

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
CURSO DE CIÊNCIA DA COMPUTAÇÃO

MATHEUS DOS SANTOS GONZAGA

**Implementação de um Framework para a  
Detecção Contextual de Anomalias em  
Dados de Sensores**

Monografia apresentada como requisito parcial  
para a obtenção do grau de Bacharel em Ciência  
da Computação

Orientadora: Profa. Dra. Renata Galante

Porto Alegre  
2017

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitora: Prof<sup>a</sup>. Jane Fraga Tutikian

Pró-Reitor de Graduação: Prof. Vladimir Pinheiro do Nascimento

Diretora do Instituto de Informática: Prof<sup>a</sup>. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Sérgio Luis Cechin

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

## RESUMO

Estou usando esse documento exemplo para estruturar o meu TCC. Esse era o texto original do Abstract: Este documento é um exemplo de como formatar documentos para o Instituto de Informática da UFRGS usando as classes L<sup>A</sup>T<sub>E</sub>X disponibilizadas pelo UTUG. Ao mesmo tempo, pode servir de consulta para comandos mais genéricos. *O texto do resumo não deve conter mais do que 500 palavras.*

**Palavras-chave:** Detecção de anomalias. Dados de sensores. Big Data.

## Using L<sup>A</sup>T<sub>E</sub>X to Prepare Documents at II/UFRGS

### ABSTRACT

I'm using this document to structure my major conclusion work. This was the original abstract text: This document is an example on how to prepare documents at II/UFRGS using the L<sup>A</sup>T<sub>E</sub>X classes provided by the UTUG. At the same time, it may serve as a guide for general-purpose commands. *The text in the abstract should not contain more than 500 words.*

**Keywords:** Anomaly Detection. Sensor Data. Big Data..

## LISTA DE FIGURAS

Figura 1.1 Anomalias em um conjunto de dados bidimensional. ....	10
Figura 1.2 Exemplo de anomalia contextual. ....	10

## **LISTA DE TABELAS**

## **LISTA DE ABREVIATURAS E SIGLAS**

SMP	Symmetric Multi-Processor
NUMA	Non-Uniform Memory Access
SIMD	Single Instruction Multiple Data
SPMD	Single Program Multiple Data
ABNT	Associação Brasileira de Normas Técnicas

## SUMÁRIO

<b>1 LEVANTAMENTO BIBLIOGRÁFICO.....</b>	<b>9</b>
<b>1.1 Introdução de Conceitos.....</b>	<b>9</b>
1.1.1 Anomalias .....	9
1.1.2 Tipos de Anomalias .....	9
1.1.2.1 Anomalia Pontual.....	9
1.1.2.2 Anomalias Contextuais .....	10
1.1.2.3 Anomalias Coletivas .....	11
<b>1.2 Trabalhos Relacionados.....</b>	<b>11</b>
<b>REFERÊNCIAS.....</b>	<b>12</b>



## 1 LEVANTAMENTO BIBLIOGRÁFICO

Neste capítulo é apresentado o levantamento bibliográfico relacionado ao desenvolvimento deste trabalho. Ele é dividido em duas partes: A seção 2.1 apresenta o conceito de anomalias, seus tipos e as principais técnicas empregadas em sua detecção. A seção 2.2 discute trabalhos relacionados e estabelece um comparativo entre eles.

### 1.1 Introdução de Conceitos

#### 1.1.1 Anomalias

Anomalias são padrões ou eventos em um conjunto de dados que não correspondem a um conceito bem definido de normalidade. Anomalias podem constar nos dados por uma série de motivos diferentes, dependendo da área de aplicação. Por exemplo, observações anômalas em dados de monitoramento de componentes industriais poderiam indicar danos. Um padrão incomum de tráfego em uma rede de computadores poderia indicar que um computador hackeado está mandando dados sensíveis para um endereço não autorizado.

Na Figura 1.1 pode-se ver anomalias em um conjunto de dados com duas dimensões. Os conjuntos  $N_1$  e  $N_2$  correspondem às regiões normais, já que a maioria dos pontos estão nessas regiões. Os pontos  $o_1$ ,  $o_2$  e o conjunto  $O_3$  são anomalias.

