



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Percepção dos desenvolvedores quanto a segurança de seus códigos: análise de falsos-positivos reportados pelo CogniCrypt-Eclipse**

Guilherme Andreúce S. Monteiro

Monografia apresentada como requisito parcial  
para conclusão do Bacharelado em Ciência da Computação

Orientador

Prof. Dr. Rodrigo Bonifacio de Almeida

Brasília  
2022



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Percepção dos desenvolvedores quanto a segurança de seus códigos: análise de falsos-positivos reportados pelo CogniCrypt-Eclipse**

Guilherme Andreúce S. Monteiro

Monografia apresentada como requisito parcial  
para conclusão do Bacharelado em Ciência da Computação

Prof. Dr. Rodrigo Bonifacio de Almeida (Orientador)  
CIC/UnB

Prof. Dr. Donald Knuth    Dr. Leslie Lamport  
Stanford University    Microsoft Research

Prof. Dr. Marcelo Grandi Mandelli  
Coordenador do Bacharelado em Ciência da Computação

Brasília, 10 de março de 2022

# Dedicatória

Na *dedicatória* o autor presta homenagem a alguma pessoa (ou grupo de pessoas) que têm significado especial na vida pessoal ou profissional. Por exemplo (e citando o poeta):  
*Eu dedico essa música a primeira garota que tá sentada ali na fila. Brigado!*

# Agradecimentos

Nos *agradecimentos*, o autor se dirige a pessoas ou instituições que contribuíram para elaboração do trabalho apresentado. Por exemplo: *Agradeço aos gigantes cujos ombros me permitiram enxergar mais longe. E a Google e Wikipédia.*

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

# Resumo

O *resumo* é um texto inaugural para quem quer conhecer o trabalho, deve conter uma breve descrição de todo o trabalho (apenas um parágrafo). Portanto, só deve ser escrito após o texto estar pronto. Não é uma coletânea de frases recortadas do trabalho, mas uma apresentação concisa dos pontos relevantes, de modo que o leitor tenha uma ideia completa do que lhe espera. Uma sugestão é que seja composto por quatro pontos: 1) o que está sendo proposto, 2) qual o mérito da proposta, 3) como a proposta foi avaliada/validada, 4) quais as possibilidades para trabalhos futuros. É seguido de (geralmente) três palavras-chave que devem indicar claramente a que se refere o seu trabalho. Por exemplo: *Este trabalho apresenta informações úteis a produção de trabalhos científicos para descrever e exemplificar como utilizar a classe L<sup>A</sup>T<sub>E</sub>X do Departamento de Ciência da Computação da Universidade de Brasília para gerar documentos. A classe UnB-CIC define um padrão de formato para textos do CIC, facilitando a geração de textos e permitindo que os autores foquem apenas no conteúdo. O formato foi aprovado pelos professores do Departamento e utilizado para gerar este documento. Melhorias futuras incluem manutenção contínua da classe e aprimoramento do texto explicativo.*

**Palavras-chave:** CogniCrypt, Eclipse, Segurança do código, Análise

# Abstract

O *abstract* é o resumo feito na língua Inglesa. Embora o conteúdo apresentado deva ser o mesmo, este texto não deve ser a tradução literal de cada palavra ou frase do resumo, muito menos feito em um tradutor automático. É uma língua diferente e o texto deveria ser escrito de acordo com suas nuances (aproveite para ler [http://dx.doi.org/10.6061/2Fclinics%2F2014\(03\)01](http://dx.doi.org/10.6061/2Fclinics%2F2014(03)01)). Por exemplo: *This work presents useful information on how to create a scientific text to describe and provide examples of how to use the Computer Science Department's L<sup>A</sup>T<sub>E</sub>X class. The UnB-CIC class defines a standard format for texts, simplifying the process of generating CIC documents and enabling authors to focus only on content. The standard was approved by the Department's professors and used to create this document. Future work includes continued support for the class and improvements on the explanatory text.*

