



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Modelagem Computacional para Avaliação da Age of Information (AoI) em Sistemas Ciberfísicos

Paulo César Prandel

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Orientadora

Prof.a Dr.a Priscila América Solís Mendez Barreto

Brasília  
2022

Ficha catalográfica elaborada automaticamente,  
com os dados fornecidos pelo(a) autor(a)

PP899m Prandel, Paulo César  
Modelagem Computacional para Avaliação da Age of  
Information (AoI) em Sistemas Ciberfísicos / Paulo César  
Prandel; orientador Priscila América Solís Mendez Barreto. -  
Brasília, 2022.  
101 p.

Dissertação (Mestrado - Mestrado Profissional em  
Computação Aplicada) -- Universidade de Brasília, 2022.

1. Age of Information. 2. Sistemas Ciberfísicos. 3.  
Simulação. I. Barreto, Priscila América Solís Mendez, orient.  
II. Título.



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Modelagem Computacional para Avaliação da Age of Information (AoI) em Sistemas Ciberfísicos

Paulo César Prandel

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Prof.a Dr.a Priscila América Solís Mendez Barreto (Orientadora)  
CIC/UnB

Prof. Dr. Edison Ishikawa    Prof. Dr. José Antonio Apolinário Jr.  
CIC/UnB                            Instituto Militar de Engenharia

Prof. Dr. Marcelo Ladeira  
Coordenador do Programa de Pós-graduação em Computação Aplicada

Brasília, 31 de maio de 2022

# Dedicatória

Dedico este trabalho em memória de minha irmã, Carla.

# Agradecimentos

À Universidade de Brasília, em especial ao Programa de Pós-Graduação em Computação Aplicada, pela oportunidade de realizar este trabalho.

Aos professores do PPCA, pela dedicação e entusiasmo com que conduziram suas aulas e trabalhos.

À professora Priscila, pela orientação, direção e contribuições ao presente trabalho e aos artigos relacionados.

Aos militares do Centro de Desenvolvimento de Sistemas do Exército Brasileiro, com os quais pude discutir algumas ideias e propostas para este trabalho.

Aos meus pais, Aldo e Maria, por entenderem, desde minha infância, a importância da educação e do autoaperfeiçoamento.

À minha esposa Laís, pela compreensão, apoio e paciência durante minhas ausências como marido e pai.

Ao meu filho Jonas, nascido no decurso deste trabalho. Você trouxe novo sentido para minha vida.

# Resumo

A *Age of Information* (AoI) é um conceito que define métricas relacionadas ao grau de atualização que um monitor possui em relação a uma ou mais fontes de informação, as quais enviam pacotes sequenciais de atualização. O presente estudo propõe a utilização da AoI para a avaliação e otimização de um Sistema Ciberfísico (SCF) caracterizado como um sistema de monitoramento e decisão no contexto da atividade de comando e controle (C2) militar. Nesse tipo de sistema, agentes enviam periodicamente suas posições georreferenciadas para um centro de operações e os cenários de uso podem ter diferentes graus de criticidade. Para atingir esse objetivo, este trabalho desenvolveu um modelo computacional para cálculo da *Age of Information* (AoI), o qual foi usado como base para a implementação de uma ferramenta de simulação. Além da referida ferramenta, este trabalho contribuiu para a melhoria das atuais técnicas de gerenciamento de pacotes para a otimização da AoI, ao propor duas novas técnicas: a *Last Generated - First Served - Conditional* (LGFS-C), para modelos de uma fonte e um servidor e a *Max Age First - Last Generated - First Served - Multiple preemption in Waiting* (MAF-LGFS-MW), para modelos de múltiplas fontes e múltiplos servidores. Ambas as técnicas, quando avaliadas, mostraram resultados iguais ou melhores do que as atuais técnicas no estado da arte, a depender dos parâmetros do modelo. Como última contribuição, o trabalho aplica uma metodologia de avaliação com o uso da ferramenta desenvolvida em três tipos de cenários de uso, em que é demonstrado o uso das simulações para melhorar o dimensionamento da infraestrutura em função dos requisitos de desempenho.

**Palavras-chave:** *age of information*, gerenciamento de pacotes, sistemas ciberfísicos, simulação

# Abstract

*Age of Information* (AoI) is a concept that defines a metric related to the degree of update that a monitor has in relation to one or more information sources, which send sequential update packets. The present study proposes the use of AoI to evaluate and optimize a Cyber Physical System characterized as a monitoring and decision-making system in the context of military command and control (C2) activity. In this type of system, agents periodically send their georeferenced positions to an operations center and common scenarios may have different degrees of criticality. To achieve this goal, this work developed a computational model to calculate the *Age of Information* (AoI), which was used as a basis for the implementation of a simulation tool. In addition to the aforementioned tool, this work contributed to the improvement of current package management techniques for the optimization of AoI, by proposing two new techniques: *Last Generated - First Served - Conditional* (LGFS-C), for single-source, single-server models, and *Max Age First - Last Generated - First Served - Multiple preemption in Waiting* (MAF-LGFS-MW), for multiple-source, multiple-server models. Both techniques, when evaluated, showed results equal or better than current state of the art techniques, depending on the model parameters. As a last contribution, this research applies an evaluation methodology, using the developed tool and AoI in three types of critical scenarios, in which the use of the tool is demonstrated to improve the infrastructure planning according to performance requirements.

**Keywords:** age of information, packet management, cyber-physical systems, simulation

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Introdução . . . . .	1
1.2	Definição do problema . . . . .	2
1.3	Contribuições . . . . .	2
1.4	Estrutura do documento . . . . .	3
<b>2</b>	<b>Fundamentação Teórica e Revisão do Estado da Arte</b>	<b>4</b>
2.1	Sistemas Ciberfísicos . . . . .	4
2.1.1	Definições . . . . .	5
2.1.2	Um sistema ciberfísico de monitoramento e decisão . . . . .	8
2.2	<i>Age of Information</i> . . . . .	9
2.2.1	Modelo canônico . . . . .	10
2.2.2	Modelos de filas básicas . . . . .	12
2.2.3	Gerenciamento de pacotes . . . . .	16
2.2.4	Modelos com múltiplas fontes e múltiplos servidores . . . . .	20
2.2.5	Funções de penalidade da AoI . . . . .	25
2.3	Revisão do estado da arte . . . . .	26
2.4	Resumo do capítulo . . . . .	29
<b>3</b>	<b>Definição do Modelo Computacional e Proposta das Técnicas LGFS-C e MAF-LGFS-MW</b>	<b>30</b>
3.1	Definição do Modelo . . . . .	30
3.1.1	Comando e Controle . . . . .	31
3.1.2	Métricas para Avaliação de Sistemas de C2 . . . . .	33
3.1.3	Descrição dos Componentes do Sistema . . . . .	34
3.2	Modelagem Computacional . . . . .	38
3.2.1	Descrição das Estruturas . . . . .	38
3.2.2	Implementação e distribuição da ferramenta . . . . .	43
3.3	A técnica LGFS-C . . . . .	44



3.3.1	Vida residual média . . . . .	44
3.3.2	Descrição da técnica . . . . .	45
3.3.3	Demonstração analítica . . . . .	48
3.3.4	Alcance e viabilidade da técnica . . . . .	50
3.4	A técnica MAF-LGFS-MW . . . . .	51
3.4.1	Descrição da técnica . . . . .	53
3.5	Resumo do capítulo . . . . .	55
<b>4</b>	<b>Resultados Experimentais</b>	<b>56</b>
4.1	Validação da Ferramenta de Simulação . . . . .	57
4.2	Resultados experimentais da técnica LGFS-C . . . . .	64
4.3	Resultados experimentais da técnica MAF-LGFS-MW . . . . .	66
4.4	Avaliação dos Cenários de Uso de Sistemas de Monitoramento e Decisão . .	69
4.4.1	Metodologia de Avaliação . . . . .	70
4.4.2	Definição dos Cenários . . . . .	72
4.5	Resultados da Avaliação dos Cenários de Uso . . . . .	73
4.6	Resumo do Capítulo . . . . .	81
<b>5</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>82</b>
	<b>Referências</b>	<b>84</b>

# Lista de Figuras

2.1	Sistema de monitoramento e decisão. . . . .	8
2.2	Publicações relacionadas à <i>Age of Information</i> (AoI) nos últimos anos. . .	10
2.3	Processo <i>Age of Information</i> (AoI) . . . . .	11
2.4	Representação do modelo de filas. . . . .	13
2.5	AoI médio para modelos de filas básicas. . . . .	16
2.6	Preempção na fila de espera (W). . . . .	18
2.7	Preempção em serviço (S). . . . .	18
2.8	AoI média para técnicas de gerenciamento de pacote . . . . .	20
2.9	Modelo básico com múltiplas fontes. . . . .	21
2.10	Modelo com múltiplos servidores. . . . .	23
3.1	Ciclo OODA. . . . .	31
3.2	Modelo genérico para um sistema de comando e controle. . . . .	32
3.3	Modelo proposto para o sistema. . . . .	35
3.4	Vida residual média para a distribuição Weibull. . . . .	45
3.5	Estrutura fila servidor com pacotes chegando fora de ordem. . . . .	46
3.6	Ilustração da técnica LGFS-C. . . . .	46
3.7	Modelo para a técnica MAF-LGFS-MW. . . . .	53
4.1	Representação do modelo computacional contendo uma <i>caixa preta</i> . . . . .	57
4.2	Validação para filas elementares. . . . .	58
4.3	Processo de convergência da AoI. . . . .	59
4.4	Validação para filas LCFS com precedência em serviço. . . . .	60
4.5	Validação para filas LCFS com precedência na espera. . . . .	61
4.6	Validação para a AoI de pico em uma fila FCFS do tipo M/M/1. . . . .	61
4.7	Validação para duas fontes com regime FCFS. . . . .	62
4.8	Validação para duas fontes com regime LCFS com precedência em serviço. . . . .	63
4.9	Validação para duas fontes com regime LCFS com precedência na espera. . . . .	63
4.10	Distribuições utilizadas para a avaliação da técnica LGFS-C. . . . .	64
4.11	Avaliação do cenário crítico. . . . .	65

4.12	Avaliação do cenário crítico. . . . .	66
4.13	Avaliação do cenário crítico. . . . .	66
4.14	Comparativo das técnicas para tempos de serviço exponenciais. . . . .	68
4.15	Comparativo das técnicas para tempos de serviço com distribuição de Weibull. . . . .	69
4.16	Abordagens para o dimensionamento do sistema. . . . .	70
4.17	Modelo como uma rede de filas. . . . .	71
4.18	Técnicas utilizadas na análise baseada em cenários. . . . .	75
4.19	Taxa média de serviço das fontes. . . . .	76
4.20	AoI e pAoI médias em função do número de agentes $N$ . . . . .	77
4.21	Avaliação do cenário crítico. . . . .	78
4.22	Avaliação do cenário pouco crítico. . . . .	79
4.23	Avaliação do cenário de normalidade. . . . .	80

# Lista de Tabelas

3.1	Comparação de técnicas no estado da arte para modelos com múltiplas fontes.	52
4.1	Métricas de desempenho para o cenário crítico. . . . .	73
4.2	Métricas de desempenho para o cenário pouco crítico. . . . .	73
4.3	Métricas de desempenho para o cenário de normalidade. . . . .	74
4.4	Parâmetros comuns para a avaliação dos cenários. . . . .	74

# Lista de Abreviaturas e Siglas

**AoI** *Age of Information.*

**AoSI** *Age of Served Information.*

**C2** Comando e Controle.

**CMTD** Cadeias de Markov de Tempo Discreto.

**COp** Centro de Operações.

**FCFS** *First Come First Served.*

**FDA** Função Distribuição Acumulada.

**FDP** Função Densidade de Probabilidade.

**GLO** Garantia da Lei e da Ordem.

**GPS** *Global Positioning System.*

**i.i.d.** Independentes e identicamente distribuídas.

**IP** *Internet Protocol.*

**LCFS** *Last Come First Served.*

**LGFS** *Last Generated First Served.*

**LTE** *Long Term Evolution.*

**MAF** *Max Age First.*

**MASIF** *Max Age of Served Information First.*

**MW** *Multiple preemption in Waiting.*

**NBU** *New Better than Used.*

**OODA** Observar, Orientar-se, Decidir e Agir.

**pAoI** *Age of Information* de pico.

**REQM** Raiz do Erro Quadrático Médio.

**SCF** Sistema Ciberfísico.

**TAI** Tempo de Autocorrelação Integrado.

**TIC** Tecnologia da Informação e Comunicações.

**VRM** Vida Residual Média.

# Capítulo 1

## Introdução

### 1.1 Introdução

A guerra moderna encontra-se cada vez mais dependente da tecnologia, sendo grande parte de suas ações travadas nos ambientes cibernético e informacional. A superioridade de informações e a rapidez das ações são fatores decisivos para o sucesso ou o fracasso uma operação. Os Sistemas de Comando e Controle (C2) são a base para o processo decisório em operações militares, garantindo aos comandantes a execução de seus ciclos com rapidez, precisão e oportunidade [1]. Esses sistemas são compostos por elementos físicos (agentes, sensores, decisores, etc) e computacionais (redes, algoritmos e outros). Nesse contexto, os sistemas de C2 podem ser estudados como Sistemas Ciberfísicos (SCFs), uma área de estudo que tem tomado grande relevância nos últimos anos.

Um Sistema Ciberfísico (SCF) contém um conjunto de interações entre processos físicos e computacionais, estando cada vez mais presentes no cenário tecnológico e sendo atualmente considerados umas das bases da Indústria 4.0 [2]. Os SCFs atuam geralmente através de modelos distribuídos complexos de larga escala, operando em ambientes multi-espaciais e multitemporais, os quais fazem uso de controles inteligentes e fusão multimodal de informações. Com a chegada da tecnologia 5G, as aplicações e a variedade dos dispositivos individuais tende a aumentar nos próximos anos, o que provoca desafios técnicos, entre os quais podem-se mencionar a definição de arquiteturas e métricas de desempenho para tais sistemas [3].

Para que a atividade de comando e controle atinja seus objetivos, torna-se necessário organizar e implementar o SCF pelo espaço de batalha. Para essa finalidade, diferentes sensores são alocados com o objetivo de enviar atualizações (*updates*) para um Centro de Operações (COp), como posição e velocidade de tropas e veículos, imagens, dados de radares, dados meteorológicos, entre outros. Dessa maneira, tal sistema militar pode ser enquadrado em uma categoria mais ampla de sistemas, a qual pode ser chamada de

sistemas de *monitoramento e decisão*. O desempenho desse tipo de sistema será função das tecnologias agregadas, as quais devem garantir a rapidez e a qualidade das informações que ali trafegam. Por esse motivo, é necessário estabelecer métricas que possam avaliar tais sistemas, especialmente aquelas relacionadas à rapidez e à oportunidade com que a informação flui entre os diferentes elementos.

As redes de Comando e Controle da guerra moderna contam com estruturas de comunicação bastante heterogêneas, o que ocorre devido à dinamicidade e à grande descentralização das ações. Esse fator tem grande impacto na velocidade com que as informações trafegam entre os nós da rede de C2, podendo ocasionar atrasos e a consequente desatualização do COp. Diferentes tipos de enlace como redes cabeadas, redes rádio em diferentes bandas, redes LTE (*Long Term Evolution*) e de satélites apresentam diferentes parâmetros de desempenho como taxa de transmissão e atraso.

Dentro desse contexto, a *Age of Information* (AoI), a qual pode ser traduzida como *Idade da Informação*, apresenta-se como um novo conceito e métrica, capaz de avaliar o quão atualizada está uma determinada informação em posse de um monitor em relação à fonte onde foi produzida. Trata-se de uma tema de pesquisa relativamente novo, que apresenta um crescente número de trabalhos publicados nos últimos anos, como uma proposta para solucionar o problema de regimes de envio de atualizações em redes de sensores [4]. O presente trabalho parte da premissa de que essa métrica pode ser aplicada para a avaliação de SCF, como por exemplo um sistema de monitoramento e decisão militar, possibilitando assim o estabelecimento de metodologias para a definição de arquiteturas e o dimensionamento da infraestrutura de sistemas de C2. Esse novo paradigma contribuirá para a melhora da qualidade da informação que é recebida pelos Centros de Operação (COp) militares.

## 1.2 Definição do problema

O problema central do presente estudo é utilizar o conceito de *Age of Information* (AoI) para avaliar e otimizar sistemas ciberfísicos, em especial um sistema de monitoramento e decisão voltado para a atividade de comando e controle militar.

## 1.3 Contribuições

As contribuições deste trabalho são as seguintes:

- Definição de um modelo computacional para a avaliação da AoI em sistemas ciberfísicos.



- Desenvolvimento de uma ferramenta de simulação e avaliação da AoI para sistemas ciberfísicos, disponibilizada para a comunidade em repositório público.
- Proposta de uma nova técnica de gerenciamento de pacotes para a otimização da AoI em modelos de filas simples: *Last Generated - First Served - Conditional* (LGFS-C).
- Proposta de uma nova técnica de gerenciamento de pacotes para a otimização da AoI em modelos de múltiplas fontes e múltiplos servidores: *Max Age First - Last Generated - First Served - Multiple preemption in Waiting* (MAF-LGFS-MW).
- Definição de uma metodologia para a avaliação e dimensionamento de sistemas ciberfísicos com o uso das métricas da AoI, em cenários típicos da atividade militar.

## 1.4 Estrutura do documento

O presente trabalho inicia, no Capítulo 2, com um apanhado dos principais conceitos teóricos que servirão de base para o entendimento desta pesquisa, em que são abordados os conceitos de Sistema Ciberfísico (SCF) e de *Age of Information* (AoI), com uma apresentação dos principais trabalhos no estado da arte. O Capítulo 3 apresenta a modelagem do sistema em estudo e discute as principais contribuições deste trabalho, quais sejam a ferramenta de simulação e as duas novas técnicas para o gerenciamento de pacotes para a otimização da AoI. O Capítulo 4 apresenta os resultados experimentais obtidos para cada uma das contribuições do trabalho, incluindo ainda resultados de avaliações do sistema de monitoramento e decisão baseadas em cenários típicos da atividade militar. Finalmente, o Capítulo 5 apresenta as conclusões e trabalhos futuros desta pesquisa.

# Capítulo 2

## Fundamentação Teórica e Revisão do Estado da Arte

Este capítulo tem por finalidade apresentar os conceitos teóricos necessários para o entendimento deste trabalho de pesquisa. Será introduzido, inicialmente, o conceito de Sistema Ciberfísico, apresentando-se sua definição e particularizando-se um sistema de monitoramento e decisão genérico. Na sequência, será apresentado o tema *Age of Information*, definindo-se desde o modelo canônico até os modelos mais complexos presentes na literatura. Ao final do capítulo, será realizada uma revisão do estado da arte, apresentando-se os principais e mais atuais estudos na área.

### 2.1 Sistemas Ciberfísicos

Os Sistemas Ciberfísicos (SCF) formam uma área de estudo que abrange várias disciplinas computacionais e de engenharia, tais como redes e comunicações complexas, controle, fusão de informações, processamento de sinais e otimização. Observam-se hoje uma infinidade de aplicações em número crescente, que conformam aplicativos de sistemas complexos e que apresentam diversos requisitos e demandas de inteligência, flexibilidade, adaptabilidade, robustez e resiliência. Um aspecto chave dos SCF é a integração de tecnologias de informação, tecnologias operacionais em termos de sistemas embarcados, sistemas de controle e sistemas físicos.

Conforme o número de aplicações e a quantidade de dispositivos em uso aumenta, observa-se também um aumento na complexidade dos sistemas ciberfísicos. Esses sistemas configuram em muitos casos sistemas distribuídos complexos de grande escala, que integram sistemas de comunicação, redes de sensores, sistemas multiagentes, sistemas humanos, entre outros, os quais operam em ambientes não estruturados e com escalas multi-espaciais e temporais, que podem fazer uso de controles inteligentes e fusão de informações

multimodais. Essa complexidade dificulta definir nesses sistemas o seu desempenho com alta precisão e confiabilidade, principalmente para os casos de aplicações em tempo real. Dessa maneira, esse tipo de sistema exige o desenvolvimento de novos métodos que possam lidar com os avanços atuais, tais como inteligência, autonomia, funcionalidade, adaptabilidade, robustez, segurança cibernética, entre outros. De maneira similar, é importante considerar que o grande número de dispositivos operacionais dá origem a vários desafios técnicos, tais como a arquitetura necessária, a administração da incerteza nas métricas de desempenho e a introdução de inteligência em tais sistemas.

Para melhor entender esses conceitos e os desafios de pesquisa, nesta seção serão apresentados os principais conceitos associados aos SCF. Será realizada, também, uma particularização para um sistema do tipo monitoramento e decisão, o qual será o foco de estudo e aplicação dos métodos desenvolvidos nesta pesquisa.

### 2.1.1 Definições

Sistemas Ciberfísicos são sistemas incorporados em rede espacialmente distribuídos, sensíveis ao tempo e em várias escalas, que conectam o mundo físico ao mundo cibernético por meio de sensores e atuadores. Em [5] são definidos como sistemas onde processos físicos interagem com processos computacionais, cada um operando em diferentes escalas de tempo e espaço, interagindo uns com os outros de diversas maneiras. Sendo assim, as características essenciais desse tipo de sistema são:

- ciber capacidades em todos os componentes físicos.
- alto grau de automação.
- uso de redes em diversas escalas.
- integração em múltiplas escalas temporais e espaciais.
- reorganização dinâmica.

As diversas aplicações dos SCFs contemplam de forma geral uma enorme quantidade de dispositivos, em áreas de diversos tamanhos, que podem interagir ou não uns com os outros, e que trabalham com um conjunto de requerimentos de desempenho, que envolvem métricas de velocidade, disponibilidade e confiabilidade. Dessa maneira, a funcionalidade dos dispositivos individuais vem aumentando ao longo do tempo, o que pode exigir arquiteturas e sistemas operacionais específicos para esses sistemas. Ainda sobre a modelagem arquitetural, ressalta-se a importância da decomposição da execução da aplicação em unidades quantizadas de comunicação ou computação, como por exemplo pacotes, bits, unidades de tempo, etc. Essa decomposição permite que o sistema seja avaliado de maneira

parametrizada, para explorar a influência que cada parâmetro exerce no comportamento da execução da aplicação.

Como um exemplo de abordagem para esse problema, o estudo de [3] apresenta o conceito de modelos compactos de sistemas, motivado pela necessidade de se aumentar o grau de abstração e de se obter modelos mais concisos que sejam capazes de interoperar. Dessa maneira, um sistema pode ser decomposto em diversos submodelos, os quais podem abordar as seguintes vertentes:

- modelo do fenômeno físico: visualizar a computação como um ato físico, o qual envolve custos de tempo, energia, confiabilidade, entre outros.
- modelo computacional: define as abstrações e as interações entre os diversos componentes, o fluxo de informações, os algoritmos e as restrições dos componentes.
- modelo da autoconsciência e adaptação: define os diversos estados do sistema e das transições entre os mesmos.
- modelo arquitetural: permite a abstração dos diferentes *hardwares* e a estimação de custos computacionais como latência, taxa de dados, entre outros.
- modelo do agendamento de tarefas: auxilia na otimização dos recursos disponíveis, alocando e ordenando as diferentes tarefas.

Um dos principais desafios na modelagem de SCF é a heterogeneidade dos sistemas no mundo real, nos quais diferentes componentes interagem e concorrem pela execução de diferentes tarefas e processos. Além disso, muitos sistemas requerem modelos que representem seu comportamento em tempo real, em diferentes níveis de granularidade e de detalhamento[6].

Um ponto interessante são as aplicações dos sistemas ciberfísicos de baixa latência. Por exemplo, imagens de câmeras de veículos ou de pontos estratégicos em uma área tais como prédios ou árvores, podem ser usadas para gerar informações que descrevem os arredores e a movimentação de seres vivos nesses lugares. Outro exemplo, é um sistema de cirurgia remota que precisa atualizar as posições das ferramentas e as imagens cirúrgicas. Em ambos os casos, os dois exemplos são caracterizados por uma fonte que gera mensagens de atualização de status com registros de data e hora e que são transmitidas através de uma rede para um ou mais monitores. Em algumas destas aplicações, um atraso de alguns milissegundos pode ser muito. Com a implantação de redes móveis de alta velocidade, tais como o 5G, vem se observando um avanço nas pesquisas direcionadas às redes de baixa latência a alta confiabilidade. Alguns cenários da rede 5G incluem a interação homem máquina e a internet tátil [7], as quais exigem atraso máximo de poucos milissegundos.

Adicionalmente, a computação em nuvem e de borda propõem a eliminação de atrasos na propagação na ordem de 40 ms para rotas transcontinentais.

Existem também aplicações que toleram alta latência. Como exemplo pode ser mencionado um projeto financiado pela União Europeia e realizado pelo *Business Experiments in Cyber Physical Production Systems* que visa implementar um sistema de monitoramento em diversas fábricas do setor de produção de calçados em Portugal [8]. O objetivo é automatizar a produção e estabelecer uma rede de sensores capaz de gerar informações que irão subsidiar decisões dos trabalhadores no tocante à resolução de problemas. Essa utilização dos SCF vem sendo vista como a base da indústria 4.0.

Algumas pesquisas em andamento nos últimos anos [6, 3] citam vários desafios a serem abordados no desenvolvimento de tecnologias de SCF, dentre os quais destaca-se aqueles de interesse neste estudo:

- Modelagem matemática e simulação: A interação dos sistemas computacionais com o mundo real impõe o acompanhamento das constantes mudanças de um mundo contínuo dentro dos processos discretos de um computador. Dessa forma, é necessária a harmonização de modelos matemáticos que capturem com a precisão necessária o mundo real para o sistema discreto de um sistema computacional e que lidem com a escala e complexidade desses sistemas para a predição de métricas de interesse. Ferramentas de simulação que integrem essas complexidades e escalas também são necessárias.
- Tempo, Comunicação e Conectividade: os dispositivos devem ser capazes de comunicar entre si, sincronizando e operando com restrições de tempo. Para algumas aplicações, o tempo é crucial e isso se aplica à comunicação, conectividade e detecção. Na detecção, eventos importantes não podem ser perdidos.
- Espaço-tempo: a representação de eventos em determinados momentos implica em problemas de sincronicidade, frequência e granularidade, que podem resultar em medições contrárias vindas de diversos dispositivos ou na combinação de resultados.
- Incerteza: a combinação de vários sistemas para formar um sistema maior provoca uma incerteza nos resultados do sistema e no seu desempenho. Alguns modelos conceituais definem como lidar com a incerteza e com o uso das informações disponíveis [9].

Existem ainda outros desafios de pesquisa nas técnicas de robustez, extensibilidade, auto-proteção, autonomia, localização, auto-consciência, segurança, privacidade e inteligência [6], os quais não serão especificamente abordados nesta pesquisa.

## 2.1.2 Um sistema ciberfísico de monitoramento e decisão

Aplicações de monitoramento e tomada de decisão através de redes de sensores são exemplos típicos de sistemas ciberfísicos, que tem aplicação direta para a atividade de comando e controle em operações militares. Nesse tipo de sistema, sensores coletam informações e as enviam para um centro de monitoramento através de uma rede de comunicações. Uma vez recebidas, essas informações alimentarão um processo de tomada de decisão. Dentro desse contexto, o presente estudo define a arquitetura de um sistema de monitoramento e decisão genérico, o qual pode ser visualizado na Fig. 2.1. O palavra *decisão* foi incluída com o intuito de se destacar a necessidade de o monitor estar sempre atualizado em relação ao ambiente físico, propiciando que o mesmo possa atuar com oportunidade. Conforme será apresentado com mais detalhes no Cap. 3, esse fluxo de decisão não será incluído no modelo principal.

A seguir, serão descritos cada um dos componentes desse sistema.

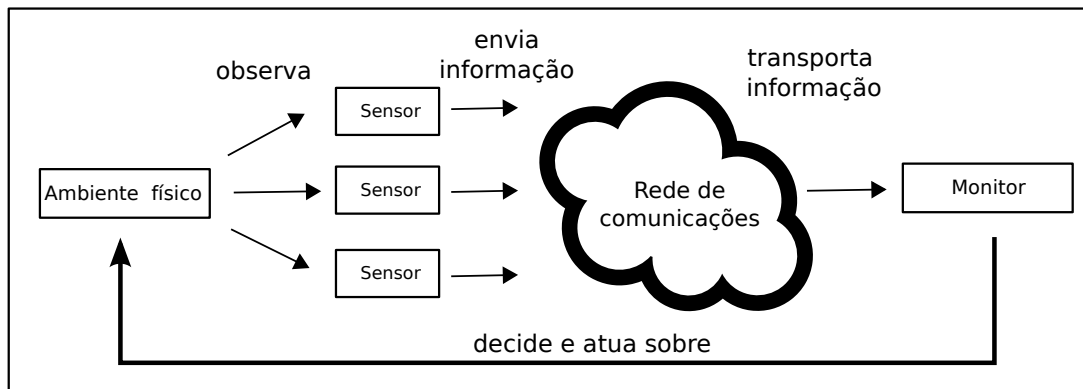


Figura 2.1: Arquitetura de um sistema ciberfísico de monitoramento e decisão.

### Ambiente físico

O ambiente físico é o local no qual ocorrem diversos eventos e processos dentro de um contexto específico. O objetivo final do sistema é ter a consciência sobre esses processos, para manter o monitor sempre atualizado e com a capacidade de modificá-los.

### Sensores

Para que se obtenha a consciência sobre o ambiente físico, um conjunto de sensores é utilizado para observar e enviar informações para o monitor. Esses sensores podem ser definidos como um subsistema ciberfísico. No componente físico, encontram-se equipamentos e *hardwares* utilizados para a coleta e preparação da informação para envio. No *ciber*, situam-se os elementos computacionais como os algoritmos de amostragem e codificação, assim como os processos de controle do fluxo da informação.

## Rede de comunicações

A rede de comunicações possui a finalidade de transportar os dados gerados pelos sensores até o monitor, e também é considerado um subsistema ciberfísico. Entre os componentes dessa rede, pode-se citar as estruturas de filas e seus regimes de serviço, bem como os algoritmos de escolha e otimização de rotas.

## Monitor

O monitor é o local onde as informações dos sensores convergem e, ao serem recebidas e processadas, permitem subsidiar um processo de tomada de decisão. No monitor encontra-se uma estrutura de servidores que processa a informação recebida, tornando-a útil para o processo de tomada de decisão. Entre os parâmetros do monitor pode-se citar as técnicas de escalonamento, a priorização da informação recebida e a carga nos servidores. O monitor deve estar sempre atualizado em relação ao ambiente físico, possibilitando que atue sobre o mesmo com eficiência e precisão.

O sistema de monitoramento e decisão aqui apresentado será retomado adiante no Cap. 3, em que será particularizado para o caso de um sistema de comando e controle militar.

