

Eficiência dos Caminhos Quase Mais Curtos em Redes Dinâmicas

Dianne S. V. Medeiros¹, Miguel Elias M. Campista¹, Marcelo Dias de Amorim²,
Nathalie Mitton³, Guy Pujolle²

¹GTA / PEE / COPPE – UFRJ – Rio de Janeiro / RJ – Brasil

²LIP6 / CNRS – UPMC – Paris – França

³FUN – INRIA Lille – Ville Neuve d’Ascq – França

{dianne,miguel}@gta.ufrj.br, {marcelo.amorim,guy.pujolle}@lip6.fr,
nathalie.mitton@inria.fr

Abstract. *Betweenness centrality metrics often underestimate the importance of nodes close to the shortest path, but that rarely participate in them. In dynamic networks, these nodes can be, for a moment, in topologically strategic positions. This paper evaluates the importance of such nodes reusing the idea of spread betweenness centrality. This metric considers, besides the shortest paths, the multiple “quasi-shortest paths”, assigning them a proportional weight. The impact of the metric on the network is evaluated through comparisons with other betweenness metrics. Results show that using the idea of spreadness can reclassify nodes, reducing the number of articulation points in the network among the most well classified nodes. Considering failure on the most central nodes, the throughput of the network generally remains higher when the spreadness is applied. This property can help to choose better the role of nodes in the network, such that the performance of networks with temporal dynamics is improved.*

Resumo. *Métricas de centralidade de intermediação frequentemente subestimam a importância dos nós próximos do caminho mais curto, mas que raramente participam deles. Em redes dinâmicas, esses nós podem se encontrar momentaneamente em posições topologicamente estratégicas. Este artigo avalia a importância desses nós reutilizando a ideia de centralidade de intermediação por espalhamento. Essa métrica considera, além dos caminhos mais curtos, os múltiplos “caminhos quase mais curtos”, atribuindo-lhes um peso proporcional. O impacto da métrica na rede é avaliado através de comparações com outras métricas de intermediação. Os resultados mostram que o uso da ideia de espalhamento pode reclassificar nós, reduzindo o número de pontos de articulação na rede que estão dentre os nós mais bem classificados. Considerando falha nos nós mais centrais, a vazão da rede em geral mantém-se mais elevada quando o espalhamento é aplicado. Essa propriedade pode ajudar a escolher melhor o papel executado pelos nós de forma a melhorar o desempenho de redes com dinâmica temporal.*

1. Introdução

A importância de um nó para a rede pode ser quantificada utilizando-se métricas de centralidade da teoria de grafos, que geralmente associam a importância do nó à sua

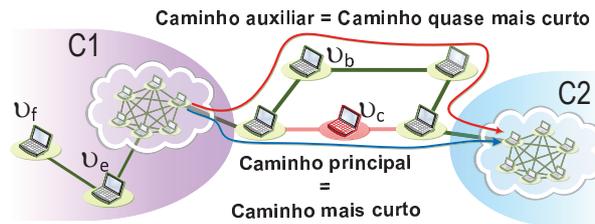


Figura 1. Exemplo de topologia de rede na qual as métricas típicas de centralidade de intermediação podem falhar quando tentam capturar a importância de nós críticos que participam mais de caminhos quase mais curtos. As nuvens representam qualquer tipo de topologia de rede conectada.

posição topológica na rede [Freeman 1978, Hui et al. 2008, Wehmuth and Ziviani 2013]. Os nós mais centrais em uma rede de computadores podem ser utilizados, por exemplo, para exercer funções de controle ou para disseminação de conteúdo [Kiss and Bichler 2008, Thilakarathna et al. 2013, Bouet et al. 2015]. Uma métrica de centralidade comumente utilizada é a intermediação, que atribui importância a um nó conforme o número de caminhos dos quais ele participa [Freeman 1978]. Dentre as métricas de intermediação existentes, uma das mais populares é a intermediação tradicional, que considera apenas a participação em caminhos mais curtos, e que pode beneficiar diversos protocolos de rede [Dolev et al. 2010, Giles et al. 2015, Yim et al. 2016, Jain 2016]. No entanto, o uso apenas desses caminhos como medida de importância dos nós pode acabar acarretando em um desperdício de potencial. Os nós em caminhos alternativos próximos aos caminhos mais curtos são subestimados e, conseqüentemente, impedidos de exercer um papel de maior importância para a rede. Na prática, esses nós podem ser bons candidatos para a manutenção da conectividade da rede em caso de falha de um nó mais central. A dúvida quanto à utilização apenas de caminhos mais curtos para quantificar a importância de um nó já foi debatida em diversos trabalhos na literatura [Freeman et al. 1991, Brandes and Fleischer 2005, Newman 2005, Borgatti and Everett 2006, Jiang et al. 2009, Opsahl et al. 2010, Medeiros et al. 2016a].

Neste artigo analisa-se uma nova versão da métrica de intermediação por espalhamento [Medeiros et al. 2016a], doravante *intermediação por espalhamento múltiplo*. De forma similar à métrica anterior, a nova versão considera os caminhos quase mais curtos dos quais um nó participa para quantificar a sua importância para a rede. A Figura 1 representa a ideia: v_b se beneficia da utilização desses caminhos, aumentando sua importância em relação a v_e , mas mantendo sua distância de v_c que é o mais central dentre os nós destacados. Caso v_c falhe, v_b assume seu papel e o tráfego entre os dois componentes de rede C1 e C2 passa a ser desviado pelo caminho anteriormente quase mais curto. Em sua versão anterior [Medeiros et al. 2016a], a métrica era capaz de elevar a importância de nós que pertenciam a poucos ou nenhum caminho mais curto. No entanto, nós que participavam de um número pequeno de caminhos quase mais curtos eram frequentemente superestimados, tornando-se mais central do que outros que participavam de muitos caminhos mais curtos. Isso ocorria porque apenas os menores caminhos quase mais curtos entre dois nós eram considerados, tornando seus pesos excessivos. Para contornar esse problema, a nova versão proposta neste trabalho extrapola a centralidade de intermediação ponderada para considerar também os múltiplos caminhos quase mais cur-

tos encontrados dentro do limite estabelecido pelo espalhamento ρ . Em suma, essa nova versão utiliza a proporção de caminhos mais curtos e quase mais curtos dos quais um nó intermediário participa e pondera essa proporção por uma relação entre os custos dos caminhos. A adição dos múltiplos caminhos encontrados permite alterar a importância do nó intermediário ao mesmo tempo em que atribui maior importância para caminhos apenas um pouco mais longos do que o mais curto. O impacto da nova intermediação por espalhamento no ranqueamento dos nós está analisado em [Medeiros et al. 2016b].

