

MIRIAN MITUSUKO IZAWA

**EFEITOS PENDULARES E GRAVITACIONAIS EM REDES
COMPLEXAS DE TRANSPORTE PÚBLICO**

Brasília

2017



Universidade de Brasília

**EFEITOS PENDULARES E GRAVITACIONAIS EM REDES
COMPLEXAS DE TRANSPORTE PÚBLICO**

Mirian Mitusuko Izawa

Orientador: Bernardo Assunção Melo
Co-orientador: Fernando Albuquerque Oliveira
Co-orientador do PDSE: Albert Diaz-Guilera

TESE DE DOUTORADO

UnB
Brasília
2017

“ If you want to go fast, go alone. If you want to go far, go together”

AFRICAN PROVERB

AGRADECIMENTOS

Agradeço a Deus em primeiro lugar. Agradeço a meus pais, Shuichi e Mitsue, a meus irmãos, Andréia, Gilberto e Sandra, pelos incentivos. Agradeço a meus sogros, dona Teresa e seu Francisco, aos cunhados Yeda, Eveline e Orion, pelo apoio. Agradeço a Bruno, por todas as discussões do projeto, quem sempre esteve presente para ajudar a consolidar as ideias, apoiar em tempo integral. Agradeço a tia Atsuko, uma pessoa que moldou meu gosto pelos estudos. Agradeço a tia Teko, que me apoia das mais diversas formas. E a todos da minha família Izawa, Nakakura e Moraes meus sinceros agradecimentos.

Agradeço imensamente ao professor-orientador Bernardo Melo, pela orientação do projeto, pela dedicação em me ensinar tanta coisa: física, computação, até redação. Agradeço pela paciência e oportunidade, por tudo. Mais que tudo, agradeço por ter acreditado em mim, obrigada! Agradeço imensamente ao professor Albert Diaz, pelo projeto de Doutorado Sanduíche na Universidade de Barcelona, por me receber com tanta gentileza e paciência, agradeço a oportunidade que me fez abrir minha mente para novos conhecimentos. Agradeço ao professor Fernando Oliveira, pela oportunidade de realizar o meu trabalho de doutorado, por todos os conselhos e dicas. Ao professor Daniel Cajueiro, pela discussões e colaboração.

Agradeço aos amigos da UnB, Evandro, Janneth, Anderson, Keila, Pedro, Felipe, Orestes, Clóvis, Sérgio, Wiliam, André, Marco Aurélio, Rogelma e demais, que me apoiaram nessa jornada. Agradeço aos amigos da Universidade de Barcelona: Marina, Adriana, Ivan, Miquel por terem me recebido com tanto carinho. Agradeço aos colegas do grupo de Complex System, Ulaguer, Pol, Kolja, Guile, Roberta, Mario, Nikos, Ruben, Antoine, Oscar e aos professores Maria Angeles Serrano e Marián Bogniá, pela discussões científicas e oportunidade. Agradeço a Laura e Martin pelo carinho e discussões.

Agradeço aos amigos e colegas do trabalho pelas boas energias e apoio: Tonim, Rodrigo, Bruno, Rita, Gustavo, Orislane, Maria do Carmo, Yara, Bolsoni, Ismael, Marcel, Diego, Cristiane, Ângela, Eder, Marcelo, Marcão, Samuel, Salerno e outros. E, principalmente, meus sinceros agradecimentos a Sérgio Marcony, Henrique Veloso, Renato Attiê e Xavier Costa, pelo fornecimento de dados brutos para minha pesquisa, sem os quais este trabalho não seria possível, pelo carinho e apoio. Agradeço a Adonis Ribeiro da SEMOB, pelos dados do PDTU 2009 e outros dados de pesquisa do transporte urbano do DF.

Agradeço ao apoio dos meus amigos de longa data, Júnior, Glécia, Waldene, Kimiko, Wagner, que acompanham minha trajetória, cada luta, cada sucesso, cada momento, e pensar que tudo começou por um simples projeto de feira de ciências do colégio (1996)

e resultou numa amizade tão forte e sincera. Agradeço aos amigos que foram agregados ao grupo e assim se tornaram queridos também, Thiago e Milton. Agradeço a Miki Saito, mesmo tão distante, lá do Japão, sempre torceu pelo meu sucesso.

E finalmente, agradeço a CAPES pela oportunidade e financiamento do estágio na Espanha pelo Programa Doutorado Sanduíche no Exterior - PDSE, a Universidade de Brasília e a Universitat de Barcelona pela minha formação e condições de trabalho. Agradeço a DFTRANS pela confiança em me conceder o afastamento para dedicar-me ao doutorado.

ABSTRACT

The complex network theory is often used to investigate the behavior of Public Transport Networks (PTN). In general, these studies are limited to the topological analysis using routes and stations/stops. However, more than routes and stations, a PTN is constituted by factors that include a provision of service, such as, number of trips, schedules, travel time and number of passengers carried. We decided to focus on PTN operations, that is, on travel and passengers, not only on the route and stations/stop as they have been studied. In this work we construct the weighted networks L -Rout, L -Trip and L -Pass derived from the L -space network. They are weighted, respectively, by number of routes, number of trips and number of passengers. The same is done with the weighted networks P -Rout, P -Trip and P -Pass derived from the P -space. We find that even for the different scales (for instance, the L -Pass is larger than the L -Trip, which is larger than the L -Rout in terms of number of edges), the nodes degree distribution function leads to an exponential decayment independent of their number of edges. From this, we have developed measures using the periodic properties of the trips in PTN, which we define as pendularity coefficients R_{CM} and h_{CM} as the center of mass of the trips in a day. We compare the distribution of these two measures between the L -Trip and L -Pass networks as a way to compare the offer of transport services and their demand. Finally, we developed an altered gravitational model for the study of the distribution of urban bus transportation trips from the Federal District from one of the properties of complex networks, the distance d_{ij} , and from the bus boarding and unboarding characteristics. We believe that the proposed methodologies for the P -Rout, P -Trip and P -Pass, and for the L -Rout, L -Trip and L -Pass, networks can contribute to the analysis and planning of public transport systems.

RESUMO

É frequente a utilização da teoria de redes complexas para investigar o comportamento de Redes de Transporte Público (PTN – do inglês, *Public Transport Network*). Em geral, são utilizadas as rotas e as estações para a construção dessas redes e geralmente as análises se limitam a topologia da rede. No entanto, mais que rotas e estações, um PTN é formado por fatores que envolvem a prestação de serviço, tais como, número de viagens, horários, tempo de viagens e número de passageiros transportados. Decidimos focar nas operações de PTN, viagens e passageiros, não apenas na sua topologia, rota e estações/ponto de paradas, como vem sendo estudadas essas redes. Neste trabalho construímos as redes ponderadas L -Rout, L -Trip e L -Pass derivadas da rede espaço- L . E ponderados, respectivamente, pelo número de rotas, de viagens e de passageiros. O mesmo é feito com as redes ponderadas P -Rout, P -Trip e P -Pass derivadas da rede espaço- P . Constatamos que mesmo para as diferentes escalas (por exemplo, a L -Pass é bem maior que a L -Trip, que, por sua vez, é bem maior que a L -Rout em termos de número de arestas), as curvas de distribuição do grau do nó levam a um decaimento exponencial independente do número de arestas das redes. A partir disso, desenvolvemos medidas usando as propriedades periódicas das viagens em redes PTN, que definimos como coeficientes de pendularidade R_{CM} e h_{CM} como o centro de “massa” das viagens num dia. Comparamos a distribuição dessas duas medidas entre as redes L -Trip e L -Pass como forma de comparar a oferta de serviços de transporte e sua demanda. E por fim, desenvolvemos um modelo gravitacional alterado para o estudo da distribuição de viagens do Transporte urbano de ônibus do Distrito Federal a partir de uma das propriedades de redes complexas, a distância d_{ij} , e a partir das características de embarque e desembarque dos ônibus. Acreditamos que as metodologias propostas para as redes P -Rout, P -Trip e P -Pass, e também para as redes L -Rout, L -Trip and L -Pass, possam contribuir na análise e planejamento de sistemas de transporte público.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 A origem da teoria de grafos: O problema das sete pontes de Königsberg. . . .	4
2.2 Ilustração de um grafo com 5 nós e 5 arestas.	16
2.3 Ilustração de redes direcionadas e não direcionadas.	17
2.4 Ilustração de rede com multiarestas arestas e rede com peso.	17
2.5 Esquema ilustrativo de coeficiente de aglomeração.	22
2.6 Esquema ilustrativo do comprimento do menor caminho.	25
3.1 Esquema ilustrativo de Rede de Transporte Público.	33
3.2 Ilustração das redes espaço- P , espaço- L , espaço- C e espaço- B	34
3.3 Esquema ilustrativo para o cálculo do menor caminho em redes espaço- P	37
3.4 Mapa de pontos de paradas e de rotas de ônibus do Distrito Federal em 2010.	41
3.5 Distribuição do grau ponderado ou força do nó da rede espaço- B	43
3.6 Esquema ilustrativo para rede espaço- P ponderada.	44
3.7 Distribuição do grau ponderado ou força do nó da rede espaço- P	44
3.8 Distribuição do grau ponderado ou força do nó da rede espaço- C	46
4.1 Ilustração da diferença topológica entre uma rede ferroviária e rodoviária. . . .	49
4.2 Esquema ilustrativo da topologia da rede não direcionada e da rede direcionada.	50
4.3 Ilustração de uma rede de ônibus, com 2 rotas com duas direções cada, $r_1 =$ (t_1, t_2) e $r_2 = (t_3, t_4)$	51
4.4 Esquema ilustrativo de construção das redes espaço- L e espaço- P	52
4.5 Rede de ônibus com peso nas arestas relativo ao número de rotas, número de viagens e de número de passageiros.	52
4.6 Resultado da Distribuição do Grau do nó da topologia do espaço- P e espaço- L .	54
4.7 Resultado da Distribuição do comprimento de menor caminho das redes espaço- L e espaço- P	56
5.1 Esquema ilustrativo das medidas de R_{CM} e θ_{CM} de viagens e de passageiros para um ponto de parada qualquer.	61
5.2 Ilustração do comportamento esperado para $R_{CM} \times \Delta h$	63
5.3 Gráfico do número de viagens e do número de passageiros que passam no ponto de parada 72122.	64
5.4 Gráfico da número de viagens e do número dos passageiros das rota de ônibus do DF com maior número de viagens.	69
5.5 Distribuição ao longo do dia do número de viagens e de passageiros, conside- rando todos os pontos de paradas do DF.	69

5.6	Gráfico da distribuição de viagens e de passageiros ao longo do dia no ponto de parada 72122 - eixinho L, próximo ao Banco Central.	71
5.7	Gráfico da distribuição de viagens e de passageiros ao longo do dia no ponto de parada 72447 - W3 sul.	72
5.8	Mapa das 30 RAs e seus respectivos pontos de paradas.	73
5.9	Resultado da distribuição de M , R_{CM} e h_{CM} das viagens das redes de nós da L -Trips e L -Pass.	74
5.10	Resultado da distribuição de M , R_{CM} e h_{CM} das viagens das redes de nós da L -Trip e da L -Pass ponderados.	75
5.11	Mapas das propriedades R_{CM} e h_{CM} dos nós das redes L -Trip e L -Pass.	76
5.12	Resultado da distribuição de M , R_{CM} e h_{CM} das viagens das redes de arestas da L -Trip e da L -Pass ponderadas.	77
5.13	Distribuição de M , R_{CM} e h_{CM} das viagens das redes de arestas da L -Trip e da L -Pass ponderadas.	77
5.14	Mapas das propriedades R_{CM} e h_{CM} das arestas das redes L -Trip e L -Pass.	79
6.1	Amostra do carregamento de viagens em rotas ônibus do DF.	93
6.2	Amostra do Embarque e Desembarques de viagens em rotas de ônibus do DF.	94
6.3	Mapas da 30 RAs do DF e seus respectivos pontos de paradas.	95
6.4	Dados censitários do DF.	96
6.5	Movimentos diários por motivo de trabalho entre as 30 Regiões Administrativas do DF.	97
6.6	Distribuição espacial dos movimentos diários por motivo de trabalho entre as 30 Regiões Administrativas do DF.	98
6.7	Esquema ilustrativo uma viagem de uma rota em espaço- L	100
6.8	Amostra do Embarque e Desembarques de viagens em rotas ônibus do DF ponto a ponto usando a distribuição de probabilidade P_{ij} como probabilidade de embarcar e probabilidade de desembarcar sejam eventos independentes.	105
6.9	Modelo de Distribuição de passageiros simulado para uma viagem com $N=70$ pontos de paradas.	108
6.10	Distribuição de passageiros por d_{ij} simulado para uma viagem com $N=70$ pontos de paradas.	109
6.11	Esquema ilustrativo da distribuição de viagens agrupados por RA	110
6.12	Distribuição pelo modelo alterado calculado para a rota 163.3.	110
6.13	Distribuição da probabilidade de viagens para todos os pontos de paradas de ônibus.	111
6.14	Distribuição de passageiros por RA.	112
6.15	Distribuição de passageiros por RA com o valores censitários da população do DF.	113