## 2.2 *Age of Information*

A AoI teve sua primeira publicação em 2011 no estudo de [10], em que surgiu como um conceito para se caracterizar a atualidade (*freshness*) do conhecimento sobre determinado processo observado remotamente. Desde então, o assunto já ultrapassou a quantidade de 1000 publicações (até 2021), com uma tendência crescente conforme mostra a Fig. 2.2. A referida área de estudo conta atualmente com vários grupos de pesquisa espalhados pelo mundo, em que os principais tópicos de interesse são os seguintes:

- Teoria da informação.
- Teoria de filas.
- Codificação de fonte e canal [11].
- Teoria dos jogos [12].
- Sistemas veiculares [13, 14].
- Sistemas com energia autônoma [15].
- Veículos Aéreos Não Tripulados [16].

- *Machine Learning* [17].
- Controle e automação [18].
- Escalonadores de tarefa [19].



Figura 2.2: Publicações relacionadas à *Age of Information* (AoI) nos últimos anos. Dados obtidos através da plataforma *Scopus*.

Em uma definição formal, a AoI pode ser definida como a quantidade de tempo decorrida desde a geração da última mensagem recebida com sucesso que contém informações a respeito de um determinado sistema [20]. O presente capítulo apresentará inicialmente o modelo canônico do processo AoI, para logo seguir pela modelagem em sistemas de fila básicos com uma única fonte e destino. Na sequência será apresentado o conceito de gerenciamento de pacotes (*packet management*) para a otimização da AoI em diferentes regimes de serviço. Por fim, serão introduzidos os principais conceitos atuais sobre o estudo da AoI em sistemas com múltiplas fontes e múltiplos servidores, semelhantes ao modelo do sistema de monitoramento e decisão definido anteriormente.

### 2.2.1 Modelo canônico

A *Age of Information* pode ser vista como o grau da atualidade de uma informação sobre determinado processo ou entidade. Em termos analíticos, um determinado pacote<sup>1</sup> com

<sup>1</sup>Os termos pacote, *update* e atualização são intercambiáveis neste estudo. *Pacote*, no presente contexto, refere-se ao nível da aplicação e não a um protocolo específico.



tempo de geração  $u$  possui uma idade (AoI) dada por  $t - u$  em um tempo  $t \geq u$ . Dado que um monitor recebe uma atualização no tempo  $t$  e que essa atualização possui um *timestamp*  $u(t)$ , a AoI é definida como o processo randômico  $\Delta(t) = t - u(t)$ .

No momento em que o monitor recebe um pacote de atualização, a AoI observa um decaimento instantâneo. Entre os recebimentos, a AoI aumenta de maneira linear com o tempo. Esse comportamento é mostrado no gráfico da Fig. 2.3, onde se observa o formato *dente-de-serra* característico do processo AoI.

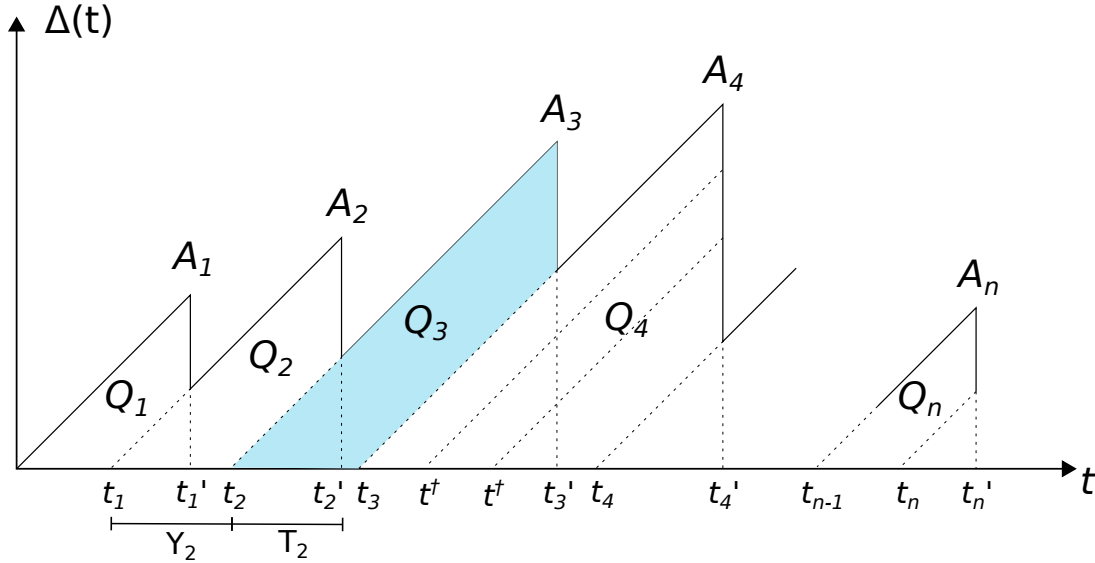


Figura 2.3: Comportamento característico do processo *Age of Information*.

Ainda na mesma figura, podem ser definidos os seguintes termos utilizados no estudo da *Age of Information*:

- $t_i$ : tempo de chegada do pacote  $i$ .
- $t'_i$ : tempo de partida do pacote  $i$ .
- $t^\dagger$ : pacote perdido ou descartado.
- $Y_i$ : tempo entre as chegadas dos pacotes  $i - 1$  e  $i$  (*interarrival time*).
- $T_i$ : tempo entre a chegada do pacote  $i$  e a partida do mesmo pacote. Também chamado de *system time*.
- $A_i$ : valores de pico da AoI.

Uma das métricas de maior interesse é a AoI média  $\langle \Delta \rangle_\tau$ , a qual pode ser calculada somando-se a área sob o gráfico e dividindo-a pelo tempo  $\tau$ . Em um regime estacionário, pode-se defini-la conforme a Eq. 2.1.

$$\langle \Delta \rangle_\tau = \frac{1}{\tau} \int_0^\tau \Delta(t) dt \quad (2.1)$$

considerando-se o limite em que  $\tau$  tende a infinito. Ao decompor a Fig. 2.3 em seções trapezoidais  $Q_i$ ,  $i = 1, 2, \dots$  e ao considerar o processo como ergódico e estacionário em um tempo  $\tau$ , pode-se obter o valor médio da AoI em função de  $Y_n$  e  $T_n$  [4], conforme mostra a Eq. 2.2.

$$\Delta = \lim_{\tau \rightarrow \infty} \frac{\sum Q_i}{\tau} = \frac{E[Q_n]}{E[Y_n]} = \frac{E[Y_n T_n] + E[Y_n^2]/2}{E[Y_n]} \quad (2.2)$$

A Eq. 2.2 é amplamente utilizada como ponto de partida para a maioria dos modelos, sendo aplicada a diversos tipos de filas e a diversos regimes de serviço. Destaca-se que as variáveis aleatórias  $Y$  (*interarrival*) e  $T$  (*system time*) não são independentes e logo o valor esperado do produto não pode ser decomposto como o produto dos valores esperados individuais. Esse fator torna a obtenção dos modelos analíticos um tanto quanto complexa.

Além da AoI média, uma outra métrica de interesse é a *Age of Information* de pico (pAoI) a qual foi introduzida por [21]. Os valores de pico do processo podem ser obtidos pela Eq. 2.3.

$$A_i = T_i + Y_i \quad (2.3)$$

para  $i \geq 0$ . Dessa maneira, o valor da AoI de pico média é dado pela média dos valores de  $A_i$ , considerando-se o número de amostras tendendo a infinito, conforme mostra a Eq. 2.4.

$$\Delta^{(p)} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N A_n \quad (2.4)$$

Em termos de expectativa, a *Age of Information* de pico pode ser obtida em função dos valores esperados de  $T_n$  e  $Y_n$  [22], conforme a Eq. 2.5.

$$\Delta^{(p)} = E[A] = E[T_n] + E[Y_n] \quad (2.5)$$

Observa-se que a AoI de pico evita o termo  $E[T_n Y_n]$ , sendo a modelagem analítica deste processo mais simples que a AoI média. Não obstante, a *Age of Information* de pico captura as características elementares do processo de envelhecimento da informação, as quais são representadas pelo tempo do sistema  $T_n$  e pelo tempo entre chegadas  $Y_n$ .

## 2.2.2 Modelos de filas básicas

Uma vez definido o modelo canônico do processo *Age of Information*, esta seção apresentará os principais modelos de obtenção da AoI média e da AoI de pico média em modelos básicos de filas [23]. A representação desses modelos segue a notação de *Kendall* simplificada [24], em que a definição de uma fila é dada por  $A/S/c$ , onde:

- $A$ : Processo de chegada dos *pacotes*.  $D$  representa um processo determinístico e  $M$  um processo de *Poisson* (*Markoviano*);
- $S$ : Processo de serviço da fila. Segue a mesma representação do processo de chegada;
- $c$ : Número de servidores.

A Fig. 2.4 mostra a representação gráfica do modelo básico de uma fila, a qual será utilizada no decorrer deste trabalho.

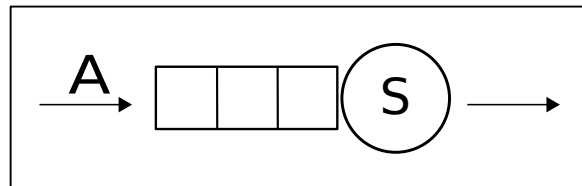


Figura 2.4: Representação gráfica de uma fila básica com processo de chegada  $A$  e processo de serviço  $S$ .

Na Seção 2.2.3 serão acrescentadas outras notações ao modelo de representação de filas, as quais podem representar regimes de serviço e estratégias de gerenciamento de pacotes.

### Modelo $M/M/1$

A modelagem analítica da AoI para modelos básicos de filas foi apresentada pela primeira vez em [4], a qual é referência para diversos trabalhos que seguiram. Considere-se inicialmente o modelo  $M/M/1$ , o qual caracteriza uma fila com chegadas de acordo com um processo de *Poisson* com taxa de chegada  $\lambda$ , e com um processo de serviço exponencial com taxa igual a  $\mu$ , possuindo a fila um único servidor. O regime da fila é o *First Come First Served* (FCFS), ou seja, os pacotes são servidos na mesma ordem de sua chegada na fila.

Uma vez que o processo de chegada é um processo de *Poisson*, a distribuição dos tempos entre chegadas (*interarrival times*) seguem uma distribuição exponencial com média  $1/\lambda$  e logo  $E[Y] = 1/\lambda$  e  $E[Y^2] = 2/\lambda^2$ . Os tempos de serviço, por sua vez, possuem média igual a  $1/\mu$  e a carga no servidor é dada pela razão entre a taxa de chegadas e a taxa de saídas, sendo  $\rho = \lambda/\mu$ .

Partindo-se da expressão obtida em 2.2, necessita-se obter o termo  $E[YT]$ , o qual pode ser encontrado decompondo-se  $T$  (tempo do sistema) na expressão da Eq. 2.6.

$$T_i = W_i + S_i \quad (2.6)$$

Nessa última equação,  $W_i$  é o tempo de espera na fila do pacote  $i$  e  $S_i$  é o tempo de serviço, sendo  $E[S] = 1/\mu$ . Dessa forma, assumindo-se que o tempo de serviço é independente do tempo de chegada, obtém-se a partir da Eq. 2.6, a Eq. 2.7.

$$E[YT] = E[Y(W + S)] = E[YW] + E[Y]E[S] \quad (2.7)$$

Após algumas manipulações envolvendo a distribuição do tempo de espera  $W$  em filas  $M/M/1$ , obtém-se a AoI média, a qual é dada pela Eq. 2.8.

$$\Delta = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} + \frac{\rho^2}{1 - \rho} \right) \quad (2.8)$$

A *Age of Information* de pico média, conforme mostrado em [25], é dada pela equação 2.9.

$$\Delta^{(p)} = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} + \frac{\rho}{1 - \rho} \right) \quad (2.9)$$

### Modelo D/M/1

Neste modelo o processo de chegada é assumido como determinístico com intervalos de chegada iguais a um tempo  $D$ . Dessa forma, a taxa de chegada é dada por  $\lambda = 1/D$  e os tempos de serviço seguem a mesma distribuição exponencial do modelo  $M/M/1$ , com média igual a  $1/\mu$ . Substituindo esses valores em 2.2, obtém-se a Eq. 2.10.

$$\Delta = \frac{1}{D} \left( \frac{D^2}{2} + DE[T] \right) \quad (2.10)$$

Para se obter o valor esperado do tempo de sistema  $T$ , o mesmo é decomposto na soma do tempo de espera e do tempo de serviço. Dessa forma, conforme mostrado por [4], chega-se à Eq. 2.11.

$$E[T] = E[S] + E[W] = \frac{1}{\mu} + \frac{\beta}{\mu(1 - \beta)} \quad (2.11)$$

Nessa equação,  $0 \leq \beta \leq 1$  é a solução da equação  $\beta = L_X(\mu(1 - \beta))$ , sendo  $L_X$  a transformada de Laplace da distribuição dos tempos entre chegadas. Considerando-se as chegadas determinísticas com período  $D$ ,  $\beta$  pode ser calculado pela Eq. 2.12.

$$\beta = e^{-\mu(1-\beta)D} = -\rho \mathcal{W} \left( -\rho^{-1} e^{(-1/\rho)} \right) \quad (2.12)$$

Ainda na Eq. 2.12,  $\mathcal{W}$  é a função W de Lambert<sup>2</sup> e  $\rho$  é a carga no servidor, dada por  $\rho = 1/(\mu D)$ . A partir de 2.10 e 2.11, obtém-se, então, a AoI média para a fila  $D/M/1$ , a qual é calculada pela Eq. 2.13.

$$\Delta = \frac{1}{\mu} \left( \frac{1}{2\rho} + \frac{1}{1-\beta} \right) \quad (2.13)$$

### Modelo M/D/1

O último modelo elementar a ser abordado é o  $M/D/1$ , o qual assume chegadas em um processo de Poisson e tempo de serviço determinístico. Assim,  $E[Y] = 1/\lambda$  e  $E[Y^2] = 2/\lambda^2$ , de maneira análoga ao modelo  $M/M/1$ . Já o tempo de serviço possui uma taxa  $\mu = 1/D$  e o tempo do sistema é obtido por  $T_i = W_i + D$ . Logo,  $E[YT]$  é dado pela Eq. 2.14.

$$E[YT] = E[YW] + DE[Y] \quad (2.14)$$

A obtenção da expressão final da AoI para este modelo, demonstrada em [26], pode ser vista na Eq. 2.15.

$$\Delta = \frac{1}{\mu} \left( \frac{2}{2(1-\rho)} + \frac{1}{2} + \frac{(1-\rho)e^\rho}{\rho} \right) \quad (2.15)$$

### Otimizando a AoI em filas elementares

Para os três modelos elementares apresentados, observa-se que a AoI é obtida em função da taxa de serviço  $\mu$  e da carga no servidor  $\rho$ . Fixando-se o valor da taxa de serviço, pode-se analisar a variação da AoI em função de  $\rho$  para cada um dos modelos. A Fig. 2.5 apresenta essa relação, variando-se o valor da carga no servidor no intervalo  $[0.1, 0.9]$ .

Observa-se também a existência de mínimos em cada um dos modelos, os quais podem ser calculados diferenciando-se a função  $\Delta$  e igualando-se o resultado a zero. Dessa maneira, os valores aproximados de  $\rho$  que minimizam a AoI são:

- $M/M/1$ : 0.53.
- $M/D/1$ : 0.625.
- $D/M/1$ : 0.515.

---

<sup>2</sup>A função W de Lambert é a função multivalorada cujos valores são dados pela função inversa de  $f(W) = We^W$ .

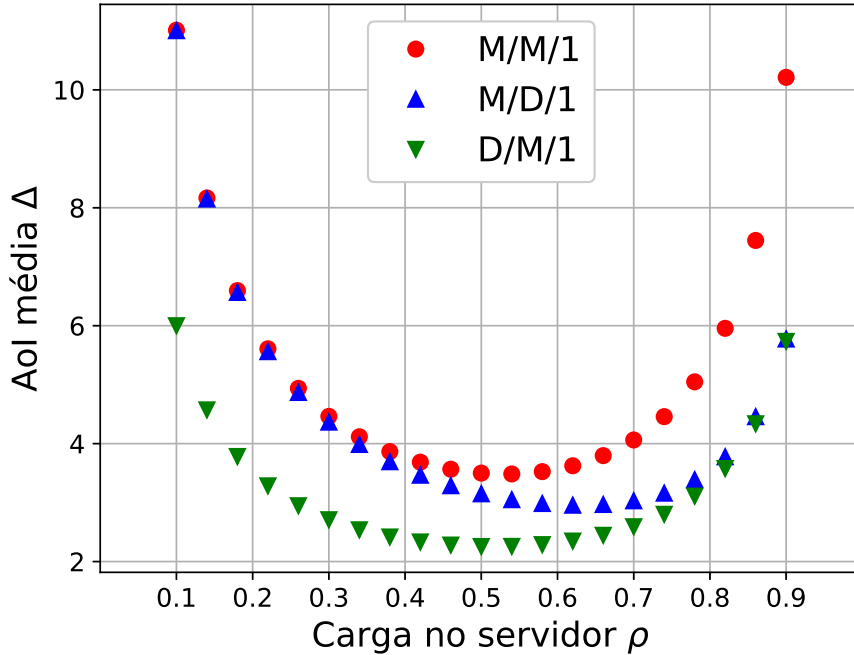


Figura 2.5: AoI médio em função da carga no servidor para os 3 modelos básicos de fila. A taxa de serviço foi fixada em  $\mu = 1$ .

### 2.2.3 Gerenciamento de pacotes

A partir dos primeiros estudos sobre otimização da AoI em sistemas de filas, verificou-se a possibilidade de se atuar ao nível dos pacotes, tanto no processo de chegada quanto no processo de saída. Nesse contexto, [21] e [27] realizaram os primeiros estudos com a utilização de técnicas de gerenciamento de pacotes (*packet management*). O gerenciamento de pacotes tem o objetivo de otimizar a AoI com base nas atualizações (*updates*) e pode ser realizado através de diversas técnicas, as quais serão apresentadas a seguir.

#### Filas com capacidade limitada

Um das maneiras de otimizar a AoI é limitar a capacidade de armazenamento das filas de modo que novas chegadas sejam bloqueadas enquanto existem pacotes em serviço e/ou na fila de espera. Essa técnica tem por finalidade evitar que os pacotes envelheçam na fila enquanto esperam a sua vez de serem servidos.

O trabalho de [21] propõe uma notação que consiste em uma extensão da notação de Kendall, apesar de ainda não haver um consenso entre os autores da área. Essa notação acrescenta um quarto termo cujo significado é a capacidade de armazenamento da fila. Dessa maneira, uma fila  $A/S/1/1$  possui a capacidade de armazenar apenas 1 pacote (aquele atualmente em serviço), sendo que quaisquer novas chegadas serão bloqueadas e

descartadas caso haja um pacote em serviço. Caso o servidor esteja livre, o pacote inicia imediatamente seu serviço e zera seu tempo de espera.

De maneira semelhante, uma fila  $A/S/1/2$  possuirá espaço para apenas 1 pacote em serviço e 1 pacote em espera. Nesse caso um novo pacote pode ser imediatamente servido, ficar em espera ou ser descartado, o que depende do estado atual da fila. De maneira geral a notação se estende para  $A/S/c/m$ , em que  $m$  é a capacidade máxima da fila, inclusive o *update* atualmente em serviço.

Conforme demonstrado em [27], os valores médios para a AoI em filas com capacidade limitada e processos de chegada e saída estocásticos (não determinísticos) podem ser obtidos pelas expressões mostradas na Eq. 2.16.

$$\Delta_{M/M/1/1} = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} + \frac{\rho}{1 + \rho} \right) \quad (2.16a)$$

$$\Delta_{M/M/1/2} = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} + \frac{2\rho^2}{1 + \rho + \rho^2} \right) \quad (2.16b)$$

Ainda de acordo com [27], os valores para a AoI de pico média para essas filas são obtidos pela Eq. 2.17.

$$\Delta_{M/M/1/1}^{(p)} = \frac{1}{\lambda} + \frac{2}{\mu} \quad (2.17a)$$

$$\Delta_{M/M/1/2}^{(p)} = \frac{1}{\lambda} + \frac{3}{\mu} - \frac{2}{\lambda + \mu} \quad (2.17b)$$

## Filas com preempção

Outra técnica de gerenciamento de pacotes é a utilização da preempção no recebimento de novos *updates*. Essa preempção ocorre quando um pacote recém chegado substitui outro mais velho que encontra-se na fila de espera ou até mesmo um pacote que já esteja em serviço. As Fig. 2.7 e 2.6 ilustram respectivamente os dois casos de preempção anteriormente citados.

No estudo de [27] a preempção foi sinalizada na notação pelo caractere asterisco (\*). Dessa forma, dois modelos principais foram objeto de diversos estudos no tocante ao gerenciamento de pacotes:  $A/S/1/2^*$  e  $A/S/1/1^*$ . O modelo  $A/S/1/2^*$  estabelece que um pacote recém chegado deverá substituir um pacote já existente na fila de espera, caso ele exista. Dessa maneira, esse regime garante que os pacotes na espera sejam sempre atualizados por pacotes mais frescos. Cabe ressaltar que alguns sistemas admitem a chegada de pacotes fora de ordem, o que ocorre por exemplo quando se modela o tempo

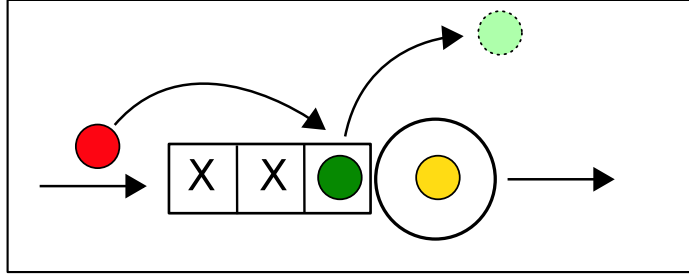


Figura 2.6: Preempção na fila de espera (W).

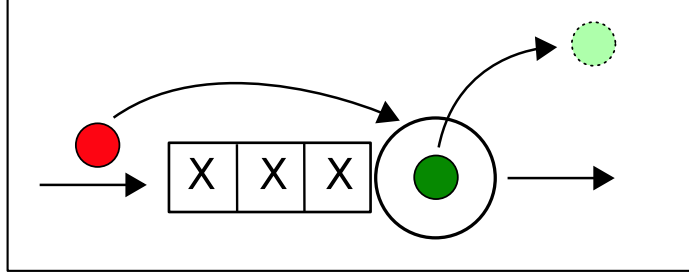


Figura 2.7: Preempção em serviço (S).

de transmissão entre dois nós de uma rede [28]. Nesse caso, é necessária uma checagem para que se verifique se o pacote recém chegado é mais fresco que o já existente na fila de espera. Caso positivo, deverá ocorrer a preempção, caso negativo, o pacote recém chegado deverá ser descartado.

O modelo  $A/S/1/1^*$ , por sua vez, admite a preempção do pacote em serviço, uma vez que não possui espaço de espera. Sendo assim, um pacote recém chegado, mais fresco, interromperá o serviço do pacote atual, caso ele exista, em que o descarta e assume sua posição no servidor. Nesse esquema, é importante ressaltar a existência do risco de saturação da fila, uma vez que a taxa de pacotes descartados pode ser muito alta, acabando por diminuir o número de pacotes efetivamente entregues.

Os modelos com preempção são abordados por outros autores [29, 30, 31] através de uma notação diferente. Essa notação alternativa utiliza as siglas LCFS-W e LCFS-S, as quais representam o regime *Last Come First Served* (LCFS), ou seja, o último pacote a chegar é o primeiro a ser servido, ocorrendo a preempção apenas na fila de espera (W - *waiting*) ou no próprio serviço (S - *service*).

A demonstração para a obtenção analítica da AoI média para os modelos de fila com preempção pode ser visualizada em [27], sendo dada para processos de chegada e saída estocásticos. Esses valores são dados pelas Eq. 2.18a e 2.18b.

$$\Delta_{M/M/1/1^*} = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} \right) \quad (2.18a)$$



$$\Delta_{M/M/1/2^*} = \frac{1}{\mu} \left( 1 + \frac{1}{\rho} + \frac{\rho^2(1 + 3\rho + \rho^2)}{(1 + \rho + \rho^2)(1 + \rho)^2} \right) \quad (2.18b)$$

Ainda para o modelo  $M/M/1/2^*$ , a AoI de pico média é dada pela Eq. 2.19.

$$\Delta_{M/M/1/2^*}^{(p)} = \frac{1}{\mu} + \frac{\lambda}{(\lambda + \mu)^2} + \frac{1}{\lambda} + \frac{\lambda}{\mu(\lambda + \mu)} \quad (2.19)$$

As expressões para a *Age of Information* média em modelos de fila com chegadas ou saídas determinísticas [26] podem ser vistas nas Eq. 2.20(a a d).

$$\Delta_{M/D/1/1^*} = \frac{1 \exp(\rho)}{\mu \rho} \quad (2.20a)$$

$$\Delta_{D/M/1/1^*} = \frac{1}{\mu} \left( 1 + \frac{1}{2\rho} \right) \quad (2.20b)$$

$$\Delta_{M/D/1/2^*} = \frac{1}{\mu} \left[ \frac{1}{1 + \rho \exp(\rho)} \left( \frac{1}{2} + \frac{1}{\rho} \right) + \frac{\exp(\rho) - (1 + \rho)}{\rho \exp(\rho)} + \frac{3}{2} \right] \quad (2.20c)$$

$$\Delta_{D/M/1/2^*} = \frac{1}{\mu} \left( \frac{1}{2\rho} + \frac{1}{1 - \rho^{-1} \exp(-\rho^{-1})} \right) \quad (2.20d)$$

## Comparativo das técnicas LCFS

Os valores da AoI média em função da carga no servidor, ao considerar uma taxa de serviço unitária, são mostrados na Fig. 2.8. Observa-se que todas as curvas, com exceção do modelo  $M/D/1/1^*$ , são funções estritamente decrescentes. Cabe ressaltar que esses resultados são válidos apenas para processos sem memória, conforme destaca [22].

O caso do modelo  $M/D/1/1^*$ , uma fila do tipo LCFS com preempção em serviço, não segue o mesmo padrão das demais. Nesse caso, um aumento da carga no servidor mantendo-se a taxa de serviço determinística,  $\mu = D$ , ocasiona um aumento sem limites da AoI a partir de  $\rho > 1$ . Isso ocorre porque a taxa de pacotes descartados pelo servidor aumenta proporcionalmente à carga, uma vez que a taxa de serviço se mantém em um valor constante. Como esses pacotes não são entregues ao monitor, a AoI tende a crescer indefinidamente. Ressalta-se que o mesmo não ocorre no modelo  $D/M/1/1^*$  pois, mesmo com o aumento da taxa de chegadas  $\lambda$ , a taxa de entrega permanece com média  $\mu$ .

## O regime LGFS

Para resolver o problema de chegadas a partir de uma mesma fonte que possam ocorrer fora de ordem, alguns trabalhos como [30] e [32] propuseram o regime de serviço conhecido

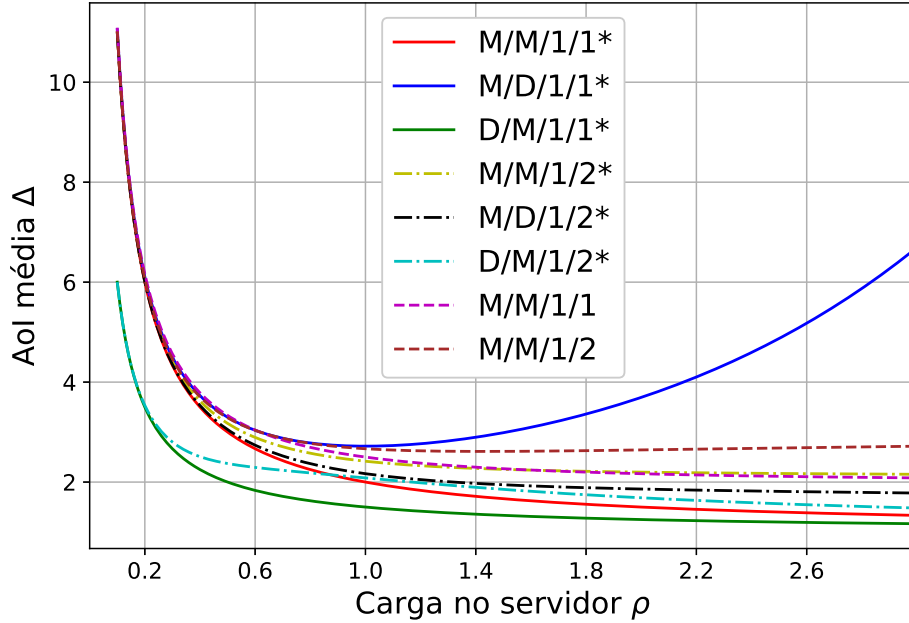


Figura 2.8: AoI média em função da carga no servidor para as diversas técnicas de gerenciamento de pacote. A taxa de serviço foi fixada em  $\mu = 1$ .

como *Last Generated First Served* (LGFS), no qual o último pacote gerado será o primeiro a ser servido. Ressalta-se que esse regime se difere do LCFS, pois a prioridade daquele são os pacotes *informativos*, independente de sua ordem de chegada. Um pacote de índice  $i > 0$  cujo tempo de geração é dado por  $s_i$  é chamado *informativo* quando, por ocasião de sua chegada, não existe um pacote já entregue ao monitor com índice  $j < i$  tal que  $s_j \geq s_i$ . Caso contrário, esse pacote será chamado de *obsoleto*.

As duas principais técnicas relacionadas com o regime de serviço LGFS são a LGFS-P, a qual realiza a preempção em serviço e a LGFS-NP ou LGFS-W, a qual realiza a preempção na fila de espera. Essas técnicas realizam o mesmo esquema de preempção mostrado nas Fig. 2.7 e 2.6. No Cap. 3 deste trabalho, será proposta uma técnica (LGFS-C) que utiliza ambos os métodos de preempção (S e W) de maneira condicional, podendo obter resultados melhores que as referidas técnicas apresentadas nesta seção.

## 2.2.4 Modelos com múltiplas fontes e múltiplos servidores

Uma vez apresentados os modelos da AoI para as filas elementares, esta seção introduz alguns conceitos relacionados à extensão desses modelos para abranger múltiplas fontes e múltiplos servidores. Conforme observa [22], este é um dos desafios atuais na área da *Age of Information* e carece de maiores estudos.

## Múltiplas fontes

Considere-se inicialmente um modelo com múltiplas fontes enviando *updates* para uma fila servida por um único servidor, o qual, por sua vez, encaminha os pacotes para um monitor. Esse esquema pode ser visualizado na Fig. 2.9.