A investigação conduzida neste trabalho é feita de forma comparativa, utilizando a centralidade de intermediação tradicional e a escalonada por distância. As métricas são aplicadas a uma rede móvel dinâmica, juntamente com a intermediação por espalhamento múltiplo. Primeiramente, a habilidade de reclassificação da intermediação por espalhamento múltiplo é analisada. Em seguida, a métrica é utilizada para estudar a conectividade de uma rede dinâmica em presença de falha nos nós mais centrais da rede. Para tanto, investiga-se os pontos de articulação que se encontram dentre os nós mais centrais e qual é o impacto provocado na vazão da rede caso ocorra uma falha em um desses nós centrais. Os resultados mostram que, (i) utilizando a nova versão da intermediação por espalhamento, um pequeno valor para o espalhamento ρ continua sendo suficiente para capturar a ideia da necessidade de utilização dos caminhos quase mais curtos, enquanto mantém-se a carga computacional baixa, sendo capaz de apontar nós cuja importância deveria ser reavaliada. Além disso, (ii) o número de pontos de articulação existentes dentre as posições classificatórias é sempre menor ou igual para a métrica estendida quando comparada à intermediação tradicional. Isso significa que menos pontos críticos existem dentre os nós mais centrais da rede. Por fim, de forma geral, (iii) a vazão da rede sofre menor redução quando ocorre falha nos nós mais bem classificados pela nova versão da intermediação por espalhamento. De forma resumida, as contribuições deste trabalho são:

- a identificação da necessidade de considerar também os múltiplos caminhos quase mais curtos existentes, além do menor deles, para determinar a importância de um nó para a rede;
- a proposta de uma nova versão da intermediação por espalhamento [Medeiros et al. 2016a] que leva em conta os múltiplos caminhos, chamada de intermediação por espalhamento múltiplo;
- a análise da conectividade de uma rede móvel dinâmica através de comparações entre a intermediação por espalhamento múltiplo e as intermediações tradicional e escalonada por distância.

Este artigo está organizado como segue. Na Seção 2 são introduzidas as definições e a notação utilizadas. A Seção 3 contextualiza este trabalho. A Seção 4 apresenta e formaliza a nova métrica de centralidade de intermediação por espalhamento. A avaliação da métrica reformulada e a sua aplicação ao estudo de conectividade são discutidas na Seção 5. Por fim, a Seção 6 conclui este artigo e apresenta os trabalhos futuros.

2. Notação e Definições

Uma rede pode ser modelada como um grafo ponderado $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$, em que \mathcal{V} e \mathcal{E} são os conjuntos de nós e enlaces, respectivamente, e o peso ω representa o custo de um enlace. Um nó v_i está conectado ao seu vizinho v_j através de um enlace $\varepsilon_{i,j}$ de custo

$\omega_{i,j} \in \mathbb{R}_+^*$. O enlace $\varepsilon_{j,i}$ existe, também, se a rede for não direcionada ou se v_j também está conectado ao vizinho v_i . Neste artigo, apenas grafos não direcionados são utilizados.

Um *caminho* $p_{i,j}$ entre uma origem v_i e um destino v_j é uma sequência ordenada de nós distintos na qual qualquer par consecutivo de nós está conectado por um enlace, não podendo conter laços. O *custo total*, $\delta_{i,j}$, do caminho $p_{i,j}$ é dado pela soma dos custos dos enlaces adjacentes $\varepsilon_{x,y}$ entre v_i e v_j . Neste trabalho utiliza-se pesos unitários para os enlaces, sem perda de generalidade, a fim de possibilitar o uso de algoritmos de menor complexidade temporal. Assim, $\delta_{i,j} \in \mathbb{N}^*$, sendo igual ao número de saltos entre v_i e v_j . O *caminho mais curto* $p_{i,j}^*$ apresenta o menor custo total $\delta_{i,j}^*$ e o número de caminhos mais curtos existentes entre um par de nós é dado por $n_{i,j}^*$. O total desses caminhos que passa por um intermediário v_k é representado por $n_{i,j}^*(v_k)$.

Os conceitos de *caminho quase mais curto* e *espalhamento* já foram introduzidos na literatura [Medeiros et al. 2016a]. Em suma, o espalhamento ρ é a diferença máxima tolerável entre os custos de um caminho qualquer e do caminho mais curto, ou seja, $\rho = \delta_{i,j} - \delta_{i,j}^*$. Assim, um caminho é dito quase mais curto se $\delta_{i,j} - \delta_{i,j}^* \leq \rho$. O espalhamento limita a profundidade de busca por caminhos, evitando a explosão do número de possibilidades, e excluindo caminhos muito mais longos do que o mais curto. Essa consideração surge do fato de que a vazão da informação se concentra em caminhos de custo não muito superior ao do caminho mais curto. O número de caminhos quase mais curtos entre dois nós é dado por $n_{i,j}$ e o total desses caminhos que passa por um intermediário v_k é dado por $n_{i,j}(v_k)$. Esses caminhos podem aumentar a importância de nós ignorados ou subestimados quando apenas os caminhos mais curtos são considerados.

