



Guilherme Gonçalves Schardong

**Uma Abordagem Visual e Interativa para a
Seleção de Conjuntos de Cenários Temporais**

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor em Ciências pelo Programa de Pós-graduação em Informática da PUC-Rio.

Orientador : Prof. Hélio Côrtes Vieira Lopes
Co-orientador: Prof^a. Simone Diniz Junqueira Barbosa

Rio de Janeiro
Setembro de 2018



Guilherme Gonçalves Schardong

**Uma Abordagem Visual e Interativa para a
Seleção de Conjuntos de Cenários Temporais**

Tese apresentada como requisito parcial para obtenção do grau de Doutor em Ciências pelo Programa de Pós-graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo.

Prof. Hélio Côrtes Vieira Lopes

Orientador

Departamento de Informática – PUC-Rio

Prof^a. Simone Diniz Junqueira Barbosa

Co-orientador

Departamento de Informática – PUC-Rio

Prof. Fernando Luiz Cyrino Oliveira

Departamento de Engenharia Industrial – PUC-Rio

Prof. Bruno Fânzeres dos Santos

Departamento de Engenharia Industrial – PUC-Rio

Prof. Regis Kruel Romeu

CENPES

Prof. Sergio Lima Netto

UFRJ

Prof. Abelardo Borges Barreto Junior

Departamento de Matemática – PUC-Rio

Prof. Alex Laier Bordignon

UFF

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro Técnico Científico – PUC-Rio

Rio de Janeiro, 13 de Setembro de 2018

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Guilherme Gonçalves Schardong

Bacharel em Ciência da Computação (2011) e Mestre em Informática com ênfase em Computação aplicada (2014) pela Universidade Federal de Santa Maria (UFSM). Atuou no Laboratório de Computação Aplicada (LaCA-UFSM) de 2009 a 2014 e no Instituto Tecgraf de 2014 a 2016 durante o doutorado. De 2016 até o momento, atua no GALGOS/ATD-Lab na PUC-Rio.

Ficha Catalográfica

Schardong, Guilherme Gonçalves

Uma Abordagem Visual e Interativa para a Seleção de Conjuntos de Cenários Temporais / Guilherme Gonçalves Schardong; orientador: Hélio Côrtes Vieira Lopes; co-orientador: Simone Diniz Junqueira Barbosa. – Rio de Janeiro: PUC-Rio, Departamento de Informática, 2018.

v., 18 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui bibliografia

1. Redução de Cenários. 2. Interatividade. 3. Visualização Científica. 4. Tomada de Decisão. 5. Conjuntos de Dados Temporais. I. Lopes, Hélio Côrtes Vieira. II. Barbosa, Simone Diniz Junqueira. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por financiar parcialmente essa pesquisa sob o contrato XXXXXX/YYYY-X. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Schardong, Guilherme Gonçalves; Lopes, Hélio Côrtes Vieira; Barbosa, Simone Diniz Junqueira. **Uma Abordagem Visual e Interativa para a Seleção de Conjuntos de Cenários Temporais**. Rio de Janeiro, 2018. 18p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O uso de abordagens de programação estocástica e redução de cenários tem se tornado imprescindível na análise e predição de comportamento de sistemas dinâmicos. Entretanto, tais técnicas não levam em conta o conhecimento prévio sobre domínio que o usuário possui. O presente trabalho tem por objetivo o desenvolvimento de uma abordagem visual e interativa para abordar o problema de redução de cenários com dados temporais. Para tanto, nós propomos a implementação de uma série de visualizações de dados temporais integradas. Também propomos a adaptação de um algoritmo de projeção multidimensional para lidar com dados temporais. Desta forma, podemos representar graficamente a evolução de um conjunto de cenários ao longo do tempo. Outra visualização proposta no presente trabalho é uma adaptação de *Bump chart* para lidar com dados temporais acumulados; através dele, um usuário pode comparar a evolução das distâncias entre os diferentes cenários e um cenário de referência. Para validar a nossa proposta, fizemos uma implementação das técnicas propostas e conduzimos um estudo com usuários de diferentes áreas do conhecimento e níveis de experiência. Os resultados obtidos até então indicam que uma abordagem visual para o problema de redução de cenários é viável, e permite a seleção de um conjunto razoável de cenários. Além disso, constatamos que essa abordagem pode ser útil em um contexto de exploração de dados visando a redução de cenários. O usuário também pode explorar visualmente os resultados de outras técnicas de redução de cenários usando nossa abordagem. Os usuários entrevistados reportaram facilidade em cumprir as tarefas propostas e comentaram positivamente sobre os mecanismos de interação fornecidos pelo nosso protótipo. Também testamos os cenários escolhidos usando nossa proposta contra outras abordagens encontradas tanto na literatura quanto em uso na indústria. Os resultados obtidos foram bons, indicando que nossa proposta é viável em casos de uso reais.

