

Estimando a Centralidade de Intermediação em Redes Ópticas de Telecomunicações

Silvana Trindade¹, Pedro Augusto Pereira Borges^{1,2}, Claunir Pavan¹

¹ Curso de Graduação em Ciência da Computação

² Curso de Graduação em Matemática

Universidade Federal da Fronteira Sul (UFFS) – Chapecó, SC-Brasil

syletri@gmail.com, {pedro.borges, claunir.pavan}@uffs.edu.br

Abstract. *From a set of real optical telecommunication network topologies, an expression is proposed to estimate the betweenness centrality of a network without prior knowledge of its topological structure, having only the number of nodes and links as requirement. Betweenness centrality shows the number of times each node is present in the shortest paths between other nodes. This information is relevant when planning networks, as the topology can have a significant impact on its costs. In this context, the analytical models can be used in evaluation tools as elimination criterion and to forecast costs, accelerating decision making. Experimental results reveal that the proposed expression obtain the value of the betweenness centrality with an average error less than 3%.*

Resumo. *A partir de um conjunto de topologias de redes ópticas reais de telecomunicações, é proposta uma expressão para a estimação da centralidade de intermediação de uma rede, sem a necessidade de conhecer a sua estrutura topológica, apenas com a informação do número de nós e enlaces. Esta medida indica a quantidade de vezes que os nós estão presentes ao longo dos caminhos mínimos entre outros nós. Esta informação é relevante no planejamento preliminar de redes, já que a topologia pode ter um impacto significativo nos custos de capital. Neste contexto, os modelos analíticos podem ser usados em ferramentas de avaliação como critério de eliminação e previsão de custos, acelerando a tomada de decisões. Resultados experimentais revelam que a expressão proposta obtêm o valor da centralidade de intermediação com erro médio menor que 3%.*

1. Introdução

O planejamento de redes ópticas de telecomunicações envolve a identificação dos componentes necessários e a sua quantidade. A partir destas informações pode-se calcular o custo de capital (CapEx). Para o caso em que todas as informações sobre a rede estão disponíveis, incluindo a estrutura topológica e a matriz de tráfego, o planejamento torna-se uma tarefa relativamente fácil, já que existem ferramentas que auxiliam a sua execução. Contudo, na falta de alguma informação, os projetistas devem considerar aproximações para estimar quantidades e custos.

A estimação de quantidades e custos é de grande relevância para os operadores no momento da avaliação preliminar de soluções alternativas. Ela serve como ferramenta para tomadas de decisão [Silva 2010]. Além da fase de planejamento inicial, as estimativas podem também ser usadas para prever o valor de parâmetros da rede sob diversas condições. É o caso da análise de tráfego multi-período. Nos últimos anos, vários pesquisadores apresentaram resultados sobre a estimação de variáveis de mérito em redes de telecomunicações. Em [Bodamer et al. 2004] os autores apresentam um método para estimar o número de transponders em redes multicamada. Em [Korotky 2004, Korotky 2005, Bhardwaj et al. 2005, Labourdette et al. 2005] os autores apresentam uma formulação semi-empírica para calcular um conjunto de variáveis de mérito, sem a necessidade de informações sobre a estrutura topológica da rede. Em [Pinto et al. 2009] é apresentado um modelo estatístico para o cálculo de custos de transmissão e gerenciamento de largura de banda. Em [Pavan et al. 2010] são identificadas, a partir de um conjunto de 29 topologias físicas de redes reais, características relevantes, como o número médio de saltos, coeficiente de proteção e coeficiente de restauro, bem como aproximações para estas variáveis. Em [Routray et al. 2013] os autores apresentam uma expressão para estimar o comprimento médio dos enlaces, baseada apenas na informação sobre a área de cobertura da rede e do número de nós. Este modelo foi utilizado para estimar vários parâmetros de redes ópticas de telecomunicações em [Routray et al. 2014]. Estes modelos contribuem para a ampliação da caracterização de redes ópticas de telecomunicações e permitem maior precisão das previsões.

