

Análises Topológicas e Interpretações Geográficas de Redes Complexas de Mobilidade Urbana - o caso da Região Metropolitana do Rio de Janeiro

Leonardo Bacelar Lima Santos¹, Antônio Miguel Vieira Monteiro²

¹Programa de Doutorado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

²Divisão de Processamento de Imagens – DPI
Instituto Nacional de Pesquisas Espaciais – INPE

santoslbl@gmail.com, miguel@dpi.inpe.br

Abstract. *According to Milton Santos, the construction of geographical space can be understood as a set of fixed and flows. The fixed represent places, that condition and are conditioned based on the properties of the flows in their relationships. In this scenario, the urban mobility is defined as the ability of moving people and goods in urban space to perform daily activities. In the study of the fixed and flows of urban mobility, an important factor to be considered is the possibility of its structure in the form of different types of (complex) networks. This paper presents an approach based on Complex Networks, graphs with large number of nodes and edges of nontrivial distribution for urban mobility. Networks are built using as a parameter to control (for edges creation) the number of people ranging from a traffic zone to another (from one vertex to another). We present the topological properties of networks, such as degree, clustering, average path and diameter. Associated with the topological issues, are discussed geography questions on the spatially embedded complex networks.*

Resumo. *Segundo Milton Santos, a construção geográfica do espaço pode ser compreendida como um conjunto de fixos e fluxos. Os fixos representam os lugares, que condicionam e são condicionados em função das propriedades dos fluxos em seus relacionamentos. Neste cenário, a mobilidade urbana é definida como a capacidade de deslocamento de pessoas e bens no espaço urbano para a realização das atividades cotidianas. No estudo dos fixos e fluxos de mobilidade urbana, um importante fator a ser considerado é a possibilidade de sua estruturação na forma de diferentes tipos de (complexas) redes. O presente trabalho traz uma abordagem baseada em Redes Complexas, grafos com grande número de vértices e distribuição não trivial de arestas, para mobilidade urbana. São construídas redes tendo como parâmetro de controle para criação de arestas o número de pessoas que vão de uma zona de tráfego a outra (de um vértice a outro). São apresentados os desenhos das redes e suas propriedades topológicas, como grau, aglomeração e mínimo caminho médio. Interligadas às questões topológicas são proferidas interpretações geográficas para os índices e apresentada a demanda por uma caracterização geográfico-topológica de Redes complexas de Base Territorializada.*

Palavras-chave: *redes complexas, mobilidade urbana, redes complexas de base territorializada*

1. INTRODUÇÃO

Mobilidade urbana é a capacidade de deslocamento de pessoas e bens no espaço urbano para a realização das atividades cotidianas [1]. A mobilidade urbana tem grande impacto na economia local e na qualidade de vida das pessoas. As taxas de mobilidade são definidas como a quantidade diária média de viagens realizadas por uma pessoa. Em estudos de mobilidade urbana, um importante fator a ser considerado é a possibilidade de sua estruturação na forma de diferentes tipos de redes. Essas redes, por sua vez, podem apresentar diferentes configurações em função dos atributos das relações espaciais considerados para sua construção [2, 3, 4].

Segundo *Cui (2006)* [5], com o aumento da utilização dos sistemas de transporte público, houve um desenvolvimento recente nas pesquisas envolvendo estimativa de Matrizes de Origem e Destino (MOD). De acordo com um relatório do METRÔ/SP (2008), “uma característica importante da pesquisa de Origem e Destino é que ela permite a localização espacial da população, dos empregos, das matrículas escolares e das origens e destinos das viagens realizadas pela população nas suas atividades diárias”.

