

Preempção Condicional de Pacotes baseada na Vida Média Residual para Otimização da *Age of Information* em Sistemas Ciberfísicos

Paulo César Prandel¹, Priscila Solis Barreto¹

¹Departamento de Ciência da Computação – Universidade de Brasília (UNB)
Brasília – DF – Brasil

paulo.prandel@aluno.unb.br, pris@unb.br

Abstract. *In the study of Cyberphysical Systems, the Age of Information (AoI) appears as a new concept to represent the degree of updating of the information that a monitor has in relation to an entity or remote process that sends updates periodically. Packet management is a set of techniques and policies that prioritize, block or preempt (drop and replace) packets that arrive at a server and that can be used for AoI optimization. State-of-the-art techniques such as LGFS (Last Generated First Served) have static behavior, performing the same action for all packages: LGFS-S preempts in service and LGFS-W and executes in service queue. The present study proposes a new conditional preemption technique called LGFS-C, which uses the concept of residual half-life to decide the policy to be applied to a packet in service or in the queue. For a single queue model, it is shown that the LGFS-C technique always achieves results equal to or superior to the other techniques, to any distribution of service time in the server. The proposed technique was validated through simulations, with the use of a computational tool specially developed for this purpose.*

Resumo. *No estudo de Sistemas Ciberfísicos, a Age of Information (AoI) surge como um novo conceito ao representar o grau de atualização da informação que um monitor possui em relação a uma entidade ou processo remoto que envia atualizações periodicamente. O gerenciamento de pacotes é um conjunto de técnicas e políticas que priorizam, bloqueiam ou executam a preempção (descarte e substituição) de pacotes que chegam a um servidor e que pode ser utilizada para otimização da AoI. As técnicas no estado da arte, como a LGFS (do inglês Last Generated First Served) possuem comportamento estático, executando a mesma ação para todos os pacotes: LGFS-S executa a preempção em serviço e LGFS-W e executa na fila de espera. O presente estudo propõe uma nova técnica de preempção condicional chamada LGFS-C, a qual utiliza o conceito de vida média residual para decidir a política a ser aplicada a um pacote em serviço ou na fila de espera. Para um modelo de fila simples, demonstra-se que a técnica LGFS-C alcança sempre resultados iguais ou superiores aos das outras técnicas, para qualquer distribuição de tempo de serviço para o servidor. A técnica proposta foi validada através de simulações, com o uso de uma ferramenta computacional especialmente desenvolvida para esse propósito.*

1. Introdução

Os Sistemas Ciberfísicos (em inglês *Cyber-Physical Systems* - CPS) englobam diversas áreas de estudo, tais como redes, comunicações complexas, controle, fusão de informações, processamento, otimização de sinais, entre outras. Os CPS geralmente fazem uso de modelos distribuídos complexos de larga escala, que operam em ambientes não estruturados em escalas multiespaciais e temporais, que usam controles inteligentes e fusão multimodal de informações. As aplicações dos CPS podem contemplar um grande número de dispositivos, que podem ou não interagir entre si. Com o advento de tecnologias como o 5G, as aplicações e variedade de dispositivos individuais tendem a aumentar consideravelmente nos próximos anos, situação que provoca vários desafios técnicos, entre os quais podem ser mencionados a definição de arquiteturas e a especificação e cálculo de métricas de desempenho nesses sistemas [Yates et al. 2021][Bhattacharyya and Wolf 2020].

Recentemente, um dos conceitos que tem se destacado na pesquisa de CPS é a *Age of Information*(AoI). Esse conceito [Kaul et al. 2012a] considera que a pontualidade das atualizações no receptor pode ser otimizada como uma métrica de ponta a ponta, que representa o frescor do conhecimento que um monitor tem sobre um processo e seu ambiente, cujas atualizações são enviadas por uma ou mais fontes através de pacotes de atualização periódicos. Nos CPS que usam o monitoramento para a tomada de decisões, o objetivo é ter em tempo hábil as informações mais úteis. Assim, a atividade de monitoramento nesses sistemas depende da obtenção de dados que pretendem informar uma percepção acurada dos fatores e condições que afetam a atividade monitorada. Nos CPS, métricas de rede clássicas como taxa de transmissão, vazão e atraso nem sempre estão alinhados com o processo de tomada de decisões. Um exemplo clássico é o desalinhamento em uma rede em que agentes ou sensores enviam atualizações usando a capacidade máxima de largura de banda. Como consequência, o servidor responsável por receber essas atualizações pode ter atrasos consideráveis ou descartar informações, entregando dados desatualizados ao tomador de decisões, apesar das métricas da rede apresentarem valores satisfatórios para vazão e taxa de transmissão. A partir deste exemplo, percebe-se a necessidade de se obter métricas mais alinhadas ao domínio da aplicação do CPS e não apenas com base nas métricas da rede.