6.16	Distribuição de passageiros em uma viagem da rota 0.308. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.	116
6.17	Simulação da distribuição de passageiros simulado para a rota 0.308 usando vários modelos.	117
C.1	Distribuição de passageiros em uma viagem da rota 0.323. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.	131
C.2	Distribuição de passageiros simulado para a rota 0.323.	132
C.3	Distribuição de passageiros em uma viagem da rota 197.3. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.	133
C.4	Distribuição de passageiros simulado para a rota 197.3.	134
C.5	Distribuição de passageiros em uma viagem da rota 200.1. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.	135
C.6	Distribuição de passageiros simulado para a rota 200.1.	136
C.7	Distribuição de passageiros em uma viagem da rota 0.411. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.	137
C.8	Distribuição de passageiros simulado para a rota 0.411.	138

LISTA DE TABELAS

	<u>Pág.</u>
4.1 Propriedade do espaço- L , valores médios de: $\langle k \rangle$ – grau do nó, $\langle s \rangle$ – força do nó; $\langle c \rangle$ – coeficiente de aglomeração; $\langle l \rangle$ – comprimento do menor caminho e $\langle \mathcal{B} \rangle$ – coeficiente de intermediação (<i>betweenness</i>).	53
4.2 Propriedade do espaço- P , valores médio de: $\langle k \rangle$ – grau do nó, $\langle s \rangle$ – força do nó; $\langle c \rangle$ – coeficiente de aglomeração; $\langle l \rangle$ – comprimento do menor caminho e $\langle \mathcal{B} \rangle$ – coeficiente de intermediação (<i>betweenness</i>).	54
6.1 Regiões Administrativas (RAs) do Distrito Federal.	98

LISTA DE ABREVIATURAS E SIGLAS

BA	–	Barabasi-Albert
CM	–	Centro de Massa
DF	–	Distrito Federal
DFTRANS	–	Transportes Urbanos do Distrito Federal
ER	–	Erdős-Rényi
GTFS	–	<i>Google Transit Feed System</i>
MG	–	Modelo Gravitacional
MGA	–	Modelo Gravitacional Alterado
PDAD-2011	–	Pesquisa Distrital por Amostra de Domicílio 2011
PDTU-2009	–	Plano Diretor de Transporte Urbano 2009
PTN	–	<i>Public Transport Network</i>
RA	–	Região Administrativa
SBA	–	Sistema de Bilhetagem Automática
SEMOB-DF	–	Secretaria de Estado de Mobilidade Urbana do Distrito Federal
WS	–	Watts-Strogatz

LISTA DE SÍMBOLOS

p	– probabilidade de estar conectado
q	– probabilidade de não estar conectado
n	– número de nós
m	– número de arestas
k	– grau (valência) do nó
$\langle k \rangle$	– valor médio de k
k_i	– grau do nó i
$D_{(i)}^{\alpha\beta}$	– matriz de distância em multiplex
G	– grafo
$N(G)$	– nós do grafo G
$E(G)$	– aresta do grafo G
$G(N, E)$	– grafo com nós N e arestas E
X	– variável aleatória
P	– probabilidade
$\mathbb{E}(X)$	– valor esperado de X
ρ	– variável de desordem
$\mathcal{L}(\rho)$	– comprimento de caminho médio
$\mathcal{C}(\rho)$	– coeficiente de aglomeração
\mathbf{A}	– matriz de adjacências
a_{ij}	– elemento da matriz \mathbf{A} , do nó i ao j
\mathbf{W}	– matriz de adjacências ponderada
w_{ij}	– elemento da matriz \mathbf{W} , do nó i ao j
k_i^{in}	– grau de entrada do nó i
k_i^{out}	– grau de saída do nó i
s_i	– força do nó i ou grau ponderado do nó i
\mathcal{C}_i	– coeficiente de aglomeração do nó i
E_i	– número de arestas entre os vizinhos do nó i
l_{ij}	– comprimento de caminho entre os nós i e j
\mathbf{L}	– matriz de menor caminho
\mathcal{B}_i	– coeficiente de intermediação do nó i
σ_{jk}	– número de menores caminhos entre os nós j e k
$\sigma_{jk}(i)$	– número desses caminhos que vão através do nó i
d_{ij}	– comprimento do menor caminho entre os nós i e j
ϵ_{ij}	– $1/d_{ij}$
\mathcal{E}	– valor médio da eficiência
\mathcal{E}_i	– valor médio da eficiência do nó i
\mathcal{E}_{glob}	– eficiência global
\mathcal{E}_{loc}	– eficiência local
\mathcal{T}	– Matriz de Transbordo
CM	– Centro de Massa
m_i	– massa da i -ésima partícula

r_i	–	raio da i -ésima partícula
x_i	–	coordenada do eixo-x da i -ésima partícula
y_i	–	coordenada do eixo-y da i -ésima partícula
θ_i	–	ângulo da i -ésima partícula
\hat{x}	–	direção unitária no eixo-x
\hat{y}	–	direção unitária no eixo-y
M	–	Massa total do nó
R_{CM}	–	raio do CM
X_{CM}	–	coordenada do eixo-x do CM
Y_{CM}	–	coordenada do eixo-y do CM
θ_{CM}	–	ângulo do CM
h_{CM}	–	horário do CM
μ	–	valor médio
σ	–	desvio padrão
\mathbf{T}	–	matriz de viagens
T_{ij}	–	elemento da matriz de viagens entre i e j
O_i	–	valor da população na origem i
D_j	–	valor da população no destino j
d_{ij}	–	distância entre i e j
P_i^E	–	probabilidade de embarcar em i
P_j^D	–	probabilidade de desembarcar em j
\mathbf{M}	–	matriz Origem-Destino
m_{ij}	–	elemento da matriz Origem-Destino
P_{ij}^P	–	probabilidade de ir de i a j da rede P -space
P_{ij}^L	–	probabilidade de ir de i a j da rede L -space
p_{ij}	–	peso da aresta i a j da rede L -space
P_{ij}	–	probabilidade de ir de i a j
b	–	fator de ajuste

SUMÁRIO

	<u>Pág.</u>
1 Introdução	1
2 Redes Complexas	3
2.1 Uma breve história das Redes Complexas	3
2.1.1 O problema das sete pontes de Königsberg	3
2.1.2 Grafos aleatórios	4
2.1.3 Um experimento social	8
2.1.4 Modelo de vantagens acumulativas	9
2.2 Redes reais	10
2.2.1 O modelo de Watts-Strogatz (WS)	11
2.2.2 O modelo de Barabási-Albert (BA)	12
2.2.3 Redes Multiplex	13
2.2.4 Exemplos de redes reais	14
2.3 Grafos e matrizes	15
2.3.1 Grafos	16
2.3.2 Estrutura de uma rede	16
2.3.3 Matriz de adjacências	18
2.4 Medidas de centralidade	19
2.4.1 Grau do nó	19
2.4.2 Distribuição do Grau	21
2.4.3 Coeficiente de aglomeração	22
2.5 Medidas de caminhos	23
2.5.1 Comprimento do menor caminho	24
2.5.2 Coeficiente de intermediação	26
2.6 Medida de Eficiência	27
2.6.1 Medida de Eficiência Topológica	27
2.6.2 Medida de Robustez Topológica	28
2.6.3 Medida de Dissimilaridade	29
3 Redes de Transporte	31
3.1 Redes de Transporte Público	32
3.2 Grau do nó em PTN	35
3.3 Comprimento do menor caminho em PTN	36

3.3.1	Topologia de PTN integrada	36
3.4	Coeficiente de aglomeração em PTN	39
3.5	Coeficiente de intermediação em PTN	39
3.6	Robustez em PTN	40
3.7	Estudo de caso: Transporte Urbano do Distrito Federal	40
3.7.1	Grau do nó da rede espaço- B	42
3.7.2	PTN com peso	43
4	Modelo alternativo para PTN	48
4.1	Dados do Transporte Urbano do Distrito Federal	50
4.2	Rede PTN direcional	50
4.3	Oferta e demanda em PTN	51
4.4	Resultados	53
5	Pendularidade em PTN	57
5.1	Conceituação Demográfica	57
5.2	Movimento Pendular	59
5.3	Estudo do centro de massa	60
5.4	Solução para sistemas periódicos	64
5.5	Dados do Transporte Urbano do Distrito Federal	66
5.6	Análise da Distribuição dos horários em Redes espaço- L	68
5.7	Resultados	73
6	Distribuição de Viagens	80
6.1	Modelo de Planejamento de Transportes	80
6.1.1	Modelo Comportamental	81
6.1.2	Modelo 4 Etapas	84
6.2	Modelo Gravitacional	85
6.2.1	Modelo de Wilson	85
6.2.2	Modelo Gravitacional com termo de competição simples	90
6.3	Redes Multiarestas	91
6.4	Dados do Transporte Urbano do Distrito Federal	91
6.5	Características do Transporte Urbano do Distrito Federal	92
6.6	Modelo de distribuição de viagens para o transporte do DF	95
6.7	Resultados	111
7	Conclusão e Perspectivas	118

REFERÊNCIAS BIBLIOGRÁFICAS	120
Apêndice - Centro de massa	126
Apêndice - Série de Fourier	128
Apêndice - Simulação MGA para viagens reais	130
Apêndice - Mineração de dados	139

CAPÍTULO 1

Introdução

Ao tratar de sistemas dinâmicos complexos facilmente podemos citar inúmeros trabalhos em Redes Complexas. Desde os trabalhos pioneiros em Teoria dos Grafos de Leonhard Euler para solucionar um difícil problema de rotas para atravessar as pontes da cidade de Königsberg, o estudo em Redes Complexas vem fascinando vários estudiosos de diversas áreas do conhecimento justamente por causa da abstração simplificada de problemas envolvendo grande quantidade de informação com base nas conexões entre os elementos que compõem certos sistemas complexos.

O final da década de 1990 é particularmente importante devido aos avanços na Teoria de Redes Complexas. Embora muitos trabalhos com redes regulares e redes aleatórias tivessem sido publicados desde os trabalhos de Euler, foi a partir da investigação de redes complexas com a utilização de dados reais que temos observado um aumento significativo no número de publicações sobre o tema. Os trabalhos particularmente importantes para o entendimento da topologia de redes foram os estudos de redes de mundos pequenos, que trouxeram luz ao problema da dinâmica em redes complexas reais. Outros trabalhos concorrente em termos de importância foram os trabalhos de redes sem escala, importante para o entendimento da formação de redes complexas reais.

Inúmeros são os sistemas complexos de interesse dentro da área de redes complexas, principalmente os sistemas que intuitivamente são formados em redes, tais como, redes de transporte, redes de comunicação, internet, cadeia alimentar, rede social, rede de negócios, epidemias, entre outros.

Assim, são inúmeros os trabalhos em redes complexas em que são usados os dados de redes de transporte para corroborar com as novas teorias de redes reais, desde o sistema de trens da Índia até a rede mundial de aeroportos, com propósitos de entender a formação, a estruturas, a dinâmica e a eficiência de redes reais.

Nesse contexto, apresentamos no Capítulo 2 uma sucinta trajetória da ciência de Redes Complexas, abrangendo estudos em diversas áreas do conhecimento que foram importantes para a atual concepção de redes reais, as principais teorias, o formalismo matemático, algumas propriedades de redes importantes e exemplos de estudos em redes reais.