Conforme *Kostakos (2010)* [6], uma importante vantagem em usar grafos como estrutura de dados para estudos de mobilidade é poder fazer uso dos algoritmos já bem estabelecidos na área para determinação das estruturas e cálculo de propriedades das (complexas) redes. A área de Redes Complexas (RC) tem sua origem na tradicional Teoria dos Grafos (TG). Um grafo (\mathbf{G}) é um conjunto de vértices (\mathbf{V}) e arestas (\mathbf{E}), denotado por $\mathbf{G} = \mathbf{G}(\mathbf{V}, \mathbf{E})$, no qual os vértices podem estar conectados uns aos outros pelas arestas [7]. Uma RC é um grafo não trivial, com um grande número de vértices e propriedades topológicas não necessariamente típicas nem de grafos regulares, como em autômatos celulares regulares, nem de grafos completamente aleatórios [8]. A abordagem de RC é utilizada para representação de conhecimento, tratamento de dados e modelagem de sistemas complexos – sistemas formados por muitas partes, estas inter-relacionadas possivelmente de forma não linear, apresentando comportamentos emergentes e em multiescalas. Exemplos de sistemas complexos podem ser encontrados especialmente em domínios biológicos, sociais e climáticos [9, 10].

A representação do espaço na teoria da RC é, contudo, topológica mas não geográfica [11] – a localização dos vértices, a proximidade entre eles e o comprimento das arestas são tratados como irrelevantes: a imensa maioria dos trabalhos versam sobre a topologia e não a geografia das redes [12]. Conforme apresentado em *Santos (2012)* [13], a incorporação de elementos espacialmente explícitos nas dinâmicas modeladas com o auxílio de RC implica na necessidade de um novo aparato conceitual – relação vértices-arestas, nova interpretação para índices topológicos tradicionais e desenvolvimento de novos índices locais e globais para as RC, que serão denominadas como Redes Complexas de Base Territorializada (RCBT)¹ – redes nas quais os vértices têm localização geográfica conhecida e a lei de criação de arestas apresenta dependência funcional frente a variáveis espaciais.

Antes de usar RCBT para tratar dados de um determinado fenômeno, entretanto, faz-se necessário: 1. caracterizar topologicamente (tradicionalmente) a rede original, sem componente espacial explícita; e 2. analisar geograficamente os valores dos índices utilizados na caracterização topológica. O presente artigo traz tais procedimentos no tratamento de

¹O termo em inglês originalmente é *Geographical Embedded Complex Networks*

um conjunto de dados reais da mobilidade urbana.

2. MATERIAIS E MÉTODOS

Na seção a seguir serão apresentados os dados reais e os recursos computacionais a serem utilizados na prova de conceito para a caracterização topológica e interpretação geográfica para um problema de mobilidade urbana.

2.1. DADOS REAIS DE MOBILIDADE URBANA

O estudo de caso investigado neste documento é referente à Região Metropolitana do Rio de Janeiro (RMRJ), contando com 10.894.756 habitantes distribuídos em 20 municípios. Do ponto de vista de transporte, a região é dividida em zonas de tráfego, as quais constituem-se o menor nível geográfico de representação das informações obtidas. A Figura 1 traz a região dividida por municípios e estes divididos por zonas de tráfego.



Figura 1. Zonas de tráfego e municípios da Região Metropolitana do Rio de Janeiro.

2.2. RECURSOS COMPUTACIONAIS

Para efetuar a caracterização topológica, abordagem tradicional, para a Rede Complexa construída com base nos dados reais supracitados será usado um conjunto de códigos escritos na linguagem *C++*, baseados no paradigma de Programação Orientada a Objetos (POO) ².

O dado de entrada para tais códigos é a matriz de fluxos, na qual cada elemento $f(i, j)$ representa o número de viagens efetuadas de uma zona de tráfego a outra: de um vértice a outro da rede. A Figura 2 traz uma representação em cores para a matriz dos fluxos. A variação das cores, em escala logarítmica, vai de valores próximos a zero (em preto) a altos valores de fluxo (de amarelo a vermelho).

²Os códigos são versões adaptadas do software gratuito e de código aberto *NetAll*, desenvolvido por Santana (2005) [14], contando com colaboração de diversos alunos e pesquisadores do grupo de pesquisa em Física Estatística e Sistemas Complexos (FESC) do Instituto de Física da Universidade Federal da Bahia (IF-UFBA).

