

Uso de um modelo matemático para análise e comparação de protocolos de encaminhamento em redes oportunísticas

Luis Armando Quintanilla Villon¹

¹Programa de Engenharia de Sistemas e Computação (PESC) – COPPE/UFRJ
Redes Móveis (CPS867)

lvillon@cos.ufrj.br

Resumo. Neste trabalho desenvolveu-se uma avaliação e comparação de diferentes estratégias de encaminhamento de mensagens em redes oportunísticas no marco de um modelo matemático "single copy" baseado em um processo semi-markoviano. Tais estratégias, que utilizam o conhecimento da estrutura social baseadas em métricas de centralidade foram: SimBet, FCER, People-Rank e uma versão modificada dessa última. A avaliação consistiu em determinar, se fosse possível, de que maneira estas estratégias se adaptam ao modelo matemático. A comparação foi realizada em três cenários diferentes, considerando taxa de encontros entre nós seguindo distribuições exponencial e de Pareto. Utilizaram-se como métricas, o valor esperado do delay e o número de encontros entre nós. Os resultados mostraram que, em geral, as três estratégias conseguem adaptar-se ao modelo matemático conseguindo uma eficiência similar entre elas com diferenças sutis, porém quantificáveis.

1. Introdução

O problema de encontrar as melhores estratégias para encaminhamento de informação em redes oportunísticas, bem como o modelo matemático mais adequado para casos gerais, foram amplamente estudados sem encontrar soluções gerais de grande aplicabilidade. Assim, a grande maioria de estratégias utilizam critérios heurísticos, uma vez que tentar encontrar a solução ótima para este problema é de complexidade *NP-difícil* [Balasubramanian et al. 2007]. No entanto, no artigo proposto por [Boldrini et al. 2014]

¹ buscou-se contribuir na construção de um modelo matemático "single copy" mais completo. Procedeu-se a utilizar este modelo para o presente trabalho.

Na seção 2, exibem-se, de forma bem sucinta, e tendo como guia de referência auxiliar o trabalho de [Dias 2018], os aspectos mais importantes do modelo matemático do Boldrini. Usando como base este modelo, escolheram-se três protocolos de encaminhamento para ser analisadas neste trabalho. As informações mais relevantes a respeito são explicadas na Seção 3, onde também são detalhadas as formas pelas quais as estratégias foram adaptadas especificamente para este projeto.

Na seção 4, explicam-se os três cenários teóricos nos quais foram aplicadas o modelo e os respectivos protocolos. Os resultados computados são mostrados na seção 5 e discutidos em maior profundidade na seção 6. Finalmente, na seção 7 abordam-se as conclusões mais ressaltantes.

¹ Aqui em diante será apenas chamado de "Modelo de Boldrini"

2. Modelo Matemático

O modelo do Boldrini é fundamentalmente baseado em um processo semi-markoviano onde a transição entre estados pode tomar um tempo aleatório, que pode ser caracterizado por uma determinada distribuição de probabilidade [Ross 2014]. Neste processo, quando uma mensagem parte de um elemento i com destino a j , com probabilidade p_{ij} , atrasos podem ser gerados se j não fosse o destino final d da mensagem. Assim, são definidos o delay D_i e o número de saltos H_i necessários para realizar a entrega ao destino final d :

$$D_i = \begin{cases} T_{id} \\ T_{ij} + D_j, \forall j \neq d \end{cases} ; H_i = \begin{cases} 1 \\ 1 + H_j, \forall j \neq d \end{cases} \quad (1)$$

