A$). Assim sendo, um conjunto clássico A , em universo de discurso X , pode ser definido pela sua função característica, que mapeia cada elemento de X em 1 ou 0 dependendo se o elemento é ou não membro do conjunto [14]. Isto pode ser representado por:

$$A(x) = \begin{cases} 0 & \text{se } x \notin A \\ 1 & \text{se } x \in A \end{cases} \quad (1)$$

Zadeh generaliza a função característica de modo que ela pudesse assumir um número infinito de valores no intervalo $[0,1]$. Um conjunto nebuloso A em um universo X é definido por uma função de pertinência $\mu_a(X) : \rightarrow [0,1]$, e representado por um conjunto de pares ordenados:

$$A = \{\mu_a(X)/x\} x \in X, \quad (2)$$

no qual $\mu_a(x)$ indica quanto x é compatível com o conjunto A . Determinado elemento pode pertencer a mais de um conjunto *fuzzy*, com diferentes graus de pertinência.

3. Metodologia

A metodologia para análise de usabilidade pode ser dividida nas seguintes etapas: seleção do software a ser avaliado; elaboração dos casos de teste (usuários do software); aplicação dos questionários e; por fim pode ser realizado o tratamento dos dados.

3.1 Seleção do Software

Neste trabalho o software do estudo de caso será sobre o Jabref que é um aplicativo para gerenciamento de referências bibliográficas. O Jabref foi escolhido por ser um software que muito utilizado para desenvolvimento de trabalhos científicos (artigos, monografias, dissertações e teses).

3.2 Elaboração dos Casos de Teste

Para os casos de teste do aplicativo foram selecionados aleatoriamente 33 usuários, sendo que 11 destes tem conhecimento sobre o software avaliado e 22 que nunca tiveram nenhum contato com o aplicativo. Foram gerados 3 casos de teste. O primeiro com 11 usuários que tem conhecimento do uso do aplicativo. O segundo com 11 usuários que não possuíam nenhum conhecimento no uso do aplicativo, mas que passaram por um treinamento prévio na sua utilização. Por fim, o terceiro caso de teste foi composto pelos 11 usuários restantes que não possuem conhecimento do aplicativo, neste caso eles não receberam nenhum tipo de treinamento no uso do aplicativo.

3.3 Aplicação dos Questionários

Os questionários foram aplicados aos usuários selecionados para que fosse realizada a identificação da usabilidade do software. Para isso, questões que objetivam indagar sobre as seis métricas de avaliação citadas anteriormente foram propostas. A seleção do público alvo foi realizada de acordo com o critério de que houvessem públicos diferenciados para que fosse satisfeitas todas as métricas estipuladas por [16, 17]. O entrevistado aplicou uma atividade a ser realizada no Jabref separadamente para cada um dos 3 conjuntos de usuários. Após a realização da atividade os questionários foram aplicados.

O questionário foi elaborado segundo as métricas propostas por [16] e é composto por vinte questões que servem de base para a avaliação do sistema. As questões foram criadas seguindo os construtos de cada métrica. Cada questão corresponde à determinada métrica e cada métrica é composta de várias perguntas, como é apresentado na Tabela 1. As métricas com o número das questões e o total de questões estão representados na Tabela 2.

Métrica	Questões	Total de questões por métrica
Facilidade de Aprender	1, 2, 3, 4	4
Facilidade de Relembrar	5, 6, 7, 8	4
Controle de Erros	9, 10, 11	3
Eficiência	12, 13, 14	3
Eficácia	15, 16, 17	3
Satisfação	18, 19, 20	3
Total de questões		20

Tabela 2. Métricas, números das questões e quantidade de questões.

Para as respostas foi adotada a escala Likert, esta é uma escala na qual os respondentes são solicitados não só a concordarem ou discordarem, mas também a informar qual seu grau de concordância ou discordância [11]. A escala Likert é composta por 5 pontos: discorda fortemente, discorda, indeciso, concorda, concorda fortemente.

As questões separadas por métricas são descritas como se segue:

Métrica - Facilidade de aprender

- 1) A navegação no sistema é fácil.
- 2) As telas apresentam apenas informações relevantes à tarefa a ser executada.
- 3) Há várias maneiras diferentes de realizar uma tarefa.
- 4) Para utilizar todas as funcionalidades disponíveis, os itens apresentados foram suficientes.

Métrica - Facilidade de lembrar

- 5) A consistência das interfaces tem a ver com a repetição de certos padrões, para sequências de ações que devem ser repetidas.
- 6) A reutilização de uma ação após um tempo reduziu o tempo para a realização da tarefa.
- 7) Tive facilidade para lembrar o uso do

sistema após um período de tempo sem utilizá-lo.

8) Não é preciso lembrar informações de uma parte do sistema para outra.

Métrica - Controle de erros

9) As mensagens de erro são expressas em linguagem simples.

10) Há facilidade na correção de erros quando a interface fornece funções desfazer e refazer.

11) A quantidade de erros provocados pelo sistema é baixa.

Métrica - Eficiência

12) Tenho controle sobre as ações no sistema.

13) O sistema é muito produtivo.

14) As tarefas são efetuadas rapidamente.

Métrica - Eficácia

15) A quantidade de passos desnecessários ou redundantes no decorrer da tarefa é baixa.