Entre as principais técnicas para a otimização da AoI em CPS, o gerenciamento de pacotes consiste em realizar ações de bloqueio, priorização e preempção (substituição e descarte) dos pacotes de informação que chegam em uma estrutura modelada como fila-servidor, visando a obtenção de melhores métricas, por exemplo, a AoI média para uma ou mais fontes. Diversos estudos abordaram esse tema nos últimos anos, como [Bedewy et al. 2017a], [Sun et al. 2018], [Kosta et al. 2019], [Akar 2021], propondo técnicas que podem ser consideradas estáticas no sentido de que a técnica utilizada na priorização ou na preempção é sempre a mesma para todos os pacotes, independente do estado atual do sistema. Dessa maneira, tais técnicas tornam-se limitadas e nem sempre conseguem produzir bons resultados em função da dinâmica e do tipo de aplicação do CPS.

Com essa motivação, este trabalho propõe uma nova técnica chamada *Last Generated First Served - Conditional* (LGFS-C), a qual toma decisões dinâmicas (pacote a pacote) com relação à técnica de preempção aplicada, baseadas no estado atual do sistema e utilizando o conceito de vida residual média (VRM) [Mugdadi and Teweldemedhin 2013]. Essa técnica é especialmente eficiente em sistemas

onde a distribuição do tempo de serviço possui valores da vida residual média inferiores à media da própria distribuição, como é o caso de algumas distribuições conhecidas, como Weibull e Lognormal. As contribuições científicas deste trabalho são as seguintes:

- Proposição de uma melhoria nas técnicas atuais de gerenciamento de pacotes para a otimização da AoI em sistemas ciberfísicos, a partir de uma abordagem dinâmica e mais abrangente em relação à distribuição do tempo de serviço;
- Contribuição à avaliação de CPS, possibilitando a obtenção de métricas superiores e mais adequadas ao domínio de aplicação do sistema.

A estrutura deste artigo é a seguinte: A Seção 2 apresenta um referencial teórico necessário ao entendimento do assunto, bem como os principais trabalhos relacionados na área. Na Seção 3 são apresentados o modelo e a técnica LGFS-C, a qual é descrita de forma analítica. A Seção 4 apresenta os resultados produzidos através de um ferramenta computacional construída pelos autores para a avaliação da AoI e, por fim, a Seção 5 apresenta as conclusões e as sugestões para trabalhos futuros.

2. Conceitos Teóricos e Trabalhos Relacionados

A presente seção tem por finalidade apresentar os principais conceitos teóricos necessários para o entendimento deste estudo, bem como os trabalhos mais recentes na pesquisa da área.

2.1. Age of Information

A *Age of Information* (AoI) [Yates et al. 2021] representa o grau de atualidade que um monitor possui sobre determinado processo ou entidade, a qual envia pacotes¹ de atualização periódicos para esse monitor. Em termos analíticos, um determinado pacote com tempo de geração u possui uma AoI dada por $t - u$ em um tempo $t \geq u$. Dessa maneira, a AoI no monitor em um tempo t pode ser definida como o processo aleatório $\Delta(t) = t - u(t)$, onde $u(t)$ é o tempo de geração do último pacote recebido. A Fig. 1 ilustra o formato *dente-de-serra* característico desse processo.

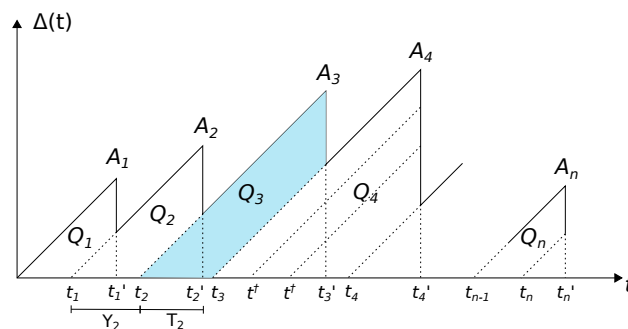


Figura 1. Comportamento característico do processo AoI.

Ainda na Fig. 1, define-se os seguintes termos:

- t_i : tempo de chegada do pacote i ;
- t'_i : tempo de partida do pacote i ;
- t^\dagger : pacote perdido ou descartado (preempção);
- Y_i : tempo entre as chegadas dos pacotes $i - 1$ e i ;

¹O termo *pacotes* no presente contexto se refere ao nível da aplicação e não a um protocolo específico.

- T_i : tempo entre a chegada do pacote i e a sua entrega ao monitor. Também chamado de *tempo de sistema*;
- A_i : valores de pico da AoI.