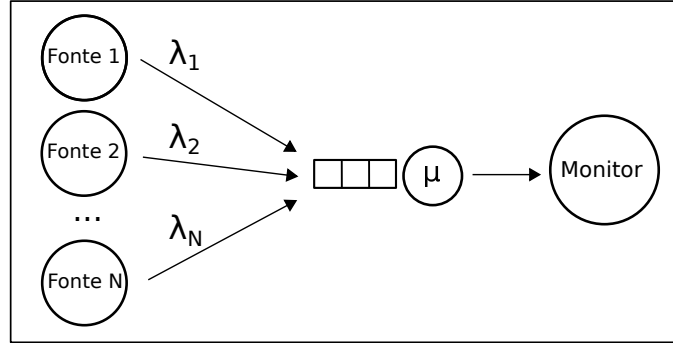


Figura 2.9: Modelo básico com múltiplas fontes enviando pacotes a uma taxa média  $\lambda_i$  para um servidor com taxa média de serviço  $\mu$ .

Na fila desse monitor, a carga total é obtida pela soma das cargas individuais de cada uma das fontes, conforme mostrado na Eq. 2.21.

$$\rho = \sum_{i=1}^N \rho_i \quad (2.21)$$

Nessa equação,  $\rho_i = \lambda_i/\mu$  é a contribuição da fonte  $i$ , cuja taxa média de chegadas é  $\lambda_i$ . De maneira complementar, a carga das demais fontes é dada por  $\rho_{-i} = \rho - \rho_i$ .

No estudo de [29], a AoI para cada uma das fontes é dada pela mesma expressão de 2.2, como mostra a Eq. 2.22.

$$\Delta_i = \frac{E[Y_i T_i] + E[Y_i^2]/2}{E[Y_i]} \quad (2.22)$$

O termo  $Y_i$  representa o processo do tempo entre chegadas e  $T_i$  representa o processo do tempo de serviço da fonte  $i$ . Para uma fila  $M/M/1$  com regime *First Come First Served* (FCFS), o mesmo estudo apresenta a AoI média para cada fonte de acordo com a Eq. 2.23.

$$\Delta_i^{M/M/1} = \frac{1}{\mu} \left[ \frac{\rho_i^2(1 - \rho\rho_{-i})}{(1 - \rho)(1 - \rho_{-i})^3} + \frac{1}{1 - \rho_{-i}} + \frac{1}{\rho_i} \right] \quad (2.23)$$

No trabalho de [33], uma outra expressão exata é apresentada para o modelo  $M/M/1$  com múltiplas fontes, sendo, no entanto, consideravelmente mais complexa que a obtida por [29]. Nesse mesmo estudo, o autor apresenta ainda expressões aproximadas para um

modelo de múltiplas fontes com um processo de serviço dado por uma distribuição geral, caracterizando uma fila do tipo  $M/G/1$ .

De maneira semelhante aos trabalhos realizados nos modelos de filas elementares com fonte única, foram apresentadas modelagens para o caso de múltiplas fontes utilizando as técnicas de gerenciamento de pacotes discutidas na Seção 2.2.3. O objetivo da utilização de tais técnicas é otimizar a AoI através da manipulação dos pacotes recebidos. Dessa maneira, [29] apresentou expressões que definem a AoI para múltiplas fontes para os modelos  $M/M/1/1^*$  e  $M/M/1/2^*$ , os quais utilizam um regime do tipo LCFS com preempção no serviço e na espera, respectivamente. Esses modelos da AoI são dados pelas Eq. 2.24a e 2.24b.

$$\Delta_i^{M/M/1/1^*} = \frac{1}{\mu} (1 + \rho) \frac{1}{\rho_i} \quad (2.24a)$$

$$\Delta_i^{M/M/1/2^*} = \frac{1}{\mu} \left[ \frac{(1 + \rho + \rho^2)^2 + 2\rho^3}{(1 + \rho + \rho^2)(1 + \rho)^2} + \left( 1 + \frac{\rho^2}{1 + \rho} \right) \frac{1}{\rho_i} \right] \quad (2.24b)$$

Ressalta-se que as Eq. 2.24a e 2.24b, quando reduzidas a um modelo de fonte única, tornam-se idênticas às equações apresentadas em 2.18a e 2.18b.

O estudo de [28] aborda um modelo de múltiplas fontes em que o tempo de transmissão das fontes até o servidor é modelado através de uma distribuição exponencial. Dessa maneira, assume-se que os pacotes de uma mesma fonte podem chegar fora de ordem na fila, sendo necessário então a utilização de técnicas de gerenciamento de pacote para que se priorize o serviço daqueles que trazem novas informações (*informativos*) em detrimento daqueles que não contribuem para essa finalidade (*obsoletos*).

## Múltiplos servidores

O modelo com múltiplos servidores pode ser observado na Fig. 2.10. Uma fonte produz pacotes a uma taxa  $\lambda$  é atendida por  $c$  servidores, os quais operam com uma taxa  $\mu_i$ , em que  $i = 1, 2, \dots, c$ . Os pacotes de todos os servidores convergem para um único monitor.

Entre os trabalhos que abordam tais modelos [34] realiza um estudo de sistemas do tipo  $M/M/2$  e  $M/M/\infty$ , para pacotes gerados a partir de uma única fonte. A motivação dos autores foi a de estudar o efeito que ocorre quando diversos *updates* provenientes de uma mesma fonte podem percorrer diversos caminhos até chegarem ao monitor. Essas chegadas podem ocorrer fora de ordem, sendo necessário avaliar se um pacote é *obsoleto* (não traz uma informação mais atualizada do que a já disponível) ou *informativo* (caso contrário). As expressões obtidas para esses modelos são bastante longas e complexas, não sendo, por esse motivo, apresentadas nesta seção.

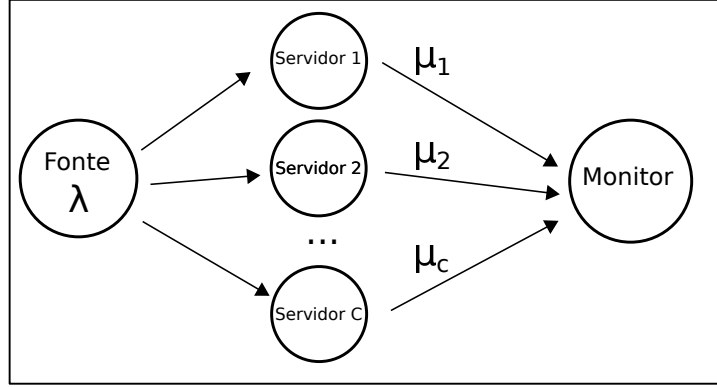


Figura 2.10: Modelo de uma fonte com taxa de geração  $\lambda$  sendo servida por  $c > 1$  servidores com taxa de serviço  $\mu_i$ ,  $i = 1, 2, \dots, c$ .

Um estudo semelhante foi realizado por [35], em que cada um dos diversos caminhos a partir de uma mesma fonte é modelado como uma fila do tipo LCFS com preempção no serviço, o que resulta em um sistema do tipo  $M/M/c^*$ , em que  $c > 1$ . Nesse modelo, um novo pacote que chega é alocado no servidor 1, enquanto os pacotes já alocados em um servidor  $i$  são realocados para o servidor  $i + 1$ . Quando todos os servidores estão ocupados, uma nova chegada ocasiona o descarte do pacote alocado no servidor de índice  $c$ . Além disso, todos os servidores possuem a mesma taxa de serviço  $\mu_i = \mu$ , ou seja, a estrutura é homogênea. Através dessas simplificações, o autor obteve uma expressão para o modelo, a qual é dada pela Eq. 2.25.

$$\Delta_{M/M/c^*} = \frac{1}{\mu} \left[ \frac{1}{c} \prod_{i=1}^{c-1} \frac{\rho}{i + \rho} + \frac{1}{\rho} + \frac{1}{\rho} \sum_{l=1}^{c-1} \prod_{i=1}^l \frac{\rho}{i + \rho} \right] \quad (2.25)$$

### Distribuições da classe NBU

O estudo de [30] realizou pela primeira vez a análise de modelos cujos servidores possuem distribuições de tempo de serviço pertencentes a uma classe especial de distribuições chamada *New Better than Used* (NBU). Nesse tipo de distribuição, uma variável aleatória  $T$ , a qual representa o tempo de serviço, possuirá a propriedade mostrada pela Eq. 2.26.

$$P(T > s + t | T > s) \leq P(T > t), \forall s, t \geq 0 \quad (2.26)$$

Isso significa que, para um pacote em serviço, com tempo decorrido igual a  $s$ , a sua probabilidade de terminar o serviço após um tempo adicional  $t$  é menor do que a probabilidade de um novo pacote (que inicia o serviço no tempo  $s$ ) ter um tempo de serviço maior que  $t$ . De forma equivalente, a mesma propriedade pode ser dada pela Eq. 2.27, onde  $\bar{F}(z) = P[Z > z]$  é a função complementar da distribuição cumulativa.

$$\bar{F}(s+t) \leq \bar{F}(s)\bar{F}(t), \forall s, t \geq 0 \quad (2.27)$$

Observa-se que a distribuição exponencial, utilizada na maioria dos modelos para o tempo de serviço é um caso especial de distribuição NBU, onde a Eq. 2.26 torna-se uma igualdade, denotando a sua propriedade de distribuição *sem memória*.

### Múltiplas fontes e múltiplos servidores

Observa-se que os modelos analíticos aumentam consideravelmente a sua complexidade quando o sistema em estudo possui diversas fontes e diversos servidores. Em muitos casos, essa modelagem abre mão de arcabouços matemáticos como Sistemas Híbridos Estocásticos [29, 36], os quais utilizam cadeias de Markov e matrizes de transição para representar mudanças de estado. Dessa maneira, estudos nesse sentido procuram o estabelecimento de teoremas que provem arquiteturas ou técnicas de otimização, ao invés de tentar obter fórmulas exatas para a AoI. Um desses estudos, realizado em [37], estabelece duas técnicas para obter resultados ótimos ou quase ótimos.

A primeira é chamada de *Maximum Age First, Last Generated First Served* (MAF-LGFS). Nessa técnica, o último pacote gerado a partir de uma fonte que possui a maior AoI entre todas as fontes será servido prioritariamente. Os autores demonstram que essa técnica, a qual não utiliza descarte de pacotes, otimiza um modelo de múltiplas fontes e um único servidor. Ressalta-se que o modelo é particularizado para o caso de fontes sincronizadas, ou seja, os pacotes de todas as fontes são gerados e recebidos sempre no mesmo tempo.

A segunda técnica chama-se *Maximum Age of Served Information First, Last Generated First Served* (MASIF-LGFS) e estabelece o conceito de *Age of Served Information* (AoSI), o qual é definido para uma fonte  $n$  conforme a Eq. 2.28.

$$\Xi(t) = t - \max\{S_{i,j} : V_{i,j} \leq t, j = 1, 2, \dots\} \quad (2.28)$$

Nessa equação,  $S_{i,j}$  representa o tempo de geração do pacote  $j$  da fonte  $i$  e  $V_{n,i}$  o tempo de início do serviço para o pacote  $j$  da fonte  $i$ . A técnica MASIF-LGFS parte do princípio que o último pacote gerado a partir de uma fonte que possui a maior AoSI entre todas as fontes será servido prioritariamente. Os autores mostram que, para o caso particular de fontes sincronizadas, essa técnica mostra-se quase ótima para o caso de múltiplos servidores com distribuições de tempo de serviço pertencentes à classe NBU. Assim como a técnica MAF-LGFS, esta técnica não utiliza descarte de pacotes.

Conforme será apresentado no Cap. 3, o presente estudo irá propor uma nova técnica para o caso do modelo com múltiplas fontes e múltiplos servidores, utilizando alguns

conceitos das técnicas apresentadas nesta seção. Ressalta-se que esse modelo é o de maior interesse neste trabalho, devido à sua semelhança com o modelo do sistema de monitoramento e decisão apresentado na Seç. 2.1.

## 2.2.5 Funções de penalidade da AoI

Conforme apresentado nas seções anteriores, as duas principais métricas de interesse relacionadas à AoI são a AoI média e a AoI de pico. Apesar das referidas métricas satisfazerem o caso de modelos com uma única fonte, as mesmas não são úteis para o caso de múltiplas fontes de pacotes, em que se necessita de uma função que represente o grau de desatualização do monitor em relação ao sistema como um todo. Nesse contexto, [37] estabelece o conceito de *função penalidade da AoI*.

Seja um sistema modelado por  $N$  fontes de atualizações, sendo o vetor  $\Delta(t) = [\Delta_1(t), \Delta_2(t), \dots, \Delta_N(t)]$  representativo da AoI em cada uma das fontes em um determinado tempo  $t$ . A função penalidade da AoI representa o grau de insatisfação com relação ao vetor  $\Delta$ , sendo calculado por  $p(\Delta) = p \circ \Delta$ . Alguns exemplos de funções de penalidade são mostrados nas Eq. 2.29 a 2.33.

1. AoI média para as  $N$  fontes:

$$p_{media}(\Delta) = \frac{1}{N} \sum_{n=1}^N \Delta_n \quad (2.29)$$

2. AoI máxima para as  $N$  fontes:

$$p_{max}(\Delta) = \max_{n=1, \dots, N} \Delta_n \quad (2.30)$$

3. AoI quadrática média para as  $N$  fontes:

$$p_{qm}(\Delta) = \frac{1}{N} \sum_{n=1}^N (\Delta_n)^2 \quad (2.31)$$

4. AoI de pico média para as  $N$  fontes:

$$p_{p-media}(\Delta) = \frac{1}{N} \sum_{n=1}^N \Delta_n^{(p)} \quad (2.32)$$

5. AoI de pico máxima para as  $N$  fontes:

$$p_{p-max}(\Delta) = \max_{n=1, \dots, N} \Delta_n^{(p)} \quad (2.33)$$

A escolha da função de penalidade da AoI a ser utilizada deve ser baseada nos requisitos negociais do sistema e no cenário em que o mesmo será inserido, assunto a ser explorado com maiores detalhes no Capítulo 3.

## 2.3 Revisão do estado da arte

Esta seção tem por objetivo apresentar uma revisão dos estudos e pesquisas mais recentes da AoI. Para cada um desses estudos, serão apresentadas também as possibilidades de contribuição em relação ao presente trabalho.

O trabalho de [22] apresenta uma introdução sobre AoI bem como um *survey*, elencando os principais estudos em diversas áreas de aplicação da *Age of Information*. Os autores apresentam a maioria dos modelos básicos de filas estudados para estimar a AoI, com os modelos analíticos correspondentes. Além disso, são definidas as principais técnicas de gerenciamento de pacotes com suas respectivas análises de desempenho.

Como contribuição para o presente trabalho, o *survey* serve como base para direcionar a pesquisa bibliográfica e a revisão da literatura, por incluir os trabalhos considerados o estado da arte na área. Entre os principais tópicos, destaca-se a referência de modelos de filas com múltiplas fontes e múltiplos servidores. Apesar das várias contribuições, o *survey* não aborda a simulação de modelos, sendo focada principalmente na obtenção de expressões analíticas.

O estudo de [26] demonstra a obtenção de diversas expressões exatas para a AoI média e AoI de pico para modelos de filas elementares, ou seja, uma única fonte e um único servidor. Partindo do modelo canônico representado pela Eq. 2.2, os autores derivam expressões para diversos modelos de filas, entre elas  $M/D/1/1^*$ ,  $D/M/1/1^*$ ,  $(M/D/1/2^*$  e  $D/M/1/2^*$ . Uma importante contribuição desse estudo é a obtenção de fórmulas para modelos com processos de chegada ou de serviço determinísticos. Esse trabalho também se destaca por obter expressões exatas considerando técnicas de gerenciamento de pacotes, como os casos de preempção em serviço e na fila de espera. Para esse último caso são ainda considerados os casos com e sem descarte dos pacotes que sofreram preempção. No contexto do presente estudo, o trabalho contribui para a validação dos modelos de simulação que serão gerados, em especial aqueles que utilizam processos de chegada determinístico. Esse tipo de modelo pode ser utilizado para representar a geração de dados sobre posições geográficas de agentes, em que as coordenadas são enviadas para o monitor em intervalos de tempo fixos (determinísticos).

Conforme será apresentado no Cap. 3, o modelo proposto para a rede de C2 a ser estudada parte do princípio de que múltiplas fontes enviam *updates* para o monitor. Dentro desse contexto, o trabalho de [29] utiliza uma abordagem analítica para modelar



um sistema composto por múltiplas fontes que enviam pacotes para uma fila básica, a qual é então analisada em termos de diferentes regimes de serviço e distribuições de chegada. Entre as técnicas utilizadas, destacam-se os regimes FCFS e LCFS, além do uso de gerenciamento de pacotes, como por exemplo a preempção em serviço. O estudo apresenta expressões exatas para a AoI média para alguns casos, como os modelos  $M/M/1$ ,  $M/M/1/1^*$  e  $M/M/1/2^*$ . Apesar desse estudo ter apresentado comparações apenas para duas fontes, é possível estender esses resultados para  $N$  fontes, bem como validá-los através de simulações. Uma limitação da aplicação do referido estudo é o fato de que o sistema assume que os pacotes gerados nas  $N$  fontes são enviados de maneira instantânea para a fila do receptor, o que não ocorre em um sistema real.

Uma abordagem semelhante é aplicada em [33], para obter uma expressão exata para a AoI para um sistema com múltiplas fontes e um servidor com regime de serviço do tipo  $M/M/1$ . Além desse resultado, os autores obtêm ainda três expressões aproximadas para o mesmo sistema ao utilizar um regime do tipo  $M/G/1$ . Por fim, o trabalho realiza algumas simulações para validar as expressões obtidas e demonstrar o comportamento do sistema sob diferentes técnicas de serviço.

No estudo de [38] é apresentada uma proposta de um sistema em que diversas fontes concorrem ao acesso de um receptor. Cada uma das fontes individuais é modelada como uma fila de capacidade infinita, a qual possui ainda a possibilidade de descartar pacotes com o intuito de otimizar a AoI no monitor. Os autores realizam uma comparação do modelo com e sem o gerenciamento de pacotes, além da comparação entre três técnicas de acesso ao meio: *round-robin*, *work-conserving* e acesso randômico.

Em um segundo momento, o estudo inclui a possibilidade de perdas nos canais de comunicação, bem como a possibilidade dos pacotes chegarem em ordens diferentes das enviadas pelas fontes. Para isso, um parâmetro de atraso na comunicação entre o receptor e o monitor é inserido (*network delay*), de forma que o modelo é representado por um processo com distribuição exponencial. A possibilidade de contribuição com o presente trabalho se dá no modelo adotado, no qual se utiliza uma fila para representar cada uma das fontes. Acredita-se que essa seja uma boa maneira de se representar os agentes em campo no caso de um sistema de C2, pois o envio dos pacotes não ocorre de maneira imediata ao depender dos tempos de acesso ao meio.

Outro estudo que aborda o caso de múltiplas fontes que convergem para um conjunto servidor-monitor foi abordado em [39], o qual utiliza modelos de tempo discreto para filas para obter resultados exatos da distribuição da AoI para cada uma das fontes de pacotes. Os algoritmos utilizados para se obter os resultados numéricos utilizam Cadeias de Markov de Tempo Discreto (CMTD) do tipo *Quasi-Birth-Death*. Os resultados obtidos são fornecidos para diversas técnicas de gerenciamento de pacotes, seja com ou sem pre-

empção e ainda com ou sem fila de espera, entre outros. O diferencial desse estudo é a utilização de métodos numéricos para a obtenção das distribuições da AoI.

Sistemas com múltiplas fontes podem possuir como requisito o estabelecimento de diferentes prioridades entre as mesmas para possibilitar o favorecimento daquelas cujos dados são mais valiosos que outras. Nesse sentido, [40] modela o sistema de múltiplas fontes e atribui prioridades a cada uma delas. Quanto às técnicas utilizadas, os autores abordam o caso de uma fila sem posição de espera, na qual pacotes com igual ou maior prioridade terão preferência, bem como o caso em que existe um fila de espera na qual o mesmo esquema de preempção é realizado, no entanto sem interferir no pacote atualmente em serviço. Esse estudo contribui com o presente trabalho ao considerar que sistemas de C2 podem priorizar o envio de dados a partir de determinados sensores ou agentes, os quais devido à sua posição ou função, são considerados de maior importância para o processo de tomada de decisão.

No estudo de [32] é definido o regime de serviço conhecido como *Last Generated First Served - Preemptive* (LGFS-P), o qual utiliza a preempção em serviço para otimizar a AoI em sistemas com múltiplos servidores e distribuição de tempo de serviço exponencial. Esse regime prioriza os pacotes que possuem maior tempo de geração, sendo especialmente eficiente para sistemas nos quais a ordem de chegada dos pacotes no monitor não é a mesma ordem em que os mesmos foram gerados.

Em sistemas reais é esperado a existência de múltiplos saltos em uma rede [41], especificamente no caminho entre uma fonte de dados e o monitor. Dentro dessa expectativa, o estudo de [31] apresenta resultados para um modelo de rede com múltiplos saltos, ao mostrar que a técnica LGFS-P não otimiza modelos com tempos de serviço cuja distribuição não é exponencial. Em um outro trabalho do mesmo autor [30], é demonstrado que, para uma distribuição de tempo de serviço pertencente à classe NBU, a técnica conhecida como LGFS-NP ou LGFS-W (não preemptiva ou com preempção na fila de espera, segundo outros autores) atinge valores quase ótimos para a AoI média, com uma pequena diferença constante de um limite inferior estabelecido. Essas duas técnicas (LGFS-P e LGFS-W) são utilizadas como base para uma nova técnica que será proposta neste trabalho, a *Last Generated - First Served - Conditional* (LGFS-C), apresentada no Capítulo 3.

O trabalho realizado por [37] aborda o caso de um modelo com múltiplas fontes e múltiplos servidores, ao definir duas técnicas de gerenciamento de pacotes sem descarte, *Maximum Age First, Last Generated First Served* (MAF-LGFS) e *Maximum Age of Served Information First, Last Generated First Served* (MASIF-LGFS). Essas técnicas, detalhadas na Seção 2.2.4, priorizam o pacote da fila que possui o maior tempo de geração (LGFS), a partir de uma fonte que possui a maior AoI ou a maior *Age of Served Information* (AoSI), respectivamente. Os experimentos, porém, são particularizados para o

caso de fontes sincronizadas, ou seja, os pacotes de todas as fontes são gerados e recebidos sempre no mesmo tempo. No caso do MAF-LGFS, a análise é feita para tempos de serviço com distribuição exponencial, em que são obtidos resultados ótimos para o caso de um único servidor. Para o caso MASIF-LGFS, a análise é ampliada para múltiplos servidores com tempos de serviço da classe NBU, mostrando-se quase ótimos. As técnicas anteriormente citadas serão a base para a proposta de uma nova técnica para modelos com múltiplas fontes e múltiplos servidores, a *Max Age First - Last Generated - First Served - Multiple preemption in Waiting* (MAF-LGFS-MW), a qual será apresentada no Capítulo 3.

Outro estudo que aborda um modelo com múltiplas fontes e múltiplos monitores foi realizado por [36], no qual o caso de múltiplas fontes foi restrito a taxas de chegadas e de serviço homogêneas, ou seja, as taxas são iguais entre as fontes e entre os servidores. O referido trabalho obtém expressões analíticas para alguns modelos como o LCFS com uma fonte e múltiplos servidores e LCFS para múltiplas fontes e múltiplos servidores (homogêneo). Alguns resultados simulados são apresentados, porém o trabalho não apresenta maiores detalhes sobre o modelo computacional utilizado.

## 2.4 Resumo do capítulo

O presente capítulo teve por finalidade apresentar os principais conceitos teóricos e a revisão da literatura referente ao estado da arte. Os dois principais conceitos abordados foram os sistemas ciberfísicos e a métrica conhecida como *Age of Information*. Os SCF formam a base para o modelo do sistema de monitoramento e decisão para aplicação na atividade de comando e controle em um cenário operacional militar, o qual será definido no próximo capítulo. Já a AoI consiste na principal métrica de avaliação que se pretende utilizar para avaliar tal modelo. Ainda sobre essa métrica, ressalta-se a importância das técnicas de gerenciamento de pacotes e dos modelos de sistemas com múltiplas fontes e múltiplos servidores no monitor.

# Capítulo 3

## Definição do Modelo Computacional e Proposta das Técnicas LGFS-C e MAF-LGFS-MW

O presente capítulo tem por finalidade definir o modelo computacional e descrever a ferramenta de simulação utilizada na avaliação experimental. Inicialmente, é definido o modelo de um SCF particularizado para um sistema de monitoramento e decisão no contexto da atividade de comando e controle militar. Para esse sistema são descritos os componentes típicos de carga de trabalho, parâmetros e principais métricas de avaliação. Com base nessa modelagem, é descrito o modelo computacional que resulta em uma ferramenta, a qual tem como objetivo simular diversos cenários de uso de sistema para o cálculo de métricas de AoI. É importante observar que, até o momento da escrita deste trabalho, não existiam ferramentas que permitissem esse tipo de simulações.

Este capítulo também apresenta duas novas técnicas de gerenciamento de pacotes para a AoI: a *Last Generated - First Served - Conditional* (LGFS-C) e a *Max Age First - Last Generated - First Served - Multiple preemption in Waiting* (MAF-LGFS-MW), as quais são contribuições deste trabalho para otimizar a AoI e garantir uma maior eficiência no dimensionamento de um SCF.

### 3.1 Definição do Modelo

Esta seção apresenta a descrição de um Sistema de Comando e Controle (C2) como uma particularização de um SCF, com definição do modelo, seus componentes, parâmetros e métricas de avaliação.

### 3.1.1 Comando e Controle

No meio militar, a atividade de Comando e Controle (C2) consiste em um conjunto de métodos e procedimentos que visam assegurar aos comandantes as informações necessárias para o processo decisório [1]. Dessa maneira, a atividade de C2 envolve a coleta de dados para se obter a chamada *consciência situacional*, a qual se caracteriza pela percepção precisa dos fatores e condições que afetam a execução das atividades, o que permite ao decisor ter condições de focar seus esforços no objetivo.

Os conflitos modernos são cada vez mais caracterizados pela presença do componente informacional, o que torna as decisões dependentes dos Sistemas de Tecnologia da Informação e Comunicações (TIC), os quais devem garantir os ciclos de C2 com rapidez, precisão e oportunidade. Um ciclo de comando e controle pode ser representado pelo modelo conhecido como ciclo OODA: Observar, Orientar-se, Decidir e Agir [1]. A Fig. 3.1 ilustra esse processo.

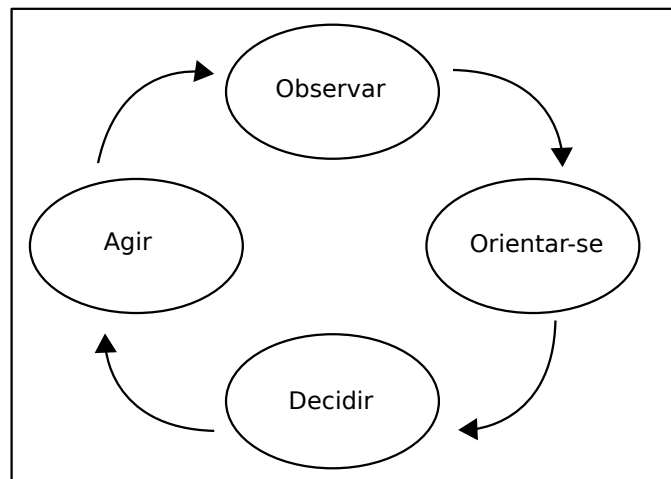


Figura 3.1: Ciclo OODA.

Para garantir a execução eficiente dos ciclos de C2, são utilizados Sistemas de Comando e Controle. Como exemplo de um sistema de C2 no Exército Brasileiro, pode-se citar o Sistema Pacificador, o qual é utilizado principalmente para operações do tipo Garantia da Lei e da Ordem (GLO), como segurança de grandes eventos, garantia das eleições e reforço ao policiamento em unidades da federação. Esse sistema foi utilizado com sucesso na Copa do Mundo de 2014 e nas Olimpíadas de 2016, considerado hoje como consolidado no âmbito do Exército.

O Pacificador opera através de um cliente móvel, o qual pode ser instalado em *smartphones* que enviam informações de campo a respeito de um determinado agente ou sensor, como por exemplo a posição geográfica ou o relato de algum incidente que necessita ser processado. Essas informações trafegam através de uma rede de comunica-

ções e são recebidas por um Centro de Operações (COp), no qual seu comandante toma decisões baseadas no conjunto das informações disponíveis.

Com base na atual arquitetura do Pacificador e na arquitetura de um sistema de monitoramento e decisão, conforme mostrado na Fig. 2.1, definiu-se um modelo genérico para um sistema de C2 militar. Esse sistema é apresentado na Fig. 3.2 e é formado pelos componentes descritos a seguir.

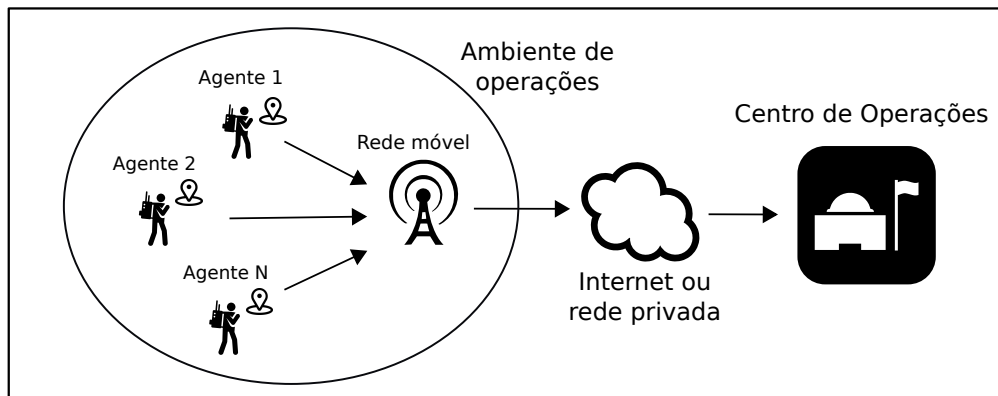


Figura 3.2: Modelo genérico para um sistema de comando e controle baseado na arquitetura de monitoramento e decisão.

### Ambiente de operações

Região em que ocorre uma determinada operação militar, seja de guerra ou de GLO. Nesse ambiente situam-se os agentes e sensores, os quais interagem com o ambiente ao coletar, atuar e enviar informações.