No Capítulo 3 mostramos o conceito de redes de transporte público (*PTN* – do inglês, *Public Transport Network*) dentro da abordagem de redes complexas, a forma-

ção das redes, estudos das propriedades topológicas de redes e estudos empíricos realizado com os dados do sistema de transporte urbano do Distrito Federal. Sugerimos um modelo de redes *PTN* no Capítulo 4, com acréscimo de alguns elementos além daqueles tradicionalmente usados (rotas e estações), tais como a quantidade de rotas diferentes, a quantidade de viagens e a quantidade de passageiros. Desenvolvemos essas redes *PTN* modificadas, usando um conceito de força do nó, simplificando uma rede de múltiplas arestas em redes ponderadas, mas tudo isso sem eliminar a simplificação do modelo de redes. Assim, construímos novas redes, os quais chamamos de redes *P-Rout*, *P-Trip* e *P-Pass* a partir da rede espaço-*P*, e também as redes *L-Rout*, *L-Trip* e *L-Pass* a partir da rede espaço-*L* do modelo de redes *PTN*.

A partir da construção de redes *PTN* modificadas estudamos no Capítulo 5 as características das distribuição das viagens e de passageiros ao longo do tempo, usando a ideia de sistemas periódicos. E no Capítulo 6, sugerimos um modelo de distribuição para as viagens de ônibus no em uma região a partir da ideia de força de interação entre duas regiões.

No Capítulo 5, portanto, construímos um modelo usando o conceito de centro de massa para descrever a distribuição de viagens ao longo do dia por meio de duas grandezas R_{CM} e θ_{CM} . Analisamos o caso da rede de transporte urbano de ônibus do Distrito Federal e apresentamos como a medida dessas grandezas podem contribuir para entendermos a distribuição de viagens em um ponto de parada específico.

E finalmente, no Capítulo 6 construímos um modelo de distribuição de viagens usando as propriedades de redes *PTN*, de modo a alterar o modelo gravitacional tão usada no planejamento de transportes como uma técnica, que a partir de dados censitários tenta prever a distribuição de viagens em um território. Aplicamos o modelo para o caso do transporte urbano de ônibus do DF e apresentamos alguns resultados.

CAPÍTULO 2

Redes Complexas

O estudo da teoria de redes complexas se intensificou no final da década de 1990 com destaque para dois modelos básicos: redes de pequenos mundos [1] e redes sem escala [2]. Antes disso, a modelagem de sistemas em redes era baseada em dois modelos principais: redes regulares e redes aleatórias.

Rede complexa é uma representação abstrata de qualquer sistema discreto formado por elementos que podem ser conectados um ao outro devido a alguma característica especial. Podem ser pessoas conectadas pelo laço de amizade; países, pela relação de comércio; lugares, pela relação de mobilidade; células, pelas funções que exercem no corpo ou computadores, pela troca de dados. Matematicamente são representados por grafos.

Newman define redes como uma representação simplificada que reduz um sistema em estruturas abstratas apreendendo apenas as bases do padrão das conexões, que podem conter informações como nome, rótulo, peso, direção, etc, que captem alguma característica do sistema. No entanto, aconselha-se que não sejam muitas as informações para não perder a simplificação do sistema. Tendo em vista que, quanto mais informações, ou variáveis, são impostas ao problema, mais complexa ela ficará e perderá a simplificação atribuída ao método [3].

2.1 Uma breve história das Redes Complexas

Nesta seção descreveremos sucintamente estudos importantes que serviram de base para a teoria de redes complexas. Desde o surgimento da teoria de grafos por Euler e de redes aleatórias de Erdős e Rényi, ambos os trabalhos realizados na matemática, passando por estudos desenvolvidos na sociologia por Milgram e na economia por Simon e Price.

2.1.1 O problema das sete pontes de Königsberg

A teoria dos grafos teve origem no século XVIII com o trabalho do matemático e físico Leonhard Euler, no qual é resolvido o famoso problema matemático das sete pontes de Königsberg, Figura (2.1). A questão a ser resolvida era sobre otimização de caminhos. Ou seja, questionava-se como percorrer por todas as porções de terra e voltar ao ponto de origem passando uma única vez em cada ponte? Naquela época, debatia-se sobre a possibilidade de existir um percurso fechado, em que um andarilho pudesse sair

de um local e voltar ao mesmo, passando uma única vez em cada uma das sete pontes sobre o rio Pregel, na Prússia [4, 5]. Em 1736 Euler provou que esse caminho não poderia existir.

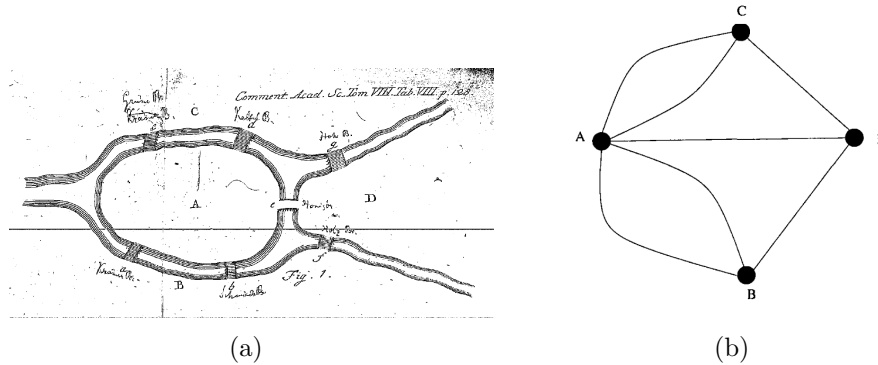


Figura 2.1 - 2.1(a): Desenho ilustrativo de Euler do problema das sete pontes de Königsberg; 2.1(b): Grafo de Euler do problema das sete pontes de Königsberg. (Fonte: [5])

Para resolver esse problema, Euler percebeu que as formas das ilhas e do rio não influenciavam na solução do problema, mas apenas as conexões entre um local e outro através das pontes realmente importavam. Assim, desenhou um diagrama formado por pontos e linhas, Figura (2.1(b)), a que denominou *grafo* [5]. Cada ponto desse diagrama é chamado de **nó** do grafo e cada linha é chamado de **aresta** do grafo [6].

Dessa forma, Euler esboçou um grafo formado por quatro nós — A, B, C e D —, que representavam as porções de terra, e sete arestas, que representavam as sete pontes. Como a regra era passar por todas as pontes uma única vez e passar por todas as ilhas um única vez e voltar ao local de origem, seria necessário que existissem duas pontes, pois uma ponte deveria somente ser utilizada uma única vez pelo andarilho. Assim, para essa viagem desejada, era necessário que cada porção de terra fosse atendida por um número par de pontes. Fato que não era atendido pelas pontes de Königsberg [7].

2.1.2 Grafos aleatórios

Muitos estudos sobre grafos são realizados desde seu surgimento. No entanto, o conceito de *grafos aleatórios* com uma abordagem estatística e com o uso de algoritmos foi introduzido somente no século XX [8].

O trabalho de Erdős e Rényi – ER (1959) é o pioneiro nessa abordagem.

Eles introduziram a ideia de que as conexões entre dois nós quaisquer são realizadas aleatoriamente e de forma equiprovável e independente da quantidade de conexões que os nós possuem [8]. Uma vez fixado o número de nós (n) e de arestas (m) da rede, as arestas entre os nós são formadas de forma aleatória entre as $\binom{n(n-1)/2}{m}$ configurações de redes equiprováveis. Esse modelo é considerado um sistema *fechado*, tal que, não há inclusão ou exclusão de nós no sistema [8]. Uma das propriedades do grafo aleatório $G(n, m)$ que pode ser medida diretamente é o grau médio, $\langle k \rangle = 2m/n$ [3], como veremos mais adiante, na seção 2.4.1, grau, também chamado de valência, de um certo nó é o número de nós vizinhos com os quais esse se conecta, que no caso de uma rede simples, é igual ao número de arestas do nó.

Uma outra definição do modelo é de que um grafo é gerado pela escolha uniformemente aleatória entre o conjunto de todos os grafos simples com exatos n nós e m arestas. Estritamente, o grafo aleatório não é definido em termos de uma rede geral simples aleatória, mas como um *ensemble* de redes, ou seja, existe uma distribuição de probabilidade sobre todos os grafos possíveis com n nós e m arestas [3].

Nesse caso, consideramos $G(n, p)$ tal que p é a probabilidade de arestas entre os n nós, mas colocamos uma aresta entre os pares distintos com a probabilidade p , fixamos o número de nós, mas o número de arestas não é fixo.

Em um grafo com n nós e m arestas, ou um par distinto de nós está conectado por uma aresta, ou não está. Supondo que a probabilidade de um nó i está conectado a outro nó dos $(n - 1)$ nós restantes do grafo seja p e de não está conectado seja $q = (1 - p)$.

Para cada nó i de $G(n, p)$ existem mais $(n - 1)$ nós com os quais i pode se conectar. $P[x_i = k]$ denota a probabilidade de que o grau do nó i seja igual a k . Temos $\binom{n-1}{k}$ possibilidades equiprováveis para escolher k diferentes nós para conectar a i . A probabilidade de i estar conectado a exatos k nós e não conectados a exatos $(n - 1 - k)$ nós é igual a $p^k(1 - p)^{(n-1-k)}$, tal que

$$P[x_i = k] = \binom{n-1}{k} p^k (1 - p)^{n-1-k}, \quad (2.1)$$

que é a distribuição Binomial [8].

Na estatística, se X é uma variável aleatória de um conjunto de $\{x_1, x_2, \dots, x_n\}$, o valor esperado $\mathbb{E}(X)$ pode ser obtido a partir da distribuição de pro-

bilidade P ,

$$\mathbb{E}(X = x_i) = \sum_{i=1}^n x_i P[X = x_i]. \quad (2.2)$$

A soma de toda probabilidade de todos os x_i deve ser igual a 1,

$$\sum_{i=1}^n P[X = x_i] = 1, \quad (2.3)$$

que é a condição de normalização. Dessa forma, se queremos o valor esperado como o grau médio $\langle k \rangle$, podemos calcular seu valor por,

$$\langle k \rangle = \sum_{k=1}^{n-1} k P[x = k] = p(n-1). \quad (2.4)$$

Para demonstrar, vejamos [9]:

$$\begin{aligned} \sum_{k=1}^{n-1} k P[x = k] &= \sum_{k=1}^{n-1} \binom{n-1}{k} k p^k (1-p)^{n-1-k} \\ &= \sum_{k=1}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} k p^k (1-p)^{n-1-k} \\ &= \sum_{k=1}^{n-1} \frac{(n-1)(n-2)!}{k(k-1)!(n-k-1)!} k p \cdot p^{k-1} (1-p)^{n-1-k} \\ &= p(n-1) \sum_{k=1}^{n-1} \frac{(n-2)!}{(k-1)!(n-k-1)!} p^{k-1} (1-p)^{n-1-k}, \end{aligned} \quad (2.5)$$

fazendo $l = k - 1$, podemos substituir $k = l + 1$, e podemos reescrever a expressão tal que,

$$\sum_{k=1}^{n-1} k P[x = k] = p(n-1) \sum_{l=0}^{n-2} \frac{(n-2)!}{l!(n-l-2)!} p^l (1-p)^{n-l-2} \quad (2.6)$$

$$= p(n-1) \sum_{l=0}^{n-2} \binom{n-2}{l} p^l (1-p)^{n-l-2}, \quad (2.7)$$

fazendo $r = n - 2$ e reescrevendo a expressão vemos que,

$$\sum_{k=1}^{n-1} k P[x = k] = p(n-1) \sum_{l=0}^r \binom{r}{l} p^l (1-p)^{r-l}. \quad (2.8)$$

Usando o binômio de Newton, desde que $q = (1 - p)$, temos,

$$\sum_{l=0}^r \frac{r!}{l!(r-l)!} p^l (1-p)^{r-l} = (p+q)^r = 1, \quad (2.9)$$

e a equação (2.8) se reduz a

$$\sum_{k=1}^{n-1} k P[x = k] = \sum_{k=1}^{n-1} k P[x = k] = p(n-1) \cdot 1. \quad (2.10)$$