Modelar redes como grafos permite analisar propriedades interessantes para redes de computadores, como a conectividade. Uma rede é *conexa* se existe pelo menos um caminho entre todos os pares de nós, e é *biconexa* se existem pelo menos dois caminhos disjuntos entre eles. O estudo da conectividade permite identificar nós críticos para a rede, que ao falharem podem desconectá-la ou aumentar o número de componentes conexos. Quando isso ocorre, esses nós são ditos *pontos de articulação*. Formalmente, v_a é um ponto de articulação se existirem dois vértices $v_i, v_j \in \mathcal{V}$, com $v_i \neq v_j \neq v_a$, tal que todos os caminhos $p_{i,j}$ contêm v_a , tenham eles custo $\delta_{i,j}$ ou $\delta_{i,j}^*$. A presença desses pontos na rede pode ser identificada determinando-se se ela é biconexa. Caso negativo, existe pelo menos um ponto de articulação. Esses pontos constituem uma grave vulnerabilidade, uma vez que uma única falha em um deles pode dividir a rede em vários componentes conexos.

3. Centralidade, Intermediação e Pontos de Articulação

O estudo da conectividade da rede é de fundamental importância para o projeto de redes tolerantes a falhas. O principal objetivo desse estudo é identificar os pontos críticos da rede para que, ou eles possam ser protegidos, ou passem a não ser mais críticos através de modificações na topologia. Uma possível forma de protegê-los é identificar todos eles, o que pode ser uma tarefa não trivial [Zhang and Sterbenz 2014]. Intuitivamente, espera-se que as métricas de centralidade tenham algum tipo de relação com a identificação desses pontos [Ausiello et al. 2013], uma vez que elas são bastante utilizadas para quantificar a importância de um nó para a rede. De fato, apesar de não identificar todos eles, a literatura confirma que, em geral, pontos de articulação apresentam valores de intermediação muito maiores do que nós ordinários [Ausiello et al. 2013]. Alternati-

vamente, em caso de um ataque externo, a não identificação desses nós de certa forma já é uma medida de proteção. Por exemplo, quando os nós mais centrais da rede exercem alguma função de controle, o ideal é que nenhum deles seja um ponto de articulação, uma vez que esses nós seriam visados por atacantes. Assim, se a localização das funções é feita utilizando alguma métrica de centralidade, quanto menos pontos de articulação estiverem dentre os nós mais centrais da rede, maior será a resistência dessa rede a falhas.

Alguns exemplos de centralidade são o grau, a proximidade e a intermediação [Freeman 1978, Hui et al. 2008, Opsahl et al. 2010, Wehmuth and Ziviani 2013]. Enquanto o grau se relaciona com a popularidade de um nó, a proximidade diz respeito à velocidade com que um nó consegue espalhar ou acessar recursos, por exemplo, informação. Este trabalho foca na intermediação, que está relacionada ao controle que um nó pode exercer sobre os fluxos entre outros nós da rede [Freeman 1978]. Essa métrica foi introduzida por Freeman [Freeman 1977], com base em intuições reveladas em trabalhos anteriores que utilizavam caminhos para determinar a importância de um nó [Bavelas 1948, Shimbel 1953, Shaw 1954, Cohn and Marriott 1958].

A ideia da *centralidade de intermediação tradicional* é a de que quanto mais um nó v_k participa de caminhos mais curtos entre outros nós, mais central ele é. Consequentemente, esses nós estão posicionados de forma estratégica para influenciar a rede através do controle do fluxo de informação entre pares. Freeman assume que uma mensagem passa por um dos caminhos mais curtos entre dois nós com probabilidade igual a $1/n_{i,j}^*$ [Freeman 1977]. Considerando todos os pares de nós na rede, a chance de v_k fazer parte de um dos caminhos mais curtos escolhido aleatoriamente entre um par qualquer de nós define matematicamente a intermediação tradicional de v_k , $B_{trad}(v_k)$, como a seguir.

$$B_{trad}(v_k) = \sum_{i \in |\mathcal{V}|} \sum_{j \in |\mathcal{V}|} \frac{n_{i,j}^*(v_k)}{n_{i,j}^*}, \text{ onde } i \neq k, j \neq i \text{ e } j \neq k.$$

A intermediação tradicional limita-se a grafos simples e não considera o peso dos enlaces, ignorando, assim, o custo fim-a-fim do caminho entre dois nós. Esse custo é importante porque caminhos mais longos são menos valiosos para serem controlados ou podem não ser realistas em algumas redes [Borgatti and Everett 2006]. Além disso, considerar também que o fluxo da informação sempre passa por caminhos mais curtos pode não ser verdade em alguns casos. Por exemplo, rumores se espalham de forma aleatória ou podem ser canalizados intencionalmente por intermediários específicos [Stephenson and Zelen 1989]. Por essa razão, diversos trabalhos já questionaram o uso apenas de caminhos mais curtos para definir a importância de um nó [Freeman et al. 1991, Brandes and Erlebach 2005, Brandes and Fleischer 2005, Newman 2005, Borgatti and Everett 2006, Geisberger et al. 2008, Jiang et al. 2009, Opsahl et al. 2010, Medeiros et al. 2016a].

