



Análise comparativa de métodos de compressão aplicados a dados textuais estruturados em SVG

Anderson Lemos da Silva

Orientador: Arthur Rodrigues Araruna

Agenda

- Introdução
- Fundamentação Teórica
- Trabalhos Relacionados
- Objetivos
- Procedimentos Metodológicos
- Cronograma de Execução

Introdução

- Necessidade da informação
- Sistemas computacionais
- Problemas de transmissão e armazenamento
 - ✓ Dados estruturados
 - ✓ Compressão de dados

Fundamentação Teórica

- Dados estruturados
 - ✓ “Dados estruturados são dados dispostos em representações rígidas, sujeitas a regras e a restrições impostas pelo esquema que os criou.” (ALMEIDA, 2002).

Fundamentação Teórica

- Dados estruturados
 - ✓ XML
 - ✓ SVG

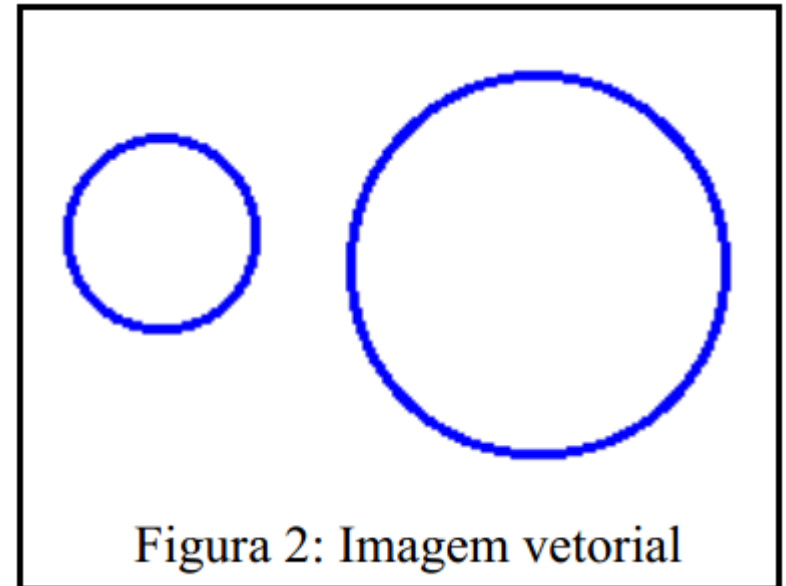
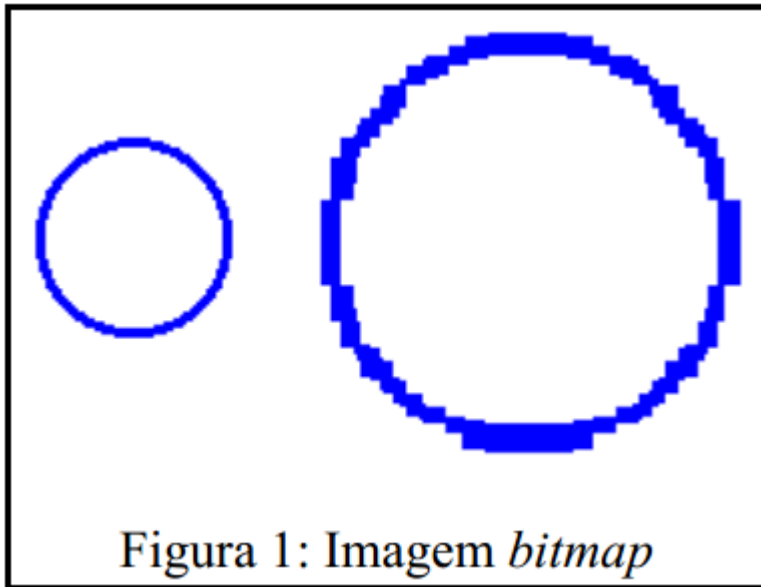
Fundamentação Teórica

- XML

```
<document>
  <report>
    <author>Video database</author >
    <date>June 12, 2000</date>
  </report >
  <paper>
    <title>XML query data model</title>
    <author>Don Robie</author>
    <source>W3C, June 2000</source>
  </paper>
</document>
```