Portanto, a melhor hipótese do valor de $\langle k \rangle$ para um dado p é:

$$\langle k \rangle = \sum_{k=1}^{n-1} k P[x = k] = p(n-1). \quad (2.11)$$

Como nos interessa verificar as propriedades de redes grandes, com $n \rightarrow \infty$, podemos observar a partir da equação (2.11), $\lambda = p(n-1)$, que $p = \lambda/(n-1)$ é muito pequeno quando $n \rightarrow \infty$, tal que podemos escrever,

$$\ln [(1-p)^{n-1-k}] = (n-1-k) \ln \left[1 - \frac{\lambda}{n-1} \right] \quad (2.12)$$

$$\simeq -(n-1-k) \frac{\lambda}{n-1} \quad (2.13)$$

$$\simeq -\lambda \quad (2.14)$$

e então podemos expandir em série de Taylor para $n \rightarrow \infty$. E elevando ambos os lados à exponencial, encontramos que $(1-p)^{1-n-k} = e^{-\lambda}$ no limite de n grande. Além disso, para n grande temos,

$$\binom{n-1}{k} = \frac{(n-1)!}{k!(n-1-k)!} \simeq \frac{(n-1)^k}{k!} \quad (2.15)$$

e a equação (2.1) se torna,

$$P[x_i = k] = \frac{(n-1)^k}{k!} p^k e^{-\lambda} \quad (2.16)$$

$$= \frac{(n-1)^k}{k!} \left(\frac{\lambda}{n-1} \right)^k e^{-\lambda} \quad (2.17)$$

$$= e^{-\lambda} \frac{\lambda^k}{k!} \quad (2.18)$$

Assim, para n grande pode ser substituída a distribuição Binomial pela distribuição de

Poisson [8]:

$$P(x_i = k) \simeq e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (2.19)$$

O estudo da distribuição do grau do nó desse tipo de redes, todavia, não condiz com redes encontradas na natureza. Segundo o modelo de redes aleatórias, para uma rede com N grande, o grau do nó apresenta uma distribuição de Poisson enquanto que para algumas redes reais, essa distribuição é do tipo lei de potência.

2.1.3 Um experimento social

Em 1969, Travers e Milgram publicaram um artigo [10] com resultados empíricos de um estudo experimental que denominaram de Estudo do Problema de Pequenos Mundos [11]. Para esse estudo foram selecionados aleatoriamente 296 pessoas de Nebraska e Boston para que, gerando uma “cadeia de conhecidos”, conseguissem entregar uma carta ao destinatário de Massachusetts. Desse total, apenas 64 correspondências conseguiram chegar ao destino.

Em resumo o documento continha as regras do jogo, os dados da pessoa alvo, uma lista e um cartão marcador. A lista era usada para evitar que o documento voltasse para as mãos de uma pessoa que já tivesse sido atingida pelo mesmo. E o cartão marcador, para coletar as informações sociais necessárias ao estudo. O primeiro remetente da carta deveria, se conhecesse o destinatário, enviar-lhe a diretamente carta, senão, enviasse a algum conhecido que intermediaria a entrega ao destinatário. Se o conhecido do primeiro conhecesse o destinatário, que enviasse a carta diretamente, senão, que enviasse a um conhecido que intermediaria a entrega ao destinatário e assim por diante, até que a carta chegue ao destino. O resultado disso, portanto, consiste de uma distribuição do comprimento da cadeia. Foram analisados apenas documentos que conseguiram chegar ao destino. Nessa distribuição, o número de intermediários entre o remetente e o destinatário foi chamado de comprimento da cadeia. Assim, o estudo mostrou que o valor médio dessa distribuição ficava em torno de 5,2 conexões.

A grande contribuição do trabalho de Travers e Milgram está relacionada ao número médio de intermediários observados, cujo valor é maior que cinco, entre o remetente e o destinatário. Daí o conceito de que qualquer pessoa está a *seis graus* de separação de outra.

2.1.4 Modelo de vantagens acumulativas

Um dos interesses de alguns estudiosos é encontrar algum tipo de padrão que possa ser observado em diferentes sistemas. O economista Simon [12] identificou em 1955 uma classe de função distribuição que decaía com lei de potência observando dados empíricos que descreviam fenômenos sociais, biológicos e econômicos. Ele observou, então, a distribuição de palavras numa prosa pela frequência de ocorrência, distribuição de cientistas pelo número de publicações, distribuições de cidades por população, distribuição de renda e outros. Ele percebeu que nesses casos, mesmo sendo tão distintos quanto à natureza que os formava, havia algo em comum, a distribuição apresentava uma curva com uma longa cauda de decaimento, que sugeria uma equação de decaimento de lei de potência.

Apesar de ter sido um grande estudo sobre processos estocásticos diversos que apresentavam um padrão de distribuição, ainda não se falava em um modelo estruturado em redes ou grafos.

Somente mais tarde, por volta dos anos 1965, o físico-historiador Price [13] observa que a distribuição da frequência de citações nas publicações do ano de 1961 seguia uma lei de potência. Nessa pesquisa, ele estuda as publicações científicas e a rede de citações, cujas publicações eram conectadas uma a outra a partir das citações em notas de rodapé ou referência bibliográfica. Podemos dizer que pela primeira vez é observada uma lei de decaimento de lei de potências numa distribuição baseado num sistema em redes.

Em 1976, Price intrigado com o fenômeno de ‘sucesso gera sucesso’ sugere, após estudo de dados empíricos com rede de citações, uma distribuição de vantagens acumulativas bastante estudada em fenômenos sociais como modelo para o fenômeno de sucessos [14]. Assim, propôs um simples e elegante modelo de formação de redes que dava origem à distribuição de grau em lei de potência [3]. Hoje esse modelo é conhecido como ‘preferência de conexão’ muito popularmente atribuído ao Modelo de Barabási de redes sem escala.

Embora tenhamos observado que os trabalhos de Price juntamente com os trabalhos de Simon tenham sido anteriores ao de Barabási, o último descreve de forma clara a construção de redes reais propondo que o crescimento e preferência de conexão são os elementos essenciais para se obter uma curva de distribuição lei de potência [2].

2.2 Redes reais

As redes reais, todavia, não obedecem uma modelagem de redes nem regulares e nem completamente aleatórias, mas se mostram com características que são meio termo dos dois modelos. Isso porque as redes reais apresentam uma série de fatores que não foram consideradas na teoria de Grafos Aleatórios como, por exemplo, a entrada de novos nós ao sistema. As descobertas nesse campo são devidas às investigações realizadas em rede de interação de sistemas complexos [1, 2, 15]

A grande disponibilidade de computadores e redes de comunicação dos últimos anos colaborou para o crescente interesse por parte dos pesquisadores acerca do estudo de redes complexas, pois permitiu-se reunir e analisar uma grande massa de dados do mundo real. Os novos movimentos nessa área, porém, desviam-se do estudo de redes pequenas para redes maiores, graças a esses recursos [16].

A evolução tecnológica de processadores mais velozes e maior capacidade de armazenamento possibilitaram estudos de redes cada vez maiores, incluindo redes reais de todos os tipos. É a partir disso que investigações de propriedades estatísticas, tais como centralidades, menor caminho, aglomerações, comunidades e outros, foram se tornando populares na expectativa de compreender a estrutura de redes reais através da topologia. Da mesma forma, houve avanços no estudo de processos dinâmicos de redes complexas, cujo objetivo está relacionado em se verificar a navegabilidade [17], a robustez [18, 19], a eficiência [20, 21, 22], a difusão em redes [23, 24], navegação ótima [25, 26] de sistemas reais.

Novos modelos alternativos aos modelos básicos de redes complexas também vêm surgindo a cada dia. Alguns estudos mostram, portanto, que uma rede sem escala pode ser também uma rede de mundo pequeno, por duas razões: redes reais sem escala (*i*) possuem coeficiente de aglomeração tão grandes quanto de redes aleatórias e (*ii*) seu diâmetro aumenta logaritmicamente com o número de nós [27].

Assim, no final da década de 1990, dois modelos de redes são adotados na análise de redes reais: Redes de Pequenos Mundos [1] e Redes sem escala [2]. E a partir desses dois modelos, muitos outros foram surgindo, tais como estudos de classes de redes de Pequenos Mundos [27], redes complexas em camadas [28], sincronização em redes complexas [29], modularidade [30] e comunidade em redes [30] e outros.

2.2.1 O modelo de Watts-Strogatz (WS)

Em 1998, Duncan J. Watts e Steven H. Strogatz, observando redes de sistemas dinâmicos acoplados, exploraram um modelo simples de redes, no qual uma rede regular era reconectada com a finalidade de introduzir uma desordem, até a obtenção de um grafo puramente aleatório. A desordem era atribuída pela variável ρ , onde $\rho = 0$ significava que a desordem era zero, tal que a rede fosse regular e $\rho = 1$ significava que a desordem era máxima, tal que a rede fosse totalmente aleatória [1].

No estudo de Watts e Strogatz duas características são levadas em conta: o Coeficiente de Aglomeração ($\mathcal{C}(\rho)$) e o Comprimento de Caminho Médio ($\mathcal{L}(\rho)$). Redes regulares apresentavam $\mathcal{C}(\rho)$ e $\mathcal{L}(\rho)$ grandes e redes puramente aleatórias $\mathcal{C}(\rho)$ e $\mathcal{L}(\rho)$ pequenos. No entanto, analisando redes de atores substitutos, redes de energia do ocidente dos Estados Unidos e redes neurais de *C-elegans*, eles encontraram sistemas altamente agregados — característica de grafos regulares — e um reduzido comprimento de caminho médio —, característica de um grafo puramente aleatório. Em analogia ao fenômeno de pequeno mundo de Travers e Milgram, Watts e Strogatz chamaram essa rede de *rede de pequeno mundo* [1].

Redes de pequenos mundos, juntamente com redes aleatórias pertencem a classe de redes complexas que apresentam uma distribuição do grau do nó $P(k)$ com um pico em $\langle k \rangle$ e decaimento exponencial para k grande. São redes nas quais cada nó tem aproximadamente o mesmo número de arestas $k \simeq \langle k \rangle$, e portanto, consideradas homogêneas [18].

O interesse deles eram redes escassa conexão e conectada, ou seja, redes cujo regime obedecia condição $n \gg k \gg \ln(n) \gg 1$, onde $k \gg \ln(n)$ garante que o grafo aleatório seja conectada. Desse regime, eles obtiveram que $\mathcal{L} \sim n/2k$ e $\mathcal{C} \sim 3/4$ para $\rho \rightarrow 0$ e $\mathcal{L} \sim \ln(n)/\ln(k)$ e $\mathcal{C} \sim k/n$ para $\rho \rightarrow 1$ [1].

\mathcal{L} é definido como o número médio de arestas do menor caminho entre dois nós, medido sobre todos os pares de nós. Bem como o número médio de pessoas na rede de relacionamento, cujos nós são pessoas.

Um exemplo disso é a propagação de um boato. Em uma rede de mundo pequeno o boato se espalhará mais e mais rapidamente comparado com os dois modelos extremos – rede regular (\mathcal{C} e \mathcal{L} grandes) e redes puramente aleatórias (\mathcal{C} e \mathcal{L} pequenos). Se por um lado, redes regulares são altamente agregados, garantindo que muitos dos amigos de X (fonte do boato) se conheçam, para o boato sair do grupo de X e chegar ao último

indivíduo Y da rede, terá de ter percorrido um longo caminho em termos de interação pessoa, de forma que o boato se propagará lentamente. Por outro lado, redes puramente aleatórias que são redes que possuem caminho médio pequeno, o boato terá de passar por poucos indivíduos de X até Y , por ser uma rede pouco agregada, todavia, nem todos saberão do boato, podendo chegar a pessoa Z e ali morrer, pois Z não conhece mais ninguém da rede.

Apesar dos trabalhos de Watts e Strogatz terem sido usados um modelo de disseminação de doenças contagiosas, os conceitos de propriedades estruturais local (\mathcal{C}) e global (\mathcal{L}) apresentados por eles foram importantes para a compreensão de redes reais como um todo. Inclusive se observa em seus trabalhos o conceito de robustez de uma rede como propriedade relacionada à estrutura da rede [1].