Onde T_{ij} é o tempo em que a mensagem se mantém em i antes de partir para j . O conjunto total de possíveis j é conhecido previamente e é determinado pelo protocolo φ que está sendo utilizado neste processo. Seja tal conjunto definido da forma:

$$\mathcal{R}_i = \{j : p_{ij}^\varphi > 0\} \quad (2)$$

Onde, p_{ij}^φ pode ter os valores $\{0, 1\}$, sendo $p_{ij} = 1$ quando j é contemplado como possível destino de encaminhamento pelo protocolo φ , e $p_{ij} = 0$, caso contrário. Levando em consideração os possíveis j de (2) e depois de algumas manipulações algébricas em (1), obtém-se os sistemas de equações para o primeiro momento do delay D_i e do salto H_i :

$$E[D_i] = E[T_i] + \sum_{j \in \mathcal{R}_i - \{d\}} p_{ij} E[D_j], \forall i \neq d \quad (3)$$

$$E[H_i] = 1 + \sum_{j \in \mathcal{R}_i - \{d\}} p_{ij} E[H_j], \forall i \neq d \quad (4)$$

A taxa de encontros μ_{ij} para o processo é definida por $\mu_{ij} = \frac{1}{E[M_{ij}]}$, onde $E[M_{ij}]$ é o valor médio do intervalo do tempo de encontro M_{ij} . Como foi mencionado no início desta seção, o arranjo dos tempos de encontro pode ser caracterizados por uma distribuição de probabilidade. Para este trabalho, considerou-se as distribuições exponencial e de Pareto. Na sequência, seguem as formulas necessárias para encontrar $E[T_i]$ e p_{ij} em cada caso, para posteriormente substituir em (3) e (4) e resolver os sistemas.

- **Distribuição exponencial.** Considerando λ_{ij} como a taxa da distribuição exponencial dos tempos M_{ij} , logo, defina-se a relação: $\mu_{ij} = \lambda_{ij}$. Assim, para esta distribuição, $E[T_i]$ e p_{ij} são calculadas da seguinte forma:

$$E[T_i] = \frac{1}{\sum_{j \in \mathcal{R}_i} \lambda_{ij}}, \quad p_{ij} = \frac{\lambda_{ij}}{\sum_{z \in \mathcal{R}_i} \lambda_{iz}} \quad (5)$$

- **Distribuição de pareto** Neste caso, dados os parâmetros α_{ij} e $t_{min_{ij}}$ da distribuição, defina-se a taxa de encontros como: $\mu_{ij} = \frac{\alpha_{ij}-1}{t_{min_{ij}}}$. Logo, as expressões para calcular $E[T_i]$ e p_{ij} são:

$$E[T_i] = \frac{t_{min}}{\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n - 1}, \quad p_{ij} = \frac{\alpha_{ij} - 1}{\sum_{z \in \mathcal{R}_i} \alpha_{iz} - n}, \quad \text{onde: } n = |\mathcal{R}_i| \quad (6)$$

3. Estratégias de encaminhamento

As estratégias que foram usadas baseiam-se no conhecimento da centralidade da rede. As métricas mais frequentes utilizadas para quantificar e ranquear a centralidade são: betweenness, closeness e pagerank. De forma intuitiva, pareceria que qualquer das três poderiam ser utilizadas para gerar, de maneira rápida, uma lista dos possíveis nós candidatos aos quais uma mensagem será encaminhada. De fato, é possível mostrar que o betweenness e closeness atingem um desempenho superior para este propósito [Borgatti 2005].

De acordo com o modelo do Boldrini, os protocolos de encaminhamento φ devem calcular, previamente, o conjunto de possíveis destinos j mediante a relação (2). Na sequência são mostrados os protocolos baseados em centralidade que foram aplicados, bem como uma breve descrição sobre a maneira em que foram adaptados ao modelo.

- **SimBet (SB).** Utiliza-se o betweenness dos nós da rede e um critério de similaridade entre nós para esta estratégia [Daly and Haahr 2008]. Realizando uma modificação, é possível adaptar este protocolo ao modelo de Boldrini. Assim, definam-se as quantidades: $\alpha u_i^{Bet} = \text{betweenness do vértice } i$, ponderado pela constante α ; e: $\beta u_{i,d}^{Sim} = \text{similaridade entre o vértice } i \text{ e o destino } d$, ponderado pela constante β . É importante enfatizar que para este trabalho, foi considerado a similaridade de Jaccard como métrica para o cálculo de $u_{i,d}^{Sim}$. Sem perda de generalidade, assumiu-se que $\alpha = \beta = 1$. Logo, o *SimBet* do vértice i com o vértice destino d , é calculado da seguinte maneira: $u_{i,d}^{SB} = u_i^{Bet} + u_{i,d}^{Sim}$