Neste artigo pretende-se estender o estudo realizado pelos autores apresentados acima, propondo uma aproximação para uma nova variável de mérito em topologias de redes de telecomunicações, a centralidade de intermediação, que mede a quantidade de vezes que um nó está presente nos caminhos mínimos entre os demais pares de nós. No âmbito das redes de telecomunicações esta medida pode representar a importância ou criticidade dos nós. Por exemplo, uma topologia pode ter a necessidade de reencaminhar um conjunto grande de canais em caso de falhas em um nó crítico. Numericamente, conhecida a topologia, a obtenção desta medida em grafos esparsos - como é o caso das redes de telecomunicações - tem complexidade $O(N^2 \log N + NL)$ [Brandes 2001]. Na literatura há estudos que buscam a obtenção de algoritmos mais eficientes para a solução deste problema [Ausiello et al. 2013].

O artigo está organizado da seguinte forma: na Seção 2, é apresentado o conceito de centralidade de intermediação e também o conjunto de redes reais considerado neste estudo. Na Seção 3, é apresentado o método utilizado para a obtenção da expressão. A avaliação do método é apresentada na Seção 4. Por fim, na Seção 5, as considerações finais são apresentadas.

2. Centralidade de Intermediação

A centralidade de intermediação, BC (*Betweenness Centrality*), avalia a quantidade de vezes que um vértice (nó da topologia) está no caminho geodésico entre dois outros vértices. Ou seja, elevados índices de centralidade indicam que o vértice está em uma grande fração dos caminhos mais curtos entre os outros vértices, revelando a sua importância na topologia. De fato, a centralidade será dependente também da estrutura topológica. Para fins de análise preliminar de topologias alternativas, projetistas de redes podem optar por avaliar com mais detalhes as redes com menor centralidade de intermediação, já que este

resultado indica que os menores caminhos estarão mais distribuídos na topologia. Uma topologia com BC alto pode sofrer maior impacto (perda de tráfego) no caso de falhas.

Para um vértice v_k , a intermediação é indicada pelo fluxo entre um vértice fonte i e o seu destino j [Tizghafam and Leon-Garcia 2010].

A centralidade de intermediação de um vértice v_k em um grafo $G = (V, E)$, onde V é o número de vértices, e E o número de arestas (enlaces entre os nós) é dada pela expressão:

$$BC(v_k) = \sum_{i \neq j \neq v_k} \frac{b_{ij}(v_k)}{b_{ij}}, \quad (1)$$

onde b_{ij} representa o número total de caminhos mínimos entre i e j e $b_{ij}(v_k)$ representa o número de caminhos mínimos entre i e j que passam pelo vértice (v_k) [Brandes 2001]. O vértice com maior centralidade de intermediação será o mais central da rede [Ufimtsev and Bhowmick 2013, Freeman 1977].

Para analisar a centralidade de intermediação em topologias de redes reais de telecomunicações, considerou-se um conjunto com 41 topologias reais sobreviventes (com grafo 2-aresta-conexo). A Tabela 1 [Routray et al. 2013] apresenta o número de nós, N , enlaces, L , grau médio, $\langle \delta \rangle$ e a média ponderada da centralidade de intermediação, BC_{mp} .

Quando analisa-se topologias como a *EON*, Figura 1, observa-se que em algumas regiões há uma maior concentração de nós e enlaces. Em redes de telecomunicações, os nós com grau acima da média, são os principais candidatos a nó central da rede. Para

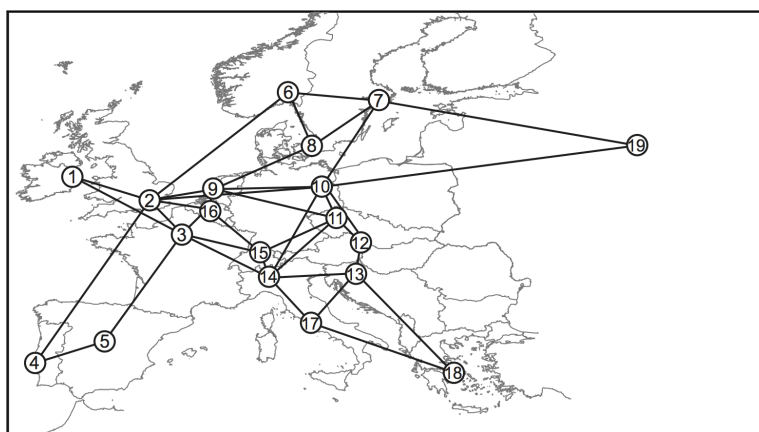


Figura 1. Topologia Física de Rede Óptica da Europa (EON) [Pavan et al. 2010].