2.2.2 O modelo de Barabási-Albert (BA)

A lei de potência da distribuição do grau do nó foi discutido pela primeira vez no trabalho de Barabási e Albert [2]. Eles investigaram o comportamento de grandes redes complexas, tais como a rede de *www* ou de citações de publicações científicas, e mostraram que a probabilidade $P(k)$ de interação entre um nó e outros k nós numa rede sofre decaimento de uma lei de potência, segundo a equação [31]:

$$P(k) \sim k^{-\gamma}. \quad (2.20)$$

Esse modelo indica que redes grandes se organizam em um estado sem escala. Para explicar isso, Barabási e Albert incorporaram dois elementos chaves que estão presentes em redes reais: o *crescimento* e a *preferência de conexão* [15]. São consideradas redes não-homogêneas [18]. Ficaram assim evidenciadas as falhas do modelo de rede aleatória como modelo de redes reais, pois a aproximação matemática para redes *fechadas*, ou seja, redes com número fixo de nós e arestas, não condiz com a estrutura de redes reais, que são *abertas* [8].

Crescimento e preferência de conexão são simultaneamente essenciais para a obtenção de uma curva de distribuição de lei de potência [8]. Para verificar isso, eles consideraram dois modelos: *Modelo A* e *Modelo B*. O primeiro modelo mantém o caráter de crescimento da rede, mas elimina a preferência de conexão dos nós. O segundo, ao contrário, a quantidade de nós do sistema é mantida constante e as conexões são feitas para um nó aleatoriamente selecionado [15]. Com isso observaram que ambos os elementos são necessários para obtenção de uma distribuição do tipo lei de potência, como era observado

empiricamente.

Segundo Barabási e outros, a distribuição do grau do modelo A é dada por um decaimento exponencial $P(k) = (e/m) \exp(-k/m)$. Esse comportamento indica ausência de preferência de conexão [8]. E para o modelo B, sugere que a distribuição do grau do nó seja uma Gaussiana, desde que $k_i \simeq (2t/N)$, com $N \gg 1$. E o modelo sugere que a rede apresente o valor igual a $k_i \simeq (2t/N)$ para todos os nós após o período de transição [8].

Acerca do desenvolvimento independente ou não dos estudos de redes complexas do fim da década de 1990 foram bastante discutidos em [32, 16]. A bem verdade, Barabási e Albert [2] não citam os trabalhos desenvolvidos por Price ou Simon (subseção 2.1.4) para o processo de formação de redes com crescimento e distribuição de vantagens acumulativas, que é semelhante a preferência de conexão em redes complexas. O estudo de redes por Barabási e Albert tiveram notoriedade no conceito de formação de redes por trazer com detalhes as propriedades de crescimento e preferência de conexão em redes reais como elementos essenciais para a obtenção de distribuição do grau em lei de potência.

Inclusive, Watts [32] ressalta que a atual geração de estudos de redes, o qual chama de “nova ciência de redes”, são pesquisas que emergiram rapidamente e com sínteses altamente interdisciplinar das novas técnicas de análise e com maior uso de dados empíricos e poder computacional.

2.2.3 Redes Multiplex

Kurant e outros introduzem o modelo de Redes em multicamadas [28]. Eles fizeram um estudo do sistema de transporte analisando uma rede com camadas distintas, onde a estrutura física, o fluxo de tráfego (dado pela tabela horária de ônibus, trem, etc) e o grafo lógico são gerados como diferentes camadas da rede. Naquele estudo são definidas propriedades que possibilitam comparar a camada do grafo lógico com a camada da estrutura física. Os autores aplicam a técnica em três grandes cidades [33] e trabalham com a tolerância a erros e os ataques em redes multicamadas [34].

Funções qualitativas acoplando matrizes adjacentes das múltiplas fatias são introduzidas para o estudo em comunidades em redes para comparar redes que evoluem no tempo. O estudo desenvolve uma estrutura formada por uma combinação de redes individuais acopladas por ligações que conectam cada nó em uma fatia ao mesmo nó que se encontra em outra fatia. A comparação é feita por meio de funções qualitativas que comparam o número de arestas intracomunidade que se deve esperar do caso aleatório

[35].

Uma rede multiplexa é um grafo de multicamadas onde os nós pertencem a várias camadas diferentes (isto é, redes monoplex) simultaneamente e estão ligados por meio de um conjunto específico de arestas em cada camada [36]. Em uma topologia multiplex cada camada é conectada bidirecionalmente a outras camadas.

Em [37] os autores mostram que os processos de difusão em multiplex podem ter um comportamento difusivo melhorado, o que significa que as escalas de tempo associadas à difusão em todo o multiplex podem ser mais curtas do que as associadas às camadas individuais. E concluem que este fenômeno está estritamente relacionado com a configuração particular do multiplex, e não tem contrapartida em redes "monoplex" clássicas.

Estudos em redes multiplex tem sido bastante usadas para a compreensão de redes reais. Em [36] os autores descrevem ferramentas teóricas para avaliar a navegabilidade em redes interconectadas sob falhas aleatórias, responsáveis pela estrutura e pela dinâmica inerentes aos sistemas reais. Eles mostram que a eficiência na exploração de redes multiplex depende criticamente da topologia das camadas, das forças de interconexão e da estratégia de caminhada.

Uma forma de quantificar o custo para mudar da camada α para a camada β para uma viagem a partir do nó i , é usada a matriz $D_{(i)}^{\alpha\beta}$. De modo que é possível distinguir entre a força $s_i = \sum_j W_i^{(\alpha)}$ de um vértice i em relação às suas conexões com outros vértices j ($j = 1, 2, 3, \dots, N$) em uma mesma camada e a força $S_{i,\alpha} = \sum_{\beta} D_{(i)}^{\alpha\beta}$ do mesmo nó com respeito a conexões a suas contrapartes em diferentes camadas [36]. Um exemplo do uso de multiplex para redes de transporte de uma cidade, pode ser feita definindo-se cada camada como a rede de diferentes modais de transporte, como metrô, ônibus, trem.

2.2.4 Exemplos de redes reais

Quase todos os tipos de sistemas reais podem ser representados por redes complexas. São inúmeros estudos empíricos de redes realizados atualmente. Alguns autores [38, 3], por questões didáticas, dividem as várias redes em grupos: Redes Tecnológicas, Redes Sociais, Redes de Informação, Redes Biológicas.

Algumas das redes tecnológicas mais estudadas são a internet, a telefonia, as redes elétricas, as redes de transporte e as redes de distribuição ou de entregas. A

www (*world wide web*), redes de citações, P2P (*peer-to-peer*), indexação de palavras-chaves e outros são agrupados em redes de informação. No grupo das redes biológicas são encontradas as redes bioquímicas, neurais e ecológicas. E nas chamadas redes sociais estão presentes todos os tipos de sistemas que utilizam o relacionamento pessoal como aresta, seja por relacionamentos amorosos, transmissão de doenças contagiosas, comércio, relações econômicas e outros [3].

De algum modo essas redes foram agrupadas levando-se em conta alguma propriedade em comum. Em redes do tipo tecnológico, apesar de rede de internet e de transportes parecerem bem distintas na sua natureza, ambas possuem a função de levar algo de um lugar a outro. A internet envia pacotes de dados de um computador ou servidor a outro, já a rede de transporte carrega pessoas ou cargas de um lugar a outro, e assim por diante.

Já Barthelemy [39] define redes de transporte de forma mais geral, tal que redes de transportes são estruturas que transmitem energia, matéria ou informação de um ponto a outro. Essas redes controlam muitos aspectos da sociedade e influenciam muitos problemas modernos, tais como o contágio de doenças, congestionamentos, espalhamento urbano e estrutura de cidades. Ele inclui as redes tecnológicas como um subgrupo das redes de transporte e não o contrário. Por serem redes cuja a distância influencia na definição da topologia, ele as considera como redes espaciais.

Ou seja, nessa classificação são agrupadas rede elétrica, rede neural, redes de comunicação e redes sociais. Em comum essas redes possuem alguma dependência da definição da topologia associada à distância. Por exemplo, na rede de relacionamentos, é comum que se faça amizade com as pessoas da vizinhança. O custo do transporte de pessoas de uma região é associada também a distância, em geral, o custo é maior quanto mais se desloca. Da mesma forma, para a distribuição de energia elétrica ou internet a cabo, os custos aumentam com a distância, visto que será necessário mais cabos, postes, e outros materiais quanto maior é a distância da fonte ao consumidor.

Embora não haja um consenso na classificação de redes reais, um ponto importante a se observar é o uso cada vez mais frequente de dados reais para investigar de forma empírica as propriedades de redes complexas de vários sistemas.

2.3 Grafos e matrizes

É natural que a representação matemática de grafos seja feita por matriz bidimensionais, pois cada elemento da matriz representa uma interação a cada dois ele-

mentos.

2.3.1 Grafos

A rigor, um grafo é um par de conjuntos $G = \{N, E\}$, onde $N = N(G)$ é o conjunto de n nós (ou vértices) e $E = E(G)$ é o conjunto de m arestas (ou arcos) que conectam nós dois a dois. Se o grafo é finito, então N e E também são finitos [4]. Uma aresta $\{i, j\}$ conecta os nós i e j e é denotada por ij . Assim, ij e ji são exatamente a mesma aresta, os nós i e j são os *nós finais* dessa aresta.

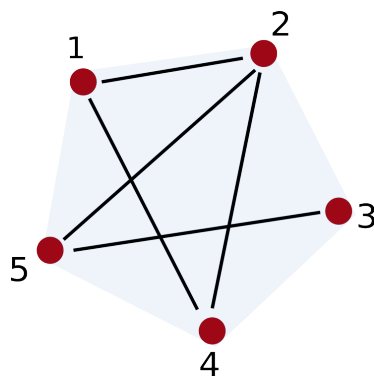


Figura 2.2 - Ilustração de um grafo com 5 nós e 5 arestas. Onde temos $N = \{1, 2, 3, 4, 5\}$ e $E = \{\{1, 2\}, \{1, 4\}, \{2, 4\}, \{2, 5\}, \{3, 5\}\}$. (Fonte: [40])

A representação ilustrativa de um grafo é feita geralmente por um conjunto de pontos ligados um ao outro quando existe algum tipo de conexão entre os nós e não ligados quando não há conexão. Na Figura (2.2), temos $N = \{1, 2, 3, 4, 5\}$ e $E = \{\{1, 2\}, \{1, 4\}, \{2, 4\}, \{2, 5\}, \{3, 5\}\}$.

2.3.2 Estrutura de uma rede

As redes podem ser classificadas conforme critérios estruturais, tais como [24]:

- Direção: as redes podem ser direcionadas ou não direcionadas. Se o sentido da conexão representa alguma informação relevante, então ela é direcionada, caso contrário, é não direcionada. Em outras palavras, isso significa que existirão nós origens e nós destinos, de forma que haverá diferença entre uma aresta do nó i ao nó j e uma aresta do nó j ao nó i .
- Peso: existem informações importantes que vão além da simples informação da

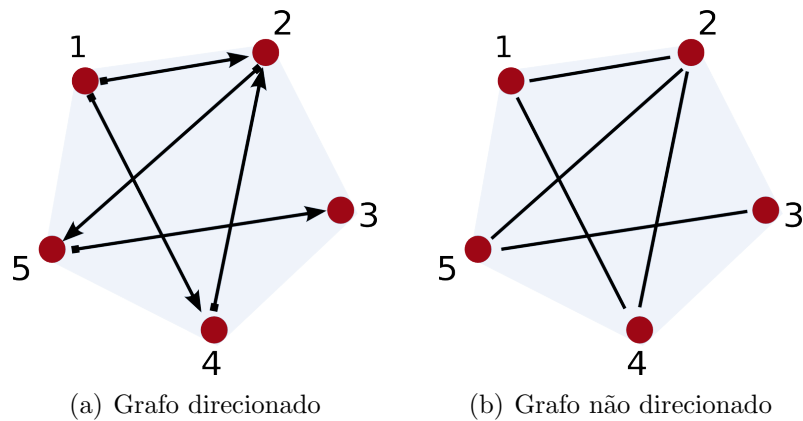


Figura 2.3 - Direção de arestas, existência ou ausência de nó alvo e nó fonte; 2.3(a): Grafo direcionado; 2.3(b): Grafo não direcionado. (Fonte: [40])

existência de aresta entre dois nós. Uma conexão pode ser mais forte que outra, por meio de algum parâmetro. Quando é levado em consideração o peso das conexões, dizemos que a rede é ponderada. O peso pode equivaler à quantidade de arestas existentes na conexão de um par de nós [41, 42]. Estudos de redes ponderadas são bastante comuns na literatura, podendo abordar desde redes de colaboradores [43] a redes *world wide web* (*www*) e transportes aéreos [44].