Fundamentação Teórica

- SVG
 - ✓ Imagens Bitmap x Imagens Vetoriais



Fundamentação Teórica

- Compressão de dados
 - ✓ “Compressão de dados é o processo de converter um conjunto de entrada de dados – dados originais – em um conjunto de dados de saída – dados comprimidos –, que possui um menor tamanho comparado aos dados originais” (TEIXEIRA, 2011)

Fundamentação Teórica

- Compressão de dados é aplicada no processamento de dados de diversos formatos:

- ✓ **Texto**
- ✓ Binário
- ✓ Pixel

Fundamentação Teórica

- Técnicas de compressão de dados
 - ✓ Binário x Texto
 - ✓ Genéricas x Dedicadas

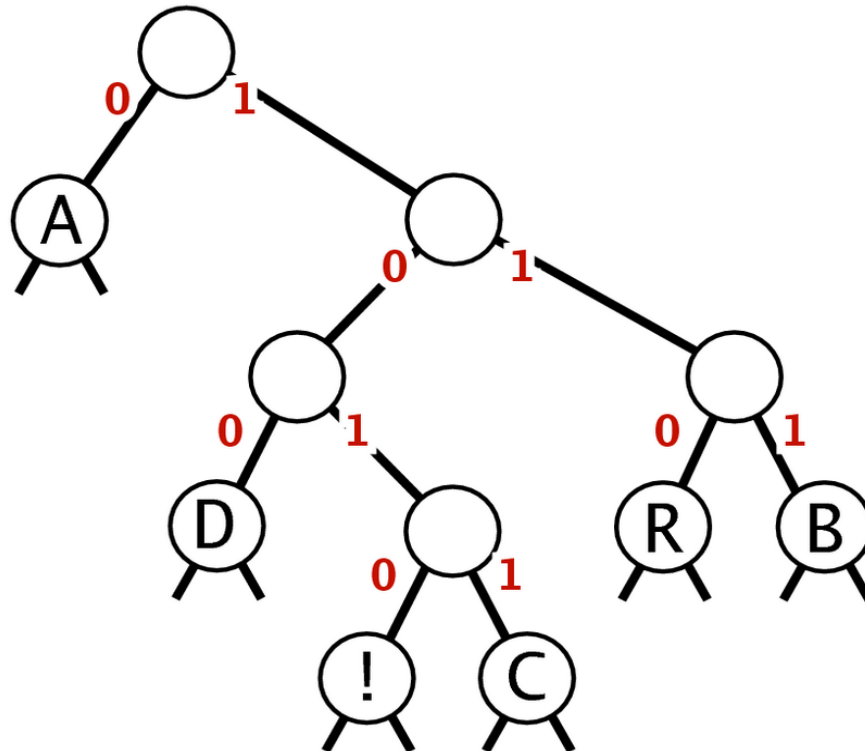