a rede exemplo, o nó com índice 10 (de grau 7) é o nó mais central. Considerando o conjunto de redes de referência apresentado na Tabela 1, observou-se que 80% das redes apresentam esta característica, onde um nó com grau máximo (relativo aos demais nós da rede) é o nó mais central na rede. Por este motivo, considerou-se a média ponderada da centralidade de intermediação para representar a rede. A média ponderada é dada pela seguinte expressão:

$$BC_{mp} = \sum_{1 \leq v_k \leq N}^N \frac{BC(v_k)}{\delta(v_k)}, \quad (2)$$

onde N é o número total de nós, sendo BC a soma das centralidades de intermediação obtidas na Equação (1) para cada nó, dividido pelo grau do mesmo. O valor da média ponderada da centralidade de intermediação BC_{mp} apresentado na coluna 6 da Tabela 1 foi obtido pela Equação (2). Na Figura 2(a) é apresentado o grau mínimo, médio e máximo de cada rede, e na Figura 2(b) é apresentada a centralidade mínima, média e máxima de cada rede.

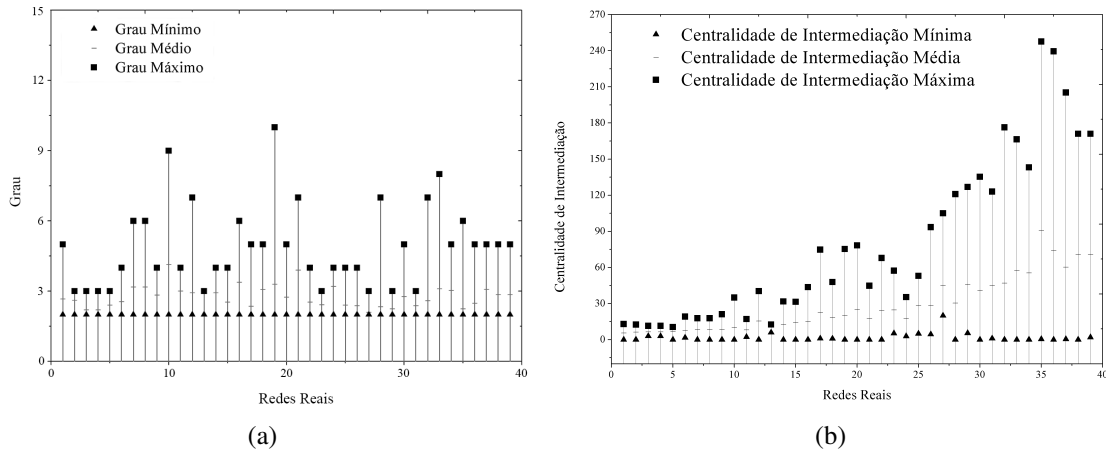


Figura 2. (a) valor Mínimo, Médio e Máximo para o grau nodal e, (b) centralidade de intermediação para 39 topologias. As topologias 40 e 41 (1) não são apresentadas no gráfico em função da escala, mas possuem centralidade de intermediação mínima de 0 e 13,5, média de 212,22 e 334,17 e máxima de 631,42 e 1720,56 respectivamente, e grau mínimo 2 e 2, grau médio 2,18 e 3,42 e grau máximo 3 e 6.

3. Método Proposto para Obtenção da Centralidade de Intermediação

O método consiste em obter os coeficientes de uma função matemática, que relacione a média ponderada da centralidade em função do número de nós e do número de enlaces em cada nó, com base em dados de topologias reais. Utilizando o método dos mínimos quadrados para obter os coeficientes da função, avalia-se a correspondência entre os dados das topologias e a função obtida, através do coeficiente de correlação.

Pretende-se ajustar os coeficientes de uma função polinomial, utilizando os dados conhecidos da média ponderada da centralidade de intermediação $BC_{mp}(k) = f(N_k, L_k)$, para $k = 0, 1, \dots, n$, para $n + 1$ pares de pontos de cada uma das k topologias, onde N_k e L_k são o número de nós e o número de enlaces de uma determinada topologia, respectivamente.