A limitação da intermediação tradicional em relação ao comprimento do caminho fim-a-fim é solucionada pela *centralidade de intermediação escalonada por distância* [Borgatti and Everett 2006] (*Distance Scaled Betweenness* – DS), que pondera a frequência com que um nó participa dos caminhos mais curtos utilizando com o inverso do comprimento do caminho. A métrica DS é definida a seguir.

$$B_{dist}(v_k) = \sum_{i \in |\mathcal{V}|} \sum_{j \in |\mathcal{V}|} \left(\frac{1}{\Delta L_{i,j}^*} \times \frac{n_{i,j}^*(v_k)}{n_{i,j}^*} \right).$$

4. Centralidade de Intermediação por Espalhamento Múltiplo

Este trabalho propõe uma nova versão da centralidade de intermediação por espalhamento [Medeiros et al. 2016a] para considerar os múltiplos caminhos existentes entre dois nós quaisquer da rede, e não apenas os caminhos mais curtos e os menores caminhos quase mais curtos. À versão estendida dá-se o nome de intermediação por espalhamento múltiplo. A inclusão desses caminhos adicionais permite capturar melhor o potencial dos nós intermediários que podem ser críticos para a rede, refinando a contribuição de cada caminho individual. Esses nós costumam ser desprezados pelas métricas de intermediação baseadas em caminhos mais curtos por não participarem de um número suficientemente grande de caminhos mais curtos. Por outro lado, alguns desses nós podem ser superestimados ao se considerar apenas os mais curtos dentre os caminhos quase mais curtos [Medeiros et al. 2016a]. Esse efeito se repete quando não são atribuídos pesos adequados aos caminhos, como é o caso da k -intermediação [Jiang et al. 2009], que não pondera os caminhos de acordo com o seu custo.

A Figura 1 ilustra o problema dos nós subestimados pelas métricas de intermediação típicas. Os nós v_c e v_b são cruciais para a manutenção da conexão entre os componentes de rede C1 e C2, enquanto v_e conecta um nó da borda ao restante da rede. Segundo a intermediação tradicional, v_c possui maior centralidade dentre os três nós analisados, seguido por v_e e finalmente por v_b . Essa métrica desconsidera o potencial de v_b de assumir um papel muito mais importante do que o de v_e considerando-se a conectividade da rede inteira, principalmente em caso de falha do nó v_c . Nós em posições semelhantes à posição topológica de v_b são bons candidatos para agir como nós substitutos de outros nós mais importantes, caso esses venham a falhar, mantendo a rede conectada a um custo apenas um pouco mais elevado. O grau de importância dos nós v_b , v_c e v_e para a rede é reforçado pela centralidade de intermediação escalonada por distância. No entanto, ao considerar caminhos quase mais curtos, é possível modificar essa classificação, aumentando a importância de v_b , que se torna mais importante do que v_e (Tabela 1).

Um segundo problema encontrado ao analisar redes utilizando métricas de intermediação baseadas apenas em caminhos mais curtos é a classificação como iguais de nós topologicamente diferentes. A Figura 2 ilustra esse problema. Considerando-se o conjunto de nós $\{v_a, v_b, v_c, v_d, v_e\}$, está claro que os nós v_a e v_e são topologicamente semelhantes, assim como v_b e v_d . Apesar do nó v_c ser diferente dos outros, tanto a intermediação tradicional quanto a escalonada por distância (DS) o classificam como igual aos nós v_a e v_e . Ao se considerar os múltiplos caminhos quase mais curtos, v_c passa a ser diferenciado, como visto na Tabela 2. Isso ocorre porque nenhum caminho mais curto passa por v_c , mas os caminhos quase mais curtos existentes entre quaisquer pares de nós na rede sempre têm v_c como intermediário.

Tabela 1. Comparação entre as classificações dos nós destacados na Figura 1 utilizando as métricas de intermediação tradicional (Trad), escalonada por distância (DS) e por espalhamento múltiplo (B_ρ).

Node	Trad	DS	B_ρ ($\rho = 3$)
v_c	63,0	12,1	117,10
v_b	9,0	2,5	39,56
v_e	17,0	3,9	35,62

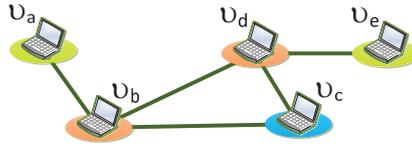


Figura 2. De acordo com sua posição topológica, os nós v_a e v_e são iguais, assim como v_b e v_d . O nó v_c é diferente de todos os outros, mas tanto a intermediação tradicional quanto a escalonada por distância consideram que v_c está posicionado de forma semelhante a v_a e v_e .

4.1. Formalização

Assim como a centralidade de intermediação por espalhamento [Medeiros et al. 2016a], o custo máximo do caminho quase mais curto considerado depende de ρ . Se $\rho = C$, com $C \in \mathbb{N}$, uma vez que neste trabalho os custos são unitários para permitir o uso de algoritmos de menor complexidade temporal, o custo máximo considerado para esses caminhos será dado por $\delta_{i,j}^* + C$. Assim, um caminho de custo $\delta_{i,j}^* + (C + \varphi)$, com $\varphi \in \mathbb{N}^*$, será ignorado na computação da métrica. Contudo, diferentemente da intermediação por espalhamento, que considerava apenas os menores caminhos quase mais curtos, neste trabalho estende-se a métrica para incluir todos os caminhos com custo $\delta_{i,j} \leq \delta_{i,j}^* + C$. A métrica estendida considera tanto o custo dos caminhos quanto a ideia de utilizar outros caminhos além dos mais curtos, além de supor que a informação conhece seu destino final e *prefere* trafegar por caminhos menos longos para chegar até ele. Para tal, a centralidade de intermediação por espalhamento múltiplo pondera a contribuição dos caminhos para a importância do nó intermediário v_k de forma proporcional aos custos desses caminhos, utilizando a razão entre os custos do caminho mais curto, $\delta_{i,j}^*$, e do caminho quase mais curto que passa por v_k , $\delta_{i,j} = \delta_{i,k} + \delta_{k,j}$. Assim, atribui-se maior importância para nós em caminhos de menor comprimento. A nova versão da métrica, chamada a partir deste trabalho de intermediação por espalhamento múltiplo, está formalizada na Equação 1. Nota-se que para $\rho = 0$, $\delta_{i,j} = \delta_{i,j}^*$, e apenas os caminhos mais curtos são considerados.

$$B_\rho(v_k) = \sum_{i \in |\mathcal{V}|} \sum_{j \in |\mathcal{V}|}^{\delta_{i,k} + \delta_{k,j} - \delta_{i,j}^* \leq \rho} \left(\frac{n_{i,j}^*(v_k) + n_{i,j}(v_k)}{n_{i,j}^* + n_{i,j}} \times \frac{\delta_{i,j}^*}{\delta_{i,k} + \delta_{k,j}} \right). \quad (1)$$

4.2. Propriedades

A centralidade de intermediação por espalhamento múltiplo possui complexidade de tempo polinomial comparável a das métricas discutidas na Seção 3 e apresenta as

Tabela 2. Comparação entre as intermediações dos nós na Figura 2. Apenas a intermediação por espalhamento múltiplo captura a importância de v_c .