Palavras-chave

Redução de Cenários; Interatividade; Visualização Científica; Tomada de Decisão; Conjuntos de Dados Temporais.

Abstract

Schardong, Guilherme Gonçalves; Lopes, Hélio Côrtes Vieira (Advisor); Barbosa, Simone Diniz Junqueira (Co-Advisor). **Visual interactive support for selecting scenarios from time-series ensembles**. Rio de Janeiro, 2018. 18p. Tese de doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Stochastic programming and scenario reduction approaches have become invaluable in the analysis and behavior prediction of dynamic systems. However, such techniques often fail to take advantage of the user's own expertise about the problem domain. This work provides visual interactive support to assist users in solving the scenario reduction problem with time-series data. We employ a series of time-based visualization techniques linked together to perform the task. By adapting a multidimensional projection algorithm to handle temporal data, we can graphically present the evolution of the ensemble. We also propose to use cumulative bump charts to visually compare the ranks of distances between the ensemble time series and a baseline series. To evaluate our approach, we developed a prototype application and conducted observation studies with volunteer users of varying backgrounds and levels of expertise. Our results indicate that a graphical approach to scenario reduction may result in a good subset of scenarios and provides a valuable tool for data exploration in this context. The users liked the interaction mechanisms provided and judged the task to be easy to perform with the tools we have developed. We tested the proposed approach against state-of-the-art techniques proposed in the literature and used in the industry and obtained good results, thus indicating that our approach is viable in a real-world scenario.

Keywords

Scenario Reduction; User Interaction; Scientific Visualization; Decision Making; Time Series Ensembles.

Sumário

1	Introdução	12
1.1	Estrutura da Tese	12
1.1.1	Subseção de exemplo	13
1.1.2	Exemplos de teoremas, lemas e provas	16
	Referências bibliográficas	17
A	Perfil de usuário para sessão de avaliação do protótipo	18

Lista de figuras

- Figura 1.1 Justaposição, Superposição and Codificação Explícita.
Imagem retirada do trabalho de (Szafir2018). 14
- Figura 1.2 Comparação entre diferentes representações do tempo
para o gráfico Time-lapsed LAMP da propriedade W_p . Super-
posição (1.2a) e Codificação Explícita (1.2b) em ordem crescente
de tamanho de glifo. 15

Lista de tabelas

Tabela 1.1 Erros dos cenários selecionados usando a abordagem da indústria.

13

Lista de abreviaturas

ANP – Agência Nacional do Petróleo

CENPES – Centro de Pesquisas Leopoldo Américo Miguez de Mello

LAMP – Local Affine Multidimensional Projection

MDS – Multidimensional Scaling

MP – Multidimensional Projections

N_p – Produção acumulada de óleo

P₁₀ – Percentil 10

P₅₀ – Percentil 50

P₉₀ – Percentil 90

PETROBRAS – Petróleo Brasileiro S.A.

Q_o – Fluxo de óleo

Q_w – Fluxo de água

W_p – Produção acumulada de água

*ALL THINGS THAT ARE, ARE OURS.
BUT WE MUST CARE. FOR IF WE DO
NOT CARE, WE DO NOT EXIST. IF
WE DO NOT EXIST, THEN THERE IS
NOTHING BUT BLIND OBLIVION. AND
EVEN OBLIVION MUST END SOMEDAY.
LORD, WILL YOU GRANT ME JUST A
LITTLE TIME? FOR THE PROPER BA-
LANCE OF THINGS. TO RETURN WHAT
WAS GIVEN. FOR THE SAKE OF PRISO-
NERS AND THE FLIGHT OF BIRDS.
LORD, WHAT CAN THE HARVEST HOPE
FOR, IF NOT FOR THE CARE OF THE
REAPER MAN?*

Terry Pratchett, *Reaper Man*.