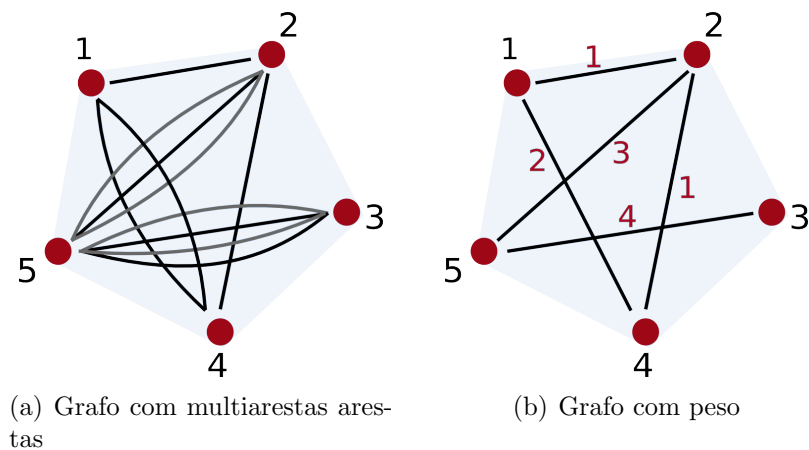


Figura 2.4 - Rede com multarestas arestas pode ser representada por um rede ponderada; 2.4(a): Rede com multarestas arestas; 2.4(b): Rede ponderada, onde cada número vinculado à aresta representa o número de arestas entre nó i e o nó j . (Fonte: [40])

- Conexão: as redes podem ser ou não conectadas. Se uma rede é conectada, então não existem nós isolados, havendo apenas um único aglomerado. Se ela é não

conectada, então existem nós isolados, podendo existir mais de um aglomerado na rede.

- Auto-arestas: não é muito raro grafos apresentarem nós que possuem arestas consigo mesmo, são os chamados auto-aresta.
- Multiarestas: quando existem múltiplas arestas conectando um par de nós.

Quando um grafo não apresenta nem auto-arestas e nem multiarestas, são chamados de grafos simples ou redes simples. Já os grafos que apresentam multiarestas são chamados de multigrafos [3].

2.3.3 Matriz de adjacências

Para a representação numérica dos nós e arestas de um grafo, são utilizados como ferramenta matemática as matrizes bidimensionais. A matriz mais conhecida é a Matriz de Adjacências, \mathbf{A} , que é uma matriz $n \times n$, cujos os elementos a_{ij} representam a conexão entre os nós i e j , tal que $a_{i,j} = 0$, se não existe nenhuma aresta entre os nós i e j , e $a_{ij} = 1$, se existem arestas entre i e j [4].

Para ilustrar, podemos construir a matriz de adjacências \mathbf{A} para representar o grafo da Figura (2.2). Percebemos que \mathbf{A} é simétrica, pois a rede não direcionada. Isso, porque num grafo não direcionado a conexão de i para j é igual ao de j para i .

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

A matriz de adjacências possui os elementos da diagonal principal, $a_{ij} = 0$ quando $i = j$, pois, em geral, representam grafos simples. Isso significa que o grafo não possui nem auto-arestas nem multiarestas. Quando o grafo é não direcionado, a matriz adjacente é simétrica em relação à diagonal principal. Se, por sua vez, o grafo é direcionado, a matriz pode ser não simétrica. Como ilustração, podemos construir a matriz \mathbf{A}' para representar o grafo direcionado na Figura (2.3(a)).

$$\mathbf{A}' = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

É possível, ainda, trabalhar com uma rede ponderada por meio de matrizes. Para isso podemos usar como exemplo a Figura (2.4). Nesse caso, basta construir uma matriz ponderada \mathbf{W} do grafo G , cujos elementos w_{ij} representam não apenas a existência de conexão entre os nós i e j , mas o peso ou força de suas conexões. Uma rede com multiarestas entre dois nós quaisquer pode também ser representada por uma matriz \mathbf{W} , desde que o valor de w_{ij} seja definido como a quantidade de arestas entre i e j .

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 4 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 3 & 4 & 0 & 0 \end{pmatrix}$$

Se $w_{ij} = 0$, significa que não há conexão, e se $w_{ij} \neq 0$, há conexão. Quando multiarestas entre dois nós são equiprováveis, podemos definir w_{ij} como a soma da quantidade de arestas entre i e j .

2.4 Medidas de centralidade

Algumas medidas de centralidade fornecem informações sobre a estrutura interna de uma rede, podendo ser usadas para se verificar sua eficiência. São várias as propriedades de centralidade, entre elas, o grau do nó, o comprimento médio, o coeficiente de aglomeração e o coeficiente de intermediação (ou *betweenness*).

2.4.1 Grau do nó

O grau (ou valência) k_i é o número de arestas do nó i ; que é igual ao número de vizinhos desse nó. Onde vizinho define-se como um nó conectado ao nó i . Um nó i de grau $k_i = 0$ é um nó isolado. Significa que não possui nenhum vizinho. Se todos os nós do grafo G têm o mesmo grau k , então G é k -regular, ou simplesmente regular [45].

Em redes espaciais, geralmente o grau do nós mostra um pico na sua distri-

buição, por causa do vínculo espacial [39].

Em uma rede não ponderada o grau do nó k_i expressa a quantidade de nós com os quais o nó i possui conexão. Esse valor pode ser calculado por meio das matrizes adjacentes (2.3.3),

$$k_i = \sum_{j=1}^n a_{ij}. \quad (2.21)$$

onde i é o i -ésimo nó analisado e a_{ij} é o elemento da matriz [4]. De forma que, $a_{ij} = 1$, se existir uma conexão de i para j e $a_{ij} = 0$ se não houver conexão [24]. Para o cálculo do grau médio $\langle k \rangle$ podemos somar todos os k_i e dividir pelo número total de nós da rede G ,

$$\langle k \rangle = \frac{1}{n} \sum_i k_i, \quad (2.22)$$

onde n é o número de nós da rede não direcionada. De outra forma, em redes não-direcionadas cada aresta tem duas 'extremidades' (*ends*) conectadas a ele. Se m é o número de arestas da rede G , o número de extremidades será $2m$. Por outro lado, o número de extremidades é também igual à soma dos graus de todos os nós, ou seja,

$$2m = \sum_i k_i. \quad (2.23)$$

Isso nos leva a uma equação simples para calcular o grau médio da rede:

$$\langle k \rangle = \frac{2m}{n}. \quad (2.24)$$

Para o caso de redes direcionadas em que a aresta é de i para j , temos k_j^{in} e k_i^{out} , respectivamente, grau das arestas que entram no nó j qualquer e que saem do nó i qualquer.

$$k_j^{in} = \sum_{i=1}^n a_{ij} \quad (2.25)$$

e

$$k_i^{out} = \sum_{j=1}^n a_{ij}. \quad (2.26)$$

O número de arestas m de uma rede direcionada é igual ao número de extremidades entrando em todos os nós, ou ainda, o número de extremidades saindo de todos os nós,

$$m = \sum_{j=1}^n k_j^{in} = \sum_{i=1}^n k_i^{out} = \sum_{ij} a_{ij}. \quad (2.27)$$

Portanto, para redes direcionadas o grau médio é,

$$\langle k^{in} \rangle = \langle k^{out} \rangle = \frac{\langle k \rangle}{2} = \frac{m}{n}. \quad (2.28)$$

A diferença do grau médio de rede direcionada e não direcionada é dada apenas pelo fator dois [3].

As propriedades de redes não ponderadas podem ser empregadas no estudo de redes ponderadas também, tais como o grau do nó [41], que recebe a denominação de força do nó (*strength*) s_i do nó i [24]. Nesse caso, não apenas a existência de conexão entre um par de nós é contada, mas também quantas formas diferentes um par de nó é conectado.

Podemos representar a operação dos nós s_i como a soma de todos os elementos matriciais pertencentes à linha i , ou seja, a soma dos valores dos elementos de cada coluna para encontrar a força do nó i , s_i :

$$s_i = \sum_{j=1}^m w_{ij}. \quad (2.29)$$

onde w_{ij} é a força da conexões entre os nós i e j , com $w_{ij} = 0$ quando não há conexão e $w_{ij} \neq 0$ quando há conexão.

As matrizes são simétricas se e somente se representarem redes não direcionadas [24].

2.4.2 Distribuição do Grau

Vimos que redes aleatórias possuem distribuição do grau típico de uma curva de distribuição Binomial e para $n \rightarrow \infty$, pode se aproximar a distribuição de Poisson. Para redes reais, no entanto, encontramos distribuições em Lei de Potência, tal como as redes sem escala de Barabási.

Isso significa matematicamente que a distribuição do grau P é uma função linear do grau k , assim:

$$\ln P_k = -\gamma \ln k + c, \quad (2.30)$$

onde α e c são constantes [3]. Tomando a exponencial de ambos os lados temos,

$$P_k = C k^{-\gamma}, \quad (2.31)$$

onde $C = e^c$ é uma constante. A distribuição varia como potência de k , daí a chamada lei de Potência. A constante γ , conhecida como expoente da lei de potência encontra-se geralmente no intervalo de $2 \leq \gamma \leq 3$. Em alguns casos pode ocorrer variações ligeiramente fora desse intervalo.

Contudo, uma verdadeira distribuição de lei de potência deveria ser monotonicamente decrescente para todo intervalo, mas observa-se em muitos casos que há um desvio da lei de potência para k pequeno e que a lei de potência manifesta-se na cauda da distribuição, ou seja, para k grande [3].

Para melhor visualizar uma distribuição lei de potência pode-se recorrer a estratégia de utilizar a função distribuição acumulada, que é definido por,

$$P_k = \sum_{k'=k}^{\infty} p_{k'} \quad (2.32)$$

Ou seja, P_k é a fração de nós que tem grau k ou maior que k . Significa dizer que é a probabilidade de escolher um nó com grau maior ou igual a k .

2.4.3 Coeficiente de aglomeração

O coeficiente de aglomeração (*clustering*) é um parâmetro local da rede, que revela a conectividade da vizinhança de um certo nó i . Essa vizinhança de i é definida como um subgrafo formado por todos os nós com os quais o nó i é conectado. A partir disso, é verificado o número de conexões existentes entre os nós pertencentes a vizinhança de i .

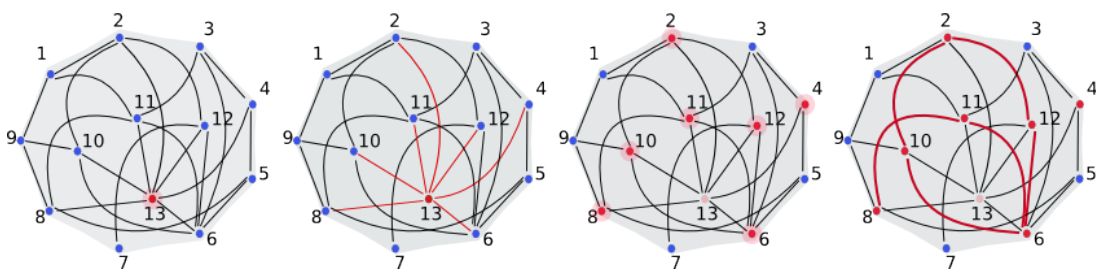


Figura 2.5 - Uma rede complexa formada por $N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$. Para calcular o coeficiente de aglomeração de 13, precisamos encontrar todos os nós com os quais 13 está conectado, se são seus vizinhos de primeira ordem, que são: $N = \{2, 4, 6, 8, 10, 11, 12\}$. Com exceção do próprio nó 13, contamos quantas arestas são formadas no grafo entre os vizinhos de 13: $E = \{(8, 11), (6, 10), (6, 11), (6, 12), (10, 2), (2, 12)\}$. Calculamos quantas arestas são possíveis e usando a equação 2.33. (Fonte: [40])

Assim, nós voltamos a atenção para um certo nó i de uma rede, que possui

k_i arestas que conectam a k_i outros nós. Se os vizinhos mais próximos do nó i — ou seja, os nós que fazem parte do aglomerado de i — são ao todo k_i nós, então teremos $k_i(k_i - 1)/2$ arestas possíveis. Assim, a razão entre o número de arestas que realmente existem entre os nós k_i da vizinhança, E_i , e número total de arestas possíveis, $k_i(k_i - 1)/2$, para um total de k_i nós, resulta no valor do coeficiente de agregação do nó i [8],

$$C_i = \frac{2 E_i}{k_i(k_i - 1)}. \quad (2.33)$$

Podemos ainda calcular o Coeficiente de Aglomeração médio da rede, por:

$$C = \frac{1}{n} \sum_i C_i, \quad (2.34)$$

onde n é o número total de nós da rede, e C_i é o Coeficiente de Aglomeração do nó i .