Node	Trad	DS	$B_\rho (\rho = 3)$
v_a	0,0	0,0	0,0
v_b	3,0	1,3	2,7
v_c	0,0	0,0	1,3
v_d	3,0	1,3	2,7
v_e	0,0	0,0	0,0

propriedades a seguir: (i) considera o número de múltiplos caminhos, tanto mais curtos quanto quase mais curtos; (ii) aumenta com a participação de v_k em ambos os caminhos mais curtos e quase mais curtos; (iii) prioriza caminhos menores diminuindo a contribuição de caminhos mais longos através de uma relação entre custos; e (iv) aumenta com a centralidade do nó. Note que, neste trabalho, um nó é mais central se ele participa de múltiplos caminhos, sejam eles mais curtos ou quase mais curtos. Essa consideração advém do fato de que nós que participam em vários caminhos quase mais curtos não devem ser descartados apenas por não estarem em caminhos mais curtos, uma vez que eles podem ganhar importância. Por exemplo, nós que estão sempre próximos ao caminho mais curto podem substituir outros nós mais importantes caso eles falhem.

5. Avaliação

O impacto e a relevância da métrica proposta é avaliado utilizando-se um conjunto de dados dinâmico. A investigação é realizada para valores de espalhamento $\rho \leq 5$, ou seja, consideram-se os múltiplos caminhos quase mais curtos que obedecem a restrição imposta pela inequação $\delta_{i,j} \leq \delta_{i,j}^* + \{1, 2, 3, 4, 5\}$.

A avaliação da métrica é realizada através de comparações com as intermediações tradicional e escalonada por distância, aplicadas ao conjunto de dados TAPAS-Cologne [Uppoor and Fiore 2011], que modela o tráfego de veículos na cidade de Colônia, na Alemanha [Uppoor and Fiore 2011], durante 24 horas. No momento em que este trabalho foi elaborado, apenas um subconjunto de 2 horas estava disponível publicamente na Internet. Desse subconjunto analisa-se 6 amostras tomadas a cada 10 segundos. Cada um dos nós representa um veículo no cenário e um enlace existe entre dois nós se eles estiverem a menos de 50 metros de distância um do outro. O número de nós nas amostras varia entre 1.584 e 1.916 e o número de enlances não direcionados, entre 1.573 e 2.044, dependendo da amostra. Note que a variação no raio de alcance rádio modifica a densidade do grafo da rede, aumentando ou diminuindo o número de enlances e, conseqüentemente, o número de possíveis caminhos entre dois nós. Inicialmente, verifica-se quão próximas da intermediação tradicional estão as intermediações escalonada por distância (DS) e por espalhamento múltiplo. Supõe-se que devido ao modo como essas métricas são formalizadas, elas devem ser semelhantes à intermediação tradicional. Contudo, deseja-se que, mesmo sendo fortemente correlacionadas, essas métricas sejam capazes de apontar nós que deveriam ser reclassificados. Em seguida, investiga-se a conectividade da rede através do estudo dos pontos de articulação existentes na rede, classificados segundo as mesmas métricas anteriores. Por fim, através de simulação, avalia-se o desempenho da rede TAPAS-Cologne quanto à vazão média alcançada em presença de falha única, considerando-se os nós mais importantes da rede de acordo com cada métrica.

Os resultados estão divididos como segue: o primeiro refere-se à habilidade de identificação de nós cuja importância deveria ser reavaliada. Já o segundo relaciona-se com o impacto das métricas no desempenho de uma rede dinâmica na presença de falhas.

5.1. Identificação de nós reclassificáveis

Investiga-se a relação entre a intermediação tradicional e as intermediações escalonada por distância e por espalhamento múltiplo, a fim de verificar como os requisitos adicionais de cada métrica influenciam a semelhança entre elas. Essa análise permite,

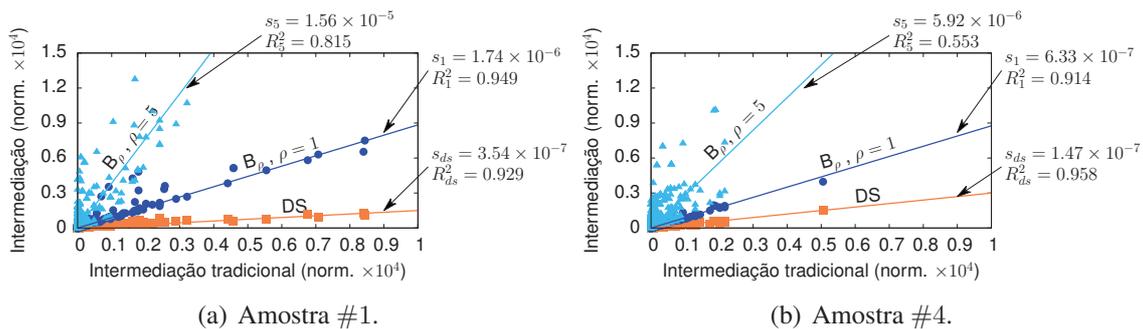


Figura 3. As métricas analisadas são fortemente correlacionadas com a intermediação tradicional, sendo essa correlação mais forte para a intermediação escalonada por distância. A intermediação por espalhamento múltiplo consegue identificar mais nós que deveriam ser reavaliados.

também, averiguar se as métricas conseguem destacar nós que deveriam ser reclassificados. Os resultados obtidos para as amostras #1 e #3 do conjunto de dados são mostrados na Figura 3, onde o eixo x representa a intermediação tradicional normalizada e cada curva refere-se a uma das outras duas métricas, também normalizadas. Note que os eixos na Figura 3 estão escalonados para melhor visualização. Apenas as curvas referentes aos valores de espalhamento 1 e 5 são mostradas, para maior clareza. As curvas para os outros valores de ρ encontram-se entre as duas curvas exibidas.