16) Perde-se pouco tempo na realização de determinada tarefa.

17) As tarefas são eficientes e podem se adaptar ao meu gosto nas ações mais frequentes ou utiliza atalhos.

Métrica - Satisfação

18) Me sinto satisfeito a respeito da interação com a interface do sistema.

19) São fornecidas informações de ajuda, e estas informações são fáceis de procurar e de focalizar nas tarefas.

20) Me sinto satisfeito a respeito da realização das tarefas no sistema, com relação à clareza das mensagens, recuperação de erros.

3.4 Tratamento dos Dados

Após especificado o software a ser analisado e o público alvo, os questionários foram aplicados e os dados obtidos tabulados. Por fim, foi desenvolvido o software para analisar a usabilidade de aplicativo.

Para a avaliação, é determinado para cada uma das seis métricas um conceito no seguinte particionamento *fuzzy*: Muito Ruim, Ruim, Média, Boa, Muito Boa. As seis métricas serão analisadas separadamente podendo ocorrer de que se chegue à conclusão de que o sistema possua Facilidade de Lembrar igual a Boa

e Controle de Erros igual a Ruim.

Para o tratamento dos dados foram utilizados o Microsoft Excel e o Matlab. O uso do Microsoft Excel foi necessário para executar as duas primeiras etapas do tratamento dos dados que são: consolidar as opiniões dos usuários e determinar a frequência das opiniões dos usuários.

Para a execução das etapas seguintes foi desenvolvido um programa em MatLab, que recebe como entrada os dados iniciais que é a Frequência das Opiniões dos Usuários (FOU) e calcula a FOUN (Frequência das Opiniões dos Usuários), NN (Número Nebuloso), NNM (Número Nebuloso Médio), NNMN (Número Nebuloso Médio Normalizado) e por fim apresenta os resultados graficamente para que seja feita a análise dos resultados. Estas etapas foram realizadas da seguinte maneira:

1. Realizar a consolidação das opiniões dos usuários;
2. Atribuir Frequência para as Opiniões dos Usuários (FOU).

3. Realizar a Normalização da Frequência das Opiniões dos Usuários (FOUN). Para que seja feita esta normalização deve ser encontrada a frequência máxima e então dividir todos os valores da FOU por esta frequência máxima.

4. Encontrar o Número Nebuloso (NN) que mais se aproxima da FOU através da semelhança entre eles. O número é obtido dentre os números triangulares nebulosos que tiverem a mesma moda da FOUN. O Número Nebuloso Triangular é o número natural que pode ser representado na forma de triângulo equilátero. A representação gráfica de um Número Triangular Nebuloso com $a = 2$, $b = 4$ e $c = 6$ encontra-se na Figura 1.

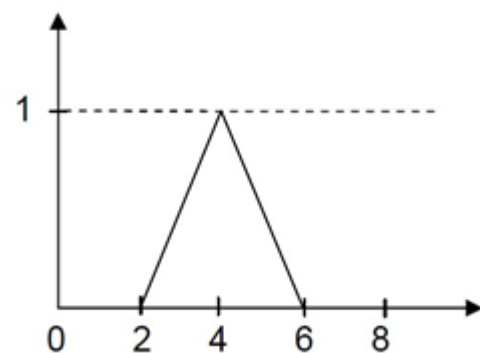


Figura 1. Número Triangular Nebuloso

Para que se possa determinar a semelhança entre o NN e a FOUN é necessário utilizar a seguinte fórmula proposta por [3]:

$$S_{A(x),B(x)} = \frac{|A(x) \cap B(x)|}{|A(x) \cup B(x)|} \quad (3)$$

Depois de encontrados os Números Nebulosos semelhantes a cada questão, a próxima etapa será determinar o conceito de usabilidade para a métrica como um todo. Devem-se selecionar os Números Triangulares Nebulosos para cada uma das perguntas referentes à métrica para posteriormente serem feitos os próximos passos.

5. Encontrar o Número Nebuloso Médio (NNM) para cada métrica, para isto será necessário o uso da seguinte fórmula proposta por [5]:

$$\mu_{NNM}(x_1, \dots, x_k) = \frac{1}{k} \sum_{i=1, \dots, k} \mu A(x_i) \quad (4)$$

6. Normalizar o Número Nebuloso Médio dividindo todos os valores de pertinência pela pertinência máxima desse mesmo número nebuloso. Desta forma é encontrado o Número Nebuloso Médio Normalizado (NNMN).

7. Encontrar o Número Nebuloso Médio Normalizado (NNMN) que mais se a próxima do Número Triangular Nebuloso (NN) de mesma moda para que se possa encontrar o Número Triangular Nebuloso final que irá representar a classificação da métrica. Com este Número Triangular Nebuloso encontrado na etapa anterior pode-se chegar a duas conclusões muito importantes:

- **Moda:** determina o conjunto em que a métrica estará classificada, da seguinte forma: 0 - Muito Ruim, 2 - Ruim, 4 - Média, 6 - Boa, 8 - Muito boa;
- **Amplitude:** determinará a dispersão média das opiniões dos usuários. Ela representa

o intervalo de confiança que se tem no valor da função de pertinência. Quanto menor a amplitude do intervalo, maior a confiança nos dados e quanto maior a amplitude do intervalo, menor a confiança que se tem nos dados [3]. A amplitude irá representar a dispersão das opiniões dos usuários da seguinte forma: 1 - Mínima, 2 - Média, 3 - Alta, 4 - Muito Alta.