Em resumo, o algoritmo para este protocolo é o seguinte:

Algorithm 1: SimBet (SB)

```

Result: O conjunto  $\mathcal{R}_i = \{j : p_{i,j}^\varphi > 0\}$ 
 $\mathcal{V}_i = \text{Vizinhos de } i;$ 
 $d = \text{Destino};$ 
 $u_{i,d}^{SB} = u_i^{Bet} + u_{i,d}^{Sim};$ 
for  $j \in \mathcal{V}_i$  do
     $u_{j,d}^{SB} = u_j^{Bet} + u_{j,d}^{Sim};$ 
    if  $j = d \vee u_{i,d}^{SB} < u_{j,d}^{SB}$  then
         $p_{i,j}^{\varphi=SB} = 1;$ 
         $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$ 
    else
         $p_{i,j}^{\varphi=SB} = 0$  //  $j$  não é incluído em  $\mathcal{R}_i$ ;
    end
end

```

- **PeopleRank (PR).** Nesta estratégia, usando como critério o pagerank de um vértice, é realizada uma comparação entre diferentes nós, considerando uma rede distribuída com mobilidade. [Mtibaa et al. 2010]. Assim, o *PeopleRank* do vértice i é definido como: $u_i^{PR} = \text{Pagerank do vértice } i$. Por causa da simplicidade do protocolo, é possível aplicar o algoritmo diretamente no modelo, considerando que a mobilidade dos nós é conhecida previamente.

Algorithm 2: PeopleRank (PR)

Result: O conjunto $\mathcal{R}_i = \{j : p_{ij}^\varphi > 0\}$
 $\mathcal{V}_i = \text{Vizinhos de } i;$
 $d = \text{Destino};$
for $j \in \mathcal{V}_i$ **do**
 if $j = d \vee u_i^{PR} \leq u_j^{PR}$ **then**
 $p_{i,j}^{\varphi=PR} = 1;$
 $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$
 else
 $p_{i,j}^{\varphi=PR} = 0$ // j não é incluído em $\mathcal{R}_i;$
 end
end

- **PeopleRank-2 (PR2).** Esta é uma versão modificada do protocolo anterior, proposta principalmente para este trabalho. Apenas são inseridas algumas linhas de código para realizar transmissão direta no caso em que algum dos nós vizinhos j seja o destino d .

Algorithm 3: PR2

Result: O conjunto $\mathcal{R}_i = \{j : p_{ij}^\varphi > 0\}$
 $\mathcal{V}_i = \text{Vizinhos de } i;$
 $d = \text{Destino};$
for $j \in \mathcal{V}_i$ **do**
 if $j = d$ **then**
 $p_{i,j}^{\varphi=PR} = 1;$
 $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$
 Break;
 // Transmissão direta
 else
 if $u_i^{PR} \leq u_j^{PR}$ **then**
 $p_{i,j}^{\varphi=PR} = 1;$
 $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$
 else
 $p_{i,j}^{\varphi=PR} = 0$ // j não é incluído em $\mathcal{R}_i;$
 end
 end
end

- **FCER.** Para este protocolo, é utilizado o closeness dos vértices para identificar os melhores encaminhamentos. No entanto, é necessário conhecer a dinâmica da rede ao longo de um determinado período de tempo, sem conhecer previamente o que ocorre na rede. [Patel and Shah 2013]. Assim, foi realizada uma adaptação para que a estratégia seja consistente com o modelo matemático do Bol-drini. Defina-se $u_i^{FCER} = \text{closeness do vértice } i$. O algoritmo para esta estratégia é aplicado de maneira similar aos casos anteriores.