A Figura 3 mostra que todas as métricas são fortemente correlacionadas com a intermediação tradicional, uma vez que o coeficiente de determinação (R^2) é elevado. A correlação é mais forte para a intermediação escalonada por distância, seguida pela por espalhamento múltiplo utilizando $\rho = 1$. A correlação diminui com o aumento do espalhamento, devido ao crescimento da participação dos nós em caminhos quase mais curtos. Isso faz com que mais nós sejam apontados pela métrica como sub ou superestimados, sugerindo a necessidade de reclassificação. Esse comportamento se repete em todas as amostras tomadas. Em todo caso, um espalhamento $\rho = 1$ já é suficiente para destacar alguns nós, o que pode ser notado pela dispersão dos pontos em torno das curvas ou, matematicamente, através da análise combinada do desvio padrão de cada ajuste de curva e do coeficiente R^2 . Quanto maior o desvio padrão e menor o R^2 , mais nós reclassificáveis podem ser identificados. No entanto, R^2 não pode ser menor do que 0,35, uma vez que, assim sendo, a correlação entre as métricas passaria a ser moderada, o que significa que elas medem duas características diferentes, o que não é do interesse deste trabalho.

Note que a possibilidade de reclassificação dos nós surge neste cenário apenas porque existem caminhos alternativos entre dois nós na rede. Caso poucos caminhos múltiplos existam, a intermediação por espalhamento múltiplo não se beneficia da existência de muitos caminhos quase mais curtos, de forma que passa a ser praticamente igual à intermediação tradicional, independentemente do valor de espalhamento.

5.2. Impacto no desempenho da rede

A conectividade da rede é analisada neste trabalho através do estudo dos pontos de articulação e do número de componentes conexos existentes. A Figura 4 mostra o número de pontos de articulação presentes em cada amostra estudada do conjunto de dados TAPASCologne, considerando-se as 5 posições que contêm os nós mais bem classifi-

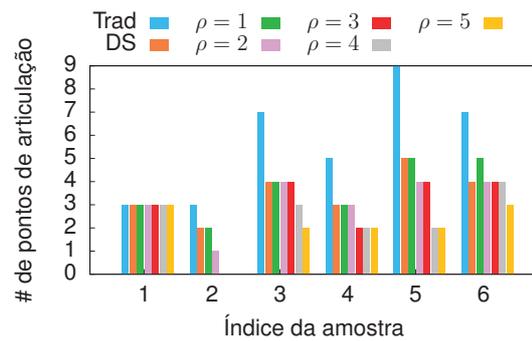


Figura 4. O número de pontos de articulação existentes em cada amostra do cenário dinâmico considerando-se as top #5 posições é, em geral, menor para a intermediação escalonada por distância (DS) e para a por espalhamento múltiplo, quando comparadas à intermediação tradicional.

cados segundo cada métrica analisada, doravante denominadas top #5 posições. Observa-se que o número de pontos de articulação dentre as top #5 posições muda dependendo da amostra, mas fica claro que tanto a intermediação escalonada por distância (DS) quanto a por espalhamento múltiplo apresentam um número menor de pontos de articulação em posições de elevada importância. A Tabela 3 mostra o número de componentes existentes anteriormente a qualquer falha na rede para cada amostra estudada, e após uma falha única em um ponto de articulação escolhido dentre as top #5 posições. Em todas as amostras, pelo menos um componente a mais é criado caso um dos pontos de articulação falhe, sendo que na amostra #4 até dois componentes a mais podem ser criados, dependendo da posição classificatória do nó.

Nós que apresentam elevada intermediação são responsáveis por uma grande quantidade de fluxos entre pares de nós na rede. Dependendo da topologia dessa rede, uma falha que ocorra em um único desses nós pode afetar inúmeros fluxos, principalmente se esse nó também for um ponto de articulação. Isso pode reduzir drasticamente o desempenho da rede. A severidade dessas falhas é investigada utilizando-se o simulador de redes NS-3. O objetivo da simulação é analisar para cada métrica, em relação à intermediação tradicional, o comportamento da vazão média obtida para a rede antes e após a ocorrência de uma falha única em um nó que apresenta intermediação elevada. Considera-se que todos os nós têm a intenção de se comunicar com todos os outros nós da rede, mesmo que não estejam em um mesmo componente conectado. Contudo, se dois nós em componentes distintos tentarem se comunicar, não serão bem sucedidos, uma vez que não existe caminho entre eles. As mensagens trocadas devem possuir o mesmo tamanho para todos os pares de nós. Contudo, o tamanho específico escolhido não altera a conclusão. Supõe-se, também, que a topologia da rede muda a cada 10 segundos apenas,

Tabela 3. Número de componentes existentes em cada amostra do conjunto de dados TAPASCologne quando não existem falhas na rede e quando um único nó ponto de articulação falha.