A AoI média $\langle \Delta \rangle_\tau$ pode ser calculada somando-se a área sob o gráfico e dividindo esse valor pelo tempo decorrido. Em um regime estacionário e ergódico, a $\langle \Delta \rangle_\tau$ é definida de acordo com a Eq. 1.

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

2.2. Gerenciamento de pacotes

O gerenciamento de pacotes consiste em um conjunto de técnicas que atuam sobre os pacotes que chegam à estrutura fila-servidor. Essa atuação pode ser realizada através de priorização, bloqueio e preempção (descarte e substituição do pacote por outro), tendo por objetivo otimizar os valores da AoI. Um dos primeiros estudos dessa técnica foi realizado por [Costa et al. 2016]. A priorização pode ser realizada através de regimes de serviço que priorizem os pacotes mais recentes ou frescos, como por exemplo o regime *Last Come First Served* (LCFS), no qual o último pacote que chega na fila é o primeiro a ser servido. Já o regime *Last Generated First Served* (LGFS), abordado em [Sun et al. 2018], é especialmente eficiente em sistemas onde os pacotes chegam fora da ordem de geração, pois serve com prioridade os pacotes que foram gerados por último, ou seja, os mais recentes.

O bloqueio de novos pacotes consiste em limitar o tamanho de uma fila, impedindo o recebimento de novos pacotes pelo conjunto fila-servidor enquanto o mesmo estiver em sua capacidade máxima. Já a preempção, conforme ilustrado pela Fig. 2, é o ato de substituir pacotes, podendo esse pacote estar em serviço (*service* - S) (Fig. 2a) ou na fila de espera (*waiting* - W) (Fig. 2b). Os principais modelos para essas técnicas são o servidor sem fila de espera e o servidor com uma única posição de espera. Alguns autores utilizam o termo *preempção* (P) para o caso de preempção em serviço e *não preempção* (NP) para o caso de preempção na fila de espera. Para fins de padronização, este trabalho adotará a nomenclatura com os sufixos *S* e *W* citados acima.

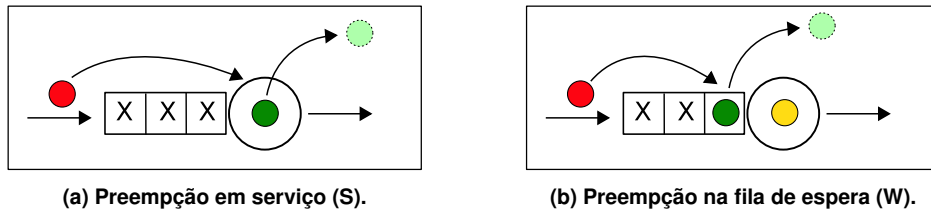


Figura 2. Técnicas de preempção de pacotes.

2.3. Vida residual média

A vida residual média (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 t . *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 de distribuição acumulada (FDA) e $\bar{F} = 1 - F$ a função sobrevivência. A função de *vida residual média* (VRM) é definida pela Eq. 2 [Gupta and Bradley 2004].

$$m(t) \equiv E(T - t | T > t) = \frac{1}{\bar{F}(t)} \int_t^{\infty} \bar{F}(u) du. \quad (2)$$

Trazendo 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 pacotes de atualização. A função $m(t)$ representa o tempo médio adicional necessário para que o servidor finalize o serviço de um 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 fresco.

2.4. Trabalhos relacionados

O primeiro trabalho a abordar a ideia de priorizar e realizar a preempção de pacotes foi [Kaul et al. 2012b]. Nesse estudo, os autores sugerem a utilização do regime de serviço *Last Come First Served - LCFS*, visando priorizar o serviço dos pacotes mais frescos (maior tempo de geração) na fila. Os autores introduzem ainda uma técnica de preempção, ou seja, interrupção do serviço de um pacote e substituição do mesmo por um mais fresco.

Em [Costa et al. 2014] é definido o conceito de gerenciamento de pacotes para a otimização da AoI, realizando-se a sua modelagem para uma fila simples e definindo-se duas técnicas base. A primeira consiste em uma fila com capacidade limitada (possuindo nenhuma ou apenas uma posição de espera), na qual novos pacotes são descartados caso o conjunto fila-servidor esteja em sua capacidade máxima. A segunda técnica consiste em um conjunto fila-servidor com apenas uma posição de espera, onde novos pacotes realizam a preempção (substituição e descarte) de um pacote que esteja nessa posição. Em um segundo trabalho, os mesmos autores [Costa et al. 2016] apresentam as mesmas técnicas com mais profundidade, apresentando expressões analíticas para o cálculo da AoI média e para a AoI de pico em filas simples.