**Keywords:** CogniCrypt, Eclipse, Code Security, Analysis

# Sumário

<b>1</b>	<b>Fundamentação Teórica</b>	<b>1</b>
1.1	Gerando o PDF . . . . .	1
1.2	Opções . . . . .	1
1.3	Informações do Trabalho . . . . .	2
1.4	Arquivos . . . . .	3
1.5	Documento . . . . .	3
1.5.1	Capítulos . . . . .	3
1.5.2	Figuras . . . . .	5
1.5.3	Equações . . . . .	7
1.5.4	Tabelas . . . . .	8
1.5.5	Abreviaturas e Siglas . . . . .	10
	<b>Referências</b>	<b>12</b>
	<b>Apêndice</b>	<b>12</b>
<b>A</b>	<b>Fichamento de Artigo Científico</b>	<b>12</b>
	<b>Anexo</b>	<b>16</b>
<b>I</b>	<b>Documentação Original UnB-CIC (parcial)</b>	<b>16</b>

# Lista de Figuras



# Lista de Tabelas

# Lista de Abreviaturas e Siglas

**BCE** Biblioteca Central.

**CIC** Departamento de Ciência da Computação.

# Capítulo 1

## Fundamentação Teórica

Este capítulo descreve a classe **UnB-CIC**, e demonstra os comandos disponíveis. A última versão foi atualizada pelo Prof. Ralha, em 2008 (vide Anexo I). A melhor forma de entender o funcionamento é observar o arquivo principal deste documento (**monografia.tex**).

### 1.1 Gerando o PDF

Para gerar corretamente as referências cruzadas, é necessário processar os arquivos mais de uma vez com a seguinte sequência de comandos (supondo que o arquivo principal seja **monografia.tex**).

```
pdflatex monografia
bibtex monografia
makeglossaries monografia
pdflatex monografia
```

O primeiro comando processa os arquivos, indicando quais referências foram citadas no texto (bibliográficas ou cruzadas), o segundo comando processa o arquivo **.bib** que contém as informações bibliográficas, o terceiro gera o índice de siglas/abreviaturas, e o último comando junta todas estas informações, produzindo um texto com referências cruzadas funcionais.

### 1.2 Opções

O documento é gerado em função do curso dado como opção [obrigatória] a classe. Os cursos disponíveis são:

**bacharelado** Bacharelado em Ciência da Computação

**licenciatura** Licenciatura em Computação

**engenharia** Engenharia de Computação

**mestrado, ppginf** Mestrado em Informática

**doutorado, ppginf** Doutorado em Informática

**mestrado, ppca** Mestrado Profissional em Computação Aplicada

No caso dos cursos de pós-graduação, há o *exame de qualificação* do discente, a qual deverá constar a definição, pertinência do projeto, a sua abrangência, comprovação da eficiência e eficácia da metodologia proposta, uma revisão bibliográfica detalhada e o cronograma para conclusão do projeto [?]. Para gerar o documento referente a este exame, use a opção **qualificacao**.

## 1.3 Informações do Trabalho

O passo seguinte é definir as informações do trabalho, identificando os autores e os membros da banca (atenção a definição do gênero!). Por exemplo, para este documento foram utilizadas as seguintes definições:

```
\orientador{\prof \dr Guilherme Novaes Ramos}{CIC/UnB}%  
%\coorientador{\prof \dr José Ralha}{CIC/UnB}  
\coordenador[a]{\prof[a] \dr[a] Ada Lovelace}{Bibliothèque universelle de Genève}%  
\diamesano{24}{dezembro}{2014}%
```

```
\membrobanca{\prof \dr Donald Knuth}{Stanford University}%  
\membrobanca{\dr Leslie Lamport}{Microsoft Research}%
```

```
\autor{Guilherme N.}{Ramos}%
```

Sobre o texto, definiu-se:

```
\titulo{UnB-CIC: Uma classe em LaTeX para textos do Departamento de  
Ciência da Computação}%
```

```
\palavraschave{LaTeX, metodologia científica}%  
\keywords{LaTeX, scientific method}%
```

O título, apesar do tamanho reduzido, deveria apresentar uma ideia clara de todo o trabalho. As palavras-chave devem indicar os conceitos genéricos mais relevantes utilizados, e servem para indexação e busca de documentos que tratam os mesmos temas.