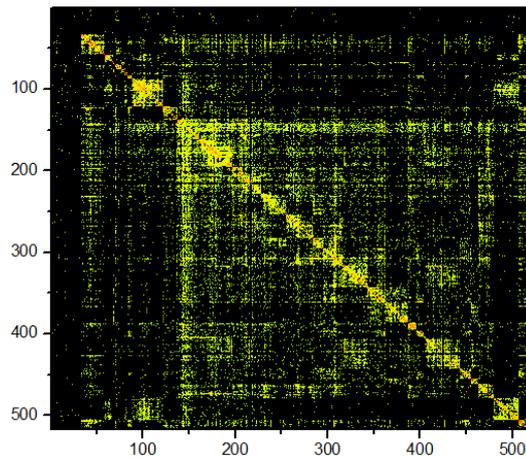


Figura 2. Representação da matriz de fluxos, na qual cada fila representa um vértice (zona de tráfego) e as cores são tão mais para o preto quanto menor o fluxo e tão mais para o amarelo quanto maior o número de viagens entre os vértices em questão.

A matriz de adjacência (*madj*) [8] para os grafos é construída utilizando um valor limiar de fluxo: se o valor de fluxo entre as zonas *i* e *j*, $f(i, j)$, for menor que o valor limiar então $madj(i, j) = 0$, se for maior ou igual então $madj(i, j) = 1$; desta forma evita-se trabalhar com grafos ponderados (nos quais as arestas tem como rótulos o valor do fluxo) [15, 16]. A partir da matriz de adjacência é possível obter os valores para os índices topológicos básicos:

- **Grau:** O grau k de um vértice i é determinado pela quantidade de vértices aos quais o vértice i está ligado, ou seja, o número de vizinhos do vértice i . O grau médio $\langle k \rangle$ de um grafo é calculado pela média aritmética dos graus de cada vértice. Fazendo um histograma dos graus k encontrados em uma rede, temos a distribuição (frequência absoluta) de graus dessa rede. Define-se o índice $p(k)$ como a probabilidade (frequência relativa) de um vértice escolhido aleatoriamente em uma rede ter grau igual a k . Para redes aleatórias, $\langle k \rangle = p(N - 1)$, onde $p = \frac{2L}{N(N-1)}$, L representando o tamanho da rede (número de arestas) e N a ordem (número de vértices).
- **Coef. de aglomeração:** O coeficiente de aglomeração de um vértice é a probabilidade de seus vizinhos serem vizinhos entre eles. Pode-se calcular o coeficiente de aglomeração de um vértice i de uma rede a partir da igualdade $C_i = 2E_i / k_i(k_i - 1)$, sendo E_i o número de arestas entre seus vizinhos e k_i o grau do vértice. O coeficiente de aglomeração médio de uma rede é a média dos coeficientes de aglomeração de cada vértice. Para redes aleatórias, $c = p$.
- **Mínimo caminho médio:** Define-se um caminho como o conjunto de vértices e arestas, alternadamente dispostos, compondo um “percurso” que vai de um vértice a outro da rede. O comprimento representativo desse caminho é o número de arestas que liga o primeiro ao último vértice do conjunto. O menor caminho que liga um vértice i a um vértice j da rede é chamado caminho mínimo entre os vértices i e j , e a média dos caminhos mínimos entre o vértice i e os demais vértices do grafo é o caminho mínimo médio – l – do vértice i . O caminho mínimo médio do grafo, por sua vez, é a média dos caminhos mínimos médios de cada

vértice. O diâmetro de uma rede é definido como o maior dos caminhos mínimos entre quaisquer dois vértices. Para redes aleatórias, $l = \frac{\log(n)}{\log(\langle k \rangle)}$.

3. RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os resultados da caracterização topológica da Rede Complexa construída com base na matriz de fluxos. Ao final são discutidas interpretações geográficas para os valores encontrados referentes aos índices tradicionais de Redes Complexas.