Algorithm 4: FCER

```
Result: O conjunto  $\mathcal{R}_i = \{j : p_{ij}^\varphi > 0\}$   
 $\mathcal{V}_i = \text{Vizinhos de } i;$   
 $d = \text{Destino};$   
for  $j \in \mathcal{V}_i$  do  
    if  $j = d$  then  
         $p_{i,j}^{\varphi=PR} = 1;$   
         $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$   
        Break;  
        // Transmissão direta  
    else  
        if  $u_i^{FCER} \leq u_j^{FCER}$  then  
             $p_{i,j}^{\varphi=PR} = 1;$   
             $\mathcal{R}_i = \mathcal{R}_i \cup \{j\};$   
        else  
             $p_{i,j}^{\varphi=PR} = 0$  // j não é incluído em  $\mathcal{R}_i;$   
        end  
    end  
end  
if  $\mathcal{R}_i = \{\}$  then  
     $\mathcal{R}_i = \{x : u_x^{FCER} = \text{Max}\{u_j^{FCER}\}\};$   
    // Transmissão direta ;  
end
```

4. Cenários de estudo

Para representar os cenários, foi utilizada a abordagem proposta no artigo do Boldrini, onde são definidos os conceito de comunidade e de "nó viajante". Ademais, são consideradas duas distribuições, exponencial e de Pareto, para definir os encontros. O valor esperado do tempo dos encontros entre nós de uma mesma comunidade é considerado equivalente nas duas distribuições e fixada em 0.4 segundos. Assim, $E[M_{ij}] = \frac{1}{\mu_{ij}} = \frac{1}{\lambda_{ij}} = \frac{t_{min_{ij}}}{\alpha_{ij}-1} = 0.4s$, portanto: $\lambda_{ij} = 2.5s^{-1}$. Fixando $t_{min_{ij}} = 1$, encontra-se: $\alpha_{ij} = 3.5$. Quando os encontros ocorrem entre um viajante e um nó que pertence a um comunidade específica, o tempo de encontro $E[M_{ij}]$ se duplica. Neste caso: $\lambda_{ij} = 1.25s^{-1}$ e $\alpha_{ij} = 2.25$. Com estas informações, e considerando que os protocolos φ calculam os conjuntos $\mathcal{R}_i = \{j : p_{ij}^\varphi > 0\}$ antecipadamente conhecendo a topologia dos cenários, é possível calcular $E[T_i]$ e p_{ij} a partir das relações (5) e (6).

A modelagem dos cenários foi realizada através da biblioteca NetworkX da linguagem python para gerar grafos nos quais as arestas representam todas as possibilidades de comunicação entre os nós da rede. Considerou-se os dois cenários estudados no trabalho do Boldrini (Figuras 1a e 1b) e um terceiro (Figura 1c), que foi desenhado para fins de avaliação do desempenho, apenas em nível teórico. Os vértices de cor verde representam comunidades onde cada elemento interage apenas com elementos da mesma comunidade ou com viajantes. Os vértices azuis representam os viajantes que se contactam com outros viajantes, ou com nós de duas comunidades diferentes.