Toda vez que a amplitude encontrada no final da métrica for maior ou igual a 2 (Média), recomenda-se que o sistema seja reavaliado sob o ponto de vista da métrica e que aumente a amostra dos usuários.

5. Resultados da Análise de Usabilidade

Esta sessão apresenta os resultados obtidos para cada um dos casos de teste. Os resultados são apresentados por meio de gráficos e tabelas. Para cada métrica é apresentado um gráfico que representa dois conjuntos. O primeiro é o Número Nebuloso Médio Normalizado (NNMN), que é a média de todos os números triangulares nebulosos encontrados para a métrica. O segundo representado pela linha verde, apresenta o número triangular nebuloso mais semelhante ao conjunto NNMN, sendo o resultado final para a métrica avaliada. Os resultados também são apresentados por meio de uma tabela.

5.1 Caso de Teste 1: Usuários Com Conhecimento Prévio do Software

- **Métrica Facilidade de Aprender:** O resultado final para a métrica Facilidade de Aprender apresentou o número triangular nebuloso (6—4), indicando uma avaliação boa, porém com uma dispersão muito alta entre as opiniões dos usuários. O que indica que mesmo que a avaliação tenha sido Boa, as opiniões dos usuários estão muito espalhadas, deixando o resultado final impreciso.

- **Métrica Facilidade de Relembrar:**

O resultado final para a métrica Facilidade de Relembrar apresentou o número triangular nebuloso (6—4), indicando uma avaliação boa, porém com uma dispersão muito alta das opiniões dos usuários.

- **Métrica Controle de Erros:** O resultado final para a métrica Controle de Erros obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Eficiência:** O resultado final para a métrica Eficiência obteve o número triangular nebuloso (6—2) como resposta, indicando uma avaliação boa, e a amplitude 2, representando uma dispersão média das opiniões dos usuários.

- **Métrica Eficácia:** O resultado final para a métrica Eficácia obteve o número triangular nebuloso (6—2) como resposta, indicando uma avaliação boa, e a amplitude 2, representando uma dispersão média das opiniões dos usuários.

- **Métrica Satisfação:** O resultado final para a métrica Satisfação obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

-
As métricas Eficiência e Eficácia apresentaram amplitude igual a 2 (média), o que representa um maior grau de confiança no resultado obtido. As métricas Controle de Erros e Satisfação apresentam amplitude igual a 3 (alta), o que representa um grau de confiança baixo. As métricas Facilidade de Aprender e Facilidade de Relembrar apresentaram a amplitude mais alta, que é a 4 (muito alta), o que representa uma dispersão das opiniões muito alta, fazendo com que esta métrica exija maior atenção na sua

reavaliação.

O resultado final da avaliação da usabilidade do Jabref para este caso de teste indica que as seis métricas de usabilidade não foram aceitas, pois apresentaram amplitude maior ou igual a 2, indicando que representam grau de confiança muito baixo. Os resultados obtidos são apresentados na Tabela 3 e Figuras 2 e 3.

Métrica	Avaliação	Amplitude	Resultado
Facilidade de Aprender	6	4	A amplitude é ≥ 2 . Reavaliar o sistema.
Facilidade de Relembrar	6	4	A amplitude é ≥ 2 . Reavaliar o sistema.
Controle de Erros	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficiência	6	2	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficácia	6	2	A amplitude é ≥ 2 . Reavaliar o sistema.
Satisfação	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.

Tabela 3. Usuários Com Conhecimento Prévio do Software.

5.2 Caso de Teste 2: Usuários Sem Conhecimento Prévio do Software com Orientação

- **Métrica Facilidade de Aprender:** O resultado final para a métrica Facilidade de Aprender obteve o número triangular nebuloso (4—3) como resposta, indicando uma avaliação média, com dispersão alta nas opiniões dos usuários.

- **Métrica Facilidade de Relembrar:** O resultado final para a métrica Facilidade de Relembrar obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com dispersão alta nas opiniões dos usuários.