	#1	#2	#3	#4	#5	#6
Sem falha	550	549	546	576	596	578
Falha única	551	550	547	577 / 578	597	579

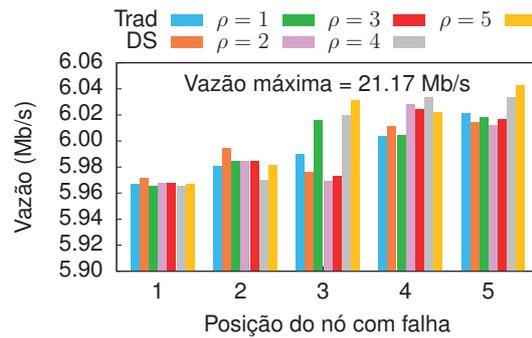


Figura 5. Em relação às intermediações tradicional e escalonada por distância (DS), a falha em nós classificados nas top #5 posições da intermediação por espalhamento múltiplo é, em geral, menos prejudicial à vazão média da rede.

que é o tempo de intervalo entre as amostras tomadas do conjunto de dados TAPAS-Cologne. Por fim, o nó que sofre a falha é escolhido aleatoriamente dentro da posição classificatória investigada. Isto é, ao se investigar o impacto da falha de um nó na posição n , e existem m nós classificados nessa posição, escolhe-se aleatoriamente um desses m nós para falhar. Os resultados relatados nesta seção são dependentes do ranqueamento obtido utilizando cada métrica. Dessa forma, redes mais esparsas, que possuem menos alternativas de caminhos entre pares de nós, devem apresentar comportamento semelhante ao encontrado para a intermediação tradicional. Caso a oferta de caminhos alternativos seja maior, a diferença entre as métricas será mais acentuada.

A Figura 5 mostra o resultado obtido para a simulação utilizando-se a classificação gerada pelas intermediações tradicional, escalonada por distância e por espalhamento múltiplo. O eixo x representa a posição classificatória do nó no qual ocorreu a falha, e o eixo y é a vazão total média da rede, em Mb/s, calculada para os 60 segundos de simulação, referentes ao tempo total das amostras. Note que o nó selecionado para falhar pode ou não ser um ponto de articulação da rede. A probabilidade de o nó em falha ser também um ponto de articulação difere para cada métrica e está mostrada na Tabela 4. Nota-se que, para cada amostra, essa probabilidade é sempre igual ou maior para a intermediação tradicional do que para as outras métricas, sendo maior na maioria dos casos. Além disso, a probabilidade de o nó em falha também ser ponto de articulação diminui com o aumento do espalhamento ρ e, em geral, é maior para a intermediação escalonada por distância (DS) do que para a intermediação por espalhamento múltiplo.

Tabela 4. A probabilidade de um nó em falha também ser ponto de articulação é maior na maioria dos casos para intermediação tradicional do que para as outras métricas.

Amostra	Trad	DS	B_ρ ($\rho = 1$)	B_ρ ($\rho = 2$)	B_ρ ($\rho = 3$)	B_ρ ($\rho = 4$)	B_ρ ($\rho = 5$)
#1	0,60	0,60	0,60	0,60	0,60	0,60	0,60
#2	0,43	0,33	0,33	0,17	0,0	0,0	0,0
#3	0,88	0,80	0,80	0,80	0,80	0,38	0,25
#4	0,72	0,60	0,60	0,60	0,40	0,40	0,40
#5	1,0	1,0	1,0	0,67	0,80	0,33	0,33
#6	1,0	0,80	1,0	0,80	0,80	0,80	0,43

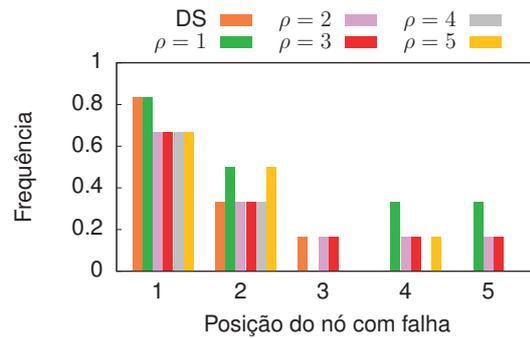


Figura 6. Os nós classificados na primeira posição frequentemente são os mesmos em relação ao ranqueamento da intermediação tradicional, considerando-se todas as amostras tomadas do conjunto de dados de TAPASCologne, enquanto nas posições seguintes essa frequência diminui consideravelmente.

Destaca-se na Figura 5 a vazão máxima alcançada, que ocorre na ausência de falhas na rede. Seu valor é aproximadamente igual a 21,17 Mb/s e é utilizado nesta análise como valor de referência. De uma forma geral, observa-se que quanto menor a importância do nó, isto é, quanto mais distante da primeira posição classificatória, menor é o impacto da falha na vazão média da rede. No entanto, essa variação é bem pequena, por exemplo, a ocorrência de falha no nó mais importante, que também é um ponto de articulação, reduz a vazão total em aproximadamente 28,18%. Já para a última posição, a redução varia entre 28,40% e 28,54%, dependendo da métrica analisada. Verifica-se ainda, que para a intermediação por espalhamento múltiplo, a perda na vazão total média é sempre muito próxima ou menor do que a perda em caso de falha de nós importantes segundo a classificação da intermediação tradicional. É importante notar que frequentemente os nós classificados como mais importantes geralmente são os mesmos e, por essa razão, a variação da vazão dentro de uma mesma posição classificatória é pequena. A Figura 6 mostra com que frequência durante os 60 segundos de simulação os nós nas top #5 posições mudam. Por exemplo, em 83% do tempo o nó classificado na primeira posição é o mesmo tanto para a intermediação tradicional quanto para a escalonada por distância. Em relação à intermediação por espalhamento múltiplo, esse valor cai para 67% do tempo, mas ainda assim não é suficiente para fazer alguma diferença na vazão média da rede. Somente para um nível de coincidência menor do que 50% é que observa-se influência na vazão média da rede.