### Agentes

Elementos em campo, a pé ou embarcados, os quais atuam no ambiente de operações ao enviar dados de maneira periódica para um Centro de Operações. Como exemplo, um desses dados pode ser a posição georreferenciada do agente.

### Infraestrutura de comunicações

Infraestrutura composta por *hardware* e *software*, a qual abrange desde aplicações até protocolos de comunicação e estruturas físicas de transmissão. Em operações militares com agentes a pé ou em viaturas, é composta geralmente por uma rede rádio ou rede móvel celular, a qual encaminha as informações dos agentes através da Internet ou de uma estrutura de rede privada.

## Centro de Operações

Local onde são recebidas e processadas as informações enviadas a partir dos agentes. Essas informações são utilizadas como subsídios para a tomada de decisão dos comandantes de uma determinada operação. As decisões, por sua vez, são transmitidas aos agentes, ação esta que fecha o ciclo de comando e controle.

Ressalta-se que o modelo de rede apresentado na Fig. 3.2 não é único, podendo ser construído de acordo com o tipo da operação, o tipo do sistema, o nível da rede na hierarquia de decisão, condicionantes geográficas, entre outros. No contexto do presente estudo, o modelo apresentado é usado como base por motivos de padronização e simplificação.

### 3.1.2 Métricas para Avaliação de Sistemas de C2

De acordo com [1], um dos princípios do comando e controle é a rapidez do processo decisório, em que os enlaces da rede devem permitir o acesso imediato às informações pelo escalão de comando. Obviamente, este acesso imediato se traduz em *o mais rápido possível*, uma vez que existe a dependência da infraestrutura física subjacente ao sistema de C2. A partir desse princípio, revela-se a necessidade de se obter métricas capazes de avaliar a eficiência das redes de comando e controle, em especial aquelas relacionadas com a rapidez e a oportunidade da informação que nela trafega.

As métricas de rede clássicas, como a vazão e atraso fim-a-fim nem sempre traduzem os requisitos negociais, os quais podem ser pensados em termos da atualidade e oportunidade das informações. Um exemplo clássico desse desalinhamento é uma rede em que os agentes ou sensores enviam pacotes de atualização (*updates*) ao utilizar suas capacidade máximas de largura de banda. Como consequência, o servidor no COP responsável por receber esses pacotes poderá entrar em colapso, ao atrasar ou descartar informações, armazenando-as em uma fila para processamento e entregando dados desatualizados para o decisor. Não obstante, as métricas da rede tais como vazão e atraso poderão apresentar resultados satisfatórios.

A partir deste último exemplo, percebe-se a necessidade de se obter métricas mais alinhadas ao domínio negocial do comando e controle e não apenas representativas de parâmetros da rede física. Particularmente em relação aos processos que enviam *updates* de forma periódica, como a posição geográfica de um agente ou a velocidade de uma viatura, por exemplo, deve-se dar especial atenção para questões relacionadas à atualidade da informação. Dessa forma, tais métricas devem ser capazes de responder aos seguintes questionamentos:

- O quão atualizada está a informação recebida sobre determinado agente ou sensor?

- Existem informações geradas na fonte mais atuais que as já recebidas? Qual a diferença de tempo entre elas?
- Qual a quantidade de informações obsoletas que um determinado monitor recebe? Como otimizar esse processo?

Além de fornecer as respostas para os questionamentos acima, a definição de métricas sobre a atualidade da informação serve de subsídio para o estabelecimento dos requisitos dessa rede, os quais por sua vez responderão às seguintes perguntas:

- Qual a taxa ótima com a qual devo transmitir os pacotes com a posição de agentes ou dados de sensores?
- Qual a influência do processo de atraso de transmissão e processamento na atualidade da informação que recebo?
- Quantos servidores necessito para receber as informações para cumprir com uma atualidade mínima definida?
- Quais técnicas de gerenciamento de pacotes são mais eficientes para o tipo de cenário e arquitetura do sistema?

Para responder a tais questionamentos este trabalho pretende utilizar os conceitos de *Age of Information* definidos na Seção 2.2. Acredita-se que a AoI pode ser aplicada como métrica em sistemas de C2 de maneira a contribuir com os processos de definição de arquitetura, definição de requisitos e metodologias de avaliação em tais sistemas. Ressalta-se que, até a data de finalização deste trabalho, não foram encontrados estudos da aplicação da AoI em sistemas C2.

### 3.1.3 Descrição dos Componentes do Sistema

Nesta seção são descritos os componentes do sistema de monitoramento e decisão que será utilizado como base para as simulações e avaliações. Para atingir esse objetivo são definidas a carga de trabalho, os parâmetros do sistema e as métricas de avaliação a serem utilizadas nas simulações.

#### Definição do sistema

Para modelar um sistema de C2 militar, considera-se a arquitetura apresentada na Fig. 3.3, na qual os agentes em campo são representados pelas fontes  $F_i$ ,  $i = 1, 2 \dots N$ , enviando periodicamente suas respectivas posições geográficas a um monitor, o qual corresponde ao COp. Cada uma das fontes é modelada por uma fila com taxa de chegada determinística



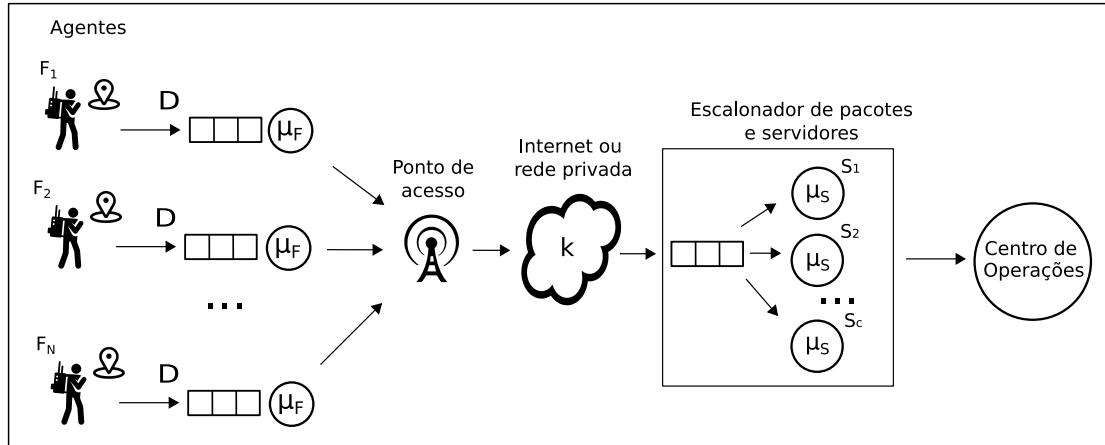


Figura 3.3: Modelo proposto para o sistema de monitoramento e decisão no contexto da atividade de C2.

$D_i$  e taxa de serviço com média  $\mu_F$ , em que essa variável tempo de serviço pode assumir qualquer tipo de distribuição. Os agentes se conectam a um ponto de acesso através de uma rede móvel sem fio, sem perdas e sem interferência, na qual os pacotes são recebidos na mesma ordem em que foram gerados. A concorrência das diferentes fontes e o tempo de serviço do ponto de acesso sem fio estão incluídos no tempo de processamento do servidor das filas de cada fonte. Para modelar essa concorrência, a taxa de serviço dos servidores das fontes é definida em função da quantidade total de fontes  $N$ , de maneira que  $\mu_F = f(N)$ . Quanto maior o número de fontes, maior será o tempo de acesso ao ponto de acesso sem fio.

A partir da rede móvel, as informações seguem imediatamente através de uma rede estruturada, a qual pode ser a Internet ou uma rede privada militar. Nessa rede, devido à possibilidade de múltiplas rotas até o servidor, o atraso de transmissão, propagação e processamento da rede é modelado através de uma distribuição exponencial com média  $1/k$ . No caso da rede, assume-se que o atraso de cada pacote independe da quantidade de fontes e da frequência de envio dos pacotes. Esse modelo de transmissão possibilita ainda que os pacotes cheguem em uma ordem diferente da qual foram enviados, o que pode causar a existência de pacotes obsoletos, os quais serão melhor descritos na seção referente à carga de trabalho.

Ao serem recebidos no sítio de destino, os pacotes oriundos das múltiplas fontes são então multiplexados em uma fila a qual atua como um escalonador e balanceador de carga. Os pacotes são então distribuídos para processamento nos servidores de aplicação  $S_i$ , sendo  $i = 1, 2, \dots, C$ . Esses servidores possuem a mesma taxa média de serviço  $\mu_S$ , em que a variável tempo de serviço pode assumir qualquer tipo de distribuição. Após o servidor de aplicação, os pacotes são considerados entregues ao monitor. Vale ressaltar que pacotes oriundos de uma mesma fonte podem ser processados simultaneamente por

dois servidores. Nesse caso, a decisão de descartar ou substituir um desses pacotes é tomada com base na técnica de gerenciamento de pacotes utilizada.

O objetivo da modelagem é avaliar o sistema em termos de seus parâmetros e da carga de trabalho, conforme requisitos mínimos de desempenho, os quais são definidos em função das métricas relacionadas à AoI. Para esse fim, torna-se necessário definir alguns conceitos que serão utilizados adiante na metodologia.

## Carga de trabalho

A carga de trabalho é definida pelos pacotes ou *updates* contendo a informação da geolocalização de um determinado agente em campo, extraída a partir de um sensor que utiliza a tecnologia *Global Positioning System* (GPS). O termo pacote, aqui definido, não deve ser confundido com o pacote que representa a unidade básica de protocolos, como o *Internet Protocol* (IP), por exemplo. Os pacotes de atualização que formam a carga de trabalho são aqueles gerados na camada de aplicação, na qual um pacote é responsável por atualizar a posição de um único agente.

Para cada um dos agentes, considera-se que os pacotes são gerados de maneira determinística, sendo a posição dos agentes coletada periodicamente com uma taxa  $D$ . Essa frequência de geração irá impactar na carga das diversas filas do sistema e, consequentemente, nos valores da AoI para cada uma das fontes. O outro parâmetro da carga de trabalho é a quantidade de agentes em campo ( $N$ ), sendo esse o principal parâmetro de interesse, pois representa a capacidade operacional do sistema. Quanto maior o seu valor, maior também será a taxa de chegadas de pacotes nas estruturas de filas do servidor no COp. Ressalta-se que o objetivo nem sempre é maximizar o envio de pacotes, o que pode causar uma sobrecarga nas filas e aumentar a AoI no COp. Dessa maneira, deve-se encontrar uma carga de trabalho que otimize a AoI de acordo com os parâmetros estabelecidos.

Os tempos de amostragem, codificação, preparação e envio dos dados com a posição GPS, bem como o tempo de acesso à rede sem fio, são modelados pela estrutura de fila do agente, conforme mostrado na Fig. 3.3. Essa fila é representada por um servidor cujo tempo de processamento pode assumir qualquer distribuição, sendo sua taxa média função do número total de agentes  $N$ , para simular as restrições do ponto de acesso sem fio. Para se obter o grau de abstração necessário para a avaliação do sistema, não foram considerados parâmetros característicos de alguns protocolos, como tamanho dos pacotes, segmentação, controle de fluxo e retransmissão.

Quando os pacotes são recebidos no COp ou em qualquer outro nó do sistema, os mesmos podem ser classificados como informativos ou obsoletos. Um pacote  $i$ , com tempo de geração  $s_i$ , é considerado obsoleto quando existe no monitor um pacote  $j$ , pertencente

à mesma fonte e com tempo de geração  $s_j > s_i$ . Caso contrário, o pacote  $i$  é considerado informativo. A decisão de descarte ou não de pacotes obsoletos é uma das técnicas de gerenciamento de pacotes descritas na Seção 2.2 e pode ser utilizada quando não existe a necessidade de se armazenar todo o histórico de pacotes recebidos.

### Parâmetros do sistema

Com o objetivo de se otimizar a AoI, além do ajuste da carga de trabalho, diversas variáveis do sistema também podem ser manipuladas, as quais serão apresentadas a seguir:

- Regime de serviço nas filas: pode assumir qualquer uma das técnicas apresentadas na Seção 2.2.3, como por exemplo LCFS ou LGFS, em que pode ainda utilizar esquemas de precedência e descarte de pacotes. Um ponto de interesse neste trabalho é encontrar um conjunto de regimes de serviço que, associado aos parâmetros da carga de trabalho, otimizem a AoI dentro de um cenário definido.
- Processo do tempo de serviço nas filas: o tempo de serviço nas diversas filas é função das atividades de preparação do pacote, acesso ao meio, transmissão para o próximo nó da rede, entre outras. Tanto nos servidores dos agentes quanto nos servidores do COP, esse tempo pode ser modelado por qualquer distribuição. Todos os servidores do COP, possuem a mesma taxa média  $\mu_s$  para o tempo de serviço, o que caracteriza um sistema homogêneo.
- Quantidade de servidores no COP e escalonamento dos pacotes: ao chegarem no COP, os pacotes serão alocados através de um escalonador para um dos  $C$  servidores, em que  $C$  é um dos parâmetros a serem ajustados. Obviamente, essa quantidade de servidores deve ser adequada a um cenário real, no qual existem limitações de custo e instalação da infraestrutura. Outro parâmetro de interesse do sistema é a técnica de escalonamento dos pacotes, a qual pode incluir também o descarte daqueles obsoletos, uma vez que o processo de transmissão pode ocasionar a chegada de pacotes fora da ordem de envio. Na Seção 3.4, será proposta uma nova técnica para esse parâmetro de interesse.
- Atraso dos pacotes na rede: o trajeto dos pacotes entre o ponto de acesso da rede móvel e local do COP (monitor) atravessa diversos ativos de rede. Para caracterizar esse atraso utiliza-se uma distribuição exponencial com média  $k$ , em que esse parâmetro independe da fonte que envia o pacote.

## Métricas

Conforme apresentado na Seção 3.1.2, o conjunto de métricas de interesse é baseado no conceito da *Age of Information* (AoI) e refere-se ao grau de atualização a respeito das posições dos agentes em campo. A aplicação das métricas será realizada através de funções de penalidade, conforme definido na Seção 2.2. A escolha da função de penalidade a ser utilizada deve ser baseada naquilo que se considere crítico para o sistema e não apenas na noção de AoI média.

Como exemplo, caso o elemento decisor no COp possua como requisito uma AoI média máxima para o conjunto de agentes, a função penalidade deve ser a *Age of Information* média para os  $N$  agentes. Caso o requisito seja em termos dos valores máximos da AoI obtidos pelos agentes, a métrica adequada seria a *Age of Information* de pico (pAoI), podendo-se utilizar a média em relação aos  $N$  agentes.

É importante observar que o objetivo é atender aos requisitos definidos para o sistema com base nos valores fornecidos pelas funções de penalidade. Nesse sentido, o Capítulo 4, define três cenários com o intuito de verificar as condições em que um determinado modelo consegue atender a tais requisitos. Atualmente não existe uma especificação padrão a respeito de valores de AoI máximos para um determinado cenário de uso. Por esse motivo, e com base na experiência em Sistemas de Comando e Controle, serão arbitrados valores hipotéticos para esses requisitos.

## 3.2 Modelagem Computacional

Conforme apresentado no Capítulo 2, as expressões analíticas para o cálculo da AoI aumentam consideravelmente sua complexidade quando se modelam sistemas com múltiplas fontes e múltiplos servidores. Por esse motivo, a simulação se apresenta como uma técnica de avaliação interessante para AoI em SCF. Sendo assim, este trabalho implementa uma ferramenta que representa o sistema de monitoramento e decisão como uma rede de filas e grafos direcionados. Diversas técnicas de gerenciamento de pacotes foram implementadas nesta ferramenta, a qual possui grande alcance e flexibilidade, tendo sido codificada com a linguagem Python. Na sequência, são descritos os elementos principais que formam o modelo computacional da ferramenta, com maiores detalhes disponíveis em [42].

### 3.2.1 Descrição das Estruturas

O modelo computacional utiliza um grafo direcionado  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , onde  $\mathcal{V}$  representa o conjunto dos nós e  $\mathcal{E}$  o conjunto de arestas, sendo  $\mathcal{E} \subseteq \{(x, y) | (x, y) \in \mathcal{V}^2, x \neq y\}$ . Cada

um dos nós do grafo é instanciado como uma fila cujos parâmetros a seguir devem ser definidos:

- Função que define o tempo de serviço  $M_i(t), i \in \mathcal{V}$ .
- Tamanho máximo de posições de espera da fila.
- Regime de serviço da fila (Ex: FCFS).
- Técnica de gerenciamento de pacote, caso utilizada.

A carga de trabalho é representada através de um conjunto de pacotes  $\mathcal{W} \subseteq \{(i, j) | i \in \mathcal{V}^*, j = 1, 2, \dots\}$ , onde  $\mathcal{V}^* \subset \mathcal{V}$  representa o subconjunto de nós que são fontes de pacotes,  $i$  representa o índice do nó fonte e  $j$  o índice do pacote. Para cada um desses nós fontes deverá ser definida uma função de geração  $G_i(t), i \in \mathcal{V}^*$ , a qual representa o processo de chegada dos pacotes no sistema. Um determinado grafo deve possuir no mínimo um nó gerador, ou seja,  $|\mathcal{V}^*| \geq 1$ .

Tanto as funções de geração quanto as funções de tempo de serviço dos pacotes podem assumir qualquer distribuição estatística desejada [43, 44] ou ainda serem definidas de maneira determinística. O regime de serviço e a técnica de gerenciamento de pacotes é definida pelo tipo da classe utilizada para se instanciar o objeto que representa um determinado nó. No decorrer da simulação, cada um dos nós (filas) armazenará as informações referentes aos pacotes que por ali trafegaram, utilizando o formato  $(i, j, s_i) : [t_1, t_2, t_3]$ , sendo:

- $(i, j, s_i^j)$  a identificação do agente (fonte e índice) e o seu tempo de geração  $s_i^j$ .
- $t_1$  o tempo de chegada no nó.
- $t_2$  o tempo de início do serviço.
- $t_3$  o tempo de partida para o próximo nó.

Para os nós geradores, o tempo de chegada  $t_1$  será igual ao tempo de geração do pacote, representado por  $s_i^j$ , em que  $i$  é a identificação da fonte e  $j$  o índice do pacote. Para os nós não geradores, o tempo de chegada do pacote é igual ao tempo de partida do mesmo pacote referente ao nó anterior. Dessa forma, para uma determinada aresta  $(x, y) \in \mathcal{E}$ ,  $t_1^y = t_3^x$ , em que  $x$  e  $y$  são respectivamente os nós origem e destino do pacote em transição. O grafo direcionado deve convergir para um único nó de destino, o qual representa o monitor do sistema. Nesse nó, serão registrados os tempos de entrega dos diversos pacotes, definidos por  $d_i^j$ , em notação semelhante aos tempos de geração. A chegada de um pacote ao nó de destino representa a entrega desse pacote ao monitor, encerrando assim o seu processamento.

A simulação inicia com todas as filas vazias e com o tempo de simulação  $t = 0$ . Um número máximo de pacotes  $N = |\mathcal{W}|$  a serem gerados é definido de antemão, sendo a simulação encerrada quando o último pacote chega ao nó de destino. Uma vez finalizada a simulação, os dados coletados são preparados e processados para o cálculo da AoI média ou pAoI média, as quais são obtidas para cada uma das fontes.

O método para o cálculo da AoI média consiste na utilização da técnica das áreas, apresentada na Seção 2.2.1 e dada pelas Eq. 2.1 e 2.2. Para um grafo com múltiplas fontes, o processo da AoI, conforme apresentado na Fig. 2.3, ocorre em paralelo para cada uma das  $i \in \mathcal{V}^*$  fontes do sistema, sendo obtida uma AoI média para cada uma delas.

Para uma determinada fonte  $i$ , cada pacote  $j$  gerado durante a simulação contribuirá com uma parcela  $Q_i^j$  da área total, sendo seu valor dado pela Eq. 3.1.

$$Q_i^j = Y_i^j T_i^j + Y_i^{j^2} / 2 \quad (3.1)$$

Nessa equação,  $Y_i^j = s_i^j - s_i^{j-1}$  é o tempo entre duas gerações consecutivas de pacotes para a fonte  $i$  e  $T_i^j = d_i^j - s_i^j$  é o tempo de sistema para o pacote  $j$  pertencente à fonte  $i$ . A AoI média para a fonte  $i$  será dada pela Eq. 3.2, onde  $J^i$  representa o índice do último pacote gerado pela fonte  $i$ .

$$\langle \Delta \rangle_\tau = \frac{1}{\tau} \int_0^\tau \Delta(t) dt = \frac{\sum_{j=1}^{J^i} Q_i^j}{d_i^{J^i} - s_i^1} \quad (3.2)$$

A AoI de pico média, por sua vez, pode ser obtida através da Eq. 3.3.

$$\langle \Delta^{(p)} \rangle = \frac{\sum_{j=1}^{J^i} A_i^j}{J^i} \quad (3.3)$$

Os valores de pico  $A_i^j$ , por sua vez, são obtidos pela Eq. 3.4.

$$A_i^j = Y_i^j + T_i^j \quad (3.4)$$

Ressalta-se que, no presente modelo, os pacotes perdidos ou descartados perdem seus índices para o próximo pacote. Dessa maneira, para um pacote descartado  $(i, j)$ , o próximo pacote  $(i, j + 1)$  assumirá imediatamente o índice  $(i, j)$  e assim por diante.

Uma vez que o presente trabalho propõe uma técnica de simulação para o cálculo da *Age of Information* em SCF, é necessário definir a precisão do modelo computacional. Essa etapa do estudo é fundamental pois, conforme apresentado anteriormente, os modelos analíticos limitam-se a estruturas simples, sem permitir de forma prática a validação de sistemas complexos. Dessa maneira, a análise de precisão do modelo computação propõe

um estimador para a Raiz do Erro Quadrático Médio (REQM) dos valores obtidos para a AoI média e para a AoI de pico média.

### Análise de precisão para a AoI média

O modelo computacional apresentado anteriormente pode ser analisado como um método Monte Carlo [45]. Seja o vetor  $\mathbf{Q} = [Q_1, Q_2, \dots, Q_k]$ , onde  $K = |\mathcal{V}^*|$  é o número de fontes de pacotes. Cada  $Q_i(\mathcal{G}, \mathcal{M})$  é uma variável aleatória que representa as áreas do processo gráfico da AoI relacionado à fonte  $i$ , conforme apresentado na Fig. 2.3. Essas variáveis são funções do conjunto de variáveis geradoras de pacotes  $\mathcal{G} = \{G_1, G_2, \dots, G_K\}$  e do conjunto de variáveis de tempo de serviço  $\mathcal{M} = \{M_1, M_2, \dots, M_{|\mathcal{V}|}\}$ . As variáveis aleatórias  $G$  e  $M$  possuem funções densidade de probabilidade (FDP) conhecidas, as quais foram definidas como parâmetros do modelo computacional. O objetivo da simulação é calcular o valor esperado  $\gamma_i = \mathbb{E}(Q_i)$  para cada uma das fontes, o qual será então utilizado para o cálculo da AoI média. Com essa finalidade, são geradas amostras independentes e identicamente distribuídas (i.i.d) das variáveis dos conjuntos  $\mathcal{G}$  e  $\mathcal{M}$ , para que se possa então, a partir do modelo computacional, calcular os valores  $Q_i^1, Q_i^2, \dots, Q_i^N$ , obtendo-se por fim o seu valor médio, o qual é obtido pela Eq. 3.5.

$$\hat{\gamma}_i = \frac{1}{N_i} \sum_{j=1}^N Q_i^j \quad (3.5)$$

Ressalta-se que  $N_i$  representa a quantidade de pacotes *entregues* pela fonte  $i$  ao monitor. Esse valor pode ser diferente da quantidade de pacotes gerados, o que ocorre em situações onde há perda ou descarte de pacotes. Pela lei dos grandes números [45], o valor estimado  $\hat{\gamma}_i$  converge para  $\gamma_i$  à medida em que  $N \rightarrow \infty$ , conforme mostra a Eq. 3.6.

$$\mathbb{P} \left( \lim_{N \rightarrow \infty} (|\hat{\gamma}_i - \gamma_i| = 0) \right) = 1 \quad (3.6)$$

Apesar de a lei dos grandes números garantir a convergência de  $\hat{\gamma}_i$ , não se pode admitir que as amostras  $Q_i^j$  são i.i.d.. Essa dependência entre as variáveis é uma característica de modelos baseados em cadeias de Markov, sendo o processo AoI incluído nesse grupo, conforme abordado por diversos estudos citados na Seção 2.3. Dessa maneira, o estimador da variância dado pelo Teorema Central do Limite (TCL),  $Var(\hat{\gamma}_i) = \frac{\sigma_i^2}{N}$ , onde  $\sigma_i^2$  é a variância de  $Q_i$ , não pode ser aplicada sem um fator de correção.

Alguns trabalhos como [46] sugerem a utilização de um ajuste na estimativa da variância em modelos de Cadeia de Markov com simulações de Monte Carlo, baseando-se na autocorrelação das amostras coletadas. Sendo assim, a estimativa da variância de  $\hat{\gamma}_i$  pode ser obtida pela Eq. 3.7, onde  $\alpha_i$  é chamado de *tempo de autocorrelação integrado* (TAI) para a fonte  $i$ .

$$Var(\hat{\gamma}_i) = \frac{\alpha_i}{N_i} \sigma_i^2 \quad (3.7)$$

Esse fator pode ser interpretado como uma correção no número de amostras consideradas independentes, o qual passa de  $N_i$ , para uma série de amostras independentes, para  $\frac{N_i}{\alpha_i}$ , para amostras dependentes entre si. O valor de  $\alpha_i$  pode ser obtido pela Eq. 3.8, sendo  $C_i(k)$  a autocovariância de  $Q_i$  para um *lag*  $k$ . O fator  $C_i(0)$  representa, então, a própria variância de  $Q_i$ .

$$\alpha_i = 1 + 2 \sum_{k=1}^{N_i} \left(1 - \frac{k}{N_i}\right) \frac{C_i(k)}{C_i(0)} \quad (3.8)$$

Essa autocovariância  $C_i(k)$  pode ser estimada pela Eq. 3.9.

$$C_i(k) = \frac{1}{N_i - k} \sum_{n=1}^{N_i - k} (Q_i^n - \hat{\gamma}_i)(Q_i^{n+k} - \hat{\gamma}_i) \quad (3.9)$$

Retomando-se a Eq. 3.7 e considerando-se o estimador não viesado  $E(s^2) = \sigma^2$ , a variância  $\sigma_i^2$  pode ser calculada pela Eq. 3.10.

$$s_i^2 = \frac{1}{N} \sum_{j=1}^N (Q_i^j - \hat{\gamma}_i)^2 \quad (3.10)$$

Uma vez obtido o valor esperado para a média de  $Q_i$ , bem como a sua variância, é possível obter o valor esperado da AoI média e o valor correspondente ao erro da estimativa. Partindo-se da Eq. 2.2, a AoI média para a fonte  $i$  com tempo de simulação  $\tau$  é calculada pela Eq. 3.11.

$$\hat{\Delta}_i = \frac{1}{\tau} \sum_{j=1}^{N_i} Q_i^j = \frac{N_i}{\tau} \hat{\gamma}_i \quad (3.11)$$

A variância para essa mesma fonte é dada pela Eq. 3.12.

$$Var(\hat{\Delta}_i) = \mathbb{E} \left( (\hat{\Delta}_i - \mathbb{E}(\Delta_i))^2 \right) = \left( \frac{N_i}{\tau} \right)^2 Var(\hat{\gamma}_i) \quad (3.12)$$

Por fim, partindo-se das Eq. 3.7, 3.10 e 3.12, a raiz do erro quadrático médio (REQM) da estimativa  $\hat{\Delta}_i$  pode ser obtida pela Eq. 3.13.

$$REQM_{\hat{\Delta}_i} = \sqrt{\frac{\alpha_i}{\tau^2} \sum_{j=1}^{N_i} (Q_i^j - \hat{\gamma}_i)^2} \quad (3.13)$$



### Análise de precisão para a AoI de pico média

A análise de precisão para a *Age of Information* de pico média pode ser realizada de maneira semelhante à análise da AoI média, substituindo-se as variáveis aleatórias  $Q_i$  pelas variáveis  $A_i$ . Nessa abordagem, o vetor de interesse passa a ser  $\mathbf{A} = [A_1, A_2, \dots, A_K]$ , sendo  $K$  o número de fontes de pacotes. O objetivo da simulação é calcular o valor esperado  $\theta = \mathbb{E}(A_i)$  para cada uma das fontes do sistema. Com base nas amostras de  $A_i$  geradas pela simulação, o valor esperado da AoI de pico para a fonte  $i$  é dado pela Eq. 3.14.

$$\hat{\theta}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} A_i^j \quad (3.14)$$

A variância de  $\hat{\theta}_i$  é obtida pela Eq. 3.15, onde  $\alpha_i$  é a TAI da variável aleatória  $A_i$ ,  $\sigma_i^2$  é a variância de  $A_i$  e  $N_i$  é a quantidade de pacotes entregues pela fonte  $i$ .

$$Var(\hat{\theta}_i) = \frac{\alpha_i}{N_i} \sigma_i^2 \quad (3.15)$$

O valor de  $\alpha_i$  pode ser obtido pela mesma fórmula da Eq. 3.8, sendo os valores da autocovariância obtidos pela Eq. 3.16.

$$C_i(k) = \frac{1}{N_i - k} \sum_{n=1}^{N_i - k} (A_i^n - \hat{\theta}_i)(A_i^{n+k} - \hat{\theta}_i) \quad (3.16)$$

A variância, por sua vez, pode ser estimada pela Eq. 3.17, assumindo-se um estimador não viesado e  $N \rightarrow \infty$ .

$$s_i^2 = \frac{1}{N_i} \sum_{j=1}^{N_i} (A_i^j - \hat{\theta}_i)^2 \quad (3.17)$$

Por fim, combinando as Eq. 3.15 e 3.17, a REQM da estimativa  $\hat{\theta}_i$  é dada pela Eq. 3.18.

$$REQM_{\hat{\theta}_i} = \sqrt{\frac{\alpha_i}{N_i^2} \sum_{j=1}^{N_i} (A_i^j - \hat{\theta}_i)^2} \quad (3.18)$$

### 3.2.2 Implementação e distribuição da ferramenta

A ferramenta implementada utiliza a linguagem Python, sendo executada através de um *script* base, o qual deve ser montado para cada cenário simulado. Por ocasião da execução do *script* base, a ferramenta instancia os diversos objetos necessários para a simulação, entre eles, cada um dos pacotes de atualização gerados nas fontes. Ao final da execução,

os resultados podem ser exibidos na tela ou ainda salvos em um arquivo para posterior análise.

O núcleo da ferramenta é o conjunto de classes e métodos que realiza a movimentação dos pacotes através dos nós do grafo. Esse *core* utiliza outro projeto Python de código aberto chamado *queueing-tool* [47], o qual consiste em uma ferramenta para a simulação de uma rede de filas genérica.