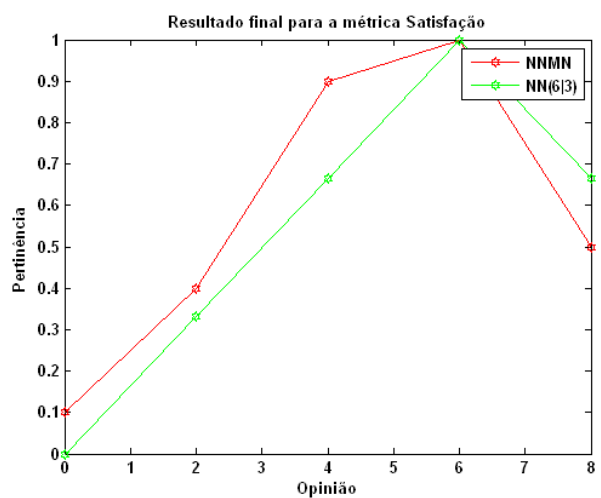
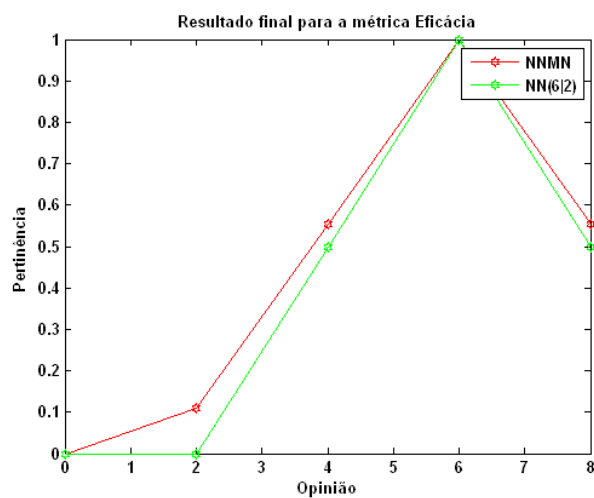
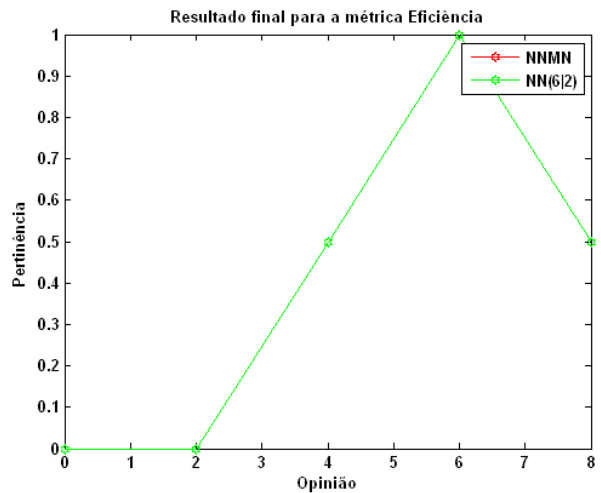
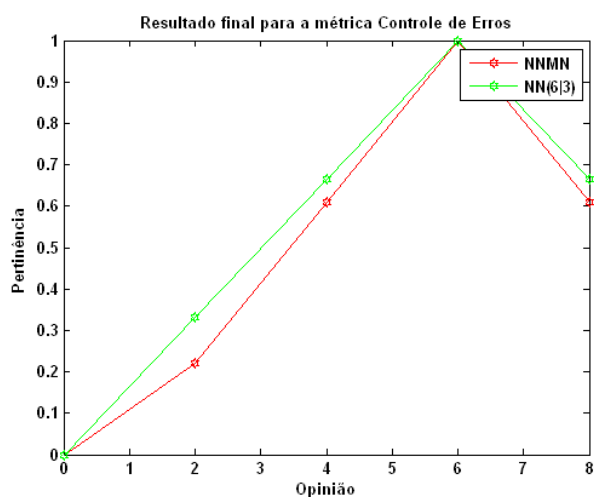
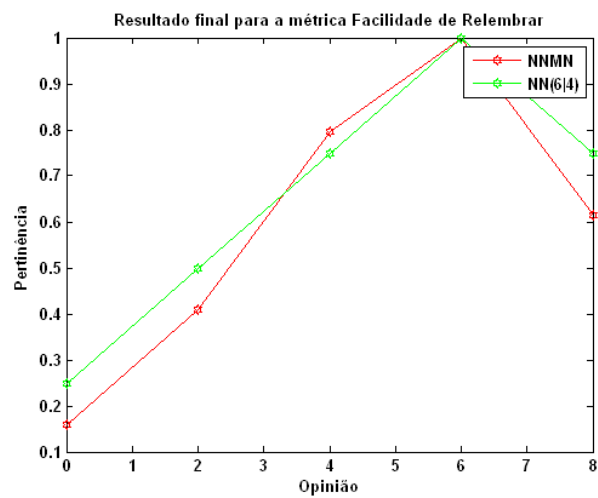
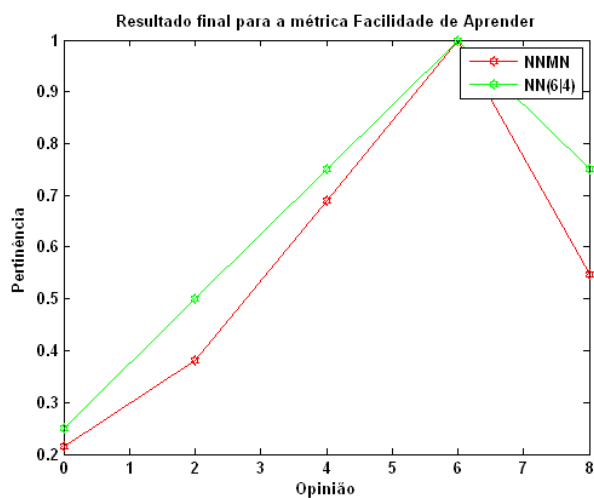


Figura 2. Usuários Com Conhecimento Prévio do Software.

Figura 3. Usuários Com Conhecimento Prévio do Software.

- **Métrica Controle de Erros:** O resultado final para a métrica Controle de Erros obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Eficiência:** O resultado final para a métrica Eficiência obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários. Observando o gráfico percebe-se que o NNMN se aproximou muito do número triangular nebuloso(6—3).