Dispondo os dados de média ponderada da centralidade (P) em função do número de nós (N) e o número de enlaces (L) em um gráfico cartesiano, foi observado que existe uma dependência levemente curva, com concavidade negativa em relação a L e fortemente positiva em relação a N . Esse tipo de dependência pode ser modelada com funções polinomiais, de grau 1 ou 2. Foram testados modelos polinomiais

Tabela 1. REDES REAIS DE REFERÊNCIA

Número	Rede	N	L	$\langle \delta \rangle$	BC_{mp}
1	Via Network	9	12	2,67	5,52
2	Abilinecore	10	13	2,60	6,32
3	BREN	10	11	2,20	6,50
4	Learn	10	11	2,20	6,50
5	RNP	10	12	2,40	6,58
6	Compuserv	11	14	2,55	7,42
7	Cesnet	12	19	3,17	8,35
8	Cesnet1	12	19	3,17	8,35
9	vBNS	12	17	2,83	8,55
10	Italy	14	29	4,14	10,30
11	NSF Network	14	21	3,00	8,04
12	Austria	15	22	2,93	15,47
13	Austria Aconet	15	22	2,93	10,50
14	Hibernia	15	23	2,93	12,81
15	Mzima	15	19	2,53	14,42
16	Garr-b Changed	16	27	3,38	15,13
17	ARNES	17	20	2,35	22,54
18	Germany	17	26	3,06	18,55
19	Redirisnet	17	28	3,29	20,28
20	CANARIE	19	26	2,74	25,29
21	EON	19	37	3,89	17,47
22	Memorex	19	24	2,53	24,12
23	NLR	19	23	2,42	24,64
24	ARPANET	20	32	3,20	17,31
25	Sweden	20	24	2,40	28,15
26	Pionier	21	25	2,38	28,56
27	Bulgaria	23	24	2,09	45,00
28	COX	24	40	3,33	30,34
29	Sanet	25	28	2,24	45,68
30	Portugal	26	36	2,77	40,84
31	New Network	26	31	2,38	44,70
32	Renater	27	35	2,59	47,05
33	CERNET	29	45	3,10	57,21
34	IBN 31	31	47	3,03	55,39
35	LONI	33	37	2,24	90,48
36	Metrona	33	41	2,48	74,33
37	COST 37	37	57	3,08	60,29
38	OMI1	38	54	2,84	70,97
39	Omicon	38	54	2,84	70,97
40	Internet 2	56	61	2,18	212,22
41	USA 100	100	171	3,42	334,17

$P(N, L) = a_0 + a_1L^p + a_2N^p$ considerando diferentes combinações de valores de p e q ($p = 1, p = 2$ e $q = 1, q = 2$). A combinação com menor coeficiente de correlação foi para $p = q = 2$ e a função é apresentada na Eq. (3).

$$P_2(N, L) = |a_0 + a_1N^2 + a_2L^2|, \quad (3)$$

onde a média ponderada da centralidade de intermediação é positiva para qualquer número de nós e enlances em uma topologia, $P_2(N, L) > 0$, para qualquer $N > 0$ e $L > 0$.

O ajuste da Equação (3) é obtido fazendo com que a diferença dada pela Equação (4), seja mínima.

$$Q = \sum_{k=0}^n \left[f(N_k, L_k) - |a_0 + a_1N_k^2 + a_2L_k^2| \right]^2, \quad (4)$$

Substituindo os $n + 1$ valores de $BC_{mp}(k) = f(N_k, L_k)$ na Equação (3) obtem-se o sistema linear dado pela Equação (5).

$$\begin{pmatrix} 1 & N_0^2 & L_0^2 \\ 1 & N_1^2 & L_1^2 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} BC_{mp}(0) \\ BC_{mp}(1) \\ \vdots \\ BC_{mp}(n) \end{pmatrix}. \quad (5)$$

Sejam A , x e b , a matriz dos coeficientes, o vetor das incógnitas e o vetor dos termos independentes da Equação (5), respectivamente.

$$A * x = b \quad (6)$$

Multiplicando a Equação (6) pela matriz A^T , obtem-se um sistema linear conhecido como equações normais do sistema (6), que expandido toma a forma da Equação (7).