Na Figura (2.5) temos uma rede complexa formada pelos nós $N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$. Para calcular o coeficiente de aglomeração de 13, por exemplo, precisamos encontrar todos os nós com os quais 13 está conectado, ou seja, se são seus vizinhos de primeira ordem. Então os vizinhos de 13 são: $N = \{2, 4, 6, 8, 10, 11, 12\}$. Com exceção do próprio nó 13, contamos quantas arestas são formadas no grafo entre os vizinhos de 13: $E = \{(8, 11), (6, 10), (6, 11), (6, 12), (10, 2), (2, 12)\}$. Calculamos quantas arestas são possíveis e usamos a equação 2.33.

Em redes espaciais, os nós tendem a se conectar com maior probabilidade com os nós mais próximos que com os nós mais distantes, tal que o valor do coeficiente de aglomeração é alto [39].

O coeficiente de aglomeração depende do número de triângulos que existem na rede e pode ser calculado a partir da matriz adjacente, $C = 1/6 / Tr(A^3)$. Uma medida bastante usual do coeficiente de aglomeração é sua distribuição por grau,

$$C(k) = \frac{1}{N(k)} \sum_{i/k_i=k} C_i. \quad (2.35)$$

2.5 Medidas de caminhos

Caminho é todo conjunto ordenado de nós tais que os pares de nós consecutivos estão conectados por uma aresta [3]. Em uma rede direcionada, cada aresta que passa por um caminho deve ser passada na direção correta. Arestas de redes não-direcionadas,

contudo, a aresta pode passar para qualquer direção. Em geral um caminho pode ser interseccionada, ou seja, um nó pode ser revisitado em algum caminho. Caminho geodésico é um caso especial onde não existe a intersecção.

O comprimento do caminho é o número de arestas que passa ao longo de um caminho [3]. Podemos calcular o comprimento l_{ij} de cada par de nós i e j usando a matriz adjacente \mathbf{A} . Se $a_{ij} = 1$, existe uma aresta entre i e j , se $a_{ij} = 0$, não existe aresta. Ou seja, se $a_{ij} = 1$ existe um caminho direto entre os dois nós i e j , assim $l_{ij} = 1$. Contudo, queremos saber para os casos em que não há conexão direta, ou seja, para os casos em que $a_{ij} = 0$. Se $a_{ik}a_{kj} = 1$, então existe um caminho de comprimento 2 entre i e j . De forma geral, a matriz de caminhos de comprimento 2 é definida como:

$$l_{ij}^{(2)} = \sum_{k=1}^n a_{ik}a_{kj} = [\mathbf{A}^2]_{ij}, \quad (2.36)$$

onde $[\dots]_{ij}$ denota o elemento de matriz ij . Para caminhos com 3 arestas,

$$l_{ij}^{(3)} = \sum_{k,m=1}^n a_{ik}a_{mj}a_{km} = [\mathbf{A}^3]_{ij}. \quad (2.37)$$

Então, de forma geral, para calcular o elemento com comprimento r :

$$l_{ij}^{(r)} = \sum_{k,\dots,m=1}^n a_{ik}\dots a_{mj} = [\mathbf{A}^r]_{ij}. \quad (2.38)$$

Para a matriz $\mathbf{L}^{(r)}$, existirá caminhos entre i e j , se $\mathbf{l}_{ij}^{(r)} \neq 0$.

2.5.1 Comprimento do menor caminho

O caminho geodésico, também chamado de menor caminho, fornece informações sobre o caminho ótimo em redes complexas. Se o objetivo é encontrar o caminho ótimo para um pacote de dados trafegar pela internet de um computador a outro, o menor caminho fornece essa informação. Por isso, o menor caminho desempenha uma importante função na caracterização da estrutura interna de um grafo [38].

Uma vez que temos uma matriz de adjacências \mathbf{A} (2.3.3), é possível calcular a matriz do comprimento do menor caminho \mathbf{L} , onde d_{ij} é o comprimento do menor caminho entre dois nós i e j quaisquer. Esse comprimento também é chamado de comprimento da geodésica, ou seja, é a menor soma das distâncias entre todos os possíveis caminhos entre i e j [20, 38]. O valor máximo de d_{ij} é chamado de diâmetro do grafo [38].

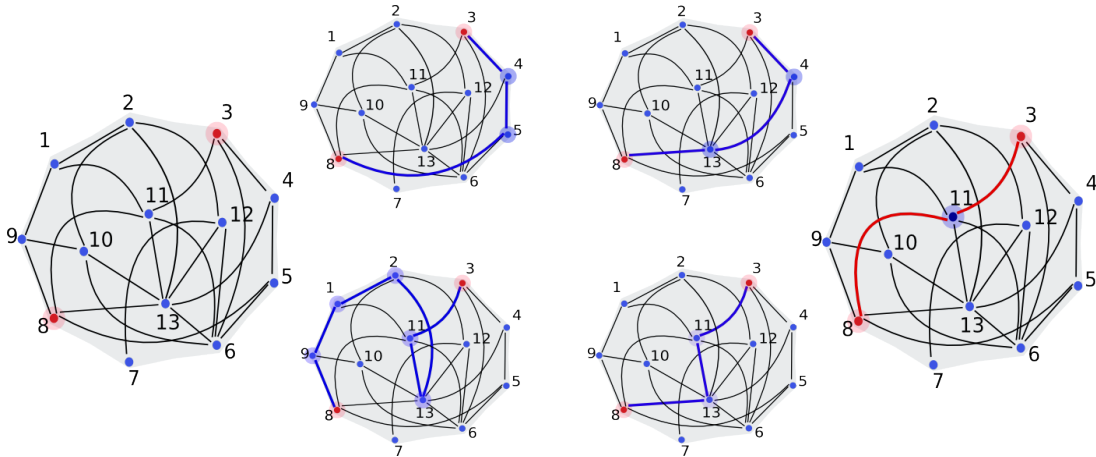


Figura 2.6 - Uma rede complexa formada por $N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$. Queremos o menor caminho entre o nó 8 e o nó 3. Vemos que há pelo menos 5 caminhos distintos sem revisitação com comprimento de caminhos $l_{83} = 3$, $l_{83} = 3$, $l_{83} = 6$, $l_{83} = 3$ e $l_{83} = 2$, ou seja, respectivamente temos o conjunto de nós ordenados $N = \{8, 5, 4, 3\}$, $N = \{8, 13, 4, 3\}$, $N = \{8, 9, 1, 2, 13, 11, 3\}$, $N = \{8, 13, 11, 3\}$, $N = \{8, 11, 3\}$. De todas as forma distintas de ir de 8 a 3, o caminho $N = \{8, 11, 3\}$, com $l_{83} = 2$, é o menor caminho pois passa apenas por um nó intermediário (nó 11), portanto com $d_{83} = 2$. (Fonte: [40])

A Figura (2.6) ilustra um exemplo do conceito de menor caminho. Nela é apresentada uma rede complexa formada por $N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$. Se queremos encontrar o menor caminho entre o nó 8 e o nó 3, devemos procurar pelo menor conjunto ordenado de nós que se inicia com o nó 8 e finaliza com o nó 3. Vemos que há pelo menos 5 caminhos distintos sem revisitação com comprimento de caminhos $l_{83} = 3$, $l_{83} = 3$, $l_{83} = 6$, $l_{83} = 3$ e $l_{83} = 2$, ou seja, respectivamente temos o conjunto de nós ordenados $N = \{8, 5, 4, 3\}$, $N = \{8, 13, 4, 3\}$, $N = \{8, 9, 1, 2, 13, 11, 3\}$, $N = \{8, 13, 11, 3\}$ e $N = \{8, 11, 3\}$. (Embora existam outros caminhos, esses são suficientes para entender o conceito.) Todavia, de todas as forma distintas de ir de 8 a 3, o caminho com $l_{83} = 2$, associado a $N = \{8, 11, 3\}$ é o menor caminho para esse par de nós, pois passa por apenas um nó intermediário (nó 11), portanto com $d_{83} = 2$.

O comprimento médio do menor caminho, \mathcal{L} , é um parâmetro global da rede, que mede a separação típica entre dois nós de uma rede [46]. Que também é conhecido como comprimento de caminho característico [38]. O comprimento médio do menor caminho está relacionado ao número médio de nós ao longo do caminho mais curto de todos os pares possíveis de nós da rede. Assim, considerando um grafo G com N nós, onde d_{ij} é o número de arestas entre o par de nós i e j , definimos $d_{ij} = 0$ quando $i = j$ e quando i não alcança

j , $d_{ij} = \infty$. Dessa forma, temos a expressão

$$\mathcal{L} = \frac{1}{n(n-1)} \sum_{i,j} d_{ij}, \quad (2.39)$$

para o comprimento médio do menor caminho [20]. O termo $n(n-1)/2$ é o total de arranjos possíveis 2 a 2 de um grafo com n nós. Ressaltamos que o somatório deverá ser dividido por 2, pois supomos um grafo não direcional, ou seja, \mathcal{L} de i a j é igual ao \mathcal{L} de j a i .

No entanto, da forma como é definido o comprimento médio do menor caminho, o resultado diverge. Para evitar essa divergência, restringe-se a análise apenas aos pares conectados, ou seja, a apenas ao componente gigante conectado [1].

Em redes reais não direcionadas é típico encontrar um grande componente, que ocupa mais a rede, em geral, mais da metade e não raramente acima de 90% dela, enquanto o restante da rede é dividido numa grande quantidade de pequenos componentes desconectados. Esse grande componente frequentemente é chamado de componente gigante. Nem sempre o maior componente é o componente gigante. Para isso, é necessário que seu tamanho cresça proporcionalmente ao número de nós da rede [3].

2.5.2 Coeficiente de intermediação

O coeficiente de intermediação (do inglês, *Betweenness*) – \mathcal{B}_i de um nó i é definido como o número de menores caminhos entre um par de nós, j e k que passam através do nó i [43]. Essa quantidade indica qual é o nó mais influente de uma rede, que não é necessariamente o mais popular (com maior grau).

Obviamente, j e k não são adjacentes, sendo que a conexão entre j e k passa através do nó i . Uma medida da relevância de um dado nó pode ser obtida contando-se o número de geodésicas que passam através do nó intermediário i , que é o nó de intermediação [38]. Aqui geodésica é definida como a menor distância entre os nós j e k de um grafo em unidades de comprimento de arestas [47]. É o comprimento do menor caminho d_{jk} , do nó j ao nó k . Quando um nó com alto índice de intermediação é removido, o resultado é o aumento da distância típica entre os outros nós [47]. O *betweenness* é calculado como:

$$\mathcal{B}_i = \sum_{j \neq i \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \quad (2.40)$$

onde σ_{jk} é o número de menores caminhos entre os nós j e k e $\sigma_{jk}(i)$ é o número desse caminhos que vão através do nó i [48]. A centralidade de intermediação de um dado nó

mede a quota de menores caminhos entre nós que este nó i intermedeia.

Além disso, existe o conceito de fluxo de intermediação que inclui tanto contribuições de menor caminho como os demais caminhos que não são menores caminhos do grafo [49]. No entanto, esse método acaba mostrando resultados contra-intuitivos em alguns casos. Newmann oferece, então, um método alternativo, intermediação da caminhada aleatória, no qual, para um dado nó i a intermediação da caminhada aleatória é igual ao número de vezes que a caminhada aleatória começa e termina em dois outros nós j e k passando pelo nó i ao longo do caminho [50]. Esse número de vezes pode ser considerado a força do link, de forma que podemos trabalhar com redes ponderadas.