- **Métrica Eficácia:** O resultado final para a métrica Eficácia obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Satisfação:** O resultado final para a métrica Satisfação obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

As métricas Facilidade de Aprender, Facilidade de Relembrar, Eficiência, Eficácia e Satisfação apresentam amplitude igual a 3 (alta), o que representa um grau de confiança baixo.

Para este caso de teste o resultado final da avaliação da usabilidade indica que as seis métricas de usabilidade não foram aceitas, pois apresentaram amplitude maior ou igual a 2, indicando que representam grau de confiança muito baixo. Os resultados obtidos através da pesquisa estão apresentados na Tabela 4 e Figuras 4 e 5.

Métrica	Avaliação	Amplitude	Resultado
Facilidade de Aprender	4	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Facilidade de Relembrar	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Controle de Erros	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficiência	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficácia	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Satisfação	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.

Tabela 4. Usuários Sem Conhecimento Prévio do Software e Com Orientação

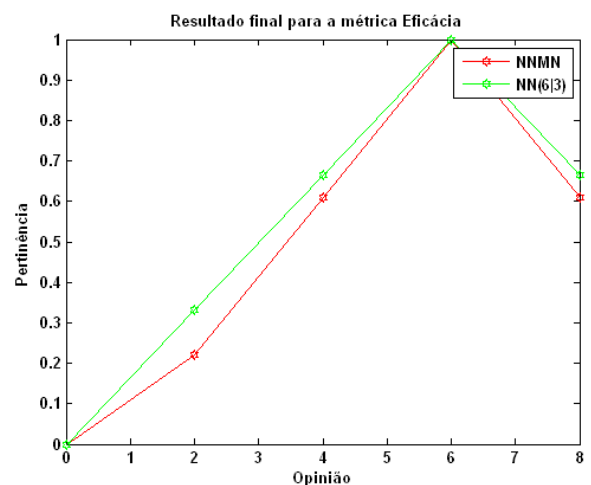
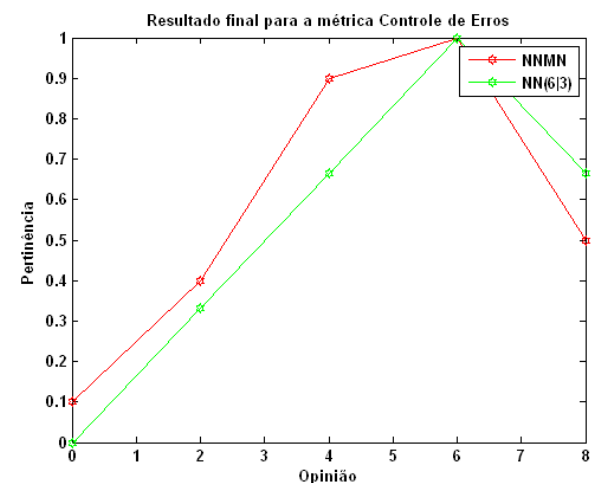
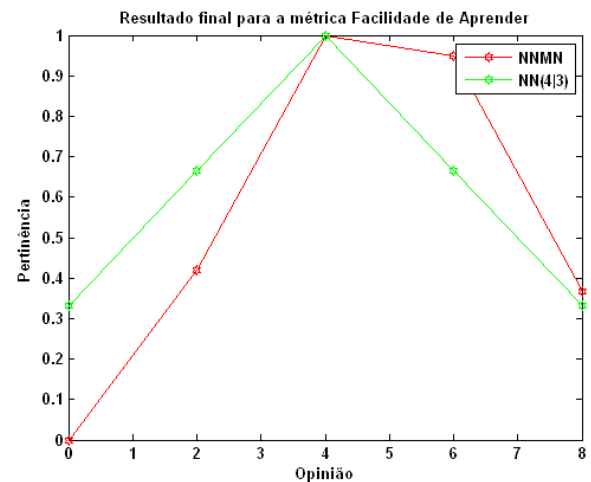


Figura 4. Usuários Sem Conhecimento Prévio do Software e Com Orientação.