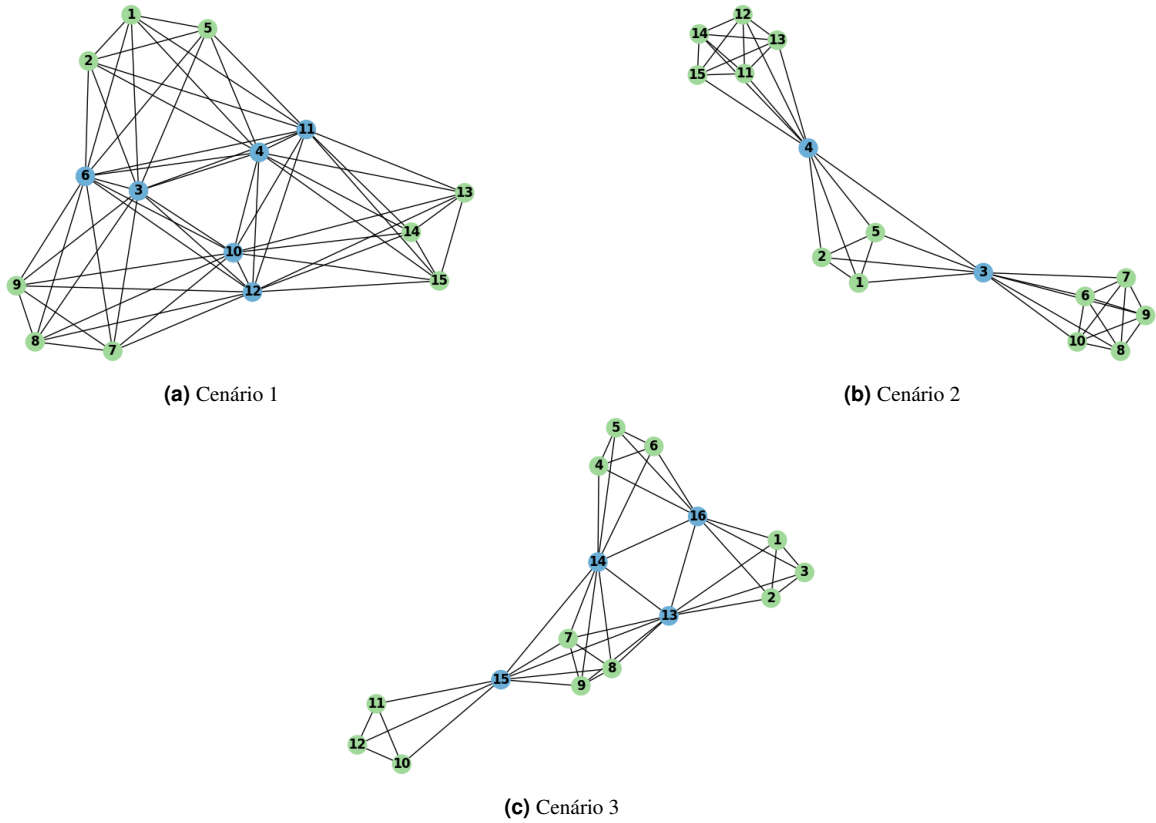


Figura 1. Cenários de estudo

5. Resultados

Os valores de $E[T_i]$ e p_{ij} , após serem identificados, foram substituídos nas equações (3) e (4). Assim, calcularam-se o valor esperado do delay² $E[D_i]$ e o número de saltos $E[H_i]$ para todos os encontros executados em cada cenário. Logo depois, determinou-se o valor médio de $E[D_i]$ e $E[H_i]$ para cada situação (Tabelas 1 até 6).

Com o objetivo de realizar uma comparação mais ampla dos protocolos, considerou-se a estratégia AF (Always Forward), baseada em epidemia, que está descrita no artigo do Boldrini. Assim, reiterando, as estratégias avaliadas foram: AF, SB (SimBet), PR (PeopleRank), PR2 (PeopleRank 2) e FCER. Elaboraram-se histogramas de frequência absoluta das distribuições do Delay (Figuras 2, 3, 4) e do número de saltos (Figura 5), com o objetivo de exibir e visualizar os resultados, de maneira mais pormenorizada.

Tabela 1. Delay médio(s) - Cenário 1

| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|--------|--------|--------|--------|--------|
| Exponencial | 1.4673 | 0.7700 | 1.1103 | 0.8634 | 1.0286 |
| Pareto | 1.6112 | 1.6147 | 1.2971 | 3.8052 | 4.7714 |

²Para simplificar, $E[D_i]$ será apenas chamado de delay

Tabela 2. Número médio de Saltos - Cenário 1

| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|---------|--------|--------|--------|--------|
| Expo./Pareto | 16.8745 | 2.3849 | 7.8990 | 1.8607 | 1.5857 |

Tabela 3. Delay médio(s) - Cenário 2

| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|--------|--------|--------|--------|--------|
| Exponencial | 3.3520 | 1.4292 | 1.8610 | 1.3981 | 1.3981 |
| Pareto | 3.7167 | 6.7113 | 5.3998 | 5.1407 | 6.5079 |

