
# CHANGE POINT DETECTION IN END-TO-END MEASUREMENTS TIME SERIES

Diego Ximenes Mendes

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientador: Edmundo Albuquerque de Souza e Silva

Rio de Janeiro
Janeiro de 2017

CHANGE POINT DETECTION IN END-TO-END MEASUREMENTS TIME SERIES

Diego Ximenes Mendes

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

RIO DE JANEIRO, RJ – BRASIL
JANEIRO DE 2017

Ximenes Mendes, Diego

Change Point Detection in End-to-End Measurements Time Series/Diego Ximenes Mendes. – Rio de Janeiro: UFRJ/COPPE, 2017.

IX, 9 p.: il.; 29, 7cm.

Orientador: Edmundo Albuquerque de Souza e Silva

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2017.

Bibliography: p. 9 – 9.

1. Change Point Detection. 2. Time Series. 3. Machine Learning. I. Albuquerque de Souza e Silva, Edmundo. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

CHANGE POINT DETECTION IN END-TO-END MEASUREMENTS TIME SERIES

Diego Ximenes Mendes

Janeiro/2017

Orientador: Edmundo Albuquerque de Souza e Silva

Programa: Engenharia de Sistemas e Computação

oioioi

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

# CHANGE POINT DETECTION IN END-TO-END MEASUREMENTS TIME SERIES

Diego Ximenes Mendes

January/2017

Advisor: Edmundo Albuquerque de Souza e Silva

Department: Systems Engineering and Computer Science

oioioi

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Contributions

## 1.2 Dissertation Outline

# Chapter 2

# Literature Review of Change Point Detection Problem

Here the change point problem will be "defined", including offline and online versions.

## 2.1 Sliding Window Techniques

Describe sliding window techniques in change point detection. Describe how two windows can be compared. Probably this technique will not be part of my solution, but I will write about it since is the simplest and most intuitive solution.

## 2.2 Dynamic Programming

Optimization problem will be formulated and the solution algorithm will be presented. Also will be presented pruning techniques such as PELT.

## 2.3 Bayesian Inference

I will erase this section if I don't use bayesian inference. Describe Fearnheard (offline) and MacKay (online) solutions. Say that there are other versions.

## 2.4 HMM

I will not describe HMM algorithms (viterbi, baum welch, etc), I will only describe how HMM have been used in change point detection. Describe Left-Right HMM, full HMM, and Regularized HMM in this problem.

## 2.5 Other Algorithms

Only cite other used algorithms and say why I chose the previous one to analyse.

## 2.6 Performance Evaluation

Describe how datasets are constructed in literature. Describe how an algorithm output is evaluated.

# Chapter 3

# Dataset

## 3.1 Description of End-to-End Packet Loss Measurements Time Series

Here will be presented the TGR dataset. Small description on how data are collected, including client informations (geographic position, routes, etc) Plots: distribution between two consecutive measures, autocorrelation after time binarization, loss distribution, hour of day x loss, day of week x loss. Maybe: clusterize clients by distribution or time series.

## 3.2 Change Points Classification Survey

Describe web system used to get "true" change points. Describe majority voting. Describe how data were divided in train/test dataset.

# Chapter 4

# Applying Change Point Detection

In each algorithm section I will: Describe adaptations and aproaches. Describe dificulties of this algorithms in real data and in the current dataset that lead to adaptation.

## 4.1 Preprocessing

Filters applied to time series before presenting to algorithms.

## 4.2 Tuning Hyperparameters

### 4.2.1 Grid Search and Randomized Grid Search

### 4.2.2 Bayesian Optimization

I don't know if I am going to use this method.

### 4.2.3 Particle Swarm Optimization

I don't know if I am going to use this method.

## 4.3 Sliding Window

## 4.4 Dynamic Programming

## 4.5 HMM

## 4.6 Bayesian Inference

I don't know I will use this: poor performance.

## 4.7 LSTM

I don't know I will use this: maybe I will not have enough data.

## 4.8 Ensembles

If there are enough models to be tested describe how to use ensembles.

# Chapter 5

# Results

## 5.1 Classification Accuracy

Present false positive/false negative/...

## 5.2 Unsupervised Analysis

Clusterize clients according with change points detected and check if latent information of clusters are also clusterized.

## 5.3 Algorithms Comparison

Compare algorithms results and computational performance.

# Chapter 6

# Conclusions

# Bibliography