Fundamentação Teórica

- Codificação de Huffman

**Tabela
palavra-chave**

Chave	Valor
!	1010
A	0
B	111
C	1011
D	100
R	110

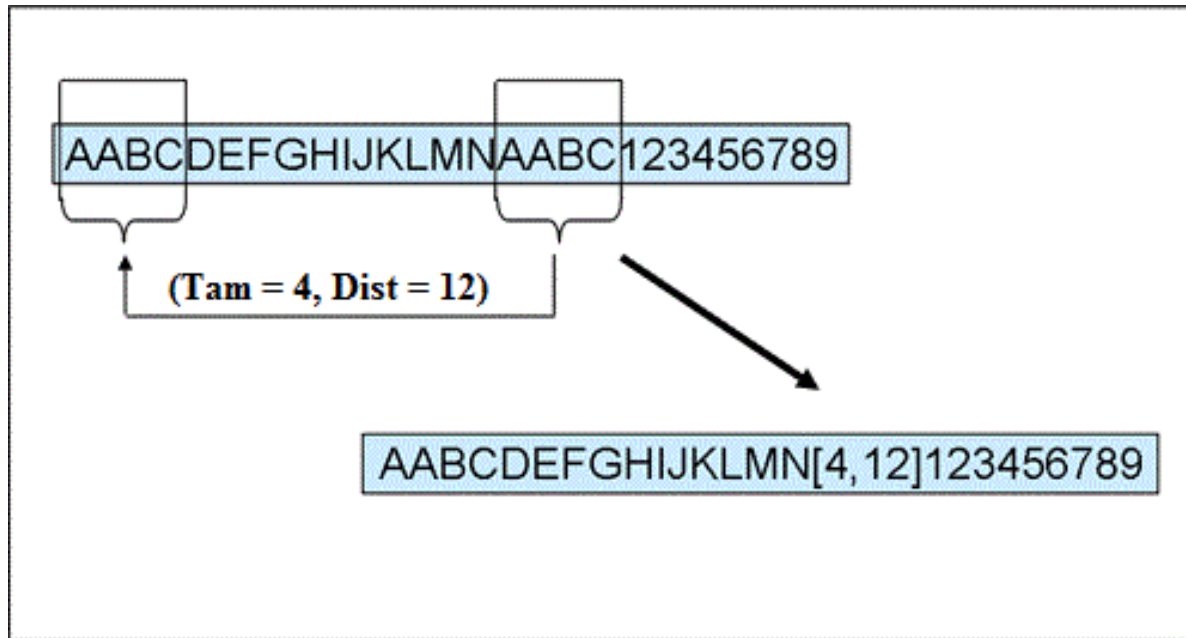
Representação em árvore



ABRACADABRA!

Fundamentação Teórica

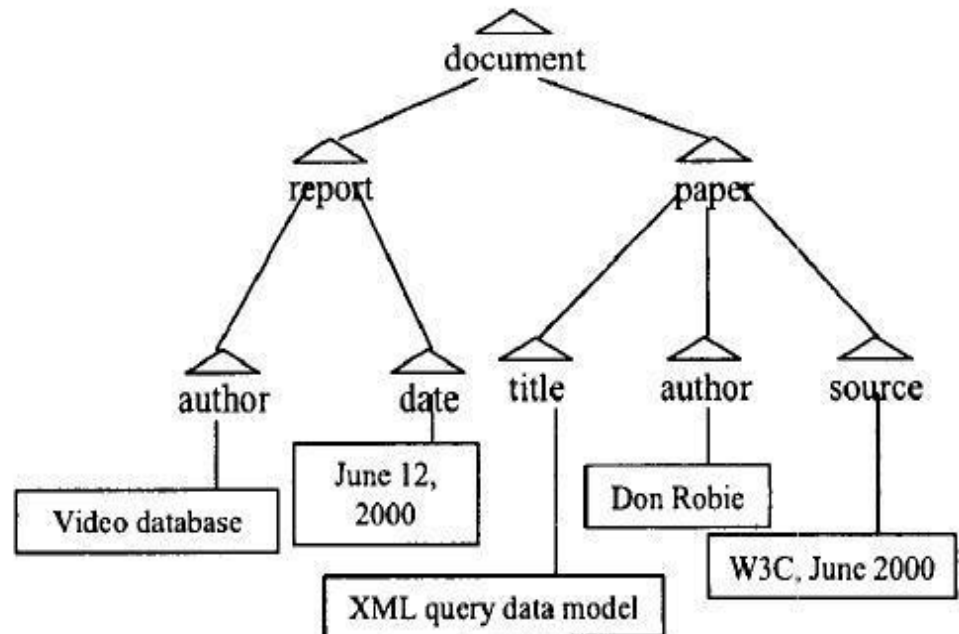
- LZ77



Fundamentação Teórica

- Compressão XML via DAGs
 - ✓ Árvores ordenadas não limitadas

```
<document>
  <report>
    <author>Video database</author>
    <date>June 12, 2000</date>
  </report>
  <paper>
    <title>XML query data model</title>
    <author>Don Robie</author>
    <source>W3C, June 2000</source>
  </paper>
</document>
```