O simulador desenvolvido neste trabalho, até a data desta publicação, se apresenta como a única ferramenta de código aberto, completa e funcional, disponível para a avaliação das métricas da AoI em SCF. O acesso pode ser realizado no endereço <https://github.com/pprandel/aoi-simulator> [48], onde constam as instruções para instalação e utilização. Espera-se continuar o seu aperfeiçoamento através de contribuições da comunidade acadêmica voltada para a área.

### 3.3 A técnica LGFS-C

As técnicas de gerenciamento de pacotes apresentadas na Seção 2.2.3 podem ser consideradas estáticas, no sentido de que a decisão de se realizar a preempção, seja em serviço (LGFS-S) ou na fila de espera (LGFS-W), é aplicada de maneira uniforme a todos os pacotes que chegam à fila. Com a motivação de melhorar os resultados no estado da arte, uma das contribuições deste estudo é a proposta de uma nova técnica de gerenciamento de pacotes, a qual toma decisões de preempção de maneira dinâmica em função do estado atual do sistema. Essa técnica é chamada *Last Generated - First Served - Conditional* (LGFS-C), em que a letra 'C' significa preempção condicional, representando um avanço em relação às técnicas LGFS-S e LGFS-W ao utilizar o conceito de vida residual média (VRM) como um dos fatores de decisão.

#### 3.3.1 Vida residual média

Para que a técnica LGFS-C possa ser apresentada, torna-se necessário definir primeiramente o conceito de Vida Residual Média (VRM). A VRM é uma métrica estatística muito utilizada na teoria da confiabilidade, fornecendo uma estimativa da vida restante de um equipamento ou componente, dado que o mesmo já possui um tempo de vida decorrido. A definição formal desse conceito será apresentada a seguir.

##### Definição formal

Seja  $T$  uma variável aleatória com média e variância finitas. Seja também  $f(t) > 0$  a sua função densidade de probabilidade (FDP),  $F(t) > 0$  a função distribuição acumulada

(FDA) e  $\bar{F} = 1 - F$  a função sobrevivência. A função de *vida residual média* (VRM) [49] é definida pela Eq. 3.19.

$$m(t) \equiv E(T - t | T > t) = \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(u) du \quad (3.19)$$

Ao aplicar o conceito de VRM para o contexto do presente estudo, suponha-se que a variável aleatória  $T$  represente o tempo de serviço necessário para o processamento de um pacote de atualização. A função  $m(t)$  representa o tempo médio adicional necessário para que o servidor finalize o serviço desse pacote, dado que o tempo já decorrido de serviço é maior que  $t$ . Como será demonstrado adiante, a VRM é o componente chave da técnica LGFS-C, a qual avalia se é mais vantajoso manter um pacote que já está em serviço ou realizar a sua preempção por outro mais recente.

Um exemplo gráfico dessa função pode ser observado na Fig. 3.4, onde a vida residual média da distribuição Weibull (à direita) é ilustrada juntamente com a função de densidade de probabilidade (à esquerda). No caso particular dessa distribuição, um pacote com qualquer tempo de serviço  $t$  já decorrido teria um valor da vida residual média menor do que um pacote com tempo de serviço igual a zero (equivalente a um pacote que ainda não iniciou o serviço).

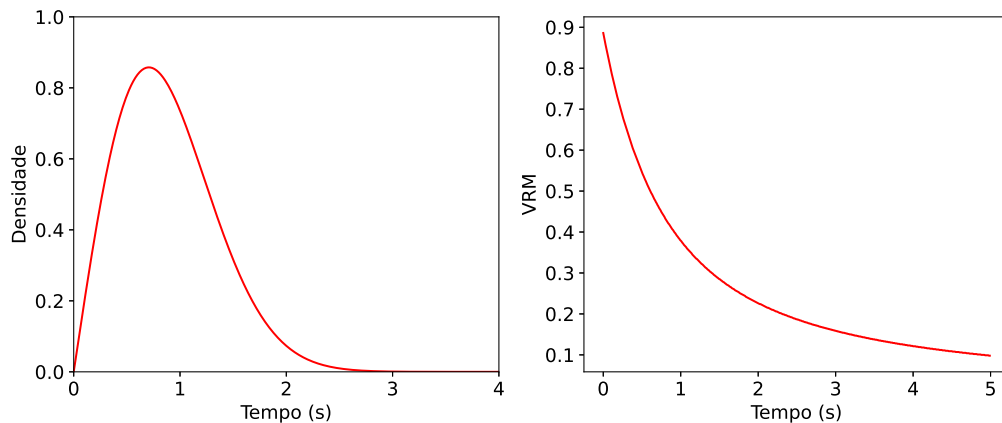


Figura 3.4: Exemplo de uma função densidade de probabilidade (Weibull) e sua respectiva vida residual média. Observa-se que, quanto maior o tempo decorrido  $t$ , menor é o valor da vida residual média.

### 3.3.2 Descrição da técnica

Considere-se inicialmente o modelo da Fig. 3.5, o qual consiste em uma fonte geradora de pacotes periódicos através de um processo de Poisson com taxa média  $\lambda$ , uma rede com atraso randômico de parâmetro  $k$  e uma estrutura fila-servidor com taxa média de serviço

$\mu$ . Devido ao atraso na rede, os pacotes podem chegar à fila fora da ordem em que foram gerados. O objetivo é otimizar a AoI da fonte vista pelo monitor.

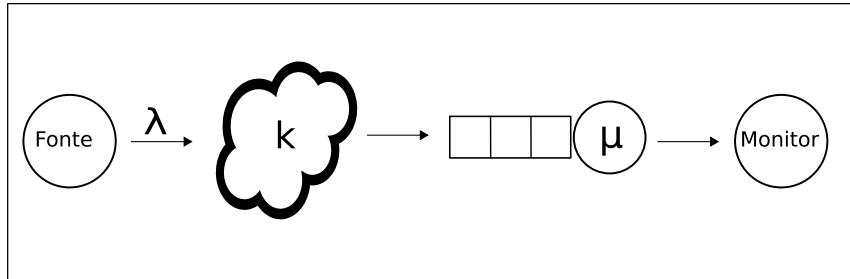


Figura 3.5: Estrutura fila servidor com pacotes chegando fora de ordem. A fila é representada de maneira genérica, podendo ter uma quantidade de posições de espera variada, a depender da técnica utilizada.

A técnica LGFS-C consiste em realizar o gerenciamento dos pacotes que chegam, ao utilizar um regime do tipo *Last Generated First Served* (LGFS) e preempção de pacotes, tanto em serviço quanto na fila de espera. Essa preempção é realizada com base em um processo decisório, o qual é ilustrado na Fig. 3.6. O mecanismo de decisão será explicado a seguir.

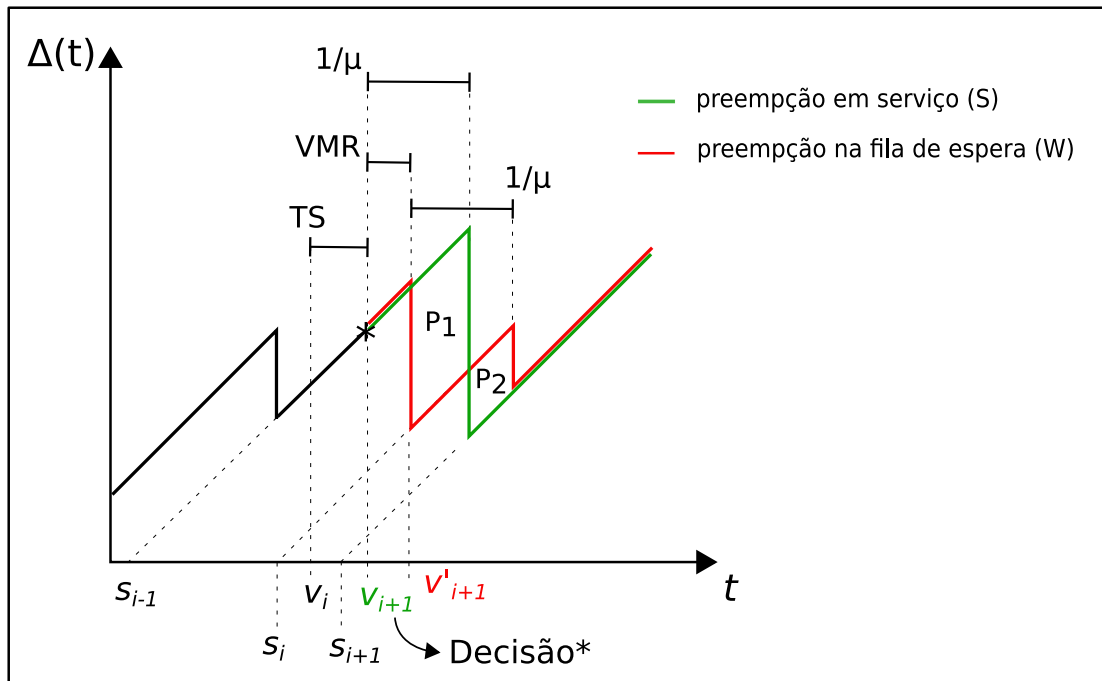


Figura 3.6: Ilustração da técnica LGFS-C. O gráfico mostra a evolução da AoI baseada na decisão de realizar a preempção em serviço (verde) ou na fila de espera (vermelho).

Sejam  $s_{i-1}$  e  $s_i$  respectivamente os tempos de geração dos pacotes de índices  $i - 1$  (último pacote entregue ao monitor) e  $i$  (pacote atualmente em serviço). Seja ainda  $v_i$  o

tempo em que o pacote  $i$  iniciou o seu serviço e  $TS$  o tempo já decorrido desse serviço até a chegada de um novo pacote de índice  $i + 1$  à fila, cujo tempo de geração é dado por  $s_{i+1}$ . Essa chegada do novo pacote à fila ocorre no instante  $v_{i+1}$ , que é exatamente o momento da tomada de decisão, a qual resultará em dois possíveis caminhos:

- Realizar a preempção do pacote em serviço (verde): o pacote  $i + 1$  realiza a preempção em serviço, substituindo o pacote  $i$ , o qual é imediatamente descartado. O pacote  $i + 1$  inicia seu serviço no tempo  $v_{i+1}$  e espera-se que seja servido em um tempo  $\frac{1}{\mu}$ , conforme mostra a Fig. 3.6.
- Realizar a preempção do pacote na fila de espera (vermelho): o pacote  $i+1$  é alocado na fila de espera e realiza a preempção na mesma caso já exista um pacote na fila. O pacote  $i$  continua o seu serviço, cujo tempo esperado é dado pela vida residual média (VRM) em função de  $TS$ , ou seja  $m(TS)$ . Após o término do serviço do pacote  $i$ , inicia-se imediatamente o serviço do pacote  $i + 1$ , o qual inicia em  $v'_{i+1}$  e cujo tempo esperado é dado pelo tempo médio do servidor, ou seja,  $\frac{1}{\mu}$ , conforme mostra a Fig. 3.6

Após o final desse processo decisório, observa-se que ambos os caminhos do processo AoI se encontram novamente e que, em função da decisão tomada, cada um desses caminhos resultará em uma área adicional específica. A área adicional  $P1$  resulta da preempção em serviço e a área adicional  $P2$  resulta da preempção na posição de espera. Uma vez que a AoI média é dada pela soma da área sob o gráfico dividida pelo tempo decorrido, conforme Eq. 2.1, a decisão que minimiza a AoI será aquela que produz a menor das áreas. Dessa maneira, o critério de decisão é dado pela Eq. 3.20.

$$\text{Preempção} = \begin{cases} S, & \text{caso } P1 \leq P2; \\ W, & \text{caso contrário.} \end{cases} \quad (3.20)$$

O indicativo  $S$  significa realizar a preempção em serviço e  $W$  significa realizar a preempção na fila de espera. Uma vez estabelecido o critério de decisão, calcula-se as áreas dos paralelogramos  $P1$  e  $P2$ , de forma a se obter o mesmo critério em função das variáveis relacionadas ao domínio do tempo, de acordo com a Eq. 3.21.

$$\begin{aligned} P1 &= (s_i - s_{i-1}) \left( \frac{1}{\mu} - VRM \right) \\ P2 &= (s_{i+1} - s_i) VRM \end{aligned} \quad (3.21)$$

Na Eq. 3.21, o valor da vida residual média é dado por  $VRM = m(TS)$ , conforme a Eq. 3.19. Por fim, após algumas manipulações, o critério de decisão da Eq. 3.20 pode ser reescrito como na Eq. 3.22.

$$\text{Preempção} = \begin{cases} S, & \text{caso } \frac{s_i - s_{i-1}}{s_{i+1} - s_{i-1}} \leq m(TS)\mu; \\ W, & \text{caso contrário.} \end{cases} \quad (3.22)$$

A técnica LGFS-C, de maneira semelhante ao que ocorre nas técnicas LGFS-S e LGFS-W, prioriza os pacotes mais recentes ou frescos, ou seja, aqueles que possuem o maior tempo de geração, independente da decisão tomada. A Fig. 3.6 ilustra um caso particular em que os pacotes chegam na mesma ordem em que foram gerados. Para os casos em que os pacotes chegam fora da ordem de geração, o Algoritmo 1 apresenta o mecanismo completo da técnica LGFS-C, o qual contempla todos os casos possíveis de chegadas bem como os possíveis estados da fila e do servidor.

### 3.3.3 Demonstração analítica

Nesta seção será demonstrado que a técnica LGFS-C apresenta sempre resultados superiores com relação às técnicas no estado da arte do gerenciamento de pacotes, quais sejam a LGFS-S e LGFS-W.

*Teorema:* Seja uma estrutura fila-servidor com uma única posição de espera, a qual recebe pacotes provenientes de uma única fonte, os quais são gerados de acordo com um processo de *Poisson* e podem ser recebidos fora da ordem de geração. A técnica LGFS-C resulta em uma AoI média  $\Delta$  igual ou menor à AoI média que seria produzida pelas técnicas LGFS-S ou LGFS-W, para qualquer distribuição de tempo de serviço.

*Demonstração:* Seja  $\delta_{P,i}$  o valor da área adicional produzida no processo AoI, conforme Fig. 3.6, resultante da decisão de se realizar a preempção em serviço ( $P1$ ) ou a preempção na fila de espera ( $P2$ ) para um pacote de índice  $i$ , de acordo com a Eq. 3.23.

$$\delta_{P,i} = \begin{cases} P_1, & \text{caso Preempção} = S; \\ P_2, & \text{caso Preempção} = W. \end{cases} \quad (3.23)$$

Como o critério de decisão da técnica LGFS-C (Eq. 3.22) é baseado na menor área adicional, pode-se afirmar que o valor esperado de  $\delta_{P,i}$  é sempre menor (ou igual) para a técnica LGFS-C do que para as técnicas LGFS-S ou LGFS-W (LGFS-S/W), conforme a Eq. 3.24.

$$E[\delta_{P,i}^{LGFS-C}] \leq E[\delta_{P,i}^{LGFS-S/W}] \quad (3.24)$$

A Eq. 3.24 é válida para todo pacote  $i > 0$  sobre o qual uma decisão de preempção foi tomada. Para os pacotes que não necessitam de decisão quanto à preempção, o processo da AoI permanece o mesmo independente da técnica. Consequentemente, pela lei dos grandes

---

**Algoritmo 1** LGFS-C

---

```
1: defina  $s_{pacote}$  como o tempo de geração do pacote  $pacote$ 
2: defina  $entregue$  como o último pacote entregue ao monitor e  $s_{entregue} \leftarrow 0$ 
3: defina  $VRM(pacote)$  como a vida residual média do pacote  $pacote$ 
4: defina  $\mu$  como a taxa de serviço
5: função PREEMPCAO(servico, novo)
6:   se  $\frac{s_{servico} - s_{entregue}}{s_{novo} - s_{entregue}} \leq VRM(servico)\mu$  então
7:     retorne Verdadeiro
8:   caso contrário
9:     retorne Falso
10:  fim se
11: fim função
12: enquanto sistema ON faça
13:   se chegou um pacote  $novo$  então
14:     se servidor e fila vazios então
15:       renomeie o pacote  $novo$  para  $servico$ 
16:       inicie o serviço do pacote  $servico$ 
17:     caso contrário
18:       se existe um pacote  $fila$  na fila de espera então
19:         se  $s_{novo} > s_{fila}$  então
20:           descarte o pacote  $fila$ 
21:         caso contrário
22:           descarte o pacote  $novo$ 
23:         fim se
24:       fim se
25:       se existe um pacote  $servico$  em serviço então
26:         se  $s_{novo} > s_{servico}$  então
27:           se PREEMPCAO( $servico, novo$ ) é Verdadeiro então
28:             interrompa o serviço e descarte o pacote  $servico$ 
29:             renomeie o pacote  $novo$  para  $servico$ 
30:             inicie o serviço do pacote  $servico$ 
31:           caso contrário
32:             renomeie o pacote  $novo$  para  $fila$ 
33:             insira o pacote  $fila$  na fila
34:           fim se
35:         caso contrário
36:           descarte o pacote  $novo$ 
37:         fim se
38:       fim se
39:     fim se
40:   caso contrário se se o pacote  $servico$  foi entregue então
41:     renomeie o pacote  $servico$  para  $entregue$ 
42:   se existe um pacote  $fila$  então
43:     renomeie o pacote  $fila$  para  $servico$ 
44:     inicie o serviço do pacote  $servico$ 
45:   fim se
46: fim se
47: fim enquanto
```

números, pode-se afirmar que a soma da área total sob o gráfico, conforme ilustrado no Cap. 2 pela Fig. 2.3, tende a ser menor para a técnica LGFS-C, quando comparada às demais técnicas, à medida em que o número de pacotes  $i$  tende a infinito. Essa afirmação é expressa na Eq. 3.25.

$$P\left(\lim_{i \rightarrow \infty} \sum Q_i^{LGFS-C} \leq \lim_{i \rightarrow \infty} \sum Q_i^{LGFS-S/W}\right) = 1 \quad (3.25)$$

Por fim, dado que a AoI média  $\Delta$  é dada pela soma das áreas sob o gráfico do processo AoI, dividida pelo tempo decorrido, conclui-se que a técnica LGFS-C tende a produzir uma AoI média igual ou inferior às demais técnicas. Ressalta-se que o referido teorema é válido para qualquer distribuição de tempo de serviço, uma vez que a decisão de preempção é realizada pacote a pacote, o que garante sempre as condições da Eq. 3.24. A análise do alcance da técnica será realizada com maiores detalhes na seção seguinte.

### 3.3.4 Alcance e viabilidade da técnica

Conforme demonstrado, a técnica LGFS-C pode produzir uma AoI média igual ou inferior do que as técnicas LGFS-S/W, a depender de outros fatores do sistema. A partir da análise da condição de decisão da técnica LGFS-C (Eq. 3.22), nota-se que o termo esquerdo da desigualdade é sempre inferior à unidade, ou seja,  $\frac{s_i - s_{i-1}}{s_{i+1} - s_{i-1}} < 1$ . Logo, verifica-se que uma condição necessária (porém não suficiente) para que ocorra a decisão de preempção na fila de espera (W) é um valor da vida residual média inferior à média da distribuição, ou seja,  $m(TS) < 1/\mu$ . Uma consequência dessa análise é que uma distribuição exponencial, onde  $m(t) = 1/\mu$ , para todo  $t > 0$ , torna as técnicas LGFS-C e LGFS-S equivalentes, pois a decisão é sempre tomada em favor da preempção em serviço (S). Adicionalmente, essas técnicas serão ótimas para a AoI média, conforme demonstrado em [32].

Para distribuições que possuem algum valor da vida média residual inferior à média da própria distribuição, ou seja,  $m(t) < 1/\mu$ , para algum  $t > 0$ , espera-se a ocorrência de preempções em serviço (S) e na fila de espera (W), a depender das demais variáveis envolvidas do processo decisório. Uma grande quantidade de distribuições satisfaz essa condição, como Weibull, Lognormal, Erlang, entre outras. Para essas distribuições, a AoI média sempre apresentará valores inferiores para a técnica LGFS-C quando comparada às técnicas LGFS-S/W.

Quanto à viabilidade, a técnica LGFS-C utiliza o valor da vida residual média do pacote em serviço para a tomada de decisão. Dessa maneira, para cada pacote em um processo decisório, faz-se necessário o cálculo da VRM em função do tempo em serviço TS, o que requer conhecimento sobre a distribuição do tempo de serviço. Para sistemas em que essa distribuição é conhecida e parametrizada, pode-se calcular a VRM a partir da



função distribuição de probabilidade (fdp), conforme Eq. 3.19. Como forma de se reduzir o custo computacional, os valores da VRM podem ser calculados *a priori*, podendo-se utilizar uma tabela estática para a obtenção da vida residual média em função do tempo em serviço do pacote.

Em sistemas reais, como por exemplo o modelo adotado neste estudo (Fig. 3.3), muitas vezes não se tem o conhecimento prévio da distribuição do tempo de serviço ou, ainda, essa distribuição é dinâmica, podendo variar conforme o estado atual do sistema. Para esses casos, pode-se utilizar estimadores não paramétricos para se obter a vida residual média em tempo real, em função dos tempos de serviço anteriores. Alguns estimadores não paramétricos para a VRM são sugeridos em [50].

O Cap. 4 apresentará os resultados experimentais obtidos através de simulações, as quais validam o modelo analítico apresentado.

### 3.4 A técnica MAF-LGFS-MW

A técnica LGFS-C, apresentada na seção anterior, mostrou-se capaz de produzir resultados superiores que as demais técnicas no estado da arte para um modelo de uma fonte e um servidor para qualquer distribuição de tempo de serviço. Com o intuito de se obter uma técnica de gerenciamento de pacotes capaz de melhorar os resultados dos atuais modelos com múltiplas fontes e múltiplos servidores, esta seção apresentará a proposta de uma outra técnica chamada MAF-LGFS-MW, a qual agrega características das técnicas MAF-LGFS e MASIF-LGFS, apresentadas na Seção 2.2.4, com o diferencial de realizar a preempção em múltiplas filas de espera. O significado da sigla adotada é o seguinte:

- MAF: *Max Age First*. Prioriza o serviço da fonte com a maior AoI, ou seja, a fonte mais *velha*.
- LGFS: *Last Generated - First Served*. Regime de serviço que prioriza os pacotes com maior tempo de geração no âmbito de uma mesma fonte.
- MW: *Multiple preemption in Waiting*. Técnica na qual cada fonte possui uma posição de espera, na qual pode ser realizada a preempção de pacotes.

Existem poucas técnicas na literatura sobre o gerenciamento de pacotes para a otimização da AoI em modelos com múltiplas fontes e múltiplos servidores. Conforme apresentado na Seção 2.2.4, o estudo de [37] apresenta duas dessas técnicas, quais sejam:

- MAF-LGFS: *Maximum Age First - LGFS*.
- MASIF-LGFS: *Maximum Age of Served Information First - LGFS*.

O foco de ambas as técnicas é otimizar a AoI média para um conjunto de fontes, de forma a obter o seu menor valor possível dado um determinado modelo e seus parâmetros. Para esse objetivo, a MAF-LGFS tem como fundamento priorizar a fonte com a maior AoI atual e a MASIF-LGFS procura priorizar a fonte com a maior *Age of Served Information* (AoSI), conceito cuja definição também foi apresentada na Seção 2.2.4. De acordo com [37], a MAF-LGFS com preempção de pacotes otimiza a AoI para um modelo com um único servidor cujos tempos de serviço são exponenciais. Já para modelos com múltiplos servidores, a MAF-LGFS não produz resultados ótimos. Um outro resultado desse estudo é que a MASIF-LGFS sem preempção produz resultados *quase ótimos* para um modelo com múltiplos servidores cujos tempos de serviço pertencem à classe de distribuições *New Better than Used* (NBU), dentro da qual também se inclui a distribuição exponencial. A Tabela 3.1 apresenta uma comparação resumida das técnicas.

Tabela 3.1: Comparação de técnicas no estado da arte para modelos com múltiplas fontes.

	MAF-LGFS c/ preempção	MASIF-LGFS s/ preempção
Servidor único	Ótima para tempos de serviço exponenciais	Não ótima
Múltiplos servidores	Não ótima	Quase ótima

Dos trabalhos citados acima, observa-se que a MASIF-LGFS sem preempção seria a principal candidata a produzir os melhores resultados para um modelo com múltiplas fontes e múltiplos servidores com tempos de serviço NBU. A referida técnica não utiliza preempção (descarte e substituição) de pacotes nos servidores, o que a obriga a alocar pacotes em uma fila de espera quando todos esses servidores estão ocupados. Pelo fato de não utilizar preempção, a carga total nos servidores deve ser inferior à unidade, ou seja,  $\rho < 1$ , o que impede o aumento da fila de maneira indefinida.

O fato da técnica MASIF-LGFS não utilizar preempção em serviço é baseado na propriedade das distribuições NBU, apresentada na Eq. 2.26. Em termos práticos, a substituição de um pacote em serviço nem sempre é compensatória devido ao fato desse pacote (já com o serviço em andamento) apresentar uma probabilidade de ser servido em um tempo menor do que um pacote recém chegado ao servidor. Já a utilização da AoSI como fator de seleção para a fonte prioritária visa evitar a alocação simultânea em serviço de vários pacotes oriundos da mesma fonte. Essa condição pode ser também alcançada limitando-se que apenas um pacote de uma determinada fonte possa estar em serviço em um determinado momento.

Quanto à técnica LGFS-C, apresentada na seção anterior, a mesma utiliza o conceito de Vida Residual Média (VRM) para dinamicamente optar pela preempção em serviço ou

na fila de espera de um determinado pacote. Para o caso de múltiplas fontes e múltiplos servidores, no entanto, a LGFS-C enfrenta algumas restrições para ser utilizada, como por exemplo o fato de não ser interessante realizar o serviço sequencial de dois pacotes oriundos da mesma fonte, uma das premissas utilizadas na modelagem dessa técnica. Por esse motivo, optou-se por não utilizar a VRM na modelagem de uma técnica para múltiplas fontes e servidores.

Com base nos aspectos levantados nesta seção, conclui-se que as características desejadas para uma técnica que visa otimizar a AoI para um modelo de múltiplas fontes e múltiplos servidores são:

- Não utilizar preempção em serviço, caso as distribuições de tempo dos servidores pertençam à classe NBU.
- Manter uma posição de espera para cada uma das fontes. Realizar a preempção nessas filas, mantendo-se apenas um pacote, o qual se manterá atualizado.
- Selecionar para serviço prioritariamente a fonte com a maior AoI, desde que não exista um pacote dessa fonte já em serviço.

A próxima seção apresenta em detalhes a técnica MAF-LGFS-MW, a qual agrega as características citadas acima. A eficiência dessa técnica é avaliada através de simulações.

### 3.4.1 Descrição da técnica

O modelo utilizado para a demonstração e avaliação da técnica MAF-LGFS-MW é mostrado na Fig. 3.7.

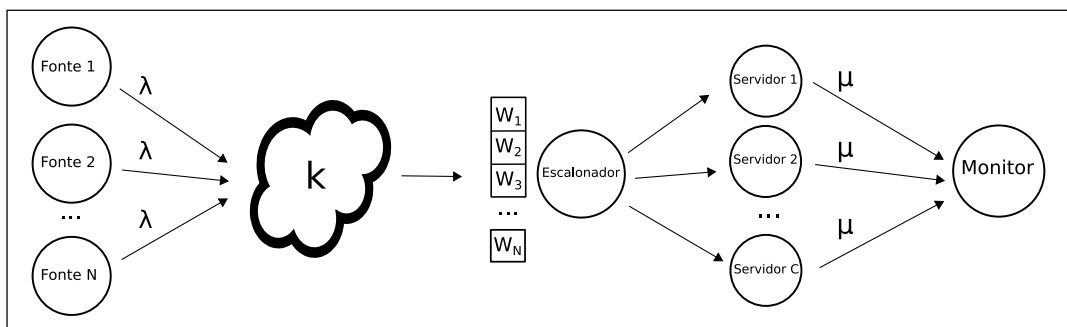


Figura 3.7: Representação do modelo para a definição e avaliação da técnica MAF-LGFS-MW.

Um conjunto de  $N$  fontes envia pacotes periódicos de atualização para um monitor de acordo com um processo de Poisson com taxa média  $\lambda$ . Esses pacotes são transmitidos através de uma rede cujo atraso segue uma distribuição exponencial com valor médio igual a  $k$ , o que possibilita que os pacotes cheguem ao conjunto de servidores fora da ordem

original de geração. Ao chegarem na estrutura de servidores, caso não haja um servidor disponível, esses pacotes são alocados em uma fila de espera  $W$ , a qual admite no máximo um pacote por fonte, ou seja, a capacidade máxima da fila será de  $N$  posições de espera. Dessa maneira, um pacote oriundo de uma fonte com índice  $i$  será identificado na fila de espera por  $W_i$ .

Por ocasião da chegada de um novo pacote  $A_i$ , oriundo da fonte  $i$ , caso já exista um pacote  $W_i$  na fila, será realizada então a preempção do mesmo, desde que o pacote  $A_i$  seja mais recente que o pacote  $W_i$ . Caso contrário, o pacote  $A_i$  será descartado. Essa condição é mostrada na Eq. 3.26.

$$A_i \begin{cases} \text{realiza a preempção de } W_i, & \text{caso } s_{A_i} > s_{W_i}; \\ \text{é descartado,} & \text{caso contrário.} \end{cases} \quad (3.26)$$

Na Eq. 3.26,  $s_P$  representa o tempo de geração do pacote  $P$ . Esse mecanismo visa manter os pacotes mais recentes na fila de espera quanto possível, sem interferir nos pacotes que já estão em serviço. Por ocasião da entrega de um pacote em serviço e a conseqüente liberação de um servidor, o escalonador selecionará para esse servidor o pacote  $W_i$  pertencente à fonte com a maior AoI atual. Essa fonte mais *velha* ou antiga pode ser selecionada com base no menor tempo de geração entre os últimos pacotes  $D_1, D_2, \dots, D_N$  entregues por cada fonte, conforme a condição da Eq. 3.27.

$$\text{Fonte com maior AoI: } i \mid D_i = \min\{s_{D_1}, s_{D_2}, \dots, s_{D_N}\} \quad (3.27)$$

Todos os servidores possuem a mesma taxa média de serviço  $\mu$  e podem assumir diferentes distribuições para o tempo de serviço. O processo completo da técnica MAF-LGFS-MW é melhor detalhado no Algoritmo 2.

Para a técnica MAF-LGFS-MW não é apresentada uma demonstração analítica comparativa com as demais técnicas no estado da arte, sendo essa técnica avaliada em função de resultados experimentais os quais podem ser vistos no Cap. 4.