Tabela 4. Número médio de Saltos - Cenário 2

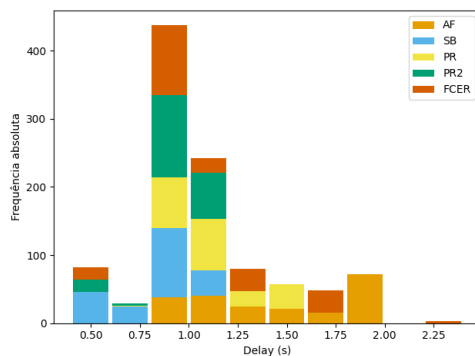
| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|---------|--------|--------|--------|--------|
| Expo./Pareto | 35.1955 | 2.0206 | 7.8333 | 5.6429 | 1.9286 |

Tabela 5. Delay médio(s) - Cenário 3

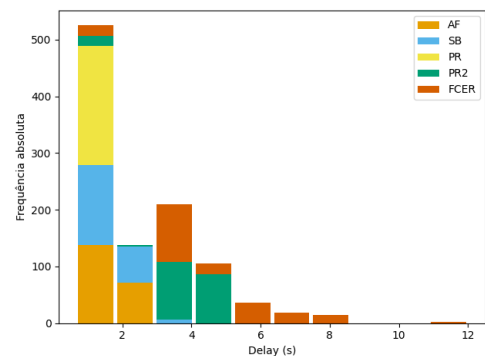
| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|--------|--------|--------|--------|--------|
| Exponencial | 3.3488 | 1.4872 | 1.6921 | 1.1104 | 1.3100 |
| Pareto | 4.0037 | 5.2175 | 5.1176 | 4.4002 | 5.6177 |

Tabela 6. Número médio de Saltos - Cenário 3

| Distribuição | AF | SB | PR | PR2 | FCER |
|--------------|---------|--------|--------|--------|--------|
| Expo./Pareto | 28.0672 | 3.0454 | 4.7686 | 2.5633 | 2.0250 |



(a) Distribuição do delay (exponencial)

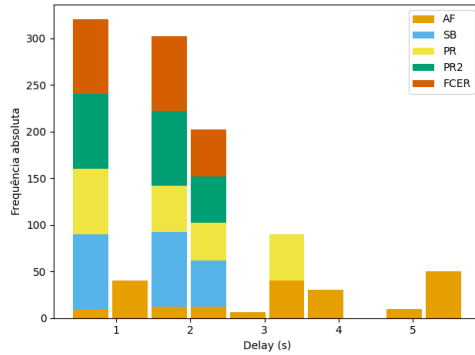


(b) Distribuição do delay (Pareto)

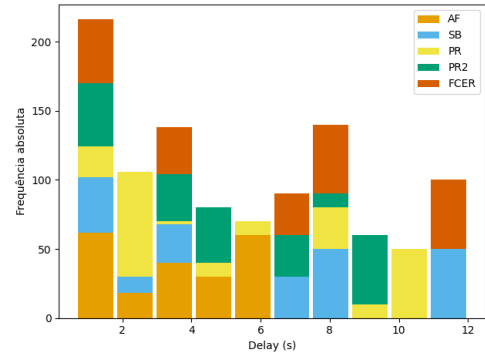
Figura 2. Cenário 1

6. Discussão

A primeira apreciação dos dados constituiu em verificar que as quatro estratégias de encaaminhamento baseadas em centralidade da rede, em geral, conseguem obter melhor desempenho, com menor tempo de delay e menor quantidade de saltos entre nós, em relação aos números alcançados pelo protocolo AF. Em particular, AF é expressivamente ineficiente em relação ao número de saltos. (Figura 5). Apenas em alguns casos particulares onde a distribuição de encontros segue uma Pareto (Tabela 3 e 5), o delay médio da estratégia AF é mais eficiente.