## 1.4 Arquivos

Os seguintes arquivos são exigidos:

**tex/abstract.tex** Contém o *abstract* do texto.

**tex/agradecimentos.tex** Contém os agradecimentos do autor.

**bibliografia.bib** Contém as referências bibliográficas no formato BibT<sub>E</sub>X<sup>1</sup>.

**tex/dedicatoria.tex** Contém a dedicatória do autor.

**tex/siglas.tex** Contém as definições de siglas/abreviaturas.

**tex/resumo.tex** Contém o resumo do texto.

Os alunos dos Programas de Pós-Graduação da Universidade de Brasília devem incluir a ficha catalográfica em seus documentos, gerada pela Biblioteca Central (BCE). Neste caso, o aluno deve substituir o arquivo PDF **doc/BDM.pdf** pelo fornecido pela BCE. *Atenção*, para que o arquivo seja incluído automaticamente pela classe, o nome deve ser *obrigatoriamente* **BDM.pdf**.

Demais arquivos não são inseridos automaticamente, mas a classe oferece comandos para inclusão, facilitando a organização destes.

## 1.5 Documento

Todo documento em L<sup>A</sup>T<sub>E</sub>X é delimitado pelo ambiente *document*. O caso aqui não é diferente, mas a interação é simplificada. Basicamente, a classe **UnB-CIC** funciona “automaticamente” em função dos comandos e dos nomes dos arquivos.

### 1.5.1 Capítulos

O texto de cada capítulo deve estar em seu próprio arquivo, dentro do diretório correto **tex**. A inclusão do texto é feita pelo comando:

```
\capitulo{arquivo}{título}%
```

Os dois argumentos são:

**arquivo** argumento obrigatório que define o nome do arquivo que contém o texto do capítulo.

---

<sup>1</sup><http://www.bibtex.org>

**título** argumento obrigatório que define o título do capítulo.

Por exemplo, este texto está no arquivo `2_UnB-CIC.tex`, e para criar os dois capítulos vistos até agora, o documento seria:

```
\begin{document}%
  \capitulo{1_Introducao}{Introdução}% inclui o arquivo 1_Introducao.tex
  \capitulo{2_UnB-CIC}{A Classe \unbcic}% inclui o arquivo 2_UnB-CIC.tex
\end{document}%
```

Para incluir um terceiro capítulo neste texto, cujo conteúdo trata de trabalhos conclusão de curso, basta criar o arquivo `tex/3_TCC.tex` e adicioná-lo com o comando descrito.

No caso de apêndices ou anexos necessários, o texto de cada um deve estar em seu próprio arquivo, também dentro do diretório `tex/capitulos`. Para facilitar as referências cruzadas, estes devem ser incluídos com os seguintes comandos (respectivamente):

```
\apendice{arquivo}{título}%
\anexo{arquivo}{título}%
```

Os dois argumentos funcionam exatamente como `\capitulo`. Desta forma, o exemplo de um documento “completo” seria:

```
\begin{document}%
  \capitulo{1_Introducao}{Introdução}%
  \capitulo{2_UnB-CIC}{A Classe \unbcic}%
  \capitulo{3_TCC}{Trabalho de Conclusão de Curso}%

  \apendice{Apendice_Fichamento}{Fichamento de Artigo Científico}%
  \anexo{Anexo1}{Parte da Documentação Original}%
\end{document}%
```

Usando estes comandos, o rótulo de cada capítulo/apêndice/anexo é criado automaticamente a partir do nome do arquivo para posterior referência cruzada. Por exemplo, este capítulo pode ser referenciado com o comando `\ref{2_UnB-CIC}` (cujo resultado é: ??), mas a classe oferece opções mais interessantes. Os comandos para referenciar capítulos são:

```
\refCap{referência}%
\refCaps{referência inicial}{referência final}%
```

Onde os argumentos são:

**referência** nome da referência do capítulo.

**referência inicial** nome da referência do capítulo inicial da sequência de capítulos.

**referência final** nome da referência do capítulo final da sequência de capítulos.

O Capítulo ?? é referenciado com o comando:

```
\refCap{1_Introducao}%
```

Considerando Capítulo ?? e também o Capítulo ??, é possível referenciar a *sequência* de Capítulos ?? a ?? com o comando:

```
\refCaps{1_Introducao}{2_UnB-CIC}%
```

Embora estes comandos não “simplifiquem” a inclusão de figuras, eles certamente facilitam a referência a elas com um padrão uniforme, e nada impede o uso dos comandos padrões.