[Pappas et al. 2015] propõe uma extensão das técnicas acima para modelos com múltiplas fontes de pacotes. Nesse modelo de um único servidor, a fila armazena no máximo um pacote de cada fonte, realizando a preempção deste sempre que um novo pacote da mesma fonte é recebido. Os autores inserem ainda um fator de atraso (*delay*) dos pacotes entre a fila e o servidor, de maneira que os pacotes possam chegar ao mesmo fora da ordem de geração, definindo assim o conceito de pacote *obsoleto*. No trabalho de [Najm and Nasser 2016] é apresentada uma modelagem para uma fila com regime LCFS e tempos de serviço com distribuição *Gamma*. Esse é um dos primeiros estudos a verificar que a técnica de preempção em serviço nem sempre otimiza a AoI, sendo que a técnica de preempção na fila de espera (ou não preempção) pode alcançar melhores resultados, o que mais tarde seria demonstrado para classes especiais de distribuição.

No estudo de [Bedewy et al. 2017b] é apresentado o regime de serviço conhecido como *Last Generated First Served - Preemptive* (LGFS-P), o qual otimiza a AoI para um sistema com múltiplos servidores e distribuição de tempo de serviço exponencial, independente da ordem de chegada dos pacotes. Em um outro trabalho, [Bedewy et al. 2017a] apresenta resultados semelhantes para um modelo de rede com múltiplos saltos, com a ressalva de que tempos de serviço com distribuição não exponencial não são otimizados pela técnica LGFS-P. Em um terceiro trabalho, [Bedewy et al. 2019] demonstra que, para uma distribuição de tempo de serviço pertencente à classe conhecida como *New Better*

Than Used (NBU) a técnica 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, estando a um *gap* constante de um limite inferior estabelecido.

O trabalho de [Sun et al. 2018] apresenta duas técnicas de gerenciamento de pacotes para modelos com múltiplas fontes e múltiplos servidores, sendo uma delas ótima para o caso de tempos de serviço com distribuição exponencial e outra quase ótima para tempos de serviço do tipo NBU. Outros trabalhos que abordam modelos com múltiplas fontes são [Yates and Kaul 2019], [Kosta et al. 2019], [Moltafet et al. 2020] e [Akar 2021]. [Inoue et al. 2019] realiza um estudo onde obtém expressões analíticas para a grande maioria das técnicas de gerenciamento anteriormente citadas. Por fim, [Yates et al. 2021] apresenta um *survey* com todos os trabalhos anteriormente citados, realizando comparações e fornecendo sugestões de trabalhos futuros.

A revisão da literatura mostra que o presente trabalho apresenta uma contribuição inovadora em relação aos trabalhos anteriormente citados, ao propor a utilização da técnica LGFS de maneira dinâmica para o gerenciamento de pacotes, em que a preempção em serviço (LGFS-S) ou na fila de espera LGFS-W (do inglês *waiting*, também abordada como não preempção) é decidida para cada pacote de acordo com o estado atual do sistema. Esse gerenciamento dinâmico é capaz de produzir resultados superiores, os quais se aplicam para qualquer distribuição de tempo de serviço.

3. Modelagem e Apresentação da Técnica LGFS-C

Esta seção tem por finalidade descrever o modelo do sistema ciberfísico utilizado neste estudo, sobre o qual será definida e demonstrada a técnica de gerenciamento de pacotes LGFS-C.

3.1. Modelo do sistema

A estrutura da Fig. 3 representa um sistema ciberfísico cuja finalidade é realizar o monitoramento de um determinado processo ou entidade. Pacotes de atualização são gerados por uma fonte através de um processo de Poisson com uma taxa λ e enviados através de uma rede cujo tempo de transmissão é dado por uma distribuição exponencial com média k . Na sequência, os pacotes são recebidos em uma fila e enviados para um servidor, cujo tempo de serviço é dado por um processo aleatório com taxa μ . Esse processo no servidor pode assumir qualquer distribuição de tempo de serviço, conforme será apresentado adiante. Após o serviço, os pacotes são considerados entregues ao monitor. Ressalta-se que, devido ao modelo da rede de transmissão, os pacotes podem chegar à fila fora da ordem em que foram gerados.

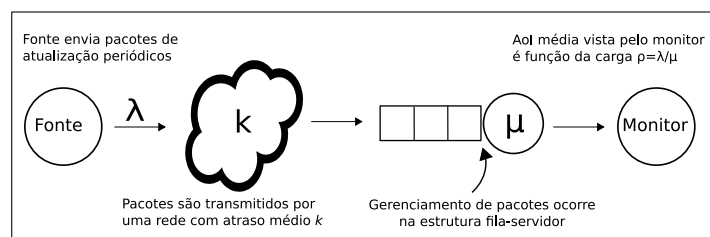


Figura 3. Modelo do sistema.