6. Conclusões

Este trabalho reutiliza a ideia da intermediação por espalhamento, estendendo-a para considerar os múltiplos caminhos quase mais curtos existentes entre pares de nós. O objetivo é melhorar a avaliação da importância dos nós para a rede, mesmo quando eles não participam de caminhos mais curtos. Esses nós podem ser críticos para a operação da rede, reduzindo a reorganização e os custos envolvidos após falha de um nó central. O impacto da métrica foi avaliado através de comparações com outras métricas de intermediação, aplicadas a um cenário de rede dinâmico. Os resultados mostraram que a nova versão da intermediação por espalhamento é capaz de apontar nós que deveriam ser reavaliados. Ao reclassificá-los, é possível reduzir o número de pontos de articulação em posições centrais na rede, considerando-se os nós mais bem classifica-

dos. Em caso de falha nesses nós, simulações mostraram que, em geral, seguindo o ranqueamento fornecido pela intermediação por espalhamento múltiplo, a vazão média total da rede é menos prejudicada do que quando se utiliza as outras métricas. A probabilidade do nó em falha também ser um ponto de articulação é frequentemente menor para a intermediação por espalhamento múltiplo e, comparada à intermediação tradicional, em geral os nós classificados na primeira posição são os mesmos, mas para as posições seguintes eles costumam ser diferentes. Como trabalhos futuros, planeja-se estender o algoritmo da nova métrica para atender redes ponderadas e comparar o impacto no desempenho da rede quando ataques ocorrerem apenas nos pontos de articulação e apenas nos nós mais centrais que não sejam críticos. Também planeja-se aplicar a métrica a redes reais através de experimentos em plataformas realísticas, e estudar a sua relevância em diferentes casos de uso.

Agradecimento

Os autores agradecem ao CNPq, à CAPES e à Faperj pelo financiamento parcial deste trabalho.

Referências

- Ausiello, G., Firmani, D., e Laura, L. (2013). The (betweenness) centrality of critical nodes and network cores. Em *IWCMC*, pp. 90–95.
- Bavelas, A. (1948). A mathematical model for group structures. *Human Organization*, 7(3):16–30.
- Borgatti, S. P. e Everett, M. G. (2006). A graph-theoretic perspective on centrality. *Social Networks*, 28(4):466 – 484.
- Bouet, M., Leguay, J., Combe, T., e Conan, V. (2015). Cost-based placement of vDPI functions in NFV infrastructures. *International Journal of Network Management*, 25(6):490–506.
- Brandes, U. e Erlebach, T. (2005). *Network Analysis: Methodological Foundations*. Springer, 1a Edição.
- Brandes, U. e Fleischer, D. (2005). Centrality measures based on current flow. Em *STACS*, pp. 533–544.
- Cohn, B. S. e Marriott, M. (1958). Networks and centres of integration in Indian civilization. *Journal of Social Research*, 1:1–9.
- Dolev, S., Elovici, Y., e Puzis, R. (2010). Routing betweenness centrality. *JACM*, 57(4):25:1–25:27.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239.
- Freeman, L. C., Borgatti, S. P., e White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, 13(2):141–154.

- Geisberger, R., Sanders, P., e Schultes, D. (2008). Better approximation of betweenness centrality. Em *ALLENEX*, pp. 90–100.
- Giles, A. P., Georgiou, O., e Dettmann, C. P. (2015). Betweenness centrality in dense random geometric networks. Em *ICC*, pp. 6450–6455.
- Hui, P., Crowcroft, J., e Yoneki, E. (2008). Bubble rap: Social-based forwarding in delay tolerant networks. Em *ACM Mobihoc*, pp. 241–250.
- Jain, A. (2016). Betweenness centrality based connectivity aware routing algorithm for prolonging network lifetime in wireless sensor networks. *Wireless Networks*, 22(5):1605–1624.
- Jiang, K., Ediger, D., e Bader, D. A. (2009). Generalizing k-betweenness centrality using short paths and a parallel multithreaded implementation. Em *ICPP*, pp. 542–549.
- Kiss, C. e Bichler, M. (2008). Identification of influencers - measuring influence in customer networks. *Decision Support Systems*, 46(1):233–253.
- Medeiros, D. S. V., Campista, M. E. M., Mitton, N., de Amorim, M. D., e Pujolle, G. (2016a). Intermediação por espalhamento: Caminhos quase mais curtos também importam. Em *SBRC*, pp. 967–980.
- Medeiros, D. S. V., Campista, M. E. M., Mitton, N., de Amorim, M. D., e Pujolle, G. (2016b). Spread betweenness centrality: The power of quasi-shortest paths. Relatório técnico, GTA – PEE / COPPE / UFRJ. <http://www.gta.ufrj.br/ftp/gta/TechReports/MCM16c.pdf>.
- Newman, M. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27(1):39 – 54.
- Opsahl, T., Agneessens, F., e Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251.
- Shaw, M. E. (1954). Group structure and the behavior of individuals in small groups. *The Journal of Psychology*, 38(1):139–149.
- Shimbel, A. (1953). Structural parameters of communication networks. *Bulletin of Mathematical Biophysics*, 15(4):501–507.
- Stephenson, K. e Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks*, 11(1):1–37.
- Thilakarathna, K., Viana, A. C., Seneviratne, A., e Petander, H. (2013). Mobile social networking through friend-to-friend opportunistic content dissemination. Em *ACM MobiHoc*, pp. 263–266.
- Uppoor, S. e Fiore, M. (2011). Large-scale urban vehicular mobility for networking research. Em *IEEE VNC '11*, pp. 62–69.
- Wehmuth, K. e Ziviani, A. (2013). DACCER: Distributed assessment of the closeness centrality ranking in complex networks. *Computer Networks*, 57(13):2536–2548.
- Yim, J., Ahn, H., e Ko, Y.-B. (2016). The betweenness centrality based geographic routing protocol for unmanned ground systems. Em *IMCOM*, pp. 74:1–74:4.
- Zhang, D. e Sterbenz, J. P. G. (2014). Modelling critical node attacks in MANETs. *LNCS Series*, 8221:127–138.