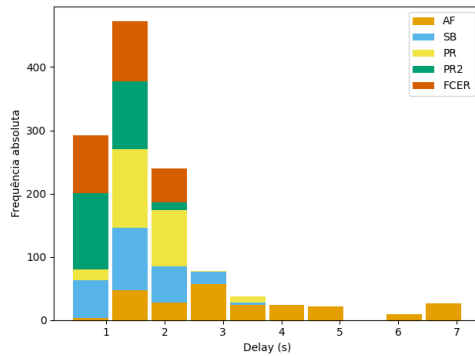


(a) Distribuição do delay (exponencial)

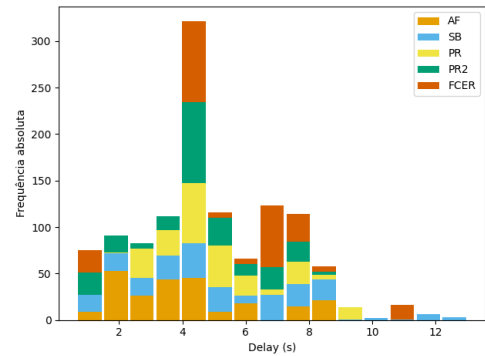


(b) Distribuição do delay (Pareto)

Figura 3. Cenário 2



(a) Distribuição do delay (exponencial)



(b) Distribuição do delay (Pareto)

Figura 4. Cenário 3

Em geral, os quatro protocolos baseados em centralidade mostraram-se muito similares a respeito das duas métricas avaliadas. Aproximadamente, o valor do delay médio encontrado oscila, entre as quatro métricas avaliadas, no intervalo $[0.5, 2]$ e $[1.5, 7]$, no caso exponencial e Pareto respectivamente. Na sequência, são relatadas as principais características da eficiência destes protocolos, bem como uma breve comparação entre eles.

No cenário 1, o protocolo SB exibe o menor delay no caso exponencial (Figura 2a). No caso da distribuição de Pareto, a estratégia PR consegue o menor delay (Figura 2b). No entanto, PR atinge o maior número de saltos (Figura 5a). Certamente, FCER e PR2 diminuem esta última métrica para os valores mais baixos

No cenário 2, as diferenças dos valores do delay entre as quatro métricas são mínimas (Figuras 3a e 3b). Apenas no número de saltos há uma leve diferença, porém não exorbitante (Figura 5b), com FCER mostrando o menor valor e PR, o máximo.

Finalmente, no cenário 3, os resultados exibem uma característica parecida com o que foi encontrado no cenário 2. Não há grandes diferenças no valor do delay dos quatro protocolos (Figuras 4a e 4b). Em relação ao número de saltos, PR exibe um maior valor e FCER, o menor (Figuras 5c).