À medida em que os pacotes vão sendo entregues ao monitor, este experimenta uma AoI média em relação ao processo observado, cujo valor depende, entre outros fatores, da carga no servidor $\rho = \lambda/\mu$. A técnica de gerenciamento de pacotes ocorre

na estrutura fila-servidor e visa otimizar a AoI vista pelo monitor para uma determinada carga. Dependendo das restrições do sistema, o objetivo pode ser minimizar a AoI, independente da quantidade de pacotes descartados pela técnica. Já em outras aplicações, onde torna-se importante receber no monitor o máximo de pacotes, ainda que obsoletos, pode-se estabelecer uma relação de compromisso entre a AoI média e a porcentagem de pacotes descartados. Ressalta-se que esse processo de ajuste não é o foco deste trabalho, o qual tem por objetivo propor uma nova técnica de gerenciamento de pacotes, capaz de produzir uma AoI média menor do que as técnicas já existentes.

3.2. A técnica LGFS-C

As técnicas de gerenciamento de pacotes anteriores a este estudo 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. O presente trabalho, por sua vez, propõe 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 de *LGFS-C*, onde a letra 'C' significa preempção condicional, e representa 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. A Fig. 4 ilustra de maneira gráfica o mecanismo de tomada de decisão para um pacote recém chegado, o qual será explicado a seguir.

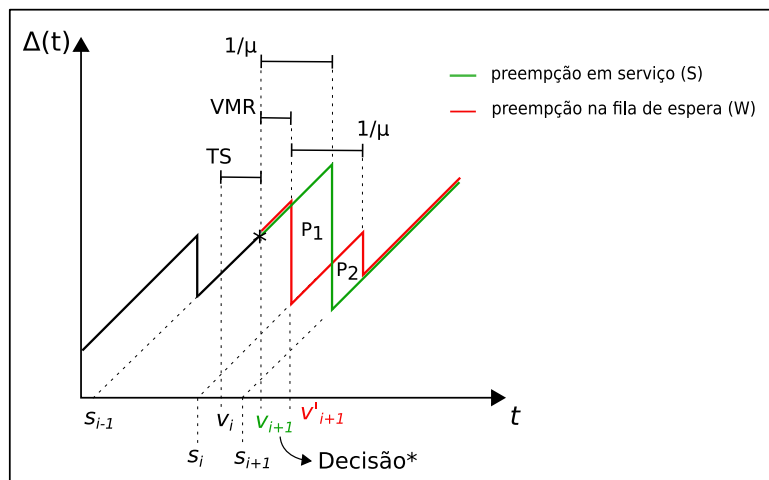


Figura 4. Ilustração da técnica LGFS-C. O gráfico mostra a evolução da AoI baseada na decisão de realizar a preempção em serviço (verde) ou na fila de espera (vermelho).

Considere-se a estrutura fila-servidor da Fig. 3 com capacidade máxima de dois pacotes, sendo um em serviço e outro na fila de espera. Seja 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} sendo exatamente o momento da tomada de decisão, a qual resultará em dois possíveis caminhos:

- Preempção em serviço (em 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 figura;

- Preempção na fila de espera (em vermelho): o pacote $i + 1$ é alocado na fila de espera, realizando 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 figura.

Ao final desse processo, observa-se que ambos os caminhos do processo AoI se encontram novamente e que, dependendo 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. 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.

$$\text{Preempção} = \begin{cases} S & \text{caso } P1 \leq P2, \\ W & \text{caso contrário.} \end{cases} \quad (3)$$

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 $P2$, obtendo-se o mesmo critério em função das variáveis relacionadas ao domínio do tempo, de acordo com a Eq. 4.

$$\begin{aligned} P1 &= (s_i - s_{i-1}) \left(\frac{1}{\mu} - VRM \right), \\ P2 &= (s_{i+1} - s_i) VRM. \end{aligned} \quad (4)$$

O valor da vida residual média, por sua vez, é dado por $VRM = m(TS)$, conforme a Eq. 2. Por fim, após algumas manipulações, o critério de decisão da Eq. 3 pode ser reescrito como na Eq. 5.

$$\text{Preempção} = \begin{cases} S & \text{caso } \frac{s_i - s_{i-1}}{s_{i+1} - s_{i-1}} \leq m(TS)\mu \\ W & \text{caso contrário.} \end{cases} \quad (5)$$

A técnica LGFS-C, semelhante ao que ocorre nas técnicas LGFS-S e LGFS-W, prioriza os pacotes mais *frescos*, ou seja, aqueles que possuem o maior tempo de geração, independente da decisão tomada. A Fig. 4 ilustrou um caso particular onde 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, contemplando todos os casos possíveis de chegadas bem como os possíveis estados da fila e do servidor.