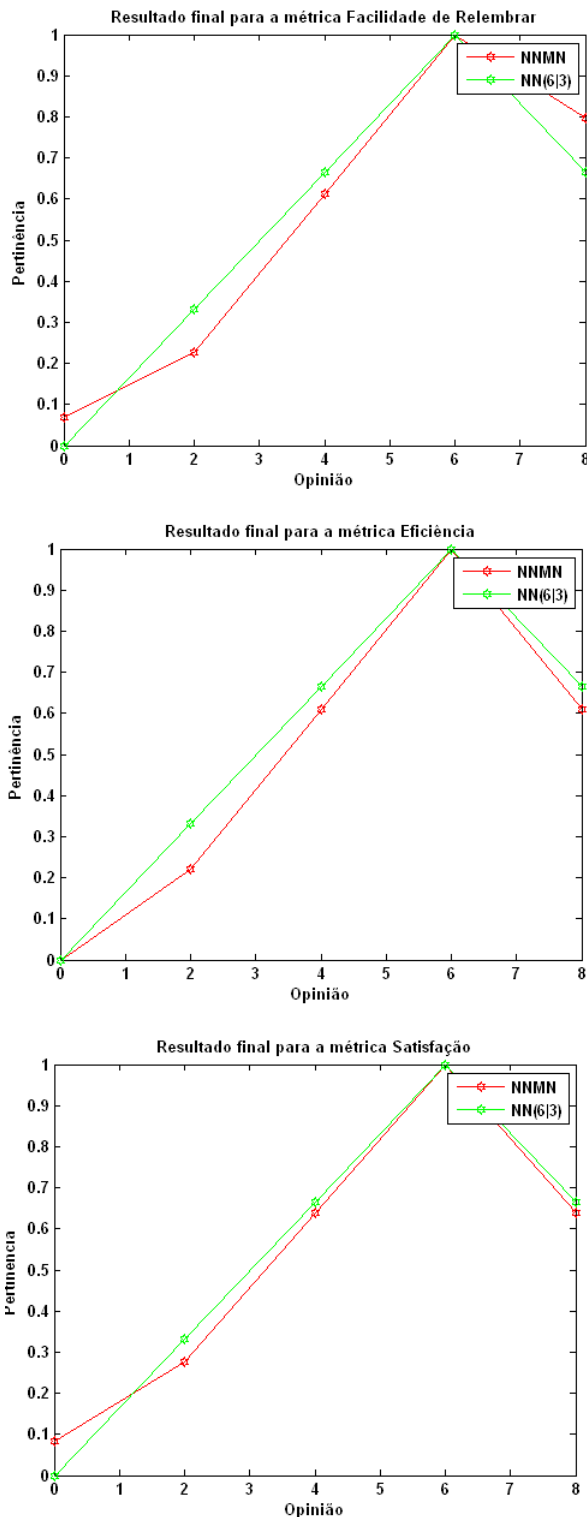


Figura 5. Usuários Sem Conhecimento Prévio do Software e Com Orientação.

5.3 Caso de Teste 3: Usuários Sem Conhecimento Prévio do Software e Sem Orientação

- **Métrica Facilidade de Aprender:** O O resultado final da avaliação da usabilidade

resultado final para a métrica Facilidade de Aprender obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Facilidade de Relembrar:** O resultado final para a métrica Facilidade de Relembrar obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Controle de Erros:** O resultado final para a métrica Controle de Erros obteve o número triangular nebuloso (4—3) como resposta, indicando uma avaliação média, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Eficiência:** O resultado final para a métrica Eficiência obteve o número triangular nebuloso (6—2) como resposta, indicando uma avaliação boa, porém com uma dispersão média das opiniões dos usuários.

- **Métrica Eficácia:** O resultado final para a métrica Eficácia obteve o número triangular nebuloso (6—3) como resposta, indicando uma avaliação boa, porém com uma dispersão alta das opiniões dos usuários.

- **Métrica Satisfação:** O resultado final para a métrica Satisfação obteve o número triangular nebuloso (8—4) como resposta, indicando uma avaliação muito boa, porém com uma dispersão muito alta das opiniões dos usuários.

A métrica eficiência apresentou amplitude igual a 2 (média), o que representa um maior grau de confiança no resultado obtido. As métricas Facilidade de Aprender, Facilidade de Relembrar, Controle de Erros e Eficácia apresentam amplitude igual a 3 (alta), o que representa um grau de confiança baixo. A métrica Satisfação apresentou a amplitude mais alta, que é a 4 (muito alta), o que representa uma dispersão das opiniões muito alta, fazendo com que esta métrica exija maior atenção na sua reavaliação.

do Jabref para este caso de teste indica que as seis métricas de usabilidade não foram aceitas, pois apresentaram amplitude maior ou igual a 2, indicando que representam grau de confiança muito baixo. Os resultados obtidos por meio da pesquisa estão apresentados na Tabela 5 e Figuras 6 e 7.

Métrica	Avaliação	Amplitude	Resultado
Facilidade de Aprender	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Facilidade de Relembrar	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Controle de Erros	4	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficiência	6	2	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficácia	6	3	A amplitude é ≥ 2 . Reavaliar o sistema.
Satisfação	8	4	A amplitude é ≥ 2 . Reavaliar o sistema.

Tabela 5. Usuários Sem Conhecimento Prévio do Software e Sem Orientação

Nos três casos de teste realizados o resultado final da avaliação da usabilidade do Jabref apresentaram amplitude maior ou igual a 2, o que representam grau de confiança muito baixo indicando que as seis métricas de usabilidade não foram aceitas. Baseado nos resultados obtidos por meio deste trabalho é recomendado que o Jabref seja repensado sobre os aspectos de usabilidade apresentados, ou que seja feita uma nova avaliação aumentando as amostras dos usuários. A comparação dos resultados é apresentada na Tabela 6.

Métrica	Caso Teste 1	Caso Teste 2	Caso Teste 3	Resultado
Facilidade de Aprender	6—4	4—3	6—3	A amplitude é ≥ 2 . Reavaliar o sistema.
Facilidade de Relembrar	6—4	6—3	6—3	A amplitude é ≥ 2 . Reavaliar o sistema.
Controle de Erros	6—3	6—3	4—3	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficiência	6—2	6—3	6—2	A amplitude é ≥ 2 . Reavaliar o sistema.
Eficácia	6—2	6—3	6—3	A amplitude é ≥ 2 . Reavaliar o sistema.
Satisfação	6—3	6—3	8—4	A amplitude é ≥ 2 . Reavaliar o sistema.