$$\begin{bmatrix} n & \sum N_k^2 & \sum L_k^2 \\ \sum N_k^2 & \sum L_k^2 & \sum L_k^2 N_k^2 \\ \sum N_k^2 & \sum N_k^2 L_k^2 & \sum L_k^4 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum BC_{mp}(k) \\ \sum N_k^2 BC_{mp}(k) \\ \sum L_k^2 BC_{mp}(k) \end{bmatrix}. \quad (7)$$

A solução do sistema da Equação (7), foi obtida neste trabalho pelo método de Gauss e dá os parâmetros ótimos, no sentido dos mínimos quadrados, da Equação (4) [Burden and Faires 2008].

A adequação do modelo pode ser feita comparando cada valor de $BC_{mp}(k)$ com o respectivo valor predito pelo modelo $P_2(N_k, L_k)$. Em uma análise ponto a ponto, a diferença entre esses dois valores é o resíduo [Barroso et al. 1987], entendido como parte do valor observado que o ajuste não foi capaz de modelar. Em uma análise mais geral, o coeficiente de correlação, calculado pela Equação (8), é mais usado para avaliar a qualidade do ajuste.

$$r_k = BC_{mp}(k) - \hat{BC}_{mp}(k), \quad (8)$$

onde, para nosso modelo, $\hat{BC}_{mp}(k)$ é dado pela equação (3). Os resíduos podem ser vistos como parte do valor observado que o ajuste não foi capaz de explicar.

Um método para medir a qualidade do ajuste linear é através do coeficiente de correlação:

$$R^2 = 1 - \sum_{k=1}^n \frac{(BC_{mp}(k) - \hat{BC}_{mp}(k))^2}{(BC_{mp}(k) - \frac{\sum BC_{mp}(k)}{n})^2}, \quad (9)$$

quanto mais próximo de 1 melhor é o ajuste, sendo que $0 < R^2 < 1$.

Aplicando o método proposto em dados das redes reais apresentadas na Tabela 1, foi obtido um ajuste com coeficiente de correlação 0,99, ou seja, com um resíduo muito baixo. Foram encontrados os seguintes valores : $a0 = 4,236$, $a1 = 0,09$ e $a2 = -0,02$, a partir da Equação(7).

Seja um grafo conexo ou não com $n \geq 3$ nós. Onde v_k de G e $BC(v_k)$ a centralidade de intermediação de v_k . Então

$$0 \leq BC(v_k) \leq \frac{N^2 - 3N + 2}{2}, \quad (10)$$

onde o limite para a centralidade de intermediação de um grafo é alcançado se e somente se G for um grafo estrela S_n e v_k o vértice central [de Freitas 2010].

Sendo que os valores obtidos na Equação (3) estão dentro dos limites, tanto para caso máximo e o mínimo da Equação (10), entretanto as redes de telecomunicações nunca chegam a este limitante pois o grau de um nó v_k será sempre maior que um. A Figura 3 apresenta diagramas de dispersão com os valores da centralidade de intermediação em função do número de nós e enlaces. Sendo a centralidade de intermediação uma métrica muito bem aceita para determinar a sobrecarga de nó em uma rede qualquer [Tizghafam and Leon-Garcia 2010].

4. Experimentos e Resultados

Em [Girolimetto et al. 2013] é apresentada uma ferramenta livre para geração de topologias sobreviventes, *NTTGen*, com as características descritas em [Pavan et al. 2010]. O presente artigo utiliza a ferramenta com o intuito de validar a função matemática obtida a partir do conjunto de redes reais apresentado na Tabela 1.

As topologias geradas foram divididas em 5 subconjuntos, com os graus médios 2, 5, 3, 0, 3, 5, 4, 0 e 4, 5. Foram gerados 50 topologias sobreviventes com número de nós entre 10 e 100. Em seguida foi obtido os valores da média ponderada da centralidade de intermediação considerando a Equação (3) para cada topologia, aplicando assim a função:

$$BC_{est}(N, L) = |4,236 + 0,09N^2 - 0,02L^2|. \quad (11)$$

Para cada topologia correspondente ao conjunto gerado, através do valor obtido pela expressão BC_{est} (11) foi calculado o resíduo através da Equação (8), resultando em um coeficiente de correlação de 0,87, Figura (4); com erro médio de 2,88%.