3.3. Demonstração analítica

Uma vez definida, torna-se necessário demonstrar que a técnica LGFS-C apresenta resultados superiores às demais técnicas no estado da arte do gerenciamento de pacotes.

Algoritmo 1 Último Gerado Primeiro Servido - preempção seletiva

```
1: defina  $s_{pacote}$  como o tempo de geração do pacote  $pacote$ 
2: defina  $s_{entregue}$  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
```

Teorema: Seja uma estrutura fila-servidor com uma única posição de espera, a qual recebe pacotes provenientes de uma fonte, os quais são gerados de acordo com um processo de *Poisson*, podendo ser recebidos fora da ordem de geração. A técnica LGFS-C, resulta em uma AoI média Δ 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. 4, 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. 6.

$$\delta_{P,i} = \begin{cases} P_1 & \text{caso Preempção} = S, \\ P_2 & \text{caso Preempção} = W. \end{cases} \quad (6)$$

Como o critério de decisão da técnica LGFS-C (Eq. 5) 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. 7.

$$E[\delta_{P,i}^{LGFS-C}] \leq E[\delta_{P,i}^{LGFS-S/W}]. \quad (7)$$

A Eq. 7 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 números, pode-se afirmar que a soma da área total sob o gráfico, conforme a Fig. 1, 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. 8.

$$P \left(\lim_{i \rightarrow \infty} \sum^i Q_i^{LGFS-C} \leq \lim_{i \rightarrow \infty} \sum^i Q_i^{LGFS-S/W} \right) = 1. \quad (8)$$

Por fim, dado que a AoI média Δ 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 inferior às demais técnicas. Ressalta-se que o referido teorema é válido para qualquer distribuição de tempo de serviço no servidor, uma vez que a decisão de preempção é realizada pacote a pacote, o que garante sempre as condições da Eq. 8. Essa análise do alcance da técnica será realizada com maiores detalhes na seção seguinte.

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. 5), 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 média residual 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 [Bedewy et al. 2017b].

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 onde essa distribuição é conhecida e parametrizada, pode-se calcular a VRM a partir da função distribuição de probabilidade (fdp), conforme Eq. 2. Como forma de se reduzir o custo computacional, os valores da VRM podem ser calculados *a priori*, utilizando-se uma tabela estática para a obtenção da vida residual média em função do tempo em serviço do pacote. Esse método foi utilizado nas simulações deste trabalho e será apresentado na Seç. 4.

Em sistemas reais, no entanto, 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 [Mugdadi and Teweldemedhin 2013].

4. Resultados Experimentais

Os resultados foram obtidos através de uma ferramenta computacional desenvolvida para simular sistemas ciberfísicos baseados em grafos, simulando o envio de pacotes e estimando a AoI de uma ou mais fontes no monitor. Essa ferramenta é descrita com detalhes em [Prandel and Barreto 2021]. O sistema da Fig. 3 foi simulado utilizando-se 3 diferentes distribuições para a variável aleatória tempo de serviço X , quais sejam, Exponencial, Weibull e Lognormal. Os parâmetros dessas distribuições, suas funções densidade de probabilidade e a função de vida média residual (VRM) podem ser visualizadas na Fig. 5. 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.

O atraso da rede foi definido como exponencial com média $k = 0,5$ segundo. Dessa maneira, os pacotes podem chegar na fila fora da ordem original de geração. Todos os gráficos são exibidos em função da carga no servidor $\rho = \lambda/\mu$, onde $\mu = 1/E(X)$. A taxa de geração de pacotes na fonte (Poisson) é obtida em função da carga através de $\lambda = \rho\mu$. A principal métrica de interesse é a AoI média da fonte tomada no monitor. 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.

Inicialmente, para um tempo de serviço exponencial (Fig. 6), observa-se que a técnica LGFS-C produz uma AoI média equivalente à técnica LGFS-S, sendo que esse resultado é ótimo conforme apresentado na Seção 3.4. A técnica LGFS-W, como esperado,