3.1. CARACTERIZAÇÃO TOPOLÓGICA

O conjunto de viagens que compõe o espaço amostral considerado é de 113.270 viagens. Dentre as zonas de tráfego 517 aparecem ao menos uma vez como origem ou destino. Considerando todos os vértices e arestas, 13% das viagens representam autoconexões (origem e destino no mesmo vértice) e 87% são de um vértice a outro. Considerando apenas arestas simples³ têm-se 16184 arestas. O número máximo de viagens entre dois vértices foi de 1920. Para um grafo de N vértices o número máximo de arestas simples é $N * (N - 1)/2$, referente ao grafo completo: todos os vértices são conectados entre eles – o grafo aqui representado tem, portanto, 12% do máximo de arestas simples possíveis.

A caracterização topológica clássica para tal rede traz um conjunto de valores para os índices básicos em função de cada valor limiar de fluxo. À medida que se aumenta o limiar de fluxo pede-se que para conectar dois vértices estes devam ter um fluxo cada vez maior, logo com o aumento do limiar deve haver diminuição do número de arestas. O ritmo de tal decaimento, para este estudo de caso, conforme a *Figura 3a*, é algébrico. Pode-se perceber que há rupturas nas redes para os limiares 6, 19, 25 e 29, pois são os valores imediatamente anteriores aos de queda do diâmetro da rede, conforme a *Figura 3b*. Na região entre os limiares 33 e 215 a rede apresenta efeito de pequeno mundo: como coef. de aglomeração maior e mínimo caminho médio menor que o correspondente para uma rede aleatória com mesmo número de vértices e arestas – tais valores de referência são: $c = 0.059$ e $l = 1.83$.

Uma rede de mobilidade, como a aqui construída, com característica de pequeno mundo para altos limiares de fluxo sugere que zonas de tráfego que são origem ou destino de um número significativo de habitantes da região são ambas provavelmente conectadas a uma outra zona de tráfego também com alto fluxo de passageiros, ou seja, que há forte redundância de caminhos muito frequentados e que os habitantes da região não precisam fazer um grande número de viagens para percorrê-la.

Todavia, tais resultados são referentes ao espaço topológico (modelado), não necessariamente ao espaço geográfico. Um mínimo caminho médio pequeno em uma rede significa que com poucos passos (poucas arestas) é possível sair de uma zona e ir a qualquer outra (ir de um vértice a qualquer outro), não há informação, entretanto, de qual o comprimento (em quilômetros, por exemplo) dos caminhos, nem mesmo se as arestas de “atalho” (redundâncias) são próximas ou não umas das outras. Em um problema com componente espacial relevante é, portanto, fundamental ter índices topológicos que levem em consideração tais componentes espaciais.

³se entre dois vértices há mais de uma aresta apenas uma é contada.