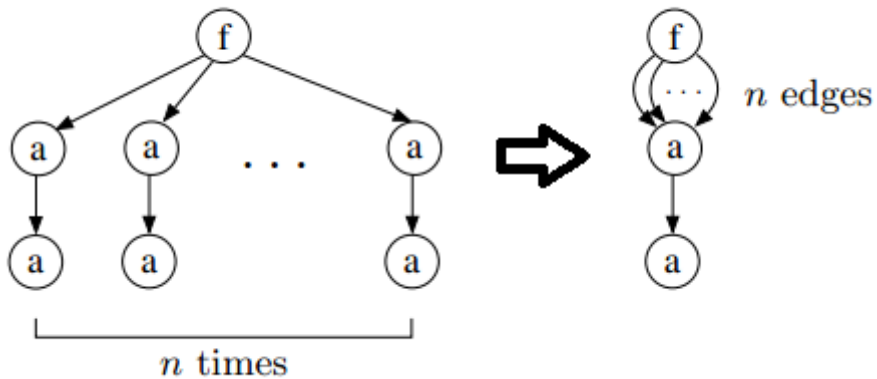


Fundamentação Teórica

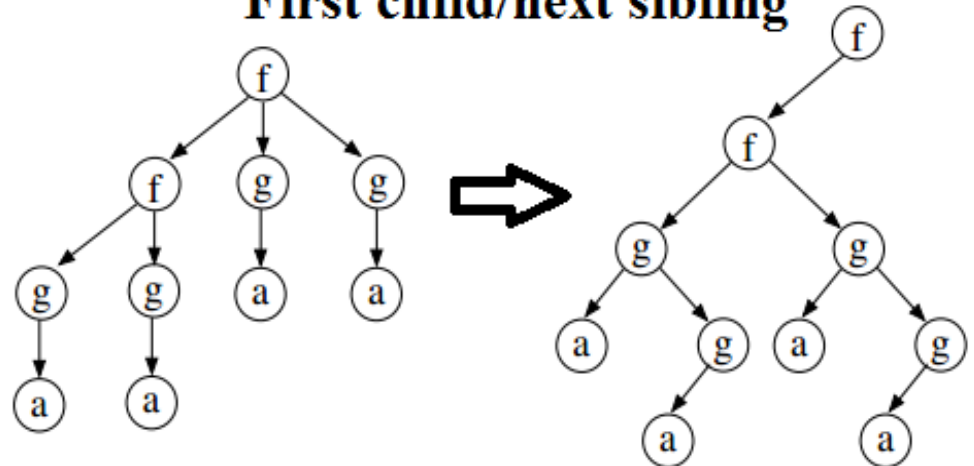
- Compressão XML via DAGs

✓ Directed Acyclic Graphs - Grafo Direcionado Acíclico

DAG

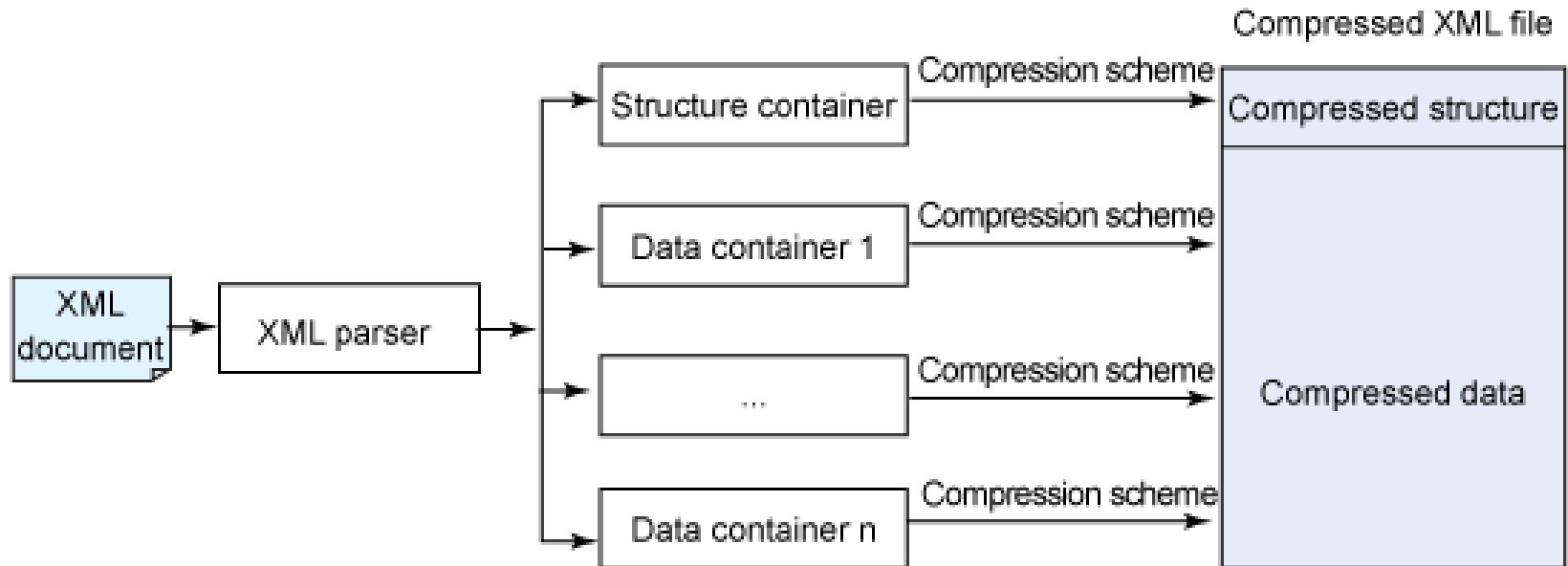


First child/next sibling



Fundamentação Teórica

- XMill



Trabalhos relacionados

- SAKR, Sherif. **XML compression techniques: A survey and comparison.** Journal of Computer and System Sciences, v. 75, n. 5, p. 303-322, 2009.

- ✓ Levantamento do estado da arte
- ✓ Comparação entre nove técnicas

Trabalhos relacionados

- **TEIXEIRA, Márlon A. C. Novas Abordagens para Compressão de Documentos XML. 2011.** Dissertação (Mestrado em Engenharia Elétrica) - Departamento de Comunicações, Universidade Estadual de Campinas, Campinas. 2011.
 - ✓ Duas novas abordagens de compressão para XML
 - ✓ Comparativo com abordagens já existentes

Objetivo geral

- Comparar métodos de compressão de dados textuais estruturados aplicando-os a arquivos SVG, levando em consideração fatores como custo de tempo, de espaço, taxa de compressão, dentre outras aplicáveis, de forma a situar uma boa opção de método.

Objetivos específicos

- Fazer estudo conceitual
- Selecionar quatro métodos de compressão de dados de categorias distintas
- Implementar os métodos selecionados
- Realizar comparativos de performance