---

**Algoritmo 2** MAF-LGFS-MW

---

```
1: defina  $s_i$  como o tempo de geração do pacote  $i$ 
2: defina  $D_i$  como o último pacote entregue ao monitor pela fonte  $i$ 
3: defina  $V_i$  como um pacote da fonte  $i$  em serviço
4: função BUSCA-FILA
5:   retorne o pacote  $W_i \mid D_i = \min\{s_{D_1}, s_{D_2}, \dots, s_{D_N}\}$ 
6: fim função
7: enquanto sistema ON faça
8:   se chegou um pacote  $A_i$  oriundo da Fonte  $i$  então
9:     se existe um servidor vazio então
10:      inicie o serviço do pacote  $A_i$ 
11:     caso contrário
12:       se existe um pacote  $W_i$  da fonte  $i$  na fila de espera então
13:         se  $s_{A_i} > s_{W_i}$  então
14:           descarte o pacote  $W_i$  da fila
15:           renomeie o pacote  $A_i$  para  $W_i$ 
16:           insira  $W_i$  na fila
17:         caso contrário
18:           descarte o pacote  $A_i$ 
19:         fim se
20:       fim se
21:     fim se
22:     caso contrário se um pacote foi entregue então
23:       se BUSCA-FILA retornou um pacote  $W_i$  então
24:         renomeie o pacote  $W_i$  para  $V_i$ 
25:         inicie o serviço do pacote  $V_i$ 
26:       fim se
27:     fim se
28: fim enquanto
```

---

### 3.5 Resumo do capítulo

O presente capítulo apresentou a modelagem do sistema de monitoramento e decisão no contexto da atividade militar de comando e controle, com destaque para a carga de trabalho, os parâmetros e as métricas de avaliação. Para possibilitar a avaliação do sistema, foi definido também um modelo computacional usado como base para o desenvolvimento de uma ferramenta de simulação. Este capítulo também apresentou a proposta de duas novas técnicas de gerenciamento de pacotes, a LGFS-C, a qual é aplicável para modelos com uma única fonte e servidor, e a MAF-LGFS-MW, a qual pode ser utilizada para modelos com múltiplas fontes e múltiplos servidores.

# Capítulo 4

## Resultados Experimentais

Este capítulo tem por objetivo apresentar os resultados experimentais desta pesquisa. Os resultados estão divididos em quatro seções, as quais correspondem às contribuições deste trabalho, a saber:

- Validação da ferramenta de simulação mediante comparações com modelos analíticos.
- Avaliação da técnica LGFS-C e comparação com outras técnicas do estado da arte.
- Avaliação da técnica MAF-LGFS-MW e comparação com outras técnicas do estado da arte.
- Avaliação de um SCF caracterizado como um sistema de monitoramento e decisão, em três cenários de uso.

### Relevância estatística dos resultados

A relevância estatística dos resultados pode ser calculada com base na Raiz do Erro Quadrático Médio (REQM) fornecida pela ferramenta para cada uma das simulações executadas. Para os experimentos realizados neste trabalho, foi considerado um intervalo de confiança de 95% para a definição dos valores da REQM máxima. Dessa maneira, a REQM retornada pela ferramenta deve ser multiplicada pelo valor de 1,96 para se obter os limites do intervalo de confiança [45]. Esse intervalo é dado pela Eq. 4.1, em que  $\hat{\Delta}$  é a estimativa retornada pela ferramenta e  $\Delta$  é o valor real, o qual seria obtido por métodos analíticos ou através da realização de um número infinito de simulações.

$$\hat{\Delta} - 1,96 \text{ REQM} \leq \Delta \leq \hat{\Delta} + 1,96 \text{ REQM} \quad (4.1)$$

Como exemplo, caso a ferramenta retorne uma AoI igual a 3 segundos e uma REQM igual a 0,1, existe uma confiança de 95% de que a verdadeira AoI está dentro do intervalo [2,804;3,196].

## 4.1 Validação da Ferramenta de Simulação

A ferramenta de simulação é um dos resultados mais importantes desta pesquisa por facilitar a avaliação de diversas técnicas e flexibilizar a configuração de diversos cenários de uso de SCFs. Conforme os conceitos de avaliação de desempenho de sistemas computacionais [51], os resultados obtidos por uma ferramenta de simulação são confiáveis se o modelo computacional é validado. Essa tarefa normalmente é realizada através da comparação dos valores obtidos nas simulações com valores obtidos por modelos analíticos. Entretanto, ressalta-se a impossibilidade de se validar a simulação de modelos de sistemas mais complexos, como por exemplo o sistema apresentado na Fig. 3.3, pois não existem modelos analíticos capazes de estimar a AoI para tais sistemas. Dessa maneira, em relação à AoI, a validação do modelo computacional será feita com modelos mais simples, para os quais já existem expressões analíticas de referência.

Apesar da limitação em se obter expressões analíticas para validar resultados em modelos mais complexos, o modelo computacional proposto pode ser considerado válido para qualquer sistema dado que o cálculo da AoI depende apenas dos tempos de geração dos pacotes nas fontes e dos tempos de entrega dos mesmos no monitor. Sendo assim, o restante do grafo passa a ser uma *caixa preta*, conforme mostra a Fig. 4.1, cujo tamanho ou complexidade não interfere na metodologia de cálculo.

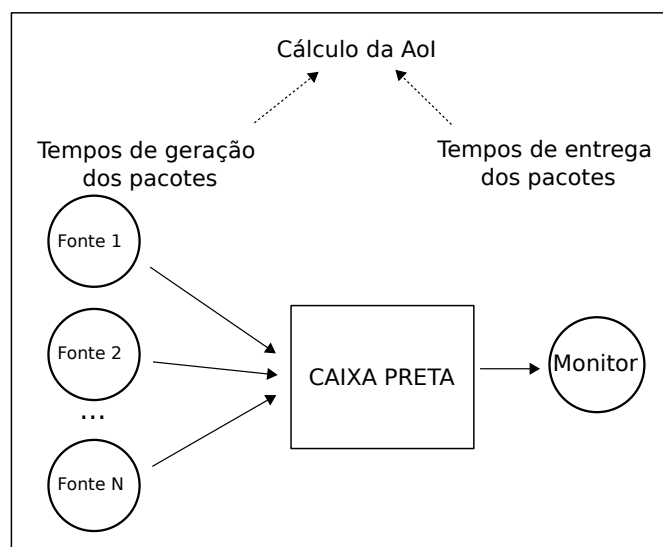


Figura 4.1: Representação do modelo computacional contendo uma *caixa preta*. Os valores utilizados para o cálculo da AoI são coletados apenas nas fontes e no nó de destino.

Uma vez que o método é válido para qualquer tipo de grafo, outro problema a ser considerado é a precisão dos resultados, uma vez que os mesmos são obtidos através de métodos estatísticos. Conforme discutido no Cap. 3, Seção 3.2.1, a precisão dos resultados é ajustada em termos da Raiz do Erro Quadrático Médio (REQM). Todos os resultados que serão apresentados nesta seção atingiram uma REQM igual ou inferior a 0,1.

## Filas elementares

Os modelos de filas elementares, sem a utilização de técnicas de gerenciamento de pacotes, foram validados através das expressões apresentadas na Seção 2.2.2. A Fig. 4.2 apresenta a comparação entre os valores obtidos pelas expressões analíticas e os valores calculados através de simulação para as filas do tipo  $M/M/1$ ,  $M/D/1$  e  $D/M/1$ . Os resultados são apresentados em função da carga  $\rho$  no servidor, fixando-se a taxa de serviço  $\mu = 1$ .

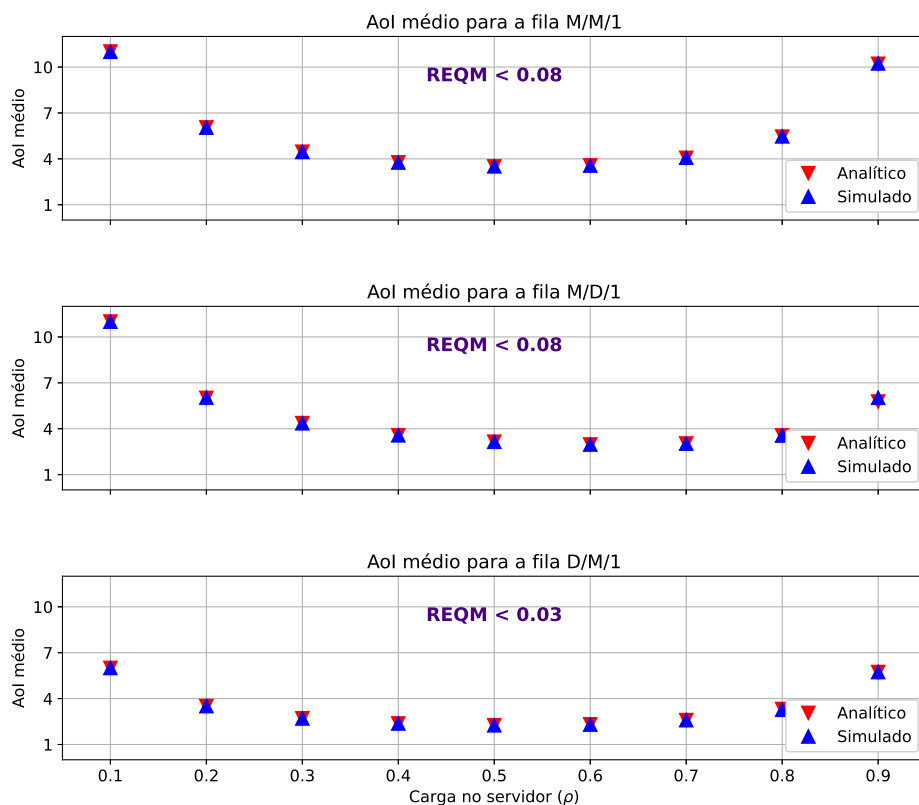


Figura 4.2: Comparação entre os resultados analíticos e simulados para filas elementares sem gerenciamento de pacotes ( $\mu = 1$ ).

Para ilustrar o processo de convergência da simulação, bem como a influência da autocorrelação na quantidade de pacotes necessários para se obter um erro quadrático



médio máximo, a Fig. 4.3 mostra os valores da AoI em função do número de pacotes gerados para uma fila do tipo  $M/M/1$  em duas situações de carga no servidor. Observa-se que para uma carga  $\rho = 0.2$  (Fig. 4.3a) a simulação atinge o valor alvo para a REQM antes dos 100.000 pacotes. Já para uma carga  $\rho = 0.9$  (Fig. 4.3b) são necessários cerca de 400.000 pacotes gerados para se obter a mesma precisão. Isso deve-se ao fato de uma maior carga ocasionar uma maior quantidade de pacotes na espera, o que aumenta a autocorrelação do sistema e exige um número maior de simulações.

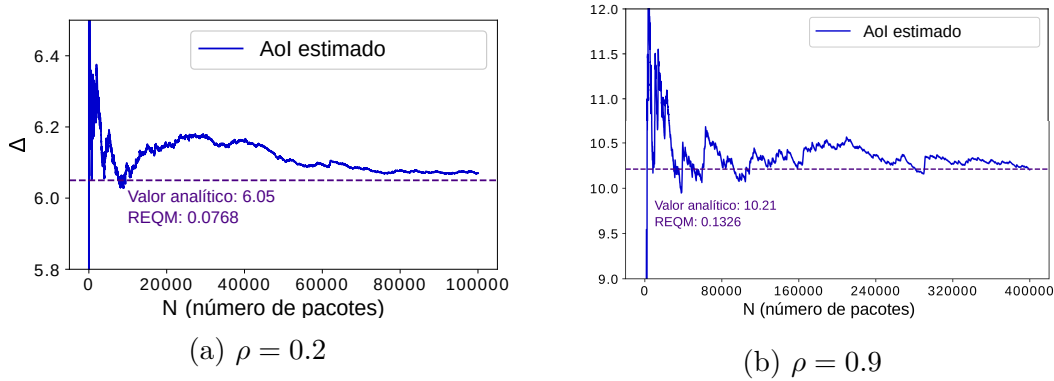


Figura 4.3: Processo de convergência da AoI para uma fila do tipo  $M/M/1$ , taxa média de serviço  $\mu = 1$  e duas cargas diferentes para o servidor.

A simulação de filas básicas com a utilização de técnicas de gerenciamento de pacotes foi validada através das expressões analíticas apresentadas na Seção 2.2.3. A Fig. 4.4 apresenta a comparação entre os valores obtidos pelas expressões analíticas e os valores calculados pela simulação para as filas do tipo  $M/M/1/1^*$ ,  $M/D/1/1^*$  e  $D/M/1/1^*$ , as quais utilizam regime do tipo LCFS e precedência do pacote em serviço.

A Fig. 4.5 apresenta a comparação entre os valores obtidos pelas expressões analíticas e os valores calculados pela simulação para as filas do tipo  $M/M/1/2^*$ ,  $M/D/1/2^*$  e  $D/M/1/2^*$ , as quais utilizam um regime do tipo LCFS com precedência na posição de espera. Em relação aos modelos para a AoI média de pico, foi realizada uma comparação entre os resultados simulados e o modelo analítico apresentado na Seção 2.2.2. Os resultados são apresentados na Fig. 4.6.

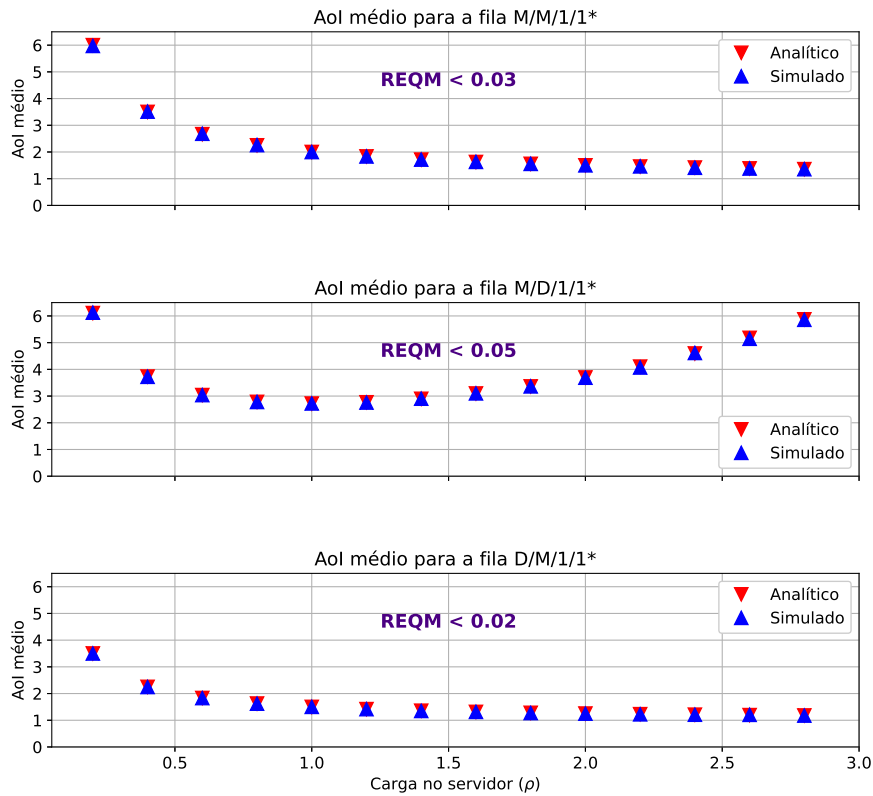


Figura 4.4: Comparação entre os resultados analíticos e simulados para filas elementares do tipo LCFS com precedência em serviço. ( $\mu = 1$ ).

## Múltiplas fontes

A literatura apresenta expressões analíticas para modelos de múltiplas fontes apenas do tipo  $M/M/1$ , para os casos sem precedência, com precedência na espera e com precedência em serviço. Essas fórmulas foram apresentadas na Seção 2.2.4. A Fig. 4.7 apresenta a comparação entre os resultados analíticos e os simulados para um modelo de duas fontes com um regime sem precedência do tipo FCFS e filas do tipo  $M/M/1$ , ou seja, processos de chegada e de serviço markovianos. O gráfico mostra a variação da AoI entre as duas fontes para diversos cenários de carga total no servidor  $\rho = \rho_1 + \rho_2$ . A taxa de serviço média foi definida em  $\mu = 1$ .

A Fig. 4.8 apresenta a comparação entre os resultados analíticos e os simulados para um modelo de duas fontes utilizando um regime com precedência em serviço ( $M/M/1/1^*$ ). A figura mostra a variação da AoI entre as duas fontes para diversos cenários de carga total no servidor  $\rho = \rho_1 + \rho_2$  com taxa de serviço média  $\mu = 1$ .

Por fim, a Fig. 4.9 apresenta a comparação entre os resultados analíticos e os simulados

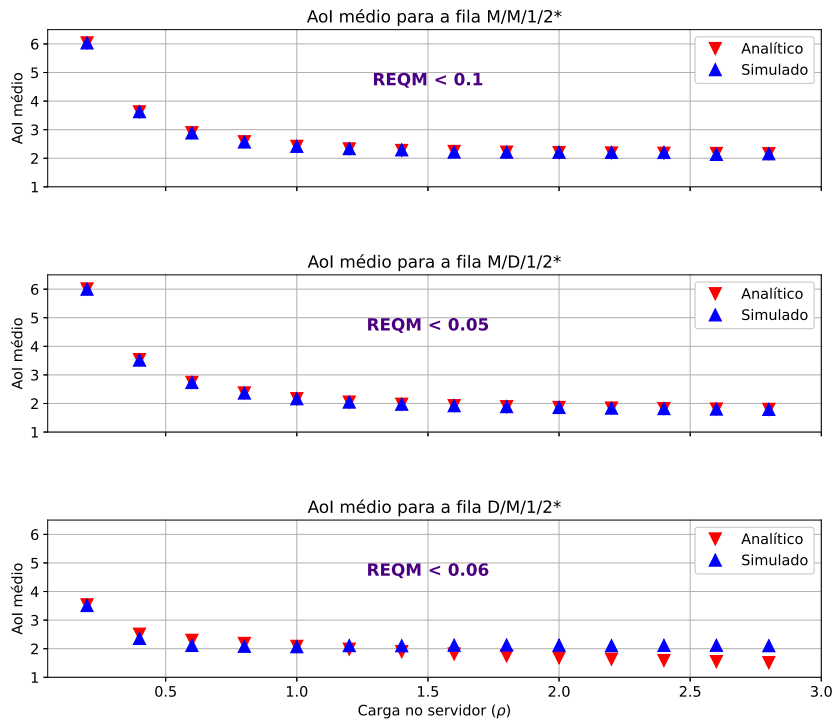


Figura 4.5: Comparação entre os resultados analíticos e simulados para filas elementares do tipo LCFS com precedência na espera. ( $\mu = 1$ ).

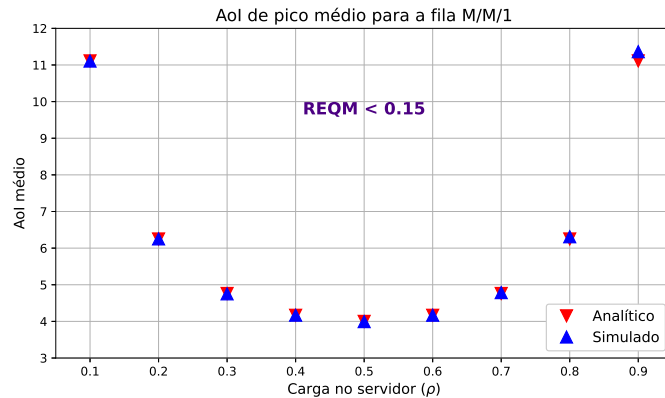


Figura 4.6: Comparação entre os resultados analíticos e simulados para a AoI de pico média em uma fila do tipo FCFS do tipo  $M/M/1$  ( $\mu = 1$ ).

para um modelo de duas fontes utilizando um regime com precedência na posição de espera ( $M/M/1/2^*$ ). A figura mostra a variação da AoI entre as duas fontes para diversos cenários de carga total no servidor  $\rho = \rho_1 + \rho_2$  com taxa de serviço média  $\mu = 1$ .

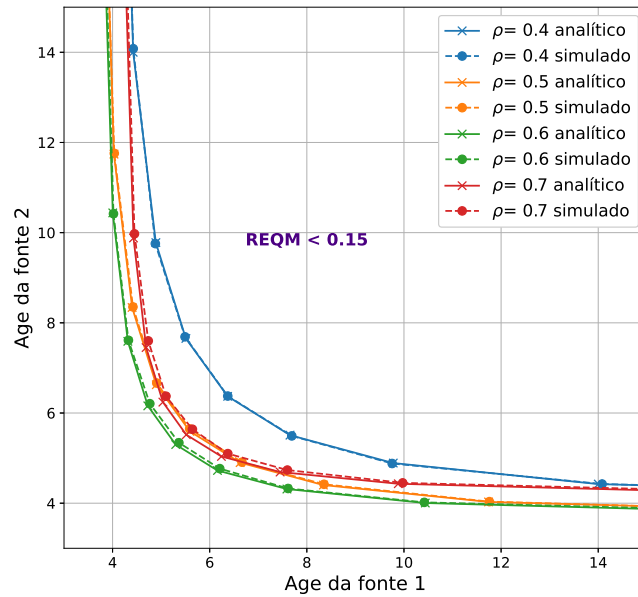


Figura 4.7: Comparação entre os resultados analíticos e simulados para o modelo com duas fontes utilizando o regime FCFS (sem precedência) do tipo  $M/M/1$  para diversos valores da carga total ( $\rho$ ) no servidor ( $\mu = 1$ ).

### Resumo das análises

A partir dos resultados apresentados nesta seção, conclui-se que o modelo computacional apresenta resultados válidos para todas os tipos de filas e técnicas avaliadas, com clara coerência entre o nível de precisão indicado pela ferramenta através da REQM e a aderência dos resultados da simulação em relação às expressões analíticas. Dessa forma, a ferramenta de simulação pode ser considerada confiável para ser utilizada na validação dos resultados das técnicas propostas e para a avaliação do sistema de monitoramento e decisão definido na Fig. 3.3.

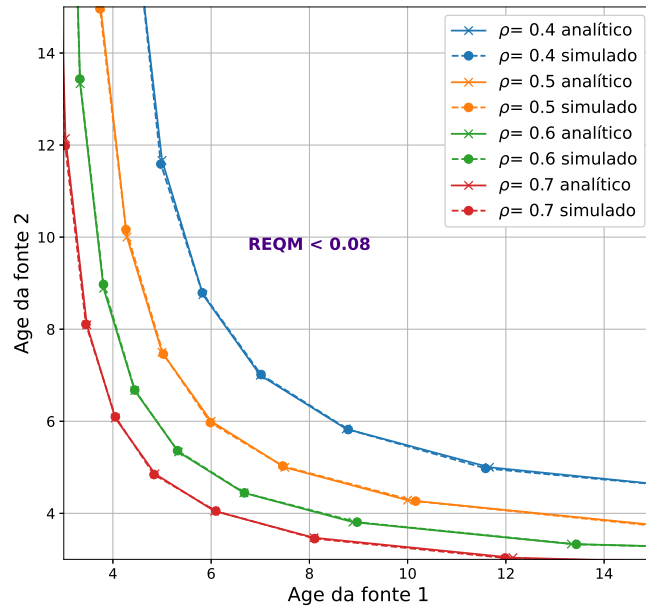


Figura 4.8: Comparação entre os resultados analíticos e simulados para o modelo com duas fontes utilizando o um regime LCFS com precedência em serviço, do tipo  $M/M/1/1^*$ , para diversos valores da carga total ( $\rho$ ) no servidor ( $\mu = 1$ ).

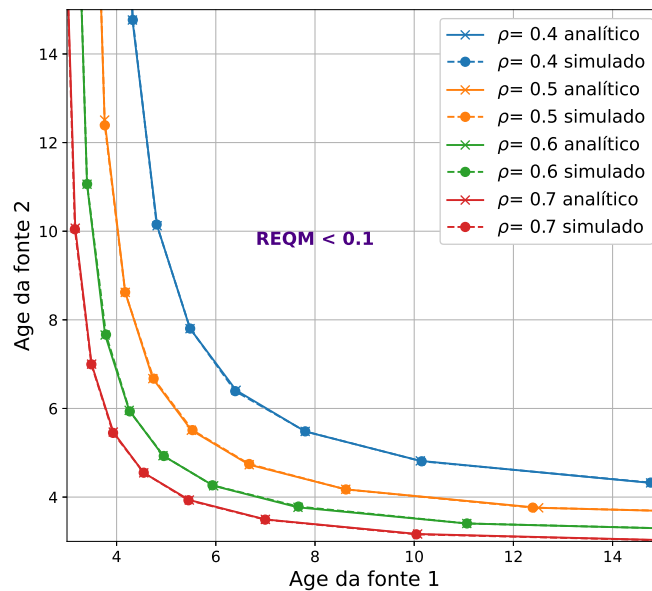


Figura 4.9: Comparação entre os resultados analíticos e simulados para o modelo com duas fontes utilizando o um regime LCFS com precedência na espera, do tipo  $M/M/1/2^*$  para diversos valores da carga total ( $\rho$ ) no servidor ( $\mu = 1$ ).

## 4.2 Resultados experimentais da técnica LGFS-C

Conforme apresentado na Seção 3.3, o modelo teórico definido para a técnica de gerenciamento de pacotes LGFS-C foi validado mediante simulações com a ferramenta descrita na Seção 3.2. Dessa maneira, o sistema da Fig. 3.5 foi simulado para 3 diferentes distribuições da variável aleatória tempo de serviço  $X$ , quais sejam, Exponencial, Weibull e Lognormal. Os parâmetros dessas distribuições, suas funções de densidade de probabilidade e a função de vida média residual (VRM) podem ser visualizadas na Fig. 4.10. Observa-se que as três distribuições utilizadas possuem funções de vida média residual diferentes. A Exponencial possui VRM constante, a Weibull apresenta VRM estritamente decrescente e a Lognormal observa um crescimento da VRM após um certo valor de tempo.

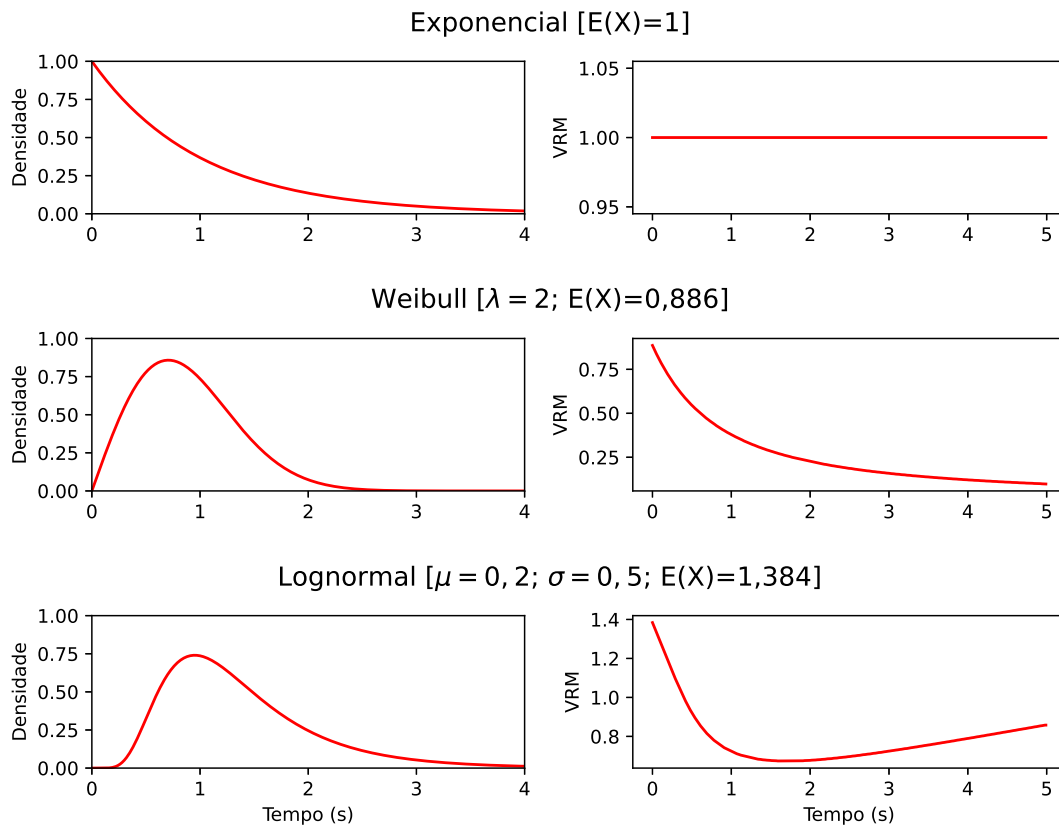


Figura 4.10: Distribuições utilizadas para a avaliação da técnica LGFS-C.

Conforme descrito no Capítulo 2, o atraso da rede foi definido como exponencial com média  $k = 0,5$  segundo, de tal forma que os pacotes podem chegar na fila fora da ordem original de geração. Para o cálculo da AoI média da fonte, medida no monitor, os cálculos são feitos em função da carga no servidor  $\rho = \lambda/\mu$ , em que  $\mu = 1/E(X)$ . A quantidade total de pacotes gerados em cada uma das simulações foi de 500.000 e a raiz do erro quadrático médio (REQM) dos resultados foi inferior a 0,05.

Os resultados para o tempo de serviço exponencial são mostrados na Fig. 4.11, em que a técnica LGFS-C mostra uma AoI média equivalente à técnica LGFS-S, resultado coerente com o apresentado na Seção 3.3.4. A técnica LGFS-W, apresenta um desempenho inferior em relação às demais, o que também é esperado para a distribuição exponencial.

O resultado para os tempos de serviço Weibull é mostrado na Fig. 4.12 e para Lognormal na Fig. 4.13a. Ambos mostram um desempenho superior da técnica LGFS-C para todas as cargas  $\rho$ . Observa-se também que a técnica LGFS-S apresenta uma AoI média maior que as demais técnicas, o que ocasiona uma degradação ainda maior à medida em que se aumenta a carga.

Ainda para a distribuição Lognormal, conforme a Fig. 4.13B, foi realizada uma comparação das quantidades de pacotes descartados entre as técnicas. Nessa análise, a LGFS-S é a que mais descarta pacotes e a LGFS-W é a que possui menos perdas. Já a LGFS-C apresenta valores de descarte de pacote intermediários entre as outras, sendo que, para cargas mais altas, esses níveis são pouco superiores ao nível de descarte da técnica LGFS-W (cerca de 10%).