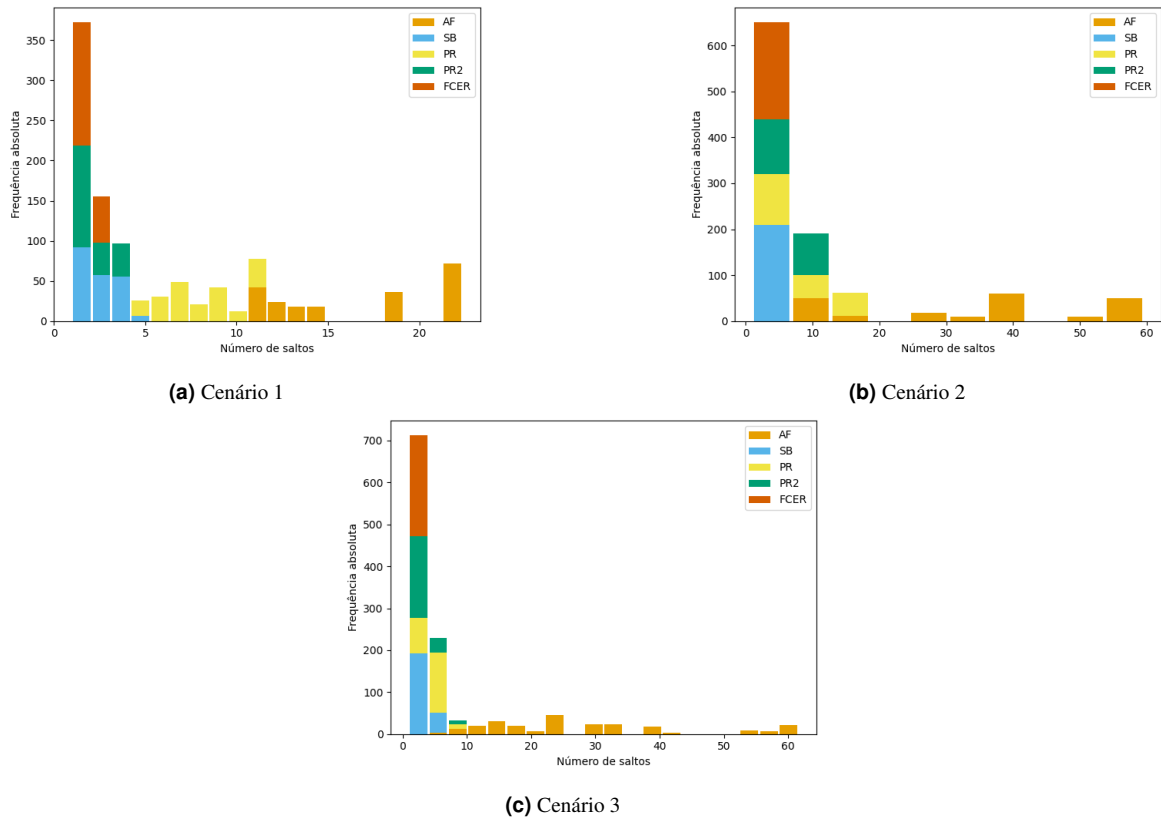


Figura 5. Distribuição do número de saltos

7. Conclusões

O modelo matemático do Boldrini é capaz de adaptar-se a diferentes cenários e contextos, bem como expor uma versatilidade quando se aplica um determinado protocolo de encaminhamento.

Inspirado em referências anteriores sobre encaminhamento em redes oportunísticas, foram empregadas várias estratégias neste trabalho. Apenas uma das três propostas, o PeopleRank pode ser aplicada, em seu formato original, utilizando o modelo de Boldrini.

De modo geral, os protocolos baseados em centralidade são muito eficientes quando a estrutura da rede é previamente conhecida, gerando baixos valores de delay no processo de encaminhamento em redes oportunísticas.

O protocolo modificado para este trabalho (PR2) mostrou melhor eficiência, em quase todas as circunstâncias comparativamente à versão original, o PeopleRank (PR). Assim, o processo de encaminhamento, mediante esta versão adaptada, é realizado com menor delay e com menor número de saltos na rede.

Referências

Balasubramanian, A., Levine, B., and Venkataramani, A. (2007). Dtn routing as a resource allocation problem. In *Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 373–384.

- Boldrini, C., Conti, M., and Passarella, A. (2014). Performance modelling of opportunistic forwarding under heterogenous mobility. *Computer Communications*, 48:56–70.
- Borgatti, S. P. (2005). Centrality and network flow. *Social networks*, 27(1):55–71.
- Daly, E. M. and Haahr, M. (2008). Social network analysis for information flow in disconnected delay-tolerant manets. *IEEE Transactions on Mobile Computing*, 8(5):606–621.
- Dias, G. M. d. S. (2018). Modelagem matemática do atraso de entrega de mensagens em redes oportunistas com taxas de encontro heterogêneas.
- Mtibaa, A., May, M., Diot, C., and Ammar, M. (2010). Peoplerank: Social opportunistic forwarding. In *2010 Proceedings IEEE INFOCOM*, pages 1–5. IEEE.
- Patel, J. and Shah, H. (2013). Fcer routing using temporal closeness centrality for delay tolerant network. In *2013 Nirma University International Conference on Engineering (NUiCONE)*, pages 1–5. IEEE.
- Ross, S. M. (2014). *Introduction to probability models*. Academic press.