1

Introdução

Esse é um texto de exemplo para demonstrar algumas funcionalidades do modelo L^AT_EX de teses e dissertações da PUC-Rio. As citações são formatadas da seguinte maneira: (Lee2010), (Hummel2013), (Dupacova2003).

A idéia desse modelo é a seguinte:

- Prover um modelo L^AT_EX mais atualizado e incorporando algumas correções feitas desde a última versão disponível, lançada em 2017;
- Prover uma forma mais ágil disponibilizar correções e melhorias do modelo.

A Tabela 1.1 apresenta o formato padrão de tabelas desse modelo. Título acima do conteúdo, basicamente. A Figura 1.1 mostra como incluir figuras com subfiguras (1.2, 1.1)

1.1

Estrutura da Tese

1.1.1

Subseção de exemplo

Tabela 1.1: Cenários escolhidos e erros ($\times 10^{10}$) para a abordagem da indústria.

Propriedade	Percentil	Cenário	SSE	MSE
N_p	P_{10}	172	1190.0	30.5
		88	663.0	17.0
		122	1660.0	42.6
		36	312.0	7.9
	P_{50}	131	162.0	4.1
		4	162.0	4.1
		26	786.0	20.1
		90	236.0	6.0
	P_{90}	96	975.0	25.0
		100	690.0	17.7
		127	925.0	23.7
		132	526.0	13.5
	P_{10}	23	7640.0	196.0
		115	42000.0	1080.0
		10	26900.0	690.0
		19	27600.0	707.0
W_p	P_{50}	29	14200.0	364.0
		153	6260.0	161.0
		84	5840.0	150.0
		88	11200.0	288.0
	P_{90}	37	14300.0	367.0
		57	22300.0	571.0
		187	14000.0	358.0
		28	2210.0	56.7