Tabela 6. Comparação dos resultados

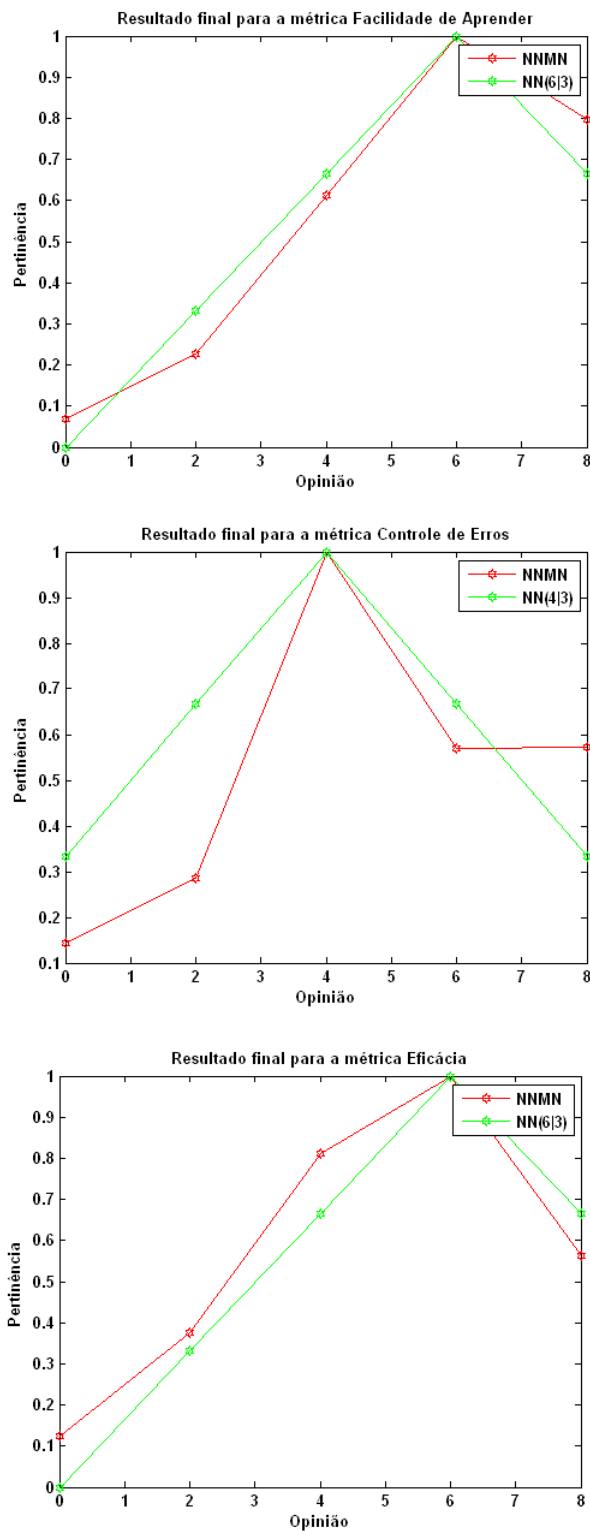


Figura 6. Usuários Sem Conhecimento Prévio do Software e Sem Orientação.