 - ✓ Custo de tempo, de espaço, taxa de compressão dentre outros aplicáveis.
- Definir a melhor opção sobre esse domínio

Procedimentos metodológicos

- Realizar revisão bibliográfica e estudar a área de compressão de dados.
- Elencar os principais métodos de compressão aplicáveis ao domínio escolhido
- Analisar técnicas e tecnologias

Procedimentos metodológicos

- Montar ou utilizar repositório de testes disponível
- Implementar os métodos
- Executar os métodos
- Eleger melhor método para o escopo escolhido segundo análise comparativa.

Cronograma de execução

ATIVIDADES	2015													
	Mai		Jun		Jul		Ago		Set		Out		Nov	
Revisão bibliográfica, definição de escopo de pesquisa e estudo da área de compressão de dados e métodos de compressão de dados textuais estruturados.	x	x	x	x	x	x								
Elencar os principais métodos de compressão aplicáveis ao domínio escolhido			x	x										
Defesa do projeto.				x										
Analisar as técnicas e tecnologias				x	x	x								
Montar ou utilizar repositório de testes disponível				x	x	x	x							
Implementar os métodos						x	x	x	x	x	x			
Iniciar redação da monografia									x	x	x			
Executar os métodos										x	x			
Eleger melhor método											x	x		
Revisão final da monografia											x	x	x	
Defesa do Trabalho Final													x	

Obrigado!