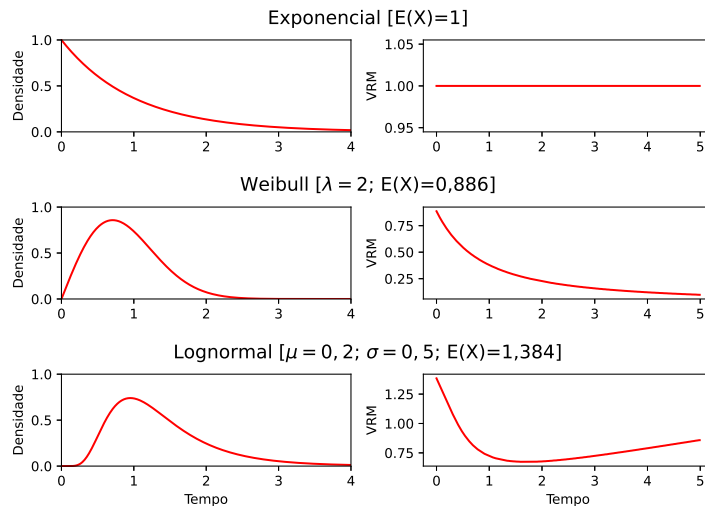


Figura 5. Distribuições e seus parâmetros utilizadas para o tempo de serviço nas simulações. A densidade é exibida à esquerda e a função VRM à direita.

apresentou desempenho inferior. Para os tempos de serviço pertencentes às distribuições Weibull (Fig. 7) e Lognormal (Fig. 8), observa-se o desempenho superior da técnica LGFS-C para todas as cargas ρ . Observa-se também que a técnica LGFS-S apresenta uma AoI média maior que as demais técnicas, ocasionando uma degradação ainda maior à medida em que se aumenta a carga. Foi realizado ainda, para a distribuição Lognormal, conforme a Fig. 8B, uma comparação das quantidades de pacotes descartados entre as técnicas. Das técnicas, 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%).

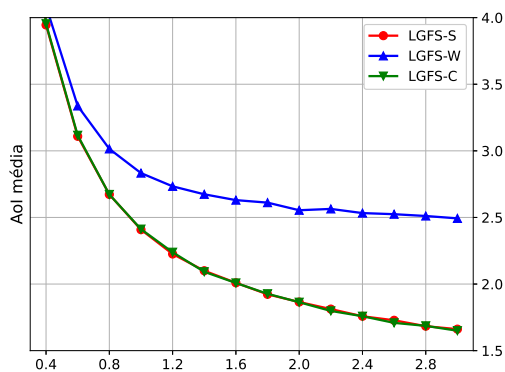


Figura 6. Comparação entre as técnicas para a distribuição exponencial.

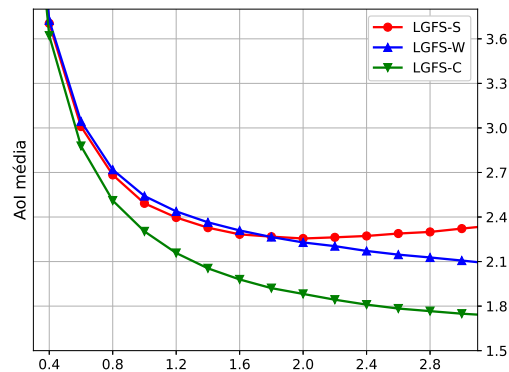


Figura 7. Comparação entre as técnicas para a distribuição Weibull.

Os resultados obtidos pelas simulações validam o modelo analítico da Seção 3.3, confirmando 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 apresenta uma leve alta e pode

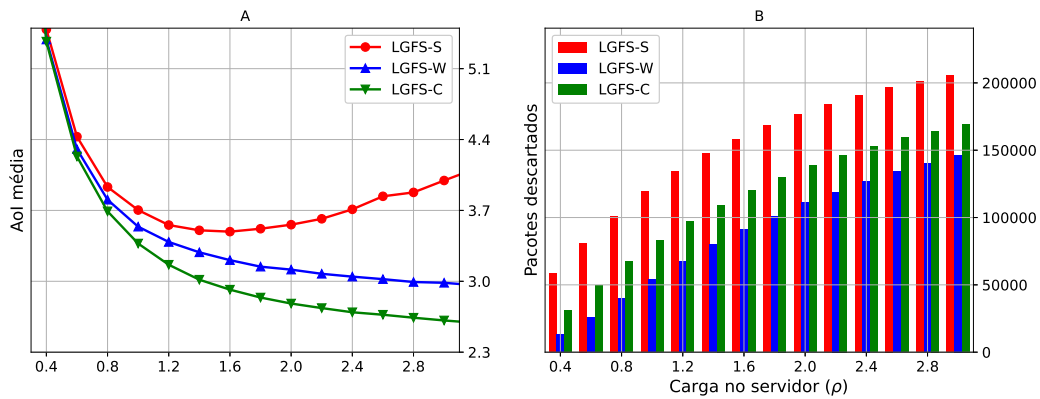


Figura 8. Comparação entre as técnicas para a distribuição Lognormal. A figura 'A' mostra a AoI média e a figura 'B' a quantidade de pacotes descartados.

ser avaliado em função do cenário de aplicação do CPS. Um sistema cujo histórico de pacotes que chega ao monitor não possui utilidade relevante pode admitir grandes taxas de descarte, de forma a priorizar a redução da AoI. Já para sistemas onde esse histórico é relevante deve-se estabelecer uma relação de compromisso entre a AoI e as taxas de descarte.