## 1.5.2 Figuras

Para manter a organização dos arquivos de seu documento, as figuras devem ficar separadas no diretório `img`. As funções de inclusão de figuras permanecem as mesmas, mas a classe `UnB-CIC` oferece uma forma mais simples de inserir uma figura (e de referenciá-la). Basta executar o comando:

```
\figura[posição]{arquivo}{legenda}{referência}{tamanho}%
```

Os 5 argumentos são:

**posição** argumento [opcional] para posicionar a figura no texto<sup>2</sup>.

**arquivo** nome do arquivo da imagem.

**legenda** legenda da figura.

**referência** nome da referência da figura para referências cruzadas.

**tamanho** tamanho da imagem<sup>3</sup>.

Por exemplo, a Figura 1.1, inserida com o seguinte comando:

```
\figura[!h]{contorno_preto}{Marca P/B}{unbPB}{width=0.5\textwidth}%
```



Figura 1.1: Marca P/B.

Os comandos para referenciar figuras são:

```
\refFig{referência}%
\refFigs{referência inicial}{referência final}%
```

Onde os argumentos são:

**referência** nome da referência da figura.

**referência inicial** nome da referência da figura inicial da sequência de figuras.

**referência final** nome da referência da figura final da sequência de figuras.

A Figura 1.1 é referenciada com o comando:

```
\refFig{unbPB}%
```



Figura 1.2: Marca colorida.

Considerando a Figura 1.2 e também a Figura 1.3, é possível referenciar a *sequência* de Figuras 1.1 a 1.3 com o comando:

```
\refFigs{unbPB}{unb2}%
```

---

<sup>2</sup>Mais informações na documentação do ambiente *figure*, mas este é um bom começo: [http://en.wikibooks.org/wiki/LaTeX/FLOATS,\\_FIGURES\\_and\\_CAPTIONS](http://en.wikibooks.org/wiki/LaTeX/FLOATS,_FIGURES_and_CAPTIONS).

<sup>3</sup>Mais informações na documentação do comando `\includegraphics`.



Algumas vezes deseja-se usar a figura de uma das referências bibliográficas. Neste caso, utilize o comando:

```
\figuraBib[posição]{arquivo}{legenda}{bib}{referência}{tamanho}%
```

Os argumentos são os mesmos do comando `\figura`, acrescidos de:

**bib** nome da referência bibliográfica que originou a figura.

Por exemplo, a Figura ?? foi gerada com o comando:

```
\figuraBib{miktex}{\LaTeX\ vs MS Word}
{pinteric_latex_2004}{latexvsword}{width=.45\textwidth}%
```

Embora estes comandos não “simplifiquem” a inclusão de figuras, eles certamente facilitam a referência a elas com um padrão uniforme, e nada impede o uso dos comandos padrões.



Figura 1.3: Outra marca colorida.

### 1.5.3 Equações

As funções de inclusão de equações permanecem as mesmas, mas a classe `UnB-CIC` oferece uma forma mais simples de inserir uma equação (e de referenciá-la). Basta executar o comando:

```
\equacao{referência}{fórmula}%
```

Os 2 argumentos são:

**referência** nome da referência da equação para referências cruzadas.

**fórmula** a equação em si.

Por exemplo, a Equação 1.1, inserida com o seguinte comando:

```
\equacao{pitagoras}{a^2 + b^2 = c^2}%
```

$$a^2 + b^2 = c^2 \quad (1.1)$$

Além disso, é possível quebrar em linhas, como na Equação 1.2, com o mesmo comando:

```
\equacao{pit2}{a = (x+y)^2\\b= (x*y)^2}%
```

$$\begin{aligned} a &= (x + y)^2 \\ b &= (x * y)^2 \end{aligned} \quad (1.2)$$

Os comandos para referenciar equações são:

```
\refEq{referência}%
```

```
\refEqs{referência inicial}{referência final}%
```