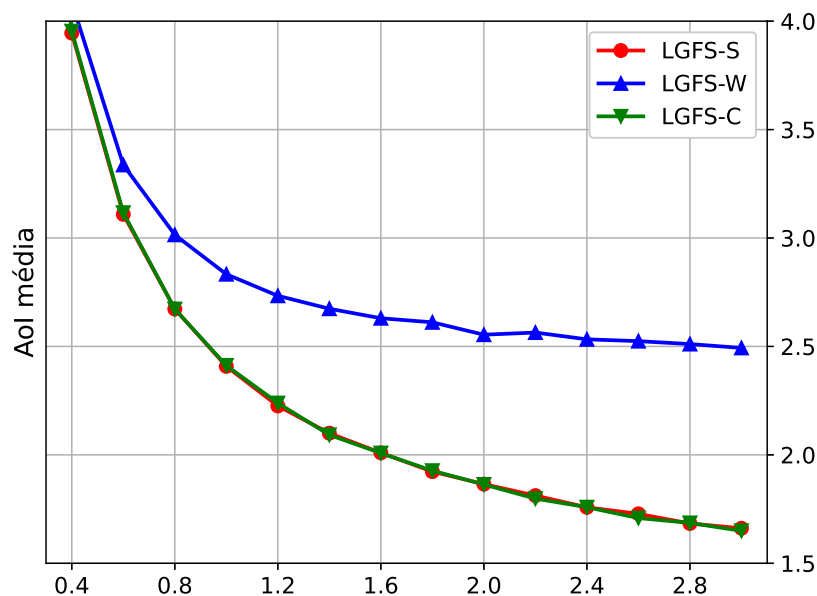


Figura 4.11: Avaliação do cenário crítico.

Os resultados obtidos pelas simulações confirmam a validade do modelo analítico da Seção 3.3.3, e ainda apontam a superioridade (ou equivalência no caso da distribuição Exponencial) da técnica LGFS-C, em termos da menor AoI média, para todas as distribuições de tempo de serviço analisadas. A AoI média atingiu valores até 18% menores para a LGFS-C em relação à LGFS-W, no caso da distribuição Weibull. O aumento da quantidade de pacotes descartados da LGFS-C em relação à LGFS-W apresentou uma

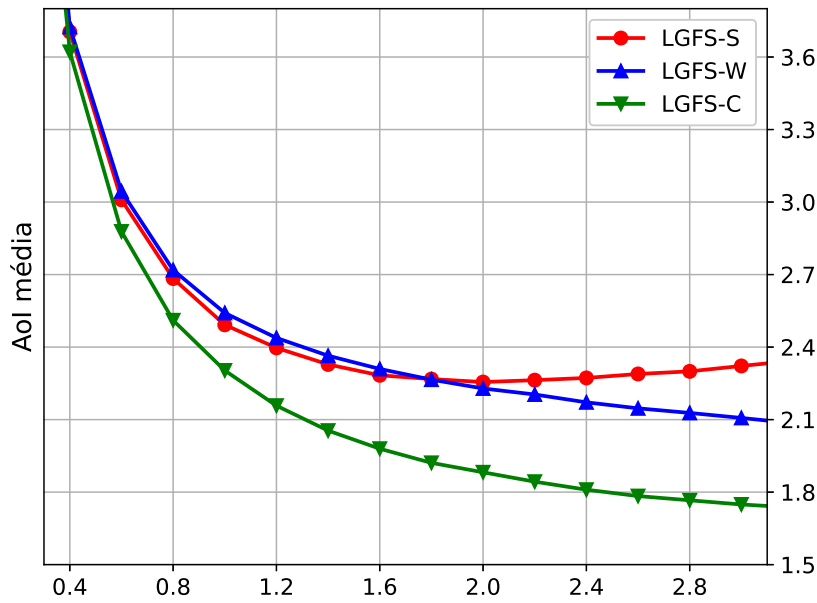


Figura 4.12: Avaliação do cenário crítico.

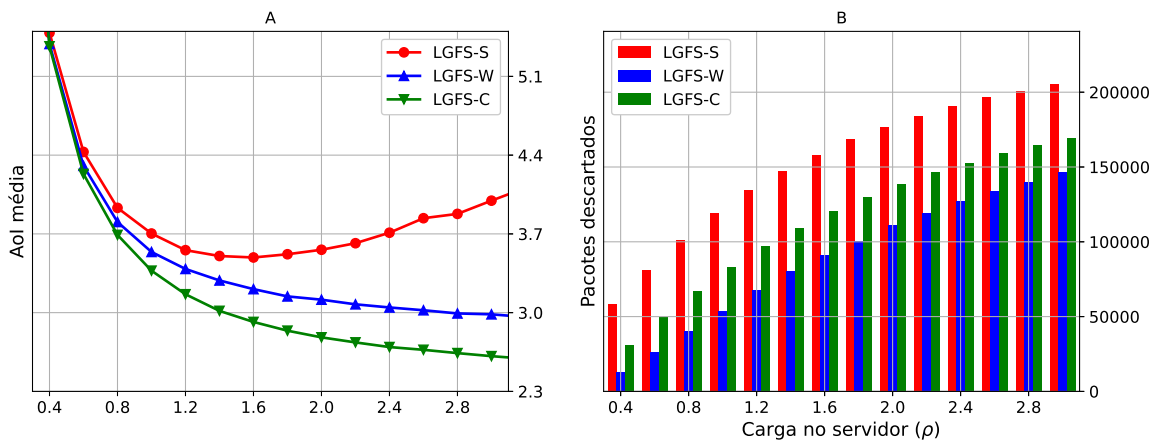


Figura 4.13: Avaliação do cenário crítico.

leve alta e pode ser avaliado em função da quantidade admissível desses descartes em um determinado cenário.

### 4.3 Resultados experimentais da técnica MAF-LGFS-MW

Esta seção apresenta os resultados da avaliação da técnica MAF-LGFS-MW (*Max Age First - Last Generated First Served - Multi preemption in Waiting*), cujos detalhes de im-



plementação foram apresentados na Seção 3.4. Como forma de se verificar o seu desempenho, a referida técnica foi comparada com as seguintes técnicas constantes na literatura:

- MAF-LGFS-S<sup>1</sup>: *Max Age First - Last Generated - First Served - preemption in Service.*
- LGFS-S<sup>1</sup>: *Last Generated - First Served - preemption in Service.*
- MASIF-LGFS (s/ preempção): *Max Age of Served Information First - Last Generated - First Served.*

As técnicas MAF-LGFS-S e MASIF-LGFS (s/ preempção) foram apresentadas nas Seções 2.2.4 e 3.4, conforme propostas em [37]. Já a técnica LGFS-S, utilizada nesta avaliação como referência, trata-se da mesma técnica LGFS-S apresentada na Seção 3.3, porém modificada para múltiplas fontes e múltiplos servidores. Nesse caso, o escalonador do modelo da Fig. 3.7 seleciona sempre o pacote com o maior tempo de geração entre todas as fontes, inclusive os pacotes já em serviço, os quais podem sofrer preempção.

Os parâmetros utilizados no modelo da Fig. 3.7 foram os seguintes:

- Taxa de geração de pacotes  $\lambda$ : definida em função da carga total nos servidores  $\rho$ , com  $\lambda = \frac{\rho C \mu}{N}$ .
- Número de fontes  $N$ : 50.
- Parâmetro do atraso na rede: 0,5 segundo (distribuição exponencial).
- Quantidade de servidores  $C$ : 3.
- Taxa média de serviço  $\mu$  para cada servidor: 5 pacotes/s.
- Distribuição do tempo de serviço para os servidores:
  - Exponencial.
  - Weibull com parâmetro  $\lambda = 2$ .

A métrica de avaliação utilizada foi a AoI média para as 50 fontes, em função da carga total  $\rho = \frac{N\lambda}{C\mu}$  nos servidores. A quantidade de pacotes utilizada foi de 200.000 para cada carga, a qual limitou o erro (REQM) a um valor máximo de 0,2.

A Fig. 4.14 apresenta os resultados para tempos de serviço com distribuição exponencial. Observa-se inicialmente que a técnica LGFS-S mostra uma AoI média mais alta para todas as cargas. Em função da técnica não priorizar a fonte mais antiga, sempre é

---

<sup>1</sup>Os sufixos 'S' não constam no trabalho original de [37], tendo sido acrescentados para indicar preempção em serviço e diferenciar do caso sem preempção.

selecionando o pacote com o maior tempo de geração. A técnica MASIF-LGFS (s/ preempção) mostra resultados semelhantes à técnica MAF-LGFS-MW para cargas inferiores à unidade. No entanto essa técnica não pode ser implementada para uma carga  $\rho > 1$ , sob pena do aumento indefinido da fila devido ao regime de não preempção.

A técnica MAF-LGFS-S mostra resultados piores que a MAF-LGFS-MW até uma carga  $\rho = 2$  e resultados praticamente iguais para cargas  $\rho > 2$ . Ressalta-se que, conforme demonstrado em [37], a MAF-LGFS-S é ótima para o caso de um servidor com regime de serviço exponencial, mas não ótima para sistemas com múltiplos servidores. Por fim, a técnica MAF-LGFS-MW mostra resultados melhores ou iguais para todas as cargas avaliadas.

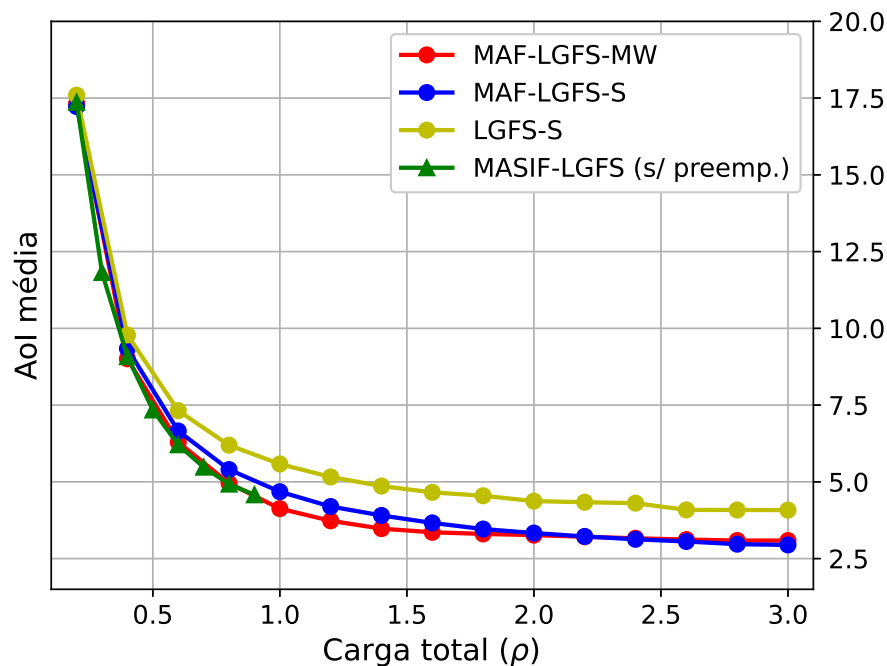


Figura 4.14: Comparativo das técnicas para tempos de serviço exponenciais.

Na Fig. 4.15 são mostrados os resultados comparativos entre as técnicas para servidores com distribuição Weibull em relação ao tempo de serviço. Novamente a técnica LGFS-S mostra uma AoI mais alta para todas as cargas analisadas, com tendência de alta para cargas  $\rho > 2,5$ . A técnica MASIF-LGFS (s/ preempção) apresenta novamente resultados iguais ao da técnica MAF-LGFS-MW, os quais ficam limitados a cargas  $\rho < 1$  pelos mesmos motivos apresentados na análise da Fig. 4.14.

O diferencial na análise dos resultados para o caso de servidores com distribuição Weibull foi o desempenho inferior da técnica MAF-LGFS-S em relação à técnica MAF-LGFS-MW para todas as cargas analisadas. Esse resultado foi previsto na Seção 3.4,

quando se abordou a principal propriedade das distribuições NBU: nem sempre compensa realizar a preempção de um pacote com tempo já decorrido de serviço em troca de um pacote que ainda não iniciou o seu serviço. Essa propriedade foi bastante utilizada na demonstração da técnica LGFS-C, na Seção 3.3, através da utilização do conceito de Vida Residual Média (VRM). No caso da técnica MAF-LGFS-MW, não foi possível definir um critério de decisão para a escolha de se realizar a preempção em serviço ou nas filas de espera, optando-se sempre por esta última. Os resultados experimentais mostram que essa escolha faz a MAF-LGFS-MW atingir um desempenho superior a todas as demais técnicas, para o caso da distribuição Weibull.

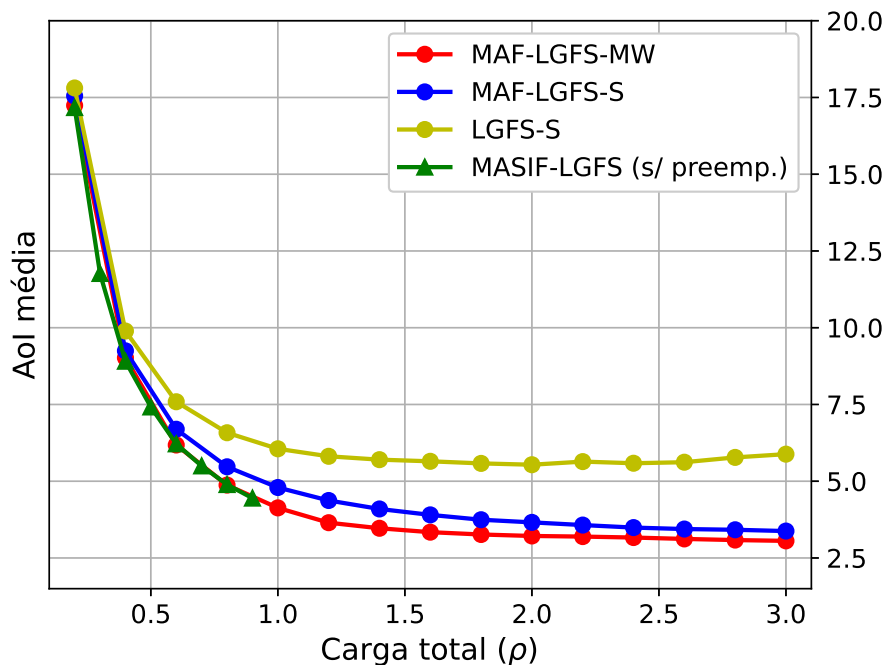


Figura 4.15: Comparativo das técnicas para tempos de serviço com distribuição de Weibull.

## 4.4 Avaliação dos Cenários de Uso de Sistemas de Monitoramento e Decisão

Esta seção tem por finalidade definir uma metodologia para avaliação de diferentes cenários de uso nos sistemas de monitoramento e decisão voltados para a atividade de C2 em operações militares. O principal critério que diferencia os possíveis cenários é a criticidade em termos da atualização das informações dos agentes no monitor. Nesse sentido, as subseções seguintes descrevem o enfoque adotado na avaliação para estabelecer a re-

lação entre carga de trabalho, infraestrutura e métricas nos diferentes tipos de cenários nesses sistemas. São definidos nesta seção, três tipos de cenários a serem considerados na avaliação.

#### 4.4.1 Metodologia de Avaliação

Um dos principais problemas enfrentados no processo de definição da arquitetura de um sistema de monitoramento e decisão, conforme discutido na Seção 3.1.3, é o dimensionamento da carga de trabalho e da infraestrutura em função de métricas de desempenho preestabelecidas. Dessa maneira, é possível abordar esse problema de duas formas, conforme a Fig. 4.16, as quais fazem um caminho inverso entre si:

1. Dada uma carga de trabalho, representada pela quantidade de agentes e a frequência de envios de pacotes, deve-se definir a infraestrutura necessária no monitor (capacidade de processamento) para que se atinja o desempenho mínimo em função das métricas estabelecidas.
2. Dada uma determinada capacidade de processamento no monitor, em função da quantidade de servidores e suas capacidades de processamento, deve-se definir o número máximo de agentes em campo suportados pelo sistema, bem como a frequência de envio de pacotes de atualização.

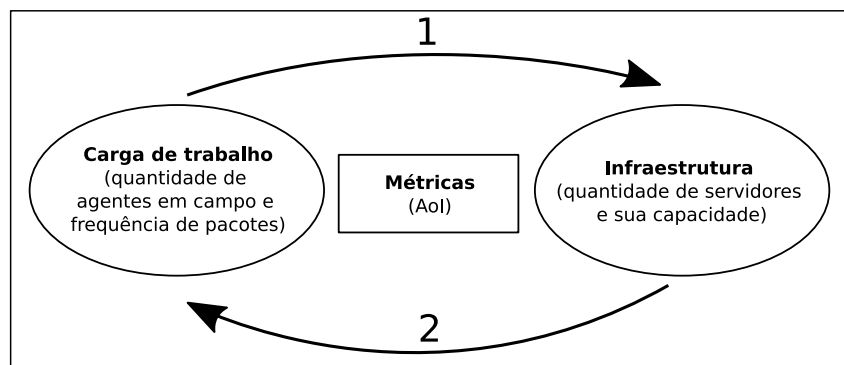


Figura 4.16: Duas abordagens distintas para o dimensionamento do sistema de monitoramento e decisão.

A abordagem (1) pode ser utilizada em um cenário com pouca ou com nenhuma restrição de recursos. Dessa forma, define-se a quantidade de agentes que uma determinada operação requer em campo, em que posteriormente é feito o ajuste da infraestrutura para atender aos requisitos de desempenho. Já a abordagem (2) leva em conta a restrição de infraestrutura, a qual possui capacidades limitadas, o que é representado em termos da quantidade de servidores no monitor bem como a capacidade de processamento de

pacotes. Nessa última abordagem, a quantidade de agentes em campo pode variar para atender aos limites de desempenho.

Entre as duas abordagens apresentadas, a número 2 é a que mais se aproxima de sistemas do mundo real, uma vez que a infraestrutura é sempre finita. Por esse motivo, neste trabalho a avaliação de cenários usa a premissa de que a capacidade de processamento no monitor possui um limite máximo, o qual não poderá ser ultrapassado. Como consequência, o objetivo da análise será calcular a quantidade máxima de agentes para um determinado cenário, em função das métricas de interesse derivadas da AoI.

O modelo de interesse foi definido na Seção 3.1.3 para todos os cenários a serem avaliados, os quais se diferenciam pelo grau de criticidade. Na simulação, o modelo é representado através de um grafo direcionado, cujos nós representam filas. Essa rede de filas é instanciada na ferramenta de simulação. A representação do grafo do modelo da Fig. 3.3 é ilustrada na Fig. 4.17.

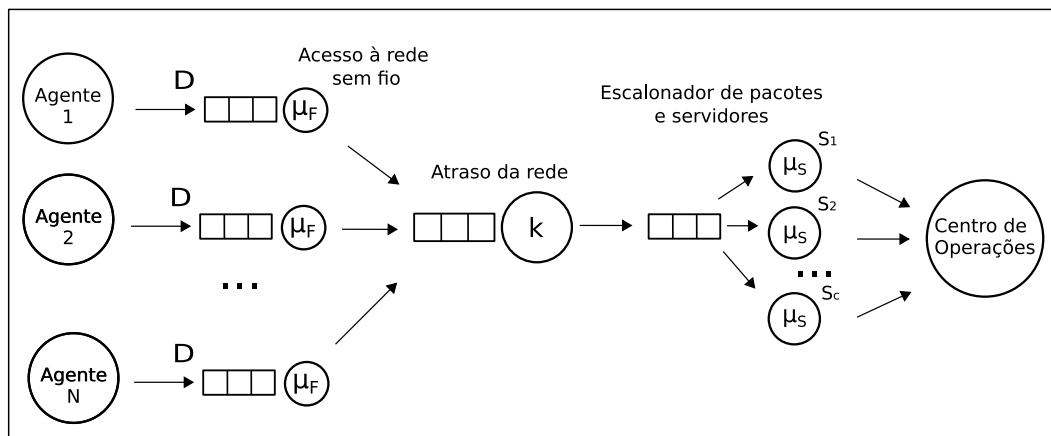


Figura 4.17: Representação do modelo do sistema em estudo através de uma rede de filas.

No sistema, são definidos os seguintes parâmetros :

- Quantidade de servidores  $C$  no monitor.
- Distribuição do tempo de serviço e taxa média  $\mu_S$ , para cada servidor no monitor.
- Valor médio do atraso na rede  $k$ , modelado por uma distribuição exponencial.
- Distribuição dos tempos de serviço e taxa média  $\mu_F$  para os servidores nas fontes, a qual engloba também o tempo de serviço ao ponto de acesso sem fio.
- Técnica de gerenciamento de pacotes nos servidores dos agentes.
- Técnica de escalonamento e gerenciamento de pacotes no monitor.

A taxa média de serviço  $\mu_F$  para as fontes é definida em função do número de fontes  $N$ , de modo que a taxa diminua exponencialmente na medida em que se aumenta o

número total de fontes. Esse modelo tem por objetivo simular a capacidade limitada de conexões simultâneas do ponto de acesso sem fio, para evitar a definição de técnicas de escalonamento para esse componente e assim diminuir a complexidade na simulação do sistema.

#### 4.4.2 Definição dos Cenários

O critério que diferencia os cenários é a sua criticidade, a qual estabelece os valores requeridos das métricas de desempenho. As métricas a serem utilizadas são a *Age of Information* (AoI) média máxima e a *Age of Information* de pico (pAoI) média máxima observadas pelo monitor em relação aos  $N$  agentes, conforme definido pelas Eq. 2.29 e 2.32.

Em termos práticos, a AoI média fornece ao monitor uma estimativa da atualidade média da informação para todos os agentes monitorados. Quanto mais crítico um cenário, menor será a AoI média máxima requerida, pois maior é o risco associado a essa desatualização do monitor. Já a pAoI fornece informações a respeito da variância da AoI, o que requer o valor médio dos picos alcançados por esse processo. É importante notar que os requisitos de desempenho de cada cenário podem ser definidos de forma variável, em função da experiência de campo ou situação a ser monitorada. Nesse sentido, com o objetivo de mostrar a aplicação das técnicas desenvolvidas neste trabalho, foi utilizada a experiência de campo na definição desses requerimentos. Entretanto, é importante ressaltar que a ferramenta de simulação oferece flexibilidade para definir esses requisitos conforme necessário e que suporta uma variedade de possíveis cenários.

Com foco nos sistemas de C2, este trabalho define três cenários, os quais são apresentados a seguir.

##### **Cenário 1: Crítico**

Este cenário se caracteriza dentro do contexto militar em operações que requerem uma atualização quase imediata das posições georreferenciadas dos agentes vistas pelo Centro de Operações. Como exemplo, pode-se citar a necessidade de apoio de fogo em uma determinada localidade onde existe a possibilidade de tropas amigas. Para minimizar o risco de fratricídio, é fundamental que o COp visualize a posição dos agentes de forma mais atualizada o possível. Com base na experiência de campo, as métricas requeridas de desempenho para esse cenário são definidas na Tabela 4.1.

Tabela 4.1: Métricas de desempenho para o cenário crítico.

Métrica	Descrição	Valor
$AoI_{lim}$	Limite máximo para a AoI média do conjunto de agentes	3 segundos
$pAoI_{lim}$	Limite máximo para a AoI de pico média do conjunto de agentes	5 segundos

### Cenário 2: Pouco crítico

O cenário pouco crítico pode ser exemplificado através de uma operação militar em que a atualização das posições dos agentes possui pouca criticidade, como uma operação de ataque a uma localidade, por exemplo. Nessa situação, o COp deseja saber a posição dos agentes com certa atualidade, para fins de acompanhamento e tomada de decisões em relação ao desenrolar das ações. A desatualização do Centro de Operações, no entanto, não oferece riscos imediatos para a tropa, apesar de influenciar no comando e controle da operação. A Tabela 4.2 apresenta os requerimentos de desempenho definidos para o cenário 2.

Tabela 4.2: Métricas de desempenho para o cenário pouco crítico.

Métrica	Descrição	Valor
$AoI_{lim}$	Limite máximo para a AoI média do conjunto de agentes	6 segundos
$pAoI_{lim}$	Limite máximo para a AoI de pico média do conjunto de agentes	10 segundos

### Cenário 3: Normalidade

Neste último cenário, a atualidade da posição dos agentes é considerada não crítica para o sucesso da missão. Um exemplo desse cenário seria uma operação de reconhecimento, em que a desatualização dos agentes não oferece riscos imediatos nem influencia diretamente no desenrolar das ações. Os valores requeridos de desempenho estão definidos na Tabela 4.3.

## 4.5 Resultados da Avaliação dos Cenários de Uso

As simulações com a ferramenta computacional descrita na Seção 3.2 tem como objetivo definir a carga de trabalho suportada pelo sistema para atender os requisitos mínimos de

Tabela 4.3: Métricas de desempenho para o cenário de normalidade.

Métrica	Descrição	Valor
$AoI_{lim}$	Limite máximo para a AoI média do conjunto de agentes	15 segundos
$pAoI_{lim}$	Limite máximo para a AoI de pico média do conjunto de agentes	20 segundos

desempenho estabelecidos através das métricas da AoI. Em termos práticos, essa carga de trabalho é definida pela quantidade de agentes  $N$  e pela taxa de envio dos pacotes com as suas posições georreferenciadas  $D$ , conforme ilustrado na Fig. 4.17. Essas informações são fundamentais para o correto dimensionamento do sistema, além de garantir que o monitor atinja os níveis mínimos de atualidade em relação aos agentes observados.

Nesta seção são apresentados os resultados da avaliação dos cenários definidos na seção anterior. O foco dos resultados é fornecer subsídios para a definição de arquitetura e o dimensionamento da infraestrutura do sistema de monitoramento e decisão. Em cada um dos cenários, a avaliação deverá calcular a carga de trabalho (agentes em campo) suportada com base nos parâmetros estabelecidos e nos requisitos definidos em termos da AoI e da pAoI. Espera-se que o emprego das técnicas propostas nos diversos componentes do sistema possam contribuir nos tópicos de SCF e *Age of Information*.

A definição dos parâmetros nos cenários de simulação foi arbitrada conforme semelhança com valores de sistemas reais, entre eles o Sistema Pacificador utilizado pelo Exército Brasileiro. Os valores dos parâmetros estão descritos na Tabela 4.4.

Tabela 4.4: Parâmetros comuns para a avaliação dos cenários.

Parâmetro	Descrição	Valor
$C$	Número de servidores	3
$\mu_S$	Taxa de serviço média dos servidores do monitor	5 pacotes/s
$k$	Tempo de atraso médio da rede	0.5 segundo
$\mu_F$	Taxa de serviço média dos servidores das fontes	$5 * 0.9^{\frac{N}{4}}$ pacotes/s
Técnica de gerenciamento de pacotes nos servidores das fontes	-	LGFS-C
Técnica de gerenciamento de pacotes no escalonador e servidores do monitor	-	MAF-LGFS-MW



As técnicas de gerenciamento de pacotes utilizadas nas fontes e no monitor podem ser visualizadas na Fig. 4.18. Cada um dos servidores das fontes (agentes) utiliza a técnica LGFS-C, a qual foi apresentada na Seção 3.3 como uma das contribuições deste trabalho. Essa técnica foi selecionada pelo fato de produzir resultados melhores ou iguais aos das técnicas no estado da arte para um modelo de uma fonte e um servidor, exatamente o que ocorre para cada um dos agentes. Já para a estrutura do escalonador e servidores do monitor, na porção destacada em azul, a técnica utilizada foi a MAF-LGFS-MW, a qual foi apresentada na Seção 3.4 e consiste em outra contribuição deste estudo. Essa técnica também apresenta resultados melhores ou equivalentes às demais técnicas no estado da arte para modelos com múltiplas fontes e múltiplos servidores.

Ainda sobre a Fig. 4.18, ressalta-se que a técnica LGFS-C é empregada de maneira independente em cada um dos servidores dos agentes, sendo que a sua execução é função apenas dos estados do próprio servidor. No caso da técnica MAF-LGFS-MW, o escalonamento e as demais tarefas executadas são funções dos estados do sistema como um todo, dependendo de fatores como a AoI atual das  $N$  fontes, o estado das filas dos servidores e a ocupação atual dos servidores.

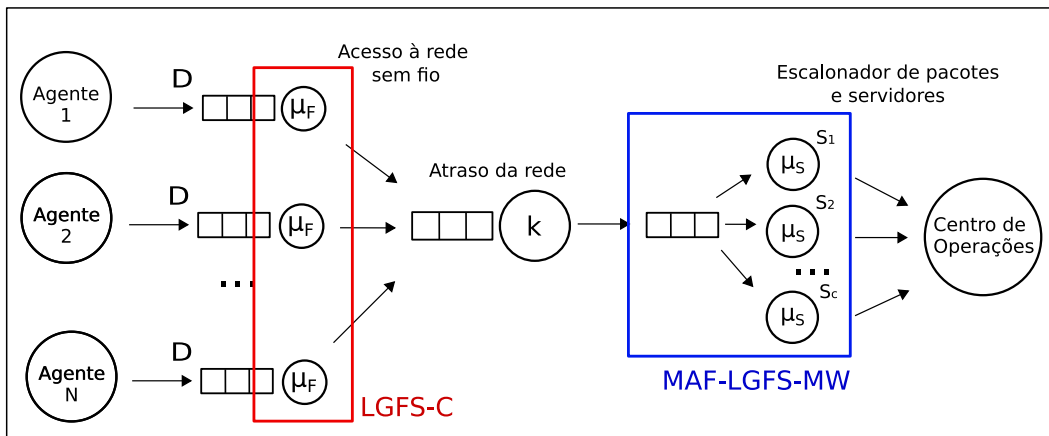


Figura 4.18: Técnicas utilizadas para simulação dos cenários: LGFS-C para os servidores dos agentes e MAF-LGFS-MW para o escalonador e servidores do monitor.

A taxa média de serviço  $\mu_F$  para cada fonte é função do número total de fontes  $N$ , conforme a Tabela 4.4. Essa função é ilustrada na Fig. 4.19. Esse serviço compreende, além da preparação e codificação do pacote, o tempo necessário para a transmissão do pacote até o ponto de acesso sem fio. Ressalta-se que essa função foi definida com base em um ponto de acesso com capacidade de servir até 100 conexões mantendo uma performance mínima. A partir desse valor, essa taxa de serviço pode apresentar uma forte degradação em seu valor, devido à sobrecarga desse componente.