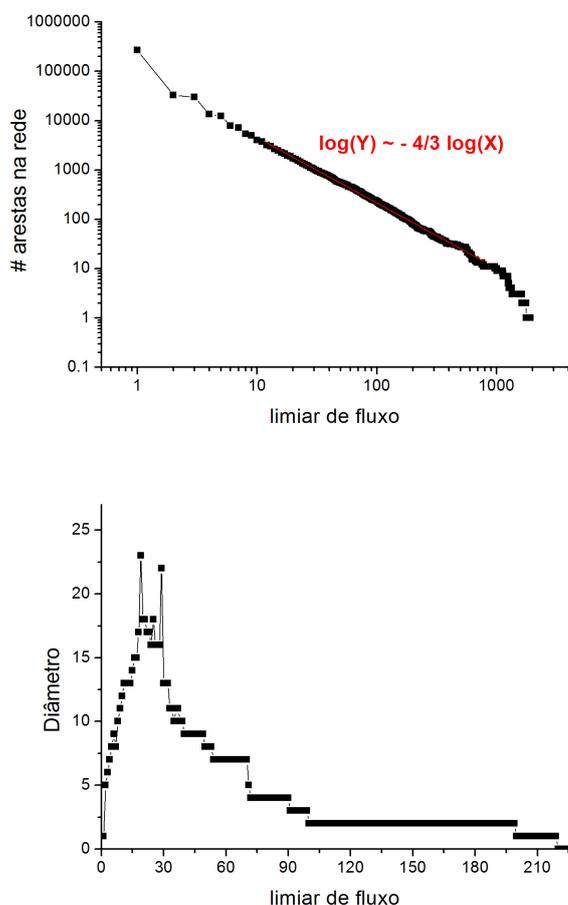


Figura 3. Comportamento do número de arestas na rede (a) e do seu diâmetro (b) em função do valor limiar de fluxo.

4. CONCLUSÕES E PERSPECTIVAS

Neste trabalho foi construída uma Rede Complexa com base em dados reais de mobilidade urbana. Partindo da matriz de fluxos obteve-se uma matriz de adjacência para cada valor de fluxo, e então analisou-se o comportamento dos índices tradicionais de caracterização topológica em funções de tal limiar (parâmetro de controle). Os resultados mostraram um comportamento com efeito de pequeno mundo para o intervalo de limiar de fluxo 33 a 215.

O plano diretor do INPE (2011-2015) traz como objetivo estratégico “Produzir dados, software e metodologias para fortalecer a atuação do INPE nas áreas da aplicações espaciais, **da saúde**, educação, segurança pública e desenvolvimento urbano.” Dentre as perspectivas do trabalho está o uso da rede construída para guiar as vizinhanças de um modelo epidemiológico, associando processos de contato aos de contágio (conforme [23]), em especial no âmbito do projeto **DengueME**, associado à rede **PRONEX** de modelagem da Dengue, doença que representa um dos mais sérios problemas das secretarias de saúde brasileiras e que vem recebendo diversas contribuições da rede em estudos e ferramentas de matemática aplicada e computação científica [17, 18, 19, 20, 21, 22].

A componente espacial, referente à localização geográfica de cada vértice, não foi utilizada explicitamente em nenhum momento da análise, apenas indiretamente (implicitamente), uma vez que o fluxo entre duas zonas de tráfego possivelmente sofre influência da localização absoluta e relativa entre as zonas. Uma Rede Complexa com componentes espaciais explícitas, uma Rede Complexa de Base Territorializada, com definições e propriedades topológicas é uma demanda da literatura [11, 12, 13].

Referências

- [1] Duarte, F.; Libardi, R.; Sanchez, K. Introdução à mobilidade urbana. Editora Juruá, 2007. ISBN-13: 9788536216737
- [2] Prager, S. D. (2008). Complex networks for representation and analysis of dynamic geographies. In: HORNSBY, K. S.; YUAN, M. (Ed.). Understanding dynamics of geographic domains. Estados Unidos da América: CRC Press.
- [3] Souza, D. D. M. R. Estimação Sintética de Matrizes Origem/Destino a partir de Contagens Volumétricas em Áreas com Controle do Tráfego em Tempo Real com o Auxílio do Queensod. Fortaleza, 2007. XV, 120 fl., Dissertação (Mestrado em Engenharia de Transportes) Programa de Mestrado em Engenharia de Transportes, Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, CE.
- [4] Câmara, G.; Monteiro, A. M. V.; Medeiros, J. S. Representações computacionais do espaço: Fundamentos epistemológicos da ciência da geoinformação. Geografia, v. 28, n. 1, p. 8396, 2003.
- [5] Cui, A. (2006). Bus Passenger Origin-Destination Matrix Estimation Using Automated Data Collection Systems. Dissertação de Mestrado, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 134.
- [6] Kostakos, V. et al. (2010). Brief encounters: sensing, modelling and visualizing urban mobility and copresence networks. ACM Transactions on Computer-Human Interaction, v. 17, n. 1. DOI = 10.1145/1721831.1721833
- [7] Boaventura Netto, P. O. (2001). Grafos: Teoria, Modelos, Algoritmos. Editora Edgard Blucher, São Paulo.
- [8] Bessa, A. D. and Santos, L. B. L. and Martinez, P. N. R. and Costa, M. C. and Cardoso, P. G. S. (2009). Introdução às Redes Complexas. Material livre, disponível em: <http://wiki.dpi.inpe.br/doku.php?id=ser301-2011:wikileonardosantos>
- [9] Reka, A. and Barabasi, A. L. (2002). Statistical mechanics of complex networks. Rev. Mod. Phys., v. 74, p. 47–97.
- [10] Newman, M. E. J. (2003). The structure and function of complex networks. SIAM Review, v.45, p. 167–256.
- [11] Kubly, M. et al. (2005). A Comparison of Geographic Information Systems, Complex Networks, and Other Models for Analyzing Transportation Network Topologies. NASA/CR, v. 213522.
- [12] Hayashi, Y. (2006). A Review of Recent Studies of Geographical Scale-Free Networks, IPSJ Digital Courier. v. 2, p. 155-164.