Onde os argumentos são:

**referência** nome da referência da equação.

**referência inicial** nome da referência da equação inicial da sequência de equações.

**referência final** nome da referência da equação final da sequência de equações.

Considerando a Equação 1.1 e também a Equação 1.3, é possível referenciar a *sequência* de Equações 1.1 a 1.3 com o comando:

```
\refEqs{pitagoras}{eq}%
```

Embora estes comandos não “simplifiquem” a inclusão de equações, eles certamente facilitam a referência a elas com um padrão uniforme e nada impede o uso dos comandos padrões.

$$d = c^3 - \frac{a}{b} \quad (1.3)$$

#### 1.5.4 Tabelas

As funções de inclusão de tabelas permanecem as mesmas, mas a classe **UnB-CIC** oferece uma forma mais simples de inserir uma tabela (e de referenciá-la). Basta executar o comando:

```
\tabela{legenda}{referência}{especificações}{tabela}%
```

Os 4 argumentos são:

Tabela 1.1: Exemplo de tabela.

Item	Descrição
1	Descrição 1
2	Descrição 2
3	Descrição 3

**legenda** legenda da tabela.

**referência** nome da referência da tabela para referências cruzadas.

**especificações** alinhamento de cada coluna da tabela.

**tabela** o conteúdo da tabela<sup>4</sup>.

Por exemplo, a Tabela 1.1, inserida com o seguinte comando:

```
\tabela{Exemplo de tabela}{exemplo}{| c | c |}%  
  {\hline  
  \textbf{Item} & \textbf{Descrição} \\\hline  
  1 & Descrição 1 \\\hline  
  2 & Descrição 2 \\\hline  
  3 & Descrição 3 \\\hline}%
```

Os comandos para referenciar tabelas são:

```
\refTab{referência}%  
\refTabs{referência inicial}{referência final}%
```

Onde os argumentos são:

**referência** nome da referência da tabela.

**referência inicial** nome da referência da tabela inicial da sequência de tabelas.

**referência final** nome da referência da tabela final da sequência de tabelas.

Considerando a Tabela 1.1 e também a Tabela 1.3, é possível referenciar a *sequência* de Tabelas 1.1 a 1.3 com o comando:

```
\refTabs{exemplo}{exemplo2}%
```

Algumas vezes deseja-se usar a tabela de uma das referências bibliográficas. Neste caso, utilize o comando:

---

<sup>4</sup>Mais informações na documentação do ambiente *tabular*.

Tabela 1.2: Matriz de Decisão de Eisenhower (Fonte: [?]).

	<b>Urgente</b>	<b>Não Urgente</b>
<b>Importante</b>	Crises	Planejamentos
<b>Não importante</b>	Interrupções	Distrações

Tabela 1.3: Outro exemplo de tabela.

#	<b>A</b>	<b>B</b>	<b>Comentário</b>
1	$a_1$	$b_1$	comentário 1
2	$a_2$	$b_2$	comentário 2
3	$a_3$	$b_3$	comentário 3

`\tabelaBib{legenda}{bib}{referência}{especificações}{tabela}%`

Os argumentos são os mesmos do comando `\tabela`, acrescidos de:

**bib** nome da referência bibliográfica que originou a tabela.

Por exemplo, a Tabela 1.2<sup>5</sup> foi gerada com o comando:

```
\tabelaBib{Matriz de Decisão de Eisenhower}
{covey_first_1995}{EisenhowerTable}{ r | c | c }{%
          & \textbf{Urgente} & \textbf{Não Urgente} \\ \hline%
\textbf{Importante}      & Crises      & Planejamentos \\ \hline%
\textbf{Não importante} & Interrupções & Distrações%
}%
```

Embora estes comandos não “simplifiquem” a inclusão de tabelas, eles certamente facilitam a referência a elas com um padrão uniforme, e nada impede o uso dos comandos padrões.