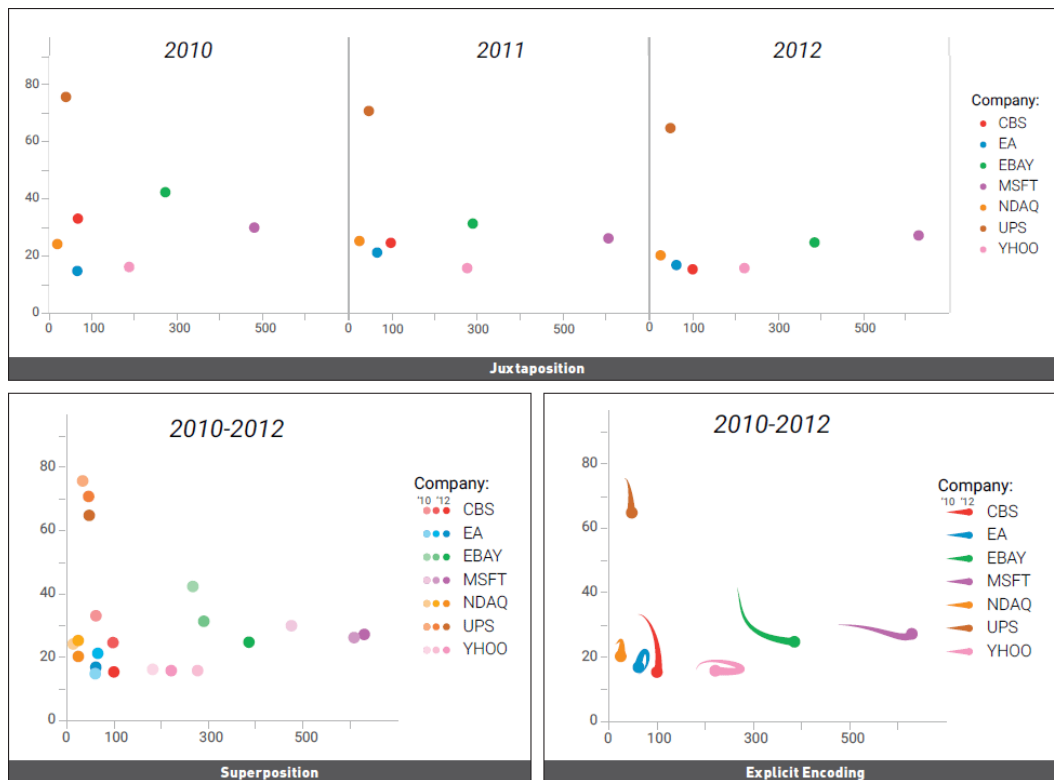
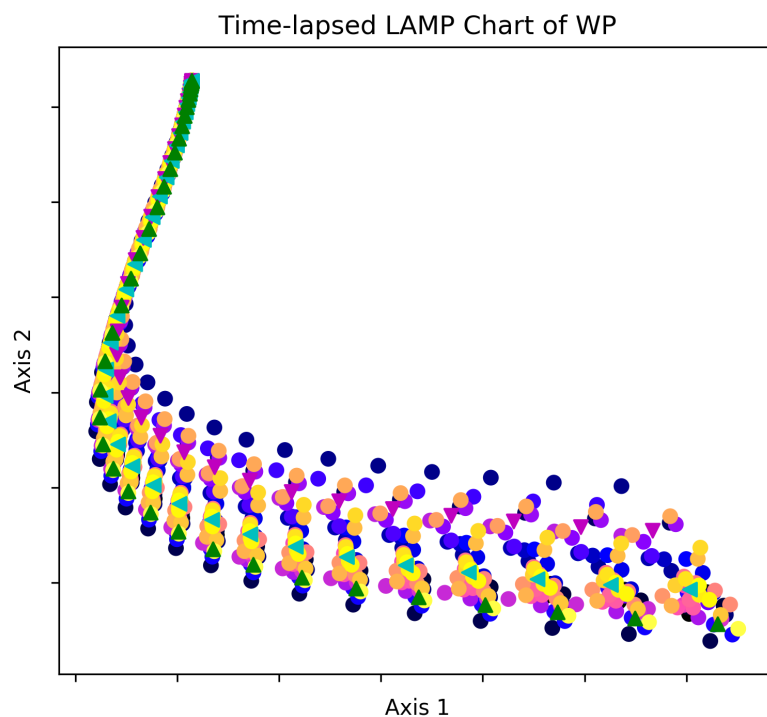
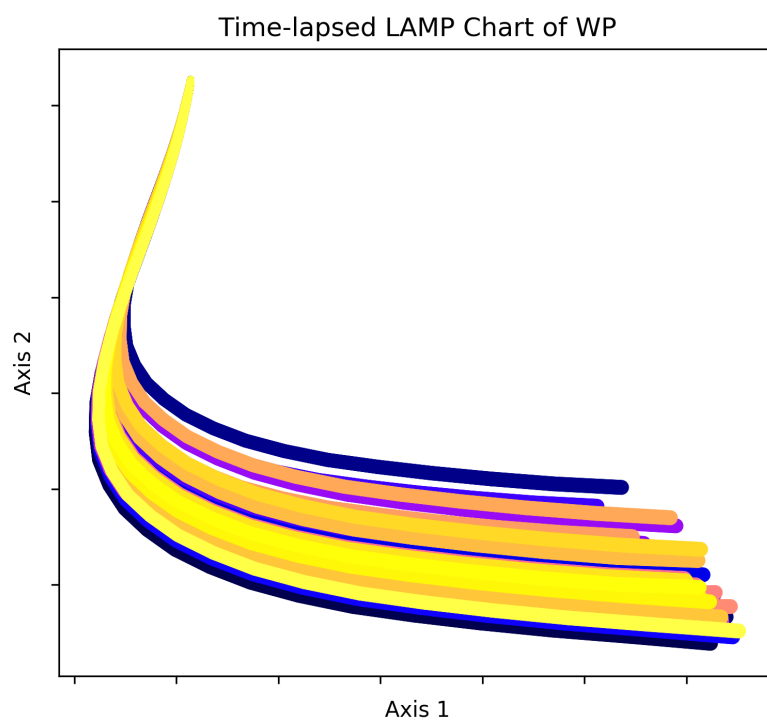


Figura 1.1: Juxtaposição, Superposição and Codificação Explícita. Imagem retirada do trabalho de (Szafir2018).