A taxa de envio definida para cada agente foi calculada de modo a se obter uma carga total nos servidores igual a 5. Dessa maneira, o cálculo para a taxa de envio  $D$  de cada

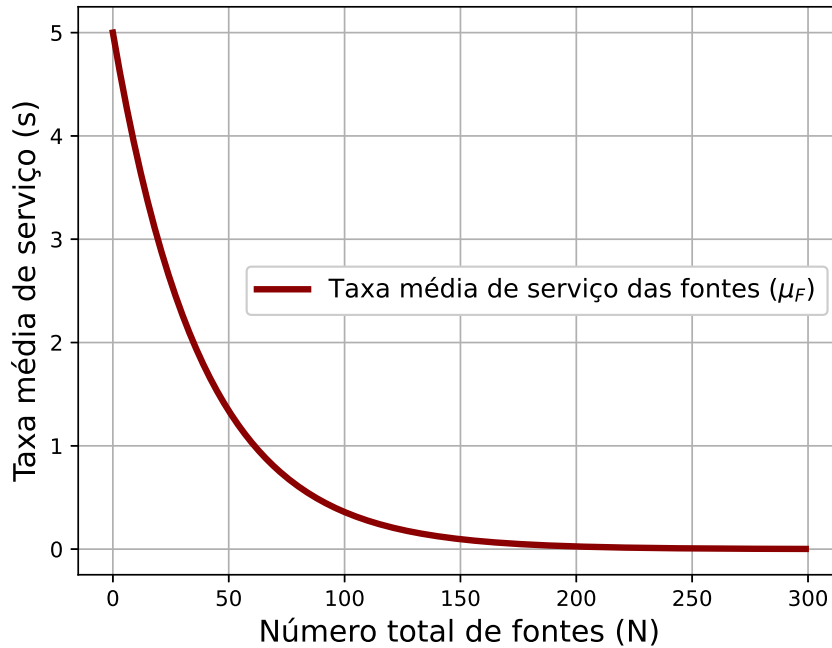


Figura 4.19: Ilustração da taxa média de serviço de uma fonte em função do número total de fontes do sistema.

agente (fonte) pode ser obtido pela Eq. 4.2, em que  $\mu_S$  é a taxa de serviço de cada servidor individualmente,  $C$  é a quantidade de servidores e  $N$  a quantidade de agentes.

$$D = \frac{5\mu_S C}{N} \quad (4.2)$$

Uma vez definidos os parâmetros e as métricas, foram realizadas simulações em que a quantidade de agentes  $N$  foi variada entre 30 e 150, obtendo-se para cada valor de  $N$  a AoI média e a pAoI média dos agentes vista pelo monitor. Dessa maneira, essas duas métricas resultam em duas funções,  $AoI(N)$  e  $pAoI(N)$  da quantidade de agentes alocados, conforme mostrado na Fig. 4.20.

Uma análise mais detalhada da Fig. 4.20 revela alguns detalhes importantes. O valor da AoI média para um determinado número de agentes representa a defasagem média com que o Centro de Operações (COp) visualiza a posição dos agentes. Como exemplo, para uma quantidade de 70 agentes em campo, o COp observará a posição dos agentes com uma defasagem média de aproximadamente 5 segundos. A avaliação quanto à grandeza desse valor dependerá do cenário considerado: uma viatura que se move a uma velocidade média de 50 km/h terá se deslocado em torno de 70 metros, o suficiente para se aproximar de uma área na qual foi solicitado um apoio de fogos de artilharia, situação que pode propiciar o chamado fratricídio.

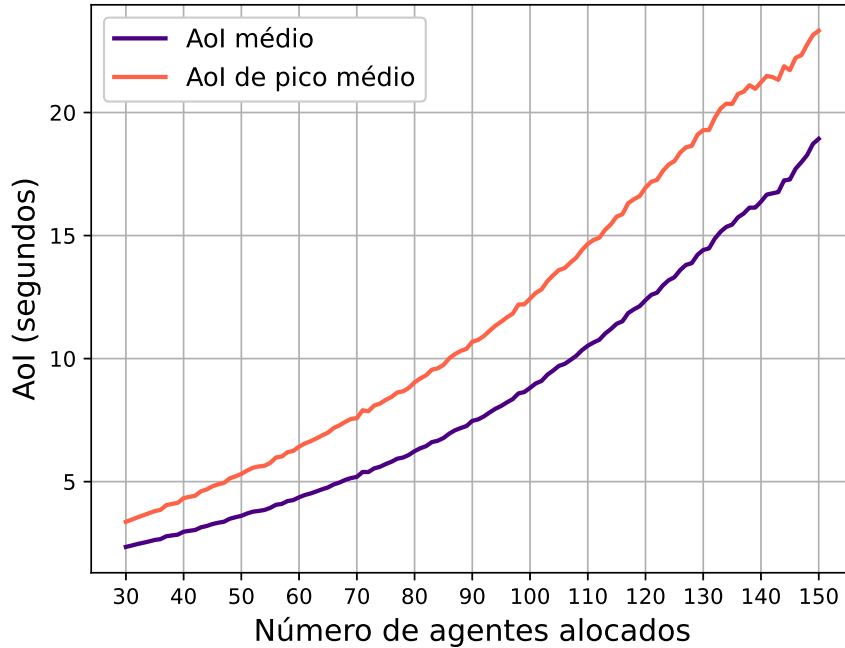


Figura 4.20: AoI média e pAoI média vistas pelo monitor em relação aos  $N$  de agentes alocados.

A pAoI média, por sua vez, fornece a média dos valores de pico da AoI, ou seja, a média dos valores de pico alcançados para a defasagem vista pelo COp. Tomando-se novamente o exemplo anterior, no qual estão alocados 70 agentes, o COp observará, em média, um valor de pico de aproximadamente de 7,5 segundos de defasagem para a posição dos agentes. Para a mesma viatura que trafega a 50 km/h, a mesma terá percorrido pouco mais de 100 metros. Ressalta-se que esse valor é a média dos valores de pico e não o valor máximo atingido pela AoI, o qual será obviamente maior que 7,5 segundos.

O dimensionamento da carga útil máxima suportada pelo sistema em cada um dos cenários, ou seja, o número máximo de agentes que podem ser alocados respeitando-se os limites máximos das métricas estabelecidas, pode ser obtida através da Eq. 4.3. Nessa equação,  $N_{AoI}$  é o número máximo de agentes alocados com base no limite máximo da AoI média ( $AoI_{lim}$ ) e  $N_{pAoI}$  o número máximo de agentes alocados com base no limite máximo da pAoI média ( $pAoI_{lim}$ ), conforme mostra a Eq. 4.4.

$$N = \min\{N_{AoI}, N_{pAoI}\} \quad (4.3)$$

$$\begin{aligned}
N_{AoI} &= \max(N) \mid AoI(N) \leq AoI_{lim} \\
N_{pAoI} &= \max(N) \mid pAoI(N) \leq pAoI_{lim}
\end{aligned}
\tag{4.4}$$

Na sequência, serão apresentados os resultados das avaliações realizadas em cada um dos cenários.

### Cenário crítico

Para o cenário crítico, o limite máximo para a AoI média do conjunto de agentes foi definido em 3 segundos e o limite para a pAoI (pico) foi definido em 5 segundos. A Fig. 4.21 mostra esses limites traçados sobre as curvas de AoI pelo número de fontes  $N$ . Observa-se na figura que o requisito para a AoI média limita o sistema de monitoramento e decisão a um máximo de 40 fontes, apesar de ainda haver certa folga para a pAoI média.

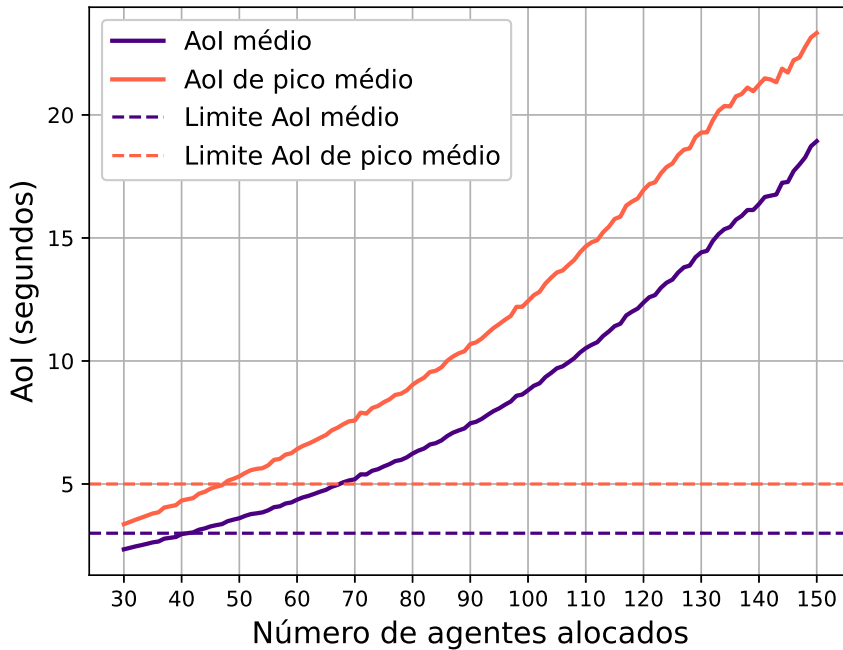


Figura 4.21: Avaliação do cenário crítico.

Para essa quantidade máxima de agentes (40), a taxa de envio de cada fonte, calculada pela Eq. 4.2, é de 1,88 pacotes por segundo, o que equivale ao envio de 1 pacote a cada 0,53 segundo. Conforme apresentado no Capítulo 3, este cenário é o mais crítico de todos, sendo que a integridade física dos agentes depende da manutenção dos valores da AoI dentro dos limites estabelecidos.

### Cenário pouco crítico

Para este cenário os limites de desempenho são de 6 e 10 segundos, respectivamente para a AoI e para a pAoI médias. Conforme se observa na Fig. 4.22 esse requisito limita o sistema a 78 fontes. Assim como no cenário crítico, o número de fontes foi limitado pela AoI média, podendo a pAoI suportar ainda alguns agentes a mais.

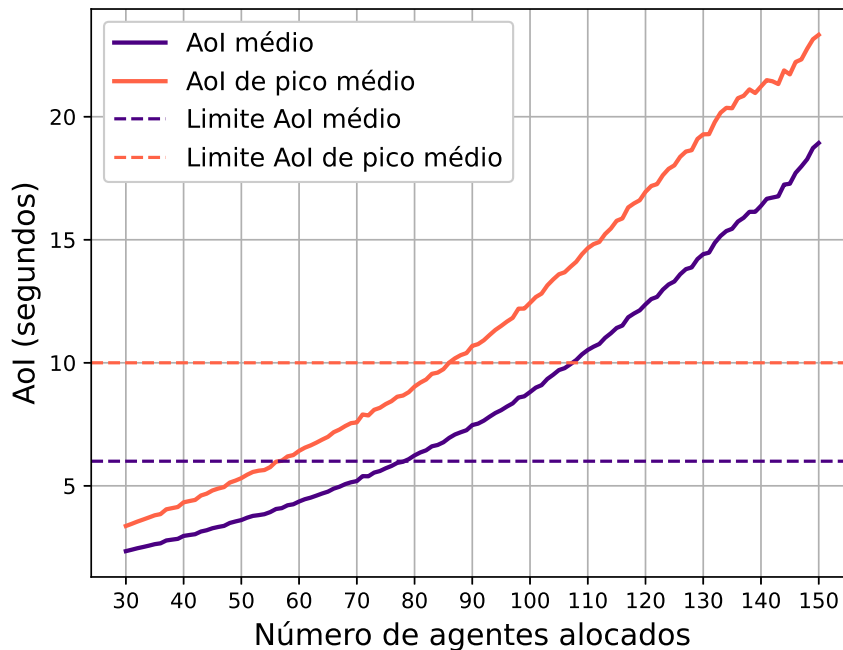


Figura 4.22: Avaliação do cenário pouco crítico.

Neste cenário, a taxa de envio de cada fonte foi calculada em 0,96 pacote por segundo, o que equivale ao envio de 1 pacote a cada 1,04 segundo. Neste cenário, apesar de não causar riscos à integridade física dos agentes, ultrapassar os limites máximos da *Age of Information* pode, por exemplo, prejudicar o desenrolar de uma operação, em que o COp possui a necessidade de conhecer a posição dos agentes para que ordens possam ser emitidas. Observa-se que este cenário, apesar de estabelecer apenas 3 segundos a mais para o limite máximo da AoI média em relação ao cenário crítico, praticamente dobrou a quantidade de agentes suportada em campo. Esse fato ressalta a importância da metodologia definida neste trabalho, a qual possibilita o correto dimensionamento da infraestrutura do sistema de acordo com a demanda do cenário, evitando o excesso (superdimensionamento) ou a falta de infraestrutura (subdimensionamento) necessária para atender aos requisitos da AoI.

## Cenário de normalidade

Neste cenário, as métricas limites de desempenho foram definidas em 15 segundos para a AoI média e 20 segundos para a pAoI média. Sendo assim, o número máximo de fontes foi limitado por ambas as métricas (AoI média e pAoI média), sendo possível alocar até 132 agentes no sistema. Ainda neste cenário, a taxa de envio para cada fonte foi calculada em 0,56 pacote por segundo, o que equivale ao envio de 1 pacote a cada 1,76 segundo. Conforme apresentado na Seção 4.4.2, os valores da AoI para este cenário não trazem grandes riscos para o COP, priorizando a alocação de uma quantidade maior de agentes.

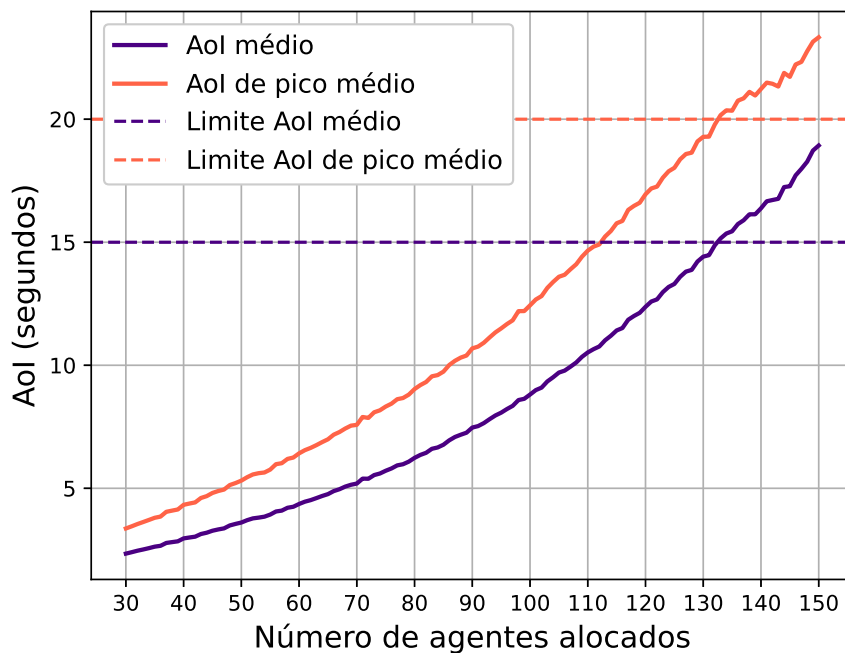


Figura 4.23: Avaliação do cenário de normalidade.

Apesar dos requisitos mais relaxados deste cenário, julga-se importante manter os valores médios da *Age of Information* dentro de um limite máximo, pois devido à dinâmica das operações militares um cenário de normalidade pode tornar-se pouco crítico ou até mesmo crítico em um curto intervalo de tempo. Essa transição entre cenários, apesar de não ser abordada em profundidade neste estudo, pode ser resolvida através de uma escalabilidade dinâmica da infraestrutura dos servidores do monitor ou ainda através da priorização dos agentes.

## 4.6 Resumo do Capítulo

Este capítulo apresentou os resultados experimentais da pesquisa em AoI, obtidos mediante simulação com uma ferramenta desenvolvida para esse propósito. Os resultados foram divididos em quatro partes, a primeira para validação da ferramenta de simulação, a segunda e terceira parte para avaliação das novas técnicas LGFS-C e MAF-LGFS-MW, propostas neste trabalho, e a quarta parte para avaliação de três cenários de uso de um SCF caracterizado como um sistema de monitoramento e decisão. Os resultados demonstram a aplicabilidade da ferramenta de simulação e a adequação das técnicas desenvolvidas para planejamento e otimização do número de fontes em SCF.

# Capítulo 5

## Conclusões e Trabalhos Futuros

A *Age of Information* surgiu nos últimos anos como um novo conceito na área de redes e pode ser aplicada como uma métrica de negócio nas atividades de dimensionamento e avaliação de sistemas ciberfísicos, como por exemplo o sistema de monitoramento e decisão em um contexto militar discutido durante este estudo. Dessa forma, o presente trabalho propôs um modelo computacional para que fosse possível avaliar sistemas complexos sob a ótica da AoI por meio de simulações. Foi definido um modelo computacional como a base para a ferramenta de simulação, o qual é uma das contribuições deste trabalho. A ferramenta está disponível para a comunidade acadêmica em um repositório público.

Também foram propostas duas técnicas de gerenciamento de pacotes, a LGFS-C e MAF-LGFS-MW, ambas com o objetivo de otimizar os valores da *AoI*. A técnica LGFS-C é voltada a modelos de filas simples e a MAF-LGFS-MW trabalha com modelos com múltiplas fontes e múltiplos servidores. A avaliação das duas técnicas mostrou resultados iguais ou melhores às técnicas identificadas no estado da arte. A avaliação das técnicas propostas foi feita com diferentes distribuições de tempo de serviço e diferentes cargas de trabalho. No trabalho também foi desenvolvida uma metodologia para aplicação das técnicas propostas no dimensionamento de sistemas ciberfísicos em três cenários com diferentes graus de criticidade, com base na definição de requerimentos de AoI. Essa metodologia possibilita a definição da arquitetura e o dimensionamento da infraestrutura de forma eficiente, ao definir níveis de desempenho mínimos a serem alcançados em função de parâmetros do sistema e do grau de criticidade do cenário avaliado.

Os resultados deste trabalho indicam que a AoI aplicada aos sistemas ciberfísicos pode trazer benefícios reais no dimensionamento e otimização. A AoI, por ser uma métrica mais voltada para o negócio e menos para as métricas de rede convencionais, sintetiza com maior propriedade os padrões de desempenho desejados para um determinado sistema baseado em atualizações periódicas. Em particular, os sistemas de Comando e Controle (C2), vetores de extrema importância no cenário militar moderno, podem ser beneficiados



diretamente conforme mostrado no decurso deste trabalho.

Apesar de promissora, a aplicação dos conceitos e métricas da AoI em sistemas ciberfísicos ainda enfrenta algumas limitações. Uma delas é o fato de que a teoria base da *Age of Information* considera os pacotes de atualização enviados da camada de aplicação, o que abstrai as demais estruturas de dados existentes nas camadas inferiores, como transporte e rede. Essa abstração pode limitar a avaliação da influência dessas outras camadas nas métricas derivadas da AoI. Outra limitação ocorre devido ao aumento da complexidade dos modelos e técnicas à medida em que se acrescenta mais componentes ao sistema. Os modelos analíticos da literatura são limitados a modelos simples e a tarefa de avaliação em sistemas mais complexos pode ser limitada apenas aos resultados experimentais de simulações.

Como trabalhos futuros, vislumbra-se o aperfeiçoamento das duas técnicas propostas neste estudo, as quais, apesar de produzirem resultados melhores que as técnicas no estado da arte, ainda não são consideradas ótimas. Em especial para a técnica LGFS-C, sugere-se a utilização de estimadores não paramétricos para a Vida Residual Média (VRM), o que pode tornar a técnica mais compatível com sistemas reais. Outra possibilidade de trabalho futuro é a melhoria da ferramenta de simulação, a qual exige habilidades técnicas do usuário na sua utilização. Uma interface mais amigável, com apresentação das simulações de maneira gráfica, pode potencializar o interesse de mais pesquisadores da área.

Além das técnicas e da ferramenta de simulação, considera-se ainda essencial a realização de mais estudos em modelos de estruturas e filas em cascata, nos quais pode-se utilizar diferentes técnicas em diferentes componentes, como ocorre no modelo do sistema de monitoramento e decisão apresentado neste estudo. A metodologia para a otimização de tais modelos ainda é uma questão em aberto na literatura, considerada de vital importância para a avaliação de sistemas complexos. Outra questão ainda pouco discutida no estudo da AoI são as técnicas de gerenciamento de pacotes para modelos de múltiplas fontes e múltiplos servidores. Existem poucas técnicas na literatura que abordam tais modelos, motivo pelo qual neste trabalho foi feita a proposta de mais uma. Ressalta-se que esse tipo de modelo com múltiplos componentes é o que mais se aproxima de sistemas do mundo real.

Por fim, acredita-se que o presente trabalho contribuiu para desenvolvimento da pesquisa em sistemas ciberfísicos e *Age of Information*. Na área militar, em especial nos sistemas e estruturas de comando e controle, pode-se utilizar diretamente os conceitos e técnicas aqui propostos, os quais podem colaborar com a melhoria dos atuais processos de avaliação e dimensionamento desses sistemas.

# Referências

- [1] Brasil: *Comando e controle - Manual de campanha*. 2015. 1, 31, 33
- [2] Arcot, Rajabhadur V.: *Cyber-physical systems: The core of industry 4.0*. <https://blog.isa.org/cyber-physical-systems-the-core-of-industry-4.0>, 2021. 1
- [3] Bhattacharyya, Shuvra S. e Marilyn C. Wolf: *Research challenges for heterogeneous cyberphysical system design*. *Computer*, 53(7):71–75, 2020. 1, 6, 7
- [4] Kaul, Sanjit, Roy Yates e Marco Gruteser: *Real-time status: How often should one update?* Em *2012 Proceedings IEEE INFOCOM*, páginas 2731–2735, 2012. 2, 12, 13, 14
- [5] Khaitan, Siddhartha Kumar e James D. McCalley: *Design techniques and applications of cyberphysical systems: A survey*. *IEEE Systems Journal*, 9(2):350–365, 2015. 5
- [6] Esterle, L. e R. Grosu: *Cyber-physical systems: challenge of the 21st century*. *Elektrotech Inftech*, (133):299–303, 2016. 6, 7
- [7] Tanwar, Sudeep, Joel Rodrigues, Sudhanshu Tyagi, Ishan Budhiraja e Neeraj Kumar: *Tactile internet for smart communities in 5g: An insight for noma-based solutions*, janeiro 2019. 6
- [8] União Europeia: *Projects about cyber physical system*, 2021. Disponível em: <<https://digital-strategy.ec.europa.eu/en/news/sensors-cyber-physical-systems-and-internet-things-revolutionise-footwear-industry-portugal>>. Acesso em: 9 de julho de 2021. 7
- [9] Zhang, M., B. Selic, S. Ali, T. Yue, O. Okariz e R. Norgren: *Understanding Uncertainty in Cyber-Physical Systems: A Conceptual Model*. Número 9764. 2016. 7
- [10] Kaul, Sanjit, Marco Gruteser, Vinuth Rai e John Kenney: *Minimizing age of information in vehicular networks*. Em *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, páginas 350–358, 2011. 9
- [11] Costa, Maice, Yalin E. Sagduyu, Tugba Erpek e Muriel Médard: *Robust improvement of the age of information by adaptive packet coding*. Em *ICC 2021 - IEEE International Conference on Communications*, páginas 1–6, 2021. 9

- [12] Badia, Leonardo: *Age of information from two strategic sources analyzed via game theory*. Em *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, páginas 1–6, 2021. 9
- [13] Inoue, Yoshiaki e Tomotaka Kimura: *Age-effective information updating over intermittently connected manets*. *IEEE Journal on Selected Areas in Communications*, 39(5):1293–1308, 2021. 9
- [14] Turcanu, Ion, Andrea Baiocchi, Nikita Lyamin e Alexey Vinel: *An age-of-information perspective on decentralized congestion control in vehicular networks*. Em *2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet)*, páginas 1–8, 2021. 9
- [15] Krikidis, Ioannis: *Average age of information in wireless powered sensor networks*. *IEEE Wireless Communications Letters*, 8(2):628–631, 2019. 9
- [16] Lin, Wensheng, Lixin Li, Jinhong Yuan, Zhu Han, Markku Juntti e Tad Matsumoto: *Cooperative lossy communications in unmanned aerial vehicle networks: Age-of-information with outage probability*. *IEEE Transactions on Vehicular Technology*, 70(10):10105–10120, 2021. 9
- [17] Ceran, Elif Tuğçe, Deniz Gündüz e András György: *A reinforcement learning approach to age of information in multi-user networks with harq*. *IEEE Journal on Selected Areas in Communications*, 39(5):1412–1426, 2021. 10
- [18] Tong, Xin, Liying Li, Guodong Zhao, Zhen Mengi, Zhi Chen e Geng Yang: *Packet management for optimizing control performance in real-time feedback control systems*. Em *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, páginas 4362–4367, 2020. 10
- [19] Chiang, Yi Han, Hai Lin e Yusheng Ji: *Information cofreshness-aware grant assignment and transmission scheduling for internet of things*. *IEEE Internet of Things Journal*, 8(19):14435–14446, 2021. 10
- [20] Kosta, Antzela, Nikolaos Pappas e Vangelis Angelakis: *Age of Information: A New Concept, Metric, and Tool*. 2017. 10
- [21] Costa, Maice, Marian Codreanu e Anthony Ephremides: *Age of information with packet management*. Em *2014 IEEE International Symposium on Information Theory*, páginas 1583–1587, 2014. 12, 16
- [22] Yates, Roy D., Yin Sun, D. Richard Brown, Sanjit K. Kaul, Eytan Modiano e Sennur Ulukus: *Age of information: An introduction and survey*. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, 2021. 12, 19, 20, 26
- [23] Bertsekas, Dimitri e Robert Gallager: *Data Networks (2nd Ed.)*. Prentice-Hall, Inc., USA, 1992, ISBN 0132009161. 12
- [24] Kendall, David G.: *Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain*. *The Annals of Mathematical Statistics*, 24(3):338 – 354, 1953. <https://doi.org/10.1214/aoms/1177728975>. 12

- [25] Huang, Longbo e Eytan Modiano: *Optimizing age-of-information in a multi-class queueing system*. páginas 1681–1685, junho 2015. 14
- [26] Inoue, Yoshiaki, Hiroyuki Masuyama, Tetsuya Takine e Toshiyuki Tanaka: *A general formula for the stationary distribution of the age of information and its application to single-server queues*. IEEE Transactions on Information Theory, 65(12):8305–8324, 2019. 15, 19, 26
- [27] Costa, Maice, Marian Codreanu e Anthony Ephremides: *On the age of information in status update systems with packet management*. IEEE Transactions on Information Theory, 62(4):1897–1910, 2016. 16, 17, 18
- [28] Pappas, Nikolaos, Johan Gunnarsson, Ludvig Kratz, Marios Kountouris e Vangelis Angelakis: *Age of information of multiple sources with queue management*. Em *2015 IEEE International Conference on Communications (ICC)*, páginas 5935–5940, 2015. 18, 22
- [29] Yates, Roy D. e Sanjit K. Kaul: *The age of information: Real-time status updating by multiple sources*. IEEE Transactions on Information Theory, 65(3):1807–1827, 2019. 18, 21, 22, 24, 26
- [30] Bedewy, Ahmed M., Yin Sun e Ness B. Shroff: *Minimizing the age of information through queues*. IEEE Transactions on Information Theory, 65(8):5215–5232, 2019. 18, 19, 23, 28
- [31] Bedewy, A., Yin Sun e N. Shroff: *The age of information in multihop networks*. IEEE/ACM Transactions on Networking, 27:1248–1257, 2019. 18, 28
- [32] Bedewy, Ahmed M., Yin Sun e Ness B. Shroff: *Optimizing data freshness, throughput, and delay in multi-server information-update systems*, 2017. 19, 28, 50
- [33] Moltafet, Mohammad, Markus Leinonen e Marian Codreanu: *On the age of information in multi-source queueing models*. IEEE Transactions on Communications, 68(8):5003–5017, 2020. 21, 27
- [34] Kam, Clement, Sastry Kompella, Gam D. Nguyen e Anthony Ephremides: *Effect of message transmission path diversity on status age*. IEEE Transactions on Information Theory, 62(3):1360–1374, 2016. 22
- [35] Yates, Roy D.: *Status updates through networks of parallel servers*. Em *2018 IEEE International Symposium on Information Theory (ISIT)*, páginas 2281–2285, 2018. 23
- [36] Javani, Alireza, Marwen Zorgui e Zhiying Wang: *Age of information for multiple-source multiple-server networks*, 2021. <https://arxiv.org/abs/2106.07247>. 24, 29
- [37] Sun, Yin, Elif Uysal-Biyikoglu e Sastry Kompella: *Age-optimal updates of multiple information flows*, 2018. 24, 25, 28, 51, 52, 67, 68

- [38] Kosta, Antzela, Nikolaos Pappas, Anthony Ephremides e Vangelis Angelakis: *Age of information performance of multiaccess strategies with packet management*. Journal of Communications and Networks, 21(3):244–255, 2019. 27
- [39] Akar, Nail: *Discrete-time queueing model of age of information with multiple information sources*. IEEE Internet of Things Journal, páginas 1–1, 2021. 27
- [40] Kaul, Sanjit K. e Roy D. Yates: *Age of information: Updates with priority*. Em *2018 IEEE International Symposium on Information Theory (ISIT)*, páginas 2644–2648, 2018. 28
- [41] Yates, Roy D.: *Age of information in a network of preemptive servers*. Em *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, páginas 118–123, 2018. 28
- [42] Prandel, Paulo César e Priscila Solis Barreto: *Computational modeling of age of information for cyber-physical systems*. Em *2021 IEEE Latin-American Conference on Communications (LATINCOM)*, páginas 1–6, 2021. 38
- [43] Yates, R.: *The age of information in networks: Moments, distributions, and sampling*. IEEE Transactions on Information Theory, 66:5712–5728, 2020. 39
- [44] Najm, Elie e Rajai Nasser: *Age of information: The gamma awakening*. Em *2016 IEEE International Symposium on Information Theory (ISIT)*, páginas 2574–2578, 2016. 39
- [45] Owen, Art B: *Monte carlo theory, methods and examples*. 2013. 41, 56
- [46] Fang, Youhan, Yudong Cao e Robert D. Skeel: *Quasi-reliable estimates of effective sample size*, 2017. 41
- [47] Jordon, Daniel: *Queueing-tool: A network simulator*. <https://queueing-tool.readthedocs.io/en/latest/index.html>, 2016. 44
- [48] Prandel, Paulo: *Uma ferramenta de simulação para a age of information*. <https://github.com/pprandel/aoi-simulator>, 2022. 44
- [49] Gupta, Ramesh C. e David M. Bradley: *Representing the mean residual life in terms of the failure rate*, 2004. 45
- [50] Mugdadi, Abdel Razzaq e Amanuel Teweldemedhin: *Two nonparametric estimators of the mean residual life*. REVSTAT – Statistical Journal, 11(3):301–315, 2013. 51
- [51] Jain, Raj: *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991, ISBN 9780471503361. 57