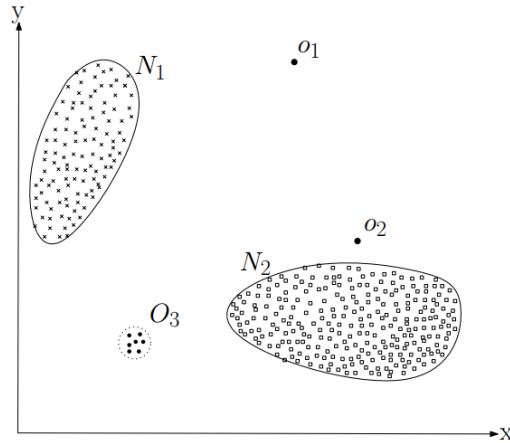
#### 1.1.2 Tipos de Anomalias

Determinar o tipo de anomalia sendo identificado é um aspecto importante das técnicas de detecção de anomalias. As anomalias podem ser classificadas em 3 tipos: Pontuais, Contextuais e Coletivas.

##### 1.1.2.1 Anomalia Pontual

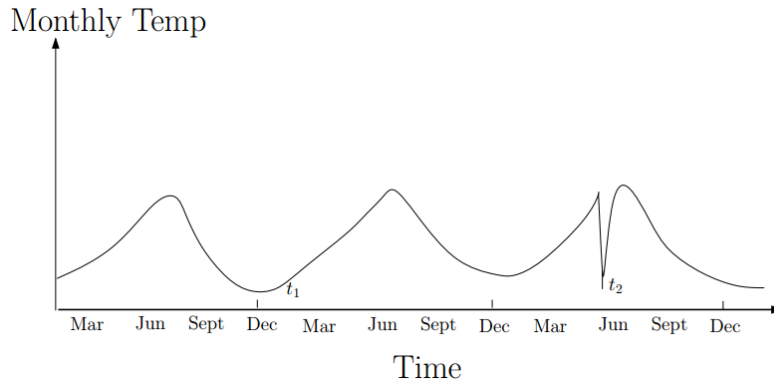
Quando uma observação individual é considerada anômala em relação ao restante dos dados, ela é chamada de **anomalia pontual**. Esse tipo de anomalia é o foco da maioria das aplicações e pesquisas na área de detecção de anomalias. Na Figura 1.1, os pontos  $o_1$  e  $o_2$  são anomalias pontuais.

Figura 1.1: Anomalias em um conjunto de dados bidimensional.



Fonte: (CHANDOLA; BARNEJEE; KUMAR, 2009)

Figura 1.2: Exemplo de anomalia contextual.



Fonte: (CHANDOLA; BARNEJEE; KUMAR, 2009)

### 1.1.2.2 Anomalias Contextuais

Quando uma instância de dado é considerada anômala dentro de um contexto específico, ela é uma **anomalia contextual**. Por exemplo, um sensor de temperatura pode ter leituras altas durante o dia e baixas durante a noite. Embora uma leitura alta não seja necessariamente uma anomalia, se ela ocorreu durante a noite, então é uma anomalia com respeito ao contexto da hora do dia.

Nesse caso, cada instância de dado é definida usando conjuntos de **atributos contextuais**, que indicam o contexto daquela instância (como *Hora do Dia*), e **atributos comportamentais**, que definem as características independentes de contexto da instância (como *Temperatura*).

A Figura 1.2 mostra um exemplo de anomalia contextual. A temperatura  $t_2$  tem o mesmo valor que  $t_1$ , porém, devido ao seu contexto,  $t_2$  é uma anomalia.

### *1.1.2.3 Anomalias Coletivas*

Quando um conjunto de observações é anômalo em relação ao restante do conjunto de dados, ele é chamado de **anomalia coletiva**. As observações individuais dentro desse conjunto podem não ser anomalias por si, mas sua ocorrência em conjunto ou sequência é anômala.

## **1.2 Trabalhos Relacionados**

## REFERÊNCIAS

CHANDOLA, V.; BARNEJEE, A.; KUMAR, V. **Anomaly Detection: A Survey**. [S.l.], 2009.