### 1.5.5 Abreviaturas e Siglas

Abreviaturas e siglas devem ser definidas no arquivo `tex/siglas.tex`, e a inserção feita com o comando:

`\sigla{sigla}{descrição}%`

Onde os argumentos são:

**sigla** a própria sigla/abreviatura.

---

<sup>5</sup>Vale a pena assistir o vídeo da palestra *Time Management* de Randy Pausch: <http://www.cs.virginia.edu/~robins/Randy/>

**descrição** definição completa do que representa a sigla/abreviatura.

Por exemplo:

```
\sigla{CIC}{Departamento de Ciência da Computação}%
```

A inserção de uma sigla/abreviatura no texto é simples, e pode ser feita de três formas diferentes:

<code>\acrshort{CIC}</code>	CIC
<code>\acrlong{CIC}</code>	Departamento de Ciência da Computação
<code>\acrfull{CIC}</code>	Departamento de Ciência da Computação (CIC)

# Referências

- [1] PPGI/UnB: *Regimento do programa de pós-graduação em informática da universidade de Brasília*. <http://ppgi.unb.br/normatives>, 2013.
- [2] Covey, Stephen R, A Roger Merrill e Rebecca R Merrill: *First things first*. Simon and Schuster, 1995.

# **Apêndice A**

## **Fichamento de Artigo Científico**



# Fichamento de Artigo Científico

*Prof. Guilherme N. Ramos*

Um fichamento reúne elementos relevantes do conteúdo, apresentando a estrutura do texto, e deve seguir a seqüência do pensamento do autor, destacando suas ideias, argumentos, justificativas, exemplos, fatos, etc.

## 1 Artigo Científico

Geralmente, um *artigo científico* é escrito com a seguinte estrutura (buscando responder algumas questões):

### I. Introdução

- Qual o contexto do problema? (O que? Onde? Quando?)
- Qual a principal questão ou problema colocado? (Por quê? Como? Qual?)
- Qual o objetivo visado? O que se pretende constatar ou demonstrar? (investigar, analisar, refletir, contribuir,...)

### II. Referencial Teórico

- Quais são os autores/teorias/conceitos que já estudaram os principais assuntos abordados e que sustentam ao texto?
- Quais os resultados mais recentes relacionados a eles?

### III. Metodologia/Desenvolvimento

- Quais os procedimentos metodológicos adotados? (natureza do trabalho: empírico, teórico, histórico) – (coleta de dados: questionário, entrevista, levantamento bibliográfico).
- Como a pesquisa foi desenvolvida? Quais as principais relações entre teoria e prática?
- Havendo artefato proposto, ele está disponível para utilização e/ou modificação?

### IV. Resultados

- Houve validação (por meio de experimentação)? Como foi feita?
- Os resultados obtidos são corretos/válidos?

### V. Conclusões

- Qual o problema atacado?
- Quais os resultados obtidos para os objetivos propostos?
- Quais conclusões podem ser tiradas destes resultados?
- Quais as limitações da metodologia utilizada?
- Quais as possibilidades de trabalhos futuros para o problema?



## 2 Fichamento

Neste contexto, um fichamento deve conter a seguinte estrutura:

1. **Identificação do aluno:** indicação precisa de quem é o autor do fichamento.
2. **Identificação do texto:** indicação precisa de quem são os autores do texto analisado e dos detalhes do documento, de modo que se possa buscá-lo para uma leitura completa.
3. **Pontos-chave:** noções mais relevantes do texto analisado. *Proposta* (o que é apresentado?), *mérito* (por que é relevante?), *validação* (como verificar a utilidade?), e *perspectivas* (o que pode ser melhorado?).
4. **Palavras-chave:** expressões que identificam o assunto abordado.
5. **Sinopse do texto:** resumo *com suas palavras*. Deve ser mais detalhado que um *abstract*, geralmente apresentando pelo menos um parágrafo por seção do texto original. No caso de inclusão de trechos, o texto deve ser identificado entre “aspas” e concatenado através de suas próprias palavras.
6. **Análise crítica:** posicionar-se em relação as seguintes questões: pertinência do assunto; forma como foi abordado; comparação com outras abordagens do mesmo assunto (caso conheça). Junto ao *resumo*, é a parte mais interessante para o leitor, pois apresenta uma avaliação do conteúdo apresentado.

