

# Sistemas Distribuídos em Larga Escala

## Trabalho Prático

-

Mestrado em Engenharia Informática  
Universidade do Minho

### Grupo nº 1

---

PG41080	João Ribeiro Imperadeiro
PG41081	José Alberto Martins Boticas
PG41091	Nelson José Dias Teixeira

19 de maio de 2020

## 1 Introdução

[1]

O presente relatório descreve o desenvolvimento do projeto de caráter prático da unidade curricular de Sistemas Distribuídos em Larga Escala. Neste trabalho prático é requerida a implementação de um dos algoritmos de agregação distribuída disponíveis na referência [2] indicada no enunciado do mesmo. Após concretizar a especificação do algoritmo escolhido, é posteriormente solicitado o teste do mesmo no simulador desenvolvido ao longo do semestre do presente ano letivo.

## 2 Algoritmo

Dos algoritmos de agregação distribuída presentes no documento referenciado no enunciado deste trabalho prático [2], os elementos deste grupo optaram por escolher o algoritmo *flow updating*. Este, ao nível de comunicação, é classificado como não estruturado, inserindo-se na categoria *gossip* que, por sua vez, diz respeito à forma como as mensagens são disseminadas pela rede de comunicação. Quanto à perspetiva computacional, este algoritmo é baseado no conceito de *averaging*, isto é, na computação iterativa de médias parciais que, ao longo do tempo, convergem para um resultado final previamente determinado. Esta última técnica permite também a

derivação de outras funções de agregação (como por exemplo, *count* ou *sum*) de acordo com as combinações dos valores inicialmente instanciados.

Uma das razões que levou este grupo a escolher este algoritmo foi a capacidade do mesmo em tolerar a injeção de falhas. Esta última característica é bastante importante sobretudo no que diz respeito à perda de mensagens trocadas na rede de comunicação. Para além desta vantagem associada ao contexto de sistemas distribuídos, este algoritmo possui não só um melhor desempenho quando comparado com os outros da mesma classe, como também possibilita uma computação precisa de valores. Por fim, a execução do algoritmo em causa é independente da topologia do roteamento de rede.

## 2.1 Conceito

## 2.2 Implementação

```

state variables:
|  $f_{ij}, \forall j \in \mathcal{D}_i$ , flows, initially  $f_{ij} = 0$ 
|  $e_{ij}, \forall j \in \mathcal{D}_i$ , estimates, initially  $e_{ij} = 0$ 
|  $v_i$ , input value

message-generation function:
|  $\text{msg}(i, j) = (f_{ij}, e_{ij}), \forall j \in \mathcal{D}_i$ 

state-transition function:
| forall  $(f_{ji}, e_{ji})$  received do
|   |  $f_{ij} \leftarrow -f_{ji}$ 
|   |  $e_{ij} \leftarrow e_{ji}$ 
|   
$$e_i \leftarrow \frac{\left(v_i - \sum_{j \in \mathcal{D}_i} f_{ij}\right) + \sum_{j \in \mathcal{D}_i} e_{ij}}{|\mathcal{D}_i| + 1}$$

|   forall  $j \in \mathcal{D}_i$  do
|     |  $f_{ij} \leftarrow f_{ij} + (e_i - e_{ij})$ 
|     |  $e_{ij} \leftarrow e_i$ 

```

Figura 1: Pseudocódigo do algoritmo *Flow Updating*

### 3 Simulador

### 4 Análise de resultados obtidos

### 5 Conclusão

### Referências

- [1] Paulo Jesus, Carlos Baquero e Paulo Almeida. «Fault-Tolerant Aggregation by Flow Updating». Em: jun. de 2009, pp. 73–86. DOI: 10.1007/978-3-642-02164-0\_6.
- [2] Paulo Jesus, Carlos Baquero e Paulo Almeida. «A Survey of Distributed Data Aggregation Algorithms». Em: *Communications Surveys & Tutorials, IEEE* 17 (out. de 2011). DOI: 10.1109/COMST.2014.2354398. URL: <https://arxiv.org/pdf/1110.0725.pdf>.