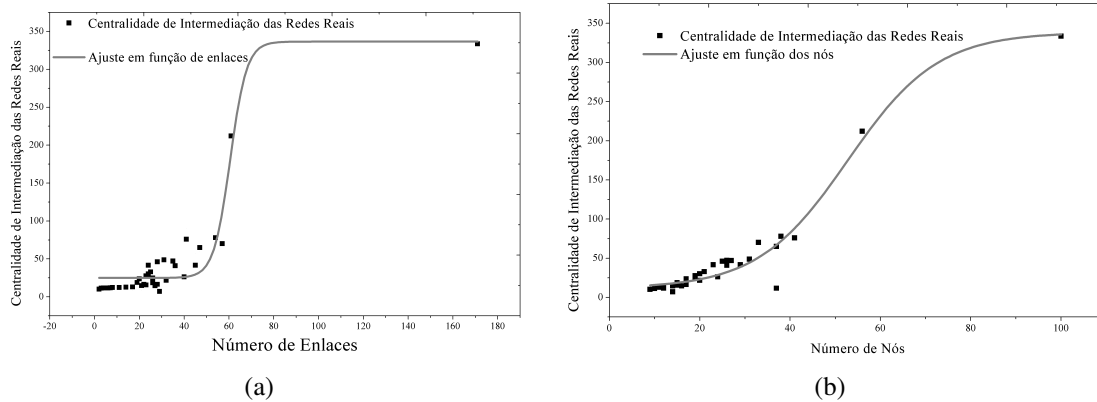


Figura 3. Centralidade de Intermediação em função do número de enlaces (a) determina o número de nós e, (b) está em função do número de nós, portanto determina o número de enlaces nas 41 topologias, sendo que o valor da centralidade de intermediação apresenta-se em ordem crescente.

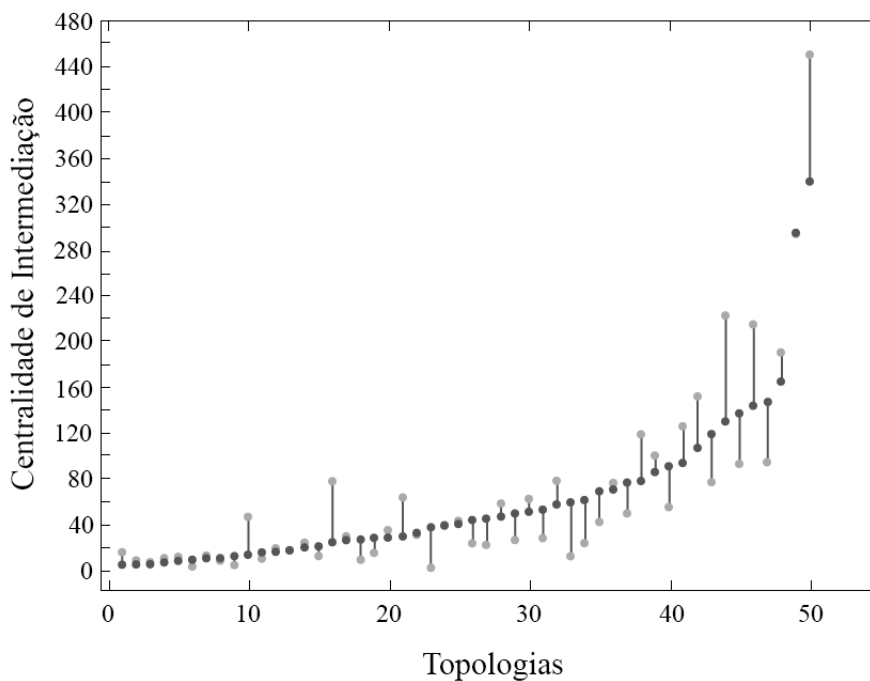


Figura 4. Topologias sobreviventes geradas pelo software *NTTGen*. Os pontos em cor cinza representam a centralidade de intermediação estimada pela função matemática e os pontos em cor preta os calculados numericamente (exatos).