### 2.1 Exemplo

1. **Identificação do aluno:** Alan Mathison Turing, 00/000000
2. **Identificação do texto:** Guilherme N. Ramos, Yutaka Hatakeyama, Fangyan Dong, and Katoru Hirota, Hyperbox clustering with Ant Colony Optimization (HACO) method and its application to medical risk profile recognition, Applied Soft Computing, Vol. 9, Issue 2, pp 632-640, 2009. (doi:10.1016/j.asoc.2008.09.004)
3. **Pontos-chave:**
  - Proposta:** HACO - método para aglomeração de dados utilizando hipercaixas com posicionamento otimizado via algoritmo de colônia de formigas.
  - Mérito:** apresenta uma nova forma de fazer agrupamentos considerando a topologia do espaço de dados e fornecendo resultados intuitivos e facilmente utilizáveis.
  - Validação:** comparação com algoritmos conhecidos em testes com dados padrões e com dados de infecção viral para diagnóstico auxiliado por computador.
  - Perspectivas:** adequação das dimensões das hipercaixas, diminuição de parâmetros.
4. **Palavras-chave:** colônia de formigas, hipercaixa, otimização, reconhecimento de padrões.

5. **Sinopse do texto:** A *Colônia de Formigas* (ACO) é um método de otimização que pode ser utilizado para agrupar dados. *Hyperbox clustering with Ant Colony Optimization* (HACO) é um método de agrupamento que utiliza ACO para tentar posicionar hipercaixas no espaço de forma a agrupar a maior quantidade de dados possível, e ainda gera uma forma simples de classificar novos dados.

ACO é baseado no comportamento de formigas reais, que otimizam o caminho percorrido entre o alimento e o formigueiro. Hipercaixas definem de forma muito simples uma região em um espaço  $n$ -dimensional, combinadas para definir regiões de topologia complexa, e utilizadas como um classificador de forma trivial.

HACO busca encontrar uma partição de dados, efetivamente definindo grupos. Primeiro, aplica ACO para tentar posicionar hipercaixas de forma que estas contenham a maior quantidade possível de dados. A seguir, se não há conhecimento prévio da quantidade de classes, considera-se que as hipercaixas que se sobrepõem representam uma mesma classe de dados, e [grupos de] hipercaixas distintas representam classes diferentes. Caso o número de classes seja conhecido, HACO aplica o algoritmo *Nearest-neighbor* (NN) para definir a quantidade correta de grupos. Uma consequência de se usar hipercaixas é que o resultado do agrupamento define também um classificador: se um novo dado está dentro de uma hipercaixa, sua classe será a mesma da definida por esta hipercaixa.

Os resultados experimentais de HACO foram, comparados a três algoritmos que têm o mesmo fim: testado em NN, *Fuzzy C-Means* (FCM), e o próprio ACO (com uma abordagem diferente para agrupamento). O primeiro teste foi em conjuntos de dados sintéticos, e serviu como prova de conceito, oferecendo diversas informações sobre o comportamento do método em função de certas configurações. Um segundo experimento foi realizado com dados reais de pacientes para agrupá-los em “saúdáveis” e “não saúdáveis”, e HACO obteve o melhor resultado dentre os algoritmos testados. A análise da estrutura do classificador gerado possibilita descobrir informações relativas às características das classes, indicando um “perfil de risco” para os pacientes.

Foi apresentado o método HACO para agrupar dados, utilizando a meta-heurística ACO e hipercaixas, que possibilita a extração de informações inerentes a estrutura dos dados. HACO foi validado com experimentos, e demonstrou grande potencial. Os resultados são muito influenciados pela configuração dos parâmetros, que será investigada.