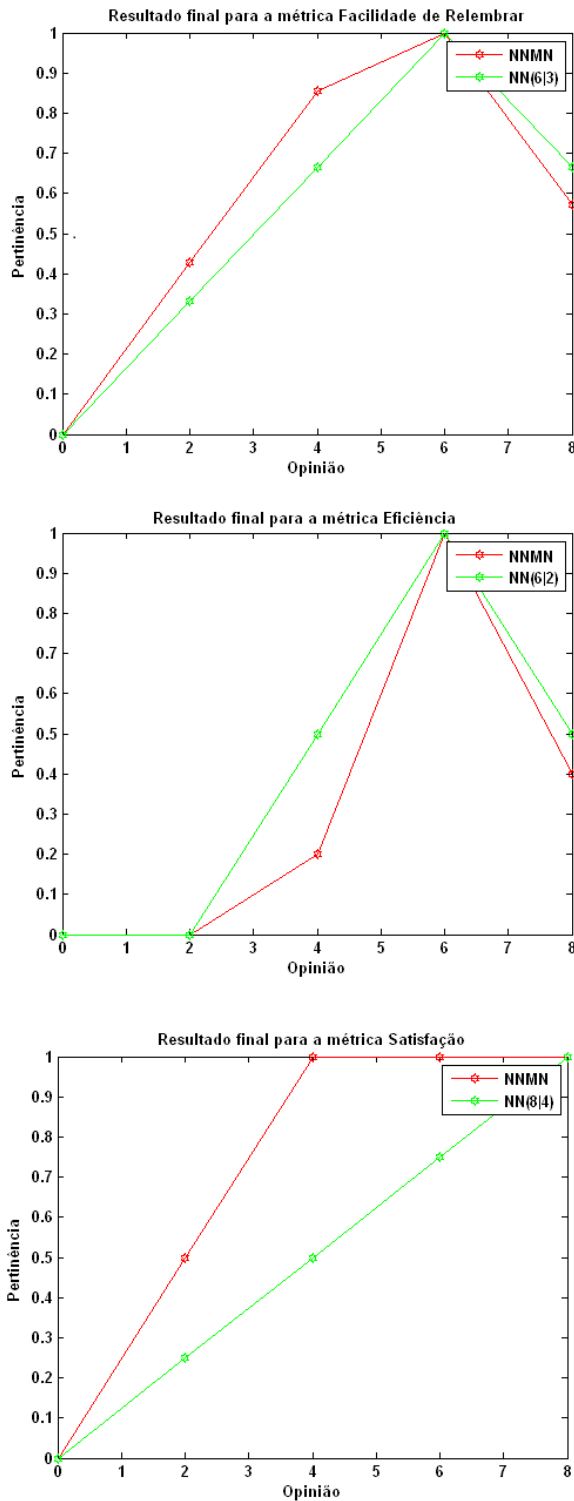


Figura 7. Usuários Sem Conhecimento Prévio do Software e Sem Orientação.

6. Considerações Finais

O presente trabalho teve como objetivo avaliar a usabilidade de software empregando pesquisa aplicada e lógica nebulosa. Os resultados obtidos por meio da metodologia proposta atenderam às expectativas de se criar um instrumento para avaliar a usabilidade do Jabref, deixando claros os pontos nos quais o software deve ser reavaliado.

Os resultados obtidos com este trabalho visam apoiar este processo de verificação da usabilidade de software. Estas verificações são necessárias para satisfazer as necessidades dos usuários. O sistema desenvolvido poderá ser utilizado para avaliação de usabilidade de outros softwares. Dependendo do software analisado, estes resultados podem ser interpretados como redução de chamadas de suporte, aumento da aceitação, aumento de lucro, retenção de clientes, entre outros benefícios.

Como sugestão para dar continuidade ao trabalho é recomendado o desenvolvimento de um aplicativo utilizando os conceitos do trabalho, onde o usuário poderá responder o questionário e o resultado da análise de usabilidade do software já será apresentado ao usuário.

Referências

- ABNT. *NBRISO/IEC9126-1 Engenharia de software - Qualidade de produto - Parte 1: Modelo de qualidade*. Associação Brasileira de Normas Técnicas, 2003.
- BARCELOS, R. F. **Uma abordagem para inspeção de documentos arquiteturais baseada em checklist**. (Tese de Mestrado) – Programa de Pós-Graduação de Engenharia da Universidade Federal do Rio de Janeiro, 2006.
- BRAGA, M. J. F.; BARRETO, J. M.; MACHADO, M. A. S. **Conceitos da Matemática Nebulosa na Análise de Risco**. Rio de Janeiro: Artes & Rabiskus, 1995.
- CYBIS, W. A. **Usabilidade de interfaces humano-computador**. [s.l], 2000.
- DUBOIS, D.; PRADE, H. **Fuzzy Sets and Systems**:

Teory and Applications. San Francisco, USA: Academic Press, 1980.

FREITAS, R. C. *et al.*. Representando a IHM de uma subestação através de modelos formais e realidade virtual. *In: SBSE – SIMPÓSIO BRASILEIRO DE SISTEMAS ELÉTRICOS*, p. 1-7, 2006.

GUELL, N.; SCHWABE, D.; BARBOSA, S. D. J. Método de avaliação de usabilidade na web baseado em modelo e padrões de comportamento. *In: Proceedings of the 7th Brazilian Symposium on Multimedia and Hypermedia Systems, SBMIDIA 2001*. Florianópolis, SC, 1:1 – 8, 2001.

ISO. *Ergonomic requirements for office work with visual display terminals (VDTs)*. ISO 9241. International Organization for Standardization, 1998.

JORDAN, P. W. **An Introducion to usability.** London, UK: Taylor & Francis, 1998.

LESSA, R. O. *et al.*. *Glist: Checklist automatizado para usabilidade*. *In: SEMINÁRIO DE INFORMÁTICA, SEMINFO-RS 2006*. Torres, RS, Brasil, 2006.

MATTAR, F. **Pesquisa de Marketing.** São Paulo: Editora Atlas, 1997.

MILLER, R. B. Human ease of use criteria and their tradeoffs. Technical report, IBM Corporation, 1971.

ONGARO, E.; CANAL, A. P. Usability techniques applied to the *softwares* of zipline technology corp. **Disc. Scientia. Série: Ciências Naturais e Tecnológicas**, 5(1):163 – 183, 2004.

PEDRYCZ, W.; GOMIDE, F. **Fuzzy Systems Engineering: Toward Human-Centric Computing.** [s.l.] John Wiley & Sons, 1 edition, 2007.

ROCHA, H. V.; BARANAUSKAS, M. C. **Design e avaliação de interfaces humano computador.** Campinas: NIED - UNICAMP, 2003.

SANTOS, R. C. **Desenvolvimento de uma metodologia para avaliação de usabilidade de**

sistemas utilizando a lógica fuzzy baseado na ISO. (Tese de Mestrado) – IBMEC Educacional S.A, Brasil, 2007.

SANTOS, R. C. *et al.* Lógica fuzzy aplicada na análise de usabilidade de sistemas computacionais embarcados. *In: SPOLM 2007 – SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA*. Rio de Janeiro: CASNAV – Marinha do Brasil, p. 1-10, 2007.