5. Considerações Finais

A centralidade de intermediação é uma variável de mérito das redes de telecomunicações. Neste trabalho, foi proposta uma expressão matemática para estimar o valor da centralidade de intermediação de uma rede, sem a necessidade de conhecer a topologia. A expressão é resultado de um método de ajuste de curvas pelos mínimos quadrados, em que se obteve um ajuste com coeficiente de correlação de 0,99. A expressão de aproximação proposta, quando aplicada à redes de telecomunicações, geradas aleatoriamente, apresenta um erro médio de 3%.

6. Agradecimentos

Este trabalho foi realizado com auxílio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), através do projeto "Desenvolvimento de Uma Ferramenta Computacional para a Geração de Topologias de Redes de Telecomunicações", referente ao edital 308/UFFS/2014.

Referências

- Ausiello, G., Firmani, D., e Laura, L. (2013). The (betweenness) centrality of critical nodes and network cores. *Wireless Communications and Mobile Computing Conference (IWCMC) - IEEE*, 9:90–95.
- Barroso, L. C., Barroso, M. M. A., Filho, F. F. C., de Carvalho, M. L. B., e Maia, M. L. (1987). *Cálculo Numérico – Com Aplicações – 2 edição*.
- Bhardwaj, M., McCaughan, L., Korotky, S. K., e Sanniec, I. (2005). Analytical description of shared restoration capacity for mesh networks. *OSA Journal of Optical Networking*, 4(3):130–141.
- Bodamer, S., Spath, J., e Glingener, C. (2004). An efficient method to estimate transponder count in multi-layer transport networks. *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, 3:1780 – 1785.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25:163–177.
- Burden, R. e Faires, J. (2008). *Análise Numérica – 8 edição*.
- de Freitas, L. Q. (2010). Medidas de centralidade em grafos. Master's thesis, Universidade Federal do Rio de Janeiro.
- Freeman, L. C. (1977). A set of centrality based on betweenness. *Sociometry*, 40:35–41.
- Girolimetto, M., Gallupo, R. A., e Pavan, C. (2013). Uma ferramenta livre para a geração de topologias físicas de redes de telecomunicações. *Salão de Ferramentas do 31º SBRC - Simpósio de Redes de Computadores e Sistemas Distribuídos*, 2:1135–1142.
- Korotky, S. K. (2004). Network global expectation model: A statistical formalism for quickly quantifying network needs and costs. *IEEE Journal of Lightwave Technology*, 22(3):703–722.
- Korotky, S. K. (2005). Analysis and optimization of two-tier networks with application to optical transport backbones. *J. Opt. Netw.*, 4(1):45–63.

- Labourdette, J. F., Bouillet, E., Ramamurthy, R., e Akyama, A. A. (2005). Fast approximate dimensioning and performance analysis of mesh optical networks. *IEEE/ACM Journal of Transactions on Networking*, 3(4):906–917.
- Pavan, C., Morais, R. M., da Rocha, J. R. F., e Pinto, A. N. (2010). Generating realistic optical transport network topologies. *IEEE/OSA Journal of Optical Communications and Networking*, 2:80–90.
- Pinto, A. N., Pavan, C., e Morais, R. M. (2009). A statistical model for capex fast calculation in optical transport networks. In *Proceedings of the 11th International Conference on Transparent Optical Networks, ICTON'09*, volume 11, páginas 1–4, Island of São Miguel, Azores, Portugal. IEEE.
- Routray, S. K., Morais, R. M., Ferreira, J. R., e Pinto, A. N. (2013). Statistical model for link lengths in optical transport networks. *Optical Communications and Networking, IEEE/OSA Journal of*, 5(7):762–773.
- Routray, S. K., Sahin, G., Ferreira, J. R., e Pinto, A. N. (2014). Estimation of link-dependent parameters of optical transport networks from the statistical models. *Optical Communications and Networking, IEEE/OSA Journal of*, 6(7):601–609.
- Silva, T. S. A. (2010). Um estudo de medidas de centralidade e confibilidade em redes. Master's thesis, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca.
- Tizghafam, A. e Leon-Garcia, A. (2010). Betweenness centrality and distance - distance in communication networks. *IEEE Network*, 24:10–16.
- Ufimtsev, V. e Bhowmick, A. (2013). Identifying high betweenness centrality vertices in large noisy networks. *IEEE 27th International Symposium on Parallel & Distributed Processing Workshops and PhD Forum*, 27:2234–2237.