- [13] Santos, L. B. L. (2012). Redes Complexas de Base Territorializada. Monografia de qualificação de doutorado, pelo curso de Computação Aplicada do Instituto Nacional de Pesquisas Espaciais.
- [14] Santana, C. N. (2005). Análise da pluviometria do nordeste brasileiro segundo modelagem em redes. Monografia de conclusão de curso apresentada ao curso de Ciência da Computação da Universidade Federal da Bahia.
- [15] Góes Neto, A; et al. (2010). Comparative protein analysis of the chitin metabolic pathway in extant organisms: A complex network approach. *BioSystems*, v. 101, p. 5966.
- [16] Andrade, R. F. S. ; Rocha-Neto, I. C. ; Santos, L. B. L. ; de Santana, C. N. ; Diniz, M. V. C. ; Lobão, T. P. ; Goés-Neto, A. ; Pinho, S. T. R. ; El-Hani, C. N. Detecting Network Communities: An Application to Phylogenetic Analysis. *PLoS Computational Biology*, v. 7, p. e1001131, 2011.
- [17] Santos, L. B. L.; Costa, M. C.; Pinho, S. T. R.; Andrade, R. F. S.; Barreto, F. R.; Teixeira, M. G.; Barreto, M. L. (2009). Periodic forcing in a three-level cellular automata model for a vector-transmitted disease. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics (Print)*, v. 80, p. 016102.
- [18] Ranck Junior, R.; Santos, L. B. L. (2010) Análise Iterativa dos Problemas de P-centros e P-medianas para um Crescente Número de Facilidades: Estudo de Caso na Epidemia de Dengue, Salvador, 1995, 9th Brazilian Conference on Dynamics, Control and Their Applications. *Anais...*, Serra Negra. Disponível em: <http://www.sbmac.org.br/dincon/trabalhos/PDF/optimization/67808.pdf>
- [19] Medeiros, L. C. D. et al. (2011). Modeling the Dynamic Transmission of Dengue Fever: Investigating Disease Persistence. *PLOS neglected tropical diseases*. v. 5, n. 1.
- [20] Lana, R. M. and Carneiro, T. G. S. and Honório, N. A. and Codeço, C. T. (2011). Multiscale analysis and modeling of *Aedes aegypti* population spatial dynamics. *Journal of Information and Data Management*. v. 2, p. 211–220.
- [21] Santos, L. B. L.; Baldo, R. (2011) Interfaces entre comunicação digital, vigilância epidemiológica e modelagem computacional. *Anais da 63ª reunião anual da Sociedade Brasileira para o Progresso da Ciência (SBPC)*.
- [22] Santos, L. B. L.; Monteiro, A. M. V. Uso e aplicação de modelagem computacional espacialmente explícita de processos epidêmicos: o exemplo da dengue. *Anais do Worcap 2011*.
- [23] Balcan, D.; Colizza, V.; Gonçalves, B.; Hu, H.; Ramasco, J. J.; Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA*, 106:21484