5. Conclusões e Trabalhos Futuros

Este trabalho apresentou uma nova técnica de gerenciamento de pacotes para otimizar a AoI média em sistemas ciberfísicos, a qual realiza decisões de preempção de pacotes com base no conceito de vida residual média. Essa técnica é particularmente eficiente para distribuições de tempo de serviço que possuem valores da vida média residual inferiores à media da própria distribuição.

A técnica proposta neste trabalho, LGFS-C, conforme demonstração analítica, produz uma AoI média sempre superior às técnicas LGFS-S e LGFS-W, independente da distribuição dos tempos de serviço. Os resultados obtidos nas simulações confirmam essa superioridade e ainda mostram que o aumento de pacotes descartado em relação à técnica LGFS-W é aceitável. Trata-se de uma técnica viável e que se posiciona no estado da arte entre as demais soluções relativas ao gerenciamento de pacotes para a otimização da AoI.

Como trabalhos futuros, pretende-se estender a técnica LGFS-C para modelos com múltiplas fontes e múltiplos servidores. Sugere-se, ainda, a avaliação da LGFS-C utilizando-se estimadores não paramétricos da vida média residual.

Referências

- Akar, N. (2021). Discrete-time queueing model of age of information with multiple information sources. *IEEE Internet of Things Journal*, pages 1–1.
- Bedewy, A. M., Sun, Y., and Shroff, N. B. (2017a). Age-optimal information updates in multihop networks. pages 576–580.
- Bedewy, A. M., Sun, Y., and Shroff, N. B. (2017b). Optimizing data freshness, throughput, and delay in multi-server information-update systems.
- Bedewy, A. M., Sun, Y., and Shroff, N. B. (2019). Minimizing the age of information through queues. *IEEE Transactions on Information Theory*, 65(8):5215–5232.

- Bhattacharyya, S. S. and Wolf, M. C. (2020). Research challenges for heterogeneous cyberphysical system design. *Computer*, 53(7):71–75.
- Costa, M., Codreanu, M., and Ephremides, A. (2014). Age of information with packet management. In *2014 IEEE International Symposium on Information Theory*, pages 1583–1587.
- Costa, M., Codreanu, M., and Ephremides, A. (2016). On the age of information in status update systems with packet management. *IEEE Transactions on Information Theory*, 62(4):1897–1910.
- Gupta, R. C. and Bradley, D. M. (2004). Representing the mean residual life in terms of the failure rate.
- Inoue, Y., Masuyama, H., Takine, T., and Tanaka, T. (2019). 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.
- Kaul, S., Yates, R., and Gruteser, M. (2012a). Real-time status: How often should one update? In *2012 Proceedings IEEE INFOCOM*, pages 2731–2735.
- Kaul, S. K., Yates, R. D., and Gruteser, M. (2012b). Status updates through queues. In *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6.
- Kosta, A., Pappas, N., Ephremides, A., and Angelakis, V. (2019). Age of information performance of multiaccess strategies with packet management. *Journal of Communications and Networks*, 21(3):244–255.
- Moltafet, M., Leinonen, M., and Codreanu, M. (2020). On the age of information in multi-source queueing models. *IEEE Transactions on Communications*, 68(8):5003–5017.
- Mugdadi, A.-R. and Teweldemedhin, A. (2013). Two nonparametric estimators of the mean residual life. *REVSTAT – Statistical Journal*, 11(3):301–315.
- Najm, E. and Nasser, R. (2016). Age of information: The gamma awakening. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 2574–2578.
- Pappas, N., Gunnarsson, J., Kratz, L., Kountouris, M., and Angelakis, V. (2015). Age of information of multiple sources with queue management. In *2015 IEEE International Conference on Communications (ICC)*, pages 5935–5940.
- Prandel, P. C. and Barreto, P. S. (2021). Computational modeling of age of information for cyber-physical systems. In *2021 IEEE Latin-American Conference on Communications (LATINCOM)*, pages 1–6.
- Sun, Y., Uysal-Biyikoglu, E., and Kompella, S. (2018). Age-optimal updates of multiple information flows.
- Yates, R. D. and Kaul, S. K. (2019). The age of information: Real-time status updating by multiple sources. *IEEE Transactions on Information Theory*, 65(3):1807–1827.
- Yates, R. D., Sun, Y., Brown, D. R., Kaul, S. K., Modiano, E., and Ulukus, S. (2021). Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210.