6. **Análise crítica:** ~~Este é o melhor artigo de todos os tempos.~~ O artigo apresenta uma forma inovadora de agrupar dados, de forma não-supervisionada (embora possa aproveitar informações se houver). O resultado pode ainda ser utilizado como classificador de novos dados, e - o mais interessante - analisado para descobrir informações sobre as classes. Além disso, explora as vantagens de cada elemento que compõe o método, obtendo melhores resultados e diminuindo o custo computacional. A aplicação em um caso real, cujos resultados podem ser utilizados para auxiliar o diagnóstico de pacientes, dá mais destaque ao trabalho.

O problema de agrupamento de dados é muito pertinente e, em tempos de excesso de dados, a possibilidade de análise intuitiva da estrutura e descoberta de conhecimento é bastante interessante. Além disso, a solução proposta é de uso geral, oferecendo mais possibilidades de uso.

Os experimentos realizados foram coerentes e suficientes para demonstrar o que foi afirmado. Entretanto, o método só foi comparado a outros algoritmos simples, seria interessante uma comparação com algoritmos mais avançados, bem como específicos para aplicação. A comparação também foi em uma única aplicação específica, seria melhor que houvesse mais testes com outros dados para conclusões melhor embasadas. Além disso, é preciso uma análise mais profunda quanto às configurações de HACO, que influenciam muito o resultado.

# Anexo I

## Documentação Original UnB-CIC (parcial)

```
% -*- mode: LaTeX; coding: utf-8; -*-
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% File      : unb-cic.cls (LaTeX2e class file)
%% Authors   : Flávio Maico Vaz da Costa
%%
%%           (based on previous versions by José Carlos L. Ralha)
%% Version   : 0.96
%% Updates   : 0.5  [??/11/2004] - Initial release. don't remember the day.
%%           : 0.75 [04/04/2005] - Fixed font problems, UnB logo
%%                                     resolution, keywords and palavras-chave
%%                                     hyphenation and generation problems,
%%                                     and a few other problems.
%%           : 0.8  [08/01/2006] - Corrigido o problema causado por
%%                                     bancas com quatro membros. O quarto
%%                                     membro agora é OPCIONAL.
%%                                     Foi criado um novo comando chamado
%%                                     bibliografia. Esse comando tem dois
%%                                     argumentos onde o primeiro especifica
%%                                     o nome do arquivo de referencias
%%                                     bibliograficas e o segundo argumento
%%                                     especifica o formato. Como efeito
%%                                     colateral, as referências aparecem no
%%                                     sumário.
%%           : 0.9  [02/03/2008] - Reformulação total, com nova estrutura
%%                                     de opções, comandos e ambientes, adequação
%%                                     do logo da UnB às normas da universidade,
%%                                     inúmeras melhorias tipográficas,
```

```

%%                                aprimoramento da integração com hyperref,
%%                                melhor tratamento de erros nos comandos,
%%                                documentação e limpeza do código da classe.
%%      : 0.91 [10/05/2008] - Suporte ao XeLaTeX, aprimorado suporte para
%%                                glossaries.sty, novos comandos \capa, \CDU
%%                                e \subtitle, ajustes de margem para opções
%%                                hyperref/impressao.
%%      : 0.92 [26/05/2008] - Melhora do ambiente {definition}, suporte
%%                                a hypcap, novos comandos \fontelogo e
%%                                \slashedzero, suporte [10pt, 11pt, 12pt].
%%                                Corrigido bug de seções de apêndice quando
%%                                usando \hypersetup{bookmarksnumbered=true}.
%%      : 0.93 [09/06/2008] - Correção na contagem de páginas, valores
%%                                load e config para opção hyperref, comandos
%%                                \ifhyperref e \SetTableFigures, melhor
%%                                formatação do quadrado CIP.
%%      : 0.94 [17/04/2014] - Inclusão da opção mpca.
%%      : 0.95 [06/06/2014] - Remoção da opção "mpca", inclusão das opções
%%                                "doutorado", "ppginf", e "ppca" para identificar
%%                                o programa de pós-graduação. Troca do teste
%%                                @mestrado por @posgraduacao.
%%      : 0.96 [24/06/2014] - Ajuste do nome do curso/nome do programa.
%%

```