(a) Superposição de tempos no gráfico Time-lapsed LAMP.



(b) Codificação Explícita do tempo no gráfico Time-lapsed LAMP. Tempos iniciais são representados por glifos menores.

Figura 1.2: Comparação entre diferentes representações do tempo para o gráfico Time-lapsed LAMP da propriedade W_p . Superposição (1.2a) e Codificação Explícita (1.2b) em ordem crescente de tamanho de glifo.

1.1.2**Exemplos de teoremas, lemas e provas**

Teorema 1.1 *Dada a função f , cuja derivada existe em todos os pontos de seu domínio, então f é chamada função contínua.*

Teorema 1.2 (Teorema de Pitágoras) *Esse é um teorema que trata de triângulos retângulos, e pode ser sumarizado pela equação a seguir:*

$$x^2 + y^2 = z^2$$

Uma consequência do Teorema 1.2 é o Corolário 1.3

Corolário 1.3 *Não existe triângulo retângulo de lados 3cm, 4cm e 6cm.*

Lema 1.4 (Operações matriciais que preservam unimodularidade)

As operações matriciais a seguir, chamadas de operações unimodulares, preservam a unimodularidade de uma matriz se esta for unimodular:

1. *Troca de duas colunas*
2. *Adição de um múltiplo inteiro de uma coluna a outra coluna*
3. *Multiplicação de uma coluna por -1*
4. *Transposição*

Lema 1.5 *Dados dois segmentos de linha de comprimentos a e b respectivamente, existe um número $r \in \mathbb{R}$ tal que $b = ra$.*

Prova. A prova do Lemma 1.5 pode ser construída por contradição. Basta assumir que a premissa é falsa e seguir a prova, em algum momento a prova cairá em contradição, portanto, o lema original é verdadeiro. ■

Referências bibliográficas

- [Dupacova2003] DUPAČOVÁ, J.; GRÖWE-KUSKA, N. ; RÖMISCH, W.. **Scenario reduction in stochastic programming**. Mathematical Programming, 95(3):493–511, 2003.
- [Hummel2013] HUMMEL, M.; OBERMAIER, H.; GARTH, C. ; JOY, K. I.. **Comparative visual analysis of lagrangian transport in cfd ensembles**. Visualization and Computer Graphics, IEEE Transactions on, 19(12):2743–2752, 2013.
- [Lee2010] LEE, H. K. H.; TADDY, M. ; GRAY, G. A.. **Selection of a Representative Sample**. Journal of Classification, 27(1):41–53, mar 2010.
- [Szafir2018] SZAFIR, D. A.. **The good, the bad, and the biased: Five ways visualizations can mislead (and how to fix them)**. Interactions, 25(4):26–33, June 2018.

A

Perfil de usuário para sessão de avaliação do protótipo

1. Informações pessoais

(a) Nome

(b) E-mail

(c) Nível de escolaridade

(d) Período do curso (apenas para cursos em andamento)

(e) Profissão

2. Marque a seguir o seu conhecimento sobre os assuntos

(a) Divisão de amostras por percentis: P10, P50, P80 ...

- ☐ Desconheço
- ☐ Conheço pouco (aprendi esses conceitos em algum momento, mas talvez tenha que aprender novamente se tiver que aplicá-los)
- ☐ Tenho conhecimento médio (talvez tenha que rever um ou outro conceito se tiver que aplicá-lo)
- ☐ Conheço bem (não aplico com frequência, mas não precisaria rever os conceitos se tivesse que aplicá-los)
- ☐ Sou especialista (aplico esses conceitos com frequência)

(b) Análise de tendências e padrões em séries temporais

- ☐ Desconheço
- ☐ Conheço pouco (aprendi esses conceitos em algum momento, mas talvez tenha que aprender novamente se tiver que aplicá-los)
- ☐ Tenho conhecimento médio (talvez tenha que rever um ou outro conceito se tiver que aplicá-lo)