2.6 Medida de Eficiência

2.6.1 Medida de Eficiência Topológica

Os primeiros estudos de eficiência de uma rede são apresentados no trabalho de redes de pequenos mundos (subseção 2.2.1) a partir do cálculo de propriedades local e global da estrutura da rede, ou seja, a partir do coeficiente de aglomeração e do comprimento médio de menor caminho.

Alguns trabalhos desenvolvem estudos de eficiência de rede baseados na análise do coeficiente de aglomeração \mathcal{C} e do comprimento médio do menor caminho \mathcal{L} [1, 20, 38, 51, 52].

Uma das técnicas interessantes usada para redes ponderadas (possivelmente não escassas e não conectadas) de medir a eficiência da rede, em vez de obterem medidas da conectividade entre os nós da rede analisa a eficácia a partir do estudo do fluxo entre os nós da rede. Tal que a eficiência global e local da rede podem ser analisadas a partir de grandezas derivadas da medida do menor caminho da rede. Nesse caso, a eficiência na comunicação entre os nós i e j é inversamente proporcional ao comprimento do menor caminho, $\epsilon_{ij} = 1/d_{ij} \forall i, j$. Se não existe nenhum caminho entre i e j , a eficiência é zero, $\epsilon_{ij} = 0$, pois $d_{ij} = \infty$. Portanto, é possível calcular a eficiência média do grafo G usando a seguinte expressão [20]:

$$\mathcal{E}(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \epsilon_{ij} = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \quad (2.41)$$

Para normalizar $\mathcal{E}(G)$ é necessário considerar o caso ideal G_{id} no qual o grafo G tem todas as $N(N-1)/2$ possíveis arestas. Situação em que a informação se propagaria pelo caminho mais eficiente desde que $d_{ij} = l_{ij} \forall i, j$, e $\mathcal{E}(G)$ assumiria o valor máximo com

$$\mathcal{E}(G_{id}) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{l_{ij}}.$$

Nesse formalismo, a definição de redes de pequenos mundos é dada pelo valor de E tal que a equação da eficiência global é dada pela equação 2.41, $\mathcal{E} = \mathcal{E}_{glob}$. E se \mathcal{E} for definido para um grafo desconectado, é possível calcular as características locais de G a partir da eficiência \mathcal{E}_i para cada nó i , tal que,

$$\mathcal{E}_{loc} = \frac{1}{n} \sum_{i \in G} \mathcal{E}_i. \quad (2.42)$$

A ideia é parecida é calcular a eficiência para cada subgrafo formado pelos vizinhos do nó i tal como é feito para o coeficiente de aglomeração \mathcal{C} , ou seja, excluindo o próprio nó i do subgrafo G_i . A medida de \mathcal{E}_{loc} revela o quanto o sistema é tolerante a falhas e, portanto, mostra o quanto é eficiente a comunicação entre os integrantes da vizinhança do nó i , quando este é removido [20].

2.6.2 Medida de Robustez Topológica

Um sistema pode sofrer falhas e mesmo assim ainda continuar eficiente? Mesmo que não pensemos em algo tão trágico como um ataque terrorista [21], será possível alterar a estrutura da rede de forma intencional e previsível e ainda manter o índice de eficiência da rede? Latora [21] desenvolve uma técnica de detecção de nós cruciais para o funcionamento eficaz da rede. Usando como exemplo as redes de comunicação e de terrorismo, tal que seja possível proteger a rede de comunicação a ataques e ainda encontrar o terrorista chave para romper a rede organizada de terrorismo. Ou seja, a partir de algum elemento topográfico que provoque a desconectividade da rede, comparar a eficiência dela após a remoção de nós e/ou arestas.

A remoção de nós ou arestas pode implicar em mal funcionamento de uma rede uma vez que pode vir a gerar a desconexão dos nós da rede, ou ainda deixar os caminhos mais longos. Se observarmos a definição de eficiência tal como dado pela equação 2.41, para uma informação se propagar entre dois nós quaisquer i e j pode ser dificultada na ausência de um nó ou aresta que os conecte.

Outros estudos verificam que há uma forte correlação entre topologia e robustez da rede [8, 18]. A tolerância a falhas ou a ataques é investigada em redes de energia [53], redes de internet [18, 53], redes de transporte [54, 55, 19, 56] e outras citadas acima.

Outros, ainda, estudam uma maneira de deixar uma rede mais robusta ou

mais frágil propondo algumas alterações na estrutura da rede [57], que mitiguem os efeitos de ataques maliciosos a redes reais. Schneider e outros [53] chamam a atenção para a melhoria na robustez de uma rede a partir de intervenções mínimas, mas que resulte em maior resistência a ataques. Eles atentam para o fato que reestruturação de redes reais está limitada ao valor dos custos dessa melhoria e ainda a operação do sistema. E portanto, sugerem uma reestruturação a partir de reconexões das arestas, mantendo o número de nós e o número de arestas constantes. Comparando-se as duas redes e verificar qual o ganho na robustez que a reconexão terá produzido. Isso tudo feito levando-se em conta apenas uma grandeza, \mathcal{R} , tal que:

$$\mathcal{R} = \frac{1}{n} \sum_{Q=1}^n s(Q), \quad (2.43)$$

onde N é o número de nós na rede e $s(Q)$ é a fração de nós na aglomeração mais conectada depois da remoção de $Q = qn$ nós. O fator de normalização $1/n$ assegura que a robustez da rede com diferentes tamanhos possam ser comparados. O intervalo de possíveis valores de R é $1/n$ e 0.5 , que correspondem a rede estrela e uma rede totalmente conectada.

A remoção de nós ou arestas para simular um ataque pode ser feita de várias maneiras: elegendo os nós aleatoriamente ou seguindo uma tipo de remoção estratégica, por exemplo, a partir de uma ordem decrescente dos nós com maior grau. Ou seja, submete-se a rede a teste de *stress* e observa-se o seu comportamento. Existem duas formas de fazer isso: 1) por meio de erros, cujas falhas são submetidas aos componentes escolhidos aleatoriamente. 2) por meio de ataques, cujas falhas são submetidas aos componentes que desempenham um papel importante no sistema [34]. Assim, redes sem escala são mais robustas que as redes aleatórias quando se trata de ataques aleatórios, contudo, são mais vulneráveis a ataques estratégicos [53].

2.6.3 Medida de Dissimilaridade

Chamada de dissimilaridade, essa métrica é capaz de identificar e quantificar as diferenças topológicas entre grafos. A ideia principal para medir a dissimilaridade de dois grafos é associar a cada estrutura um conjunto de funções de distribuição de probabilidade, representando todas as distâncias de conectividade de cada nó, e compará-las, por métricas padronizadas de informação. Esse método pode ser empregado tanto em redes contendo arestas direcionadas ou não direcionadas. A medida da dissimilaridade pode ser usada para caracterizar a evolução dos sistemas dinâmicos, podendo identificar a região do pequeno mundo no processo Watts-Strogatz (WS) e transições de fase em evolução da rede Erdős-Rényi (ER). E ainda, em redes reais, pode avaliar quão bom é o ajuste de

modelos de rede e pode prevê suas probabilidades críticas de percolação [58].

Calcula-se a medida de dissimilaridade entre redes (D), que captura as diferenças globais e locais de grafos, como [58] :

$$D(G, G') = w_1 \sqrt{\frac{\mathcal{J}(\mu_G, \mu_{G'})}{\log 2}} + w_2 \left| \sqrt{NND(G)} - \sqrt{NND(G')} \right| + \frac{w_3}{2} \left(\sqrt{\frac{\mathcal{J}(P_{\alpha,G}, P_{\alpha,G})}{\log 2}} + \sqrt{\frac{\mathcal{J}(P_{\alpha,G^c}, P_{\alpha,G^c})}{\log 2}} \right), \quad (2.44)$$

onde N e M é do tamanho de G e G' , respectivamente, e G^c indica o complemento de G . Na equação a NND é definida como a dispersão de nó da rede - NND (do inglês, network node dispersion):

$$NND(G) = \frac{\mathcal{J}(\mathbf{P}_1, \dots, \mathbf{P}_N)}{\log(d+1)}, \quad (2.45)$$

onde $\mathcal{J}(\mathbf{P}_1, \dots, \mathbf{P}_N) = \sum_{i,j} p_i(j) \log(p_i(j)/\mu_j)$ e $\mu = (\sum_{i=j} p_i(j))/N$ sendo a divergência Jensen-Shannon e a média da distribuição N , respectivamente.

Assim, o primeiro termo compara os padrões de conectividade dos nós, que corresponde à distância de distribuição de grafos. Tal que grafos que compartilham a mesma distância de distribuição apresenta o mesmo diâmetro, comprimento de caminho médio e outras propriedades de conectividade. Já o segundo termo analisa a heterogeneidade dos nós. Grafos apresentam o mesmo NND são grafos que tem o mesmo perfil de distância de conectividade. E o terceiro termo considera a centralidade de cada nó, contando-se a extensão de conectividade direta e indireta de cada nó.

E portanto, $D(G, G')$ identifica e quantifica diferentes propriedades topológicas estruturais, que afetam o fluxo de informação através das redes e pode ser usado para comparar diferentes estágios de redes reais que tenha sofrido algum tipo de alteração.

CAPÍTULO 3

Redes de Transporte

Como dito anteriormente, as redes de transporte entram no grupo de redes tecnológicas, juntamente com redes de distribuição, redes de internet, redes de telefonia, redes de energia e outras. Apesar de não ser um agrupamento rígido, percebe-se que há semelhanças nesses sistemas: em todas as redes citadas há transporte de algo de um lugar a outro, por exemplo.

Vemos também que a classificação de redes de transporte de forma mais ampla engloba redes de comunicação, neural, elétrica, carga e pessoas [39]. E ainda, são consideradas redes espaciais por terem sua topologia associada à distância.

Estudos de redes de transporte de pessoas e de cargas remetem aos anos de 1960 e 1970, sobretudo realizados por geógrafos. A pesquisa focava nos aspectos econômicos e de estruturas físicas de estradas e trilhos [3]. Todavia, estudos mais recentes em redes aérea [27, 31, 59, 60, 61], terrestres [62, 63, 64, 33, 65, 66, 67, 68, 48] ou marítima [69] usam uma nova análise de redes onde os nós representam localidades geográficas e as arestas representam as rotas entre os locais [3].

No estudo de redes aérea, onde os nós representam os aeroportos e as arestas, a conexão sem paradas, Amaral e outros [27] mostram que os resultados empíricos da distribuição de cargas e de passageiros dos maiores aeroportos do mundo é uma rede de pequeno mundo, onde dois aeroportos são conectados diretamente, ou quando não conectados por voos diretos, são conectados por meio de voos em conexão, tal que a conexão é realizada em um aeroporto ou no máximo em até quatro aeroportos. Isso significa que a distância máxima entre dois nós da rede de aeroportos é de 5 arestas.

Redes de transporte terrestre são analisadas como exemplos de redes reais para se estudar a eficiência, as propriedades de redes de mundos pequenos, a distribuição do grau do nó, entre outras propriedades de topologia de rede. Foram realizados estudos da eficiência da rede metroviária de Boston, onde cada estação foi definida como nó e cada túnel, como aresta e introduzida a metodologia de medida de eficiência global e local a partir do valor da medida de menor caminho [20, 62]. Posteriormente, foram realizados estudos das propriedades de redes de pequenos mundos na rede de trem da Índia, onde cada estação era definida um nó e a existência de uma parada por pelo menos um trem em duas estações definida como aresta entre elas [63].

Um outro trabalho que se destacou na investigação empírica de redes de

transporte foi realizado na Polônia [64], onde foram observadas que, independente do tamanho das redes de transporte público das 22 cidade polonesas, a distribuição do grau seguia uma lei de potência ou uma função exponencial. Outras propriedades da topologia de rede também foram investigadas, tais como coeficiente de aglomeração, intermediação e outros. Nesse caso, os autores definem dois tipos de redes, a rede espaço- L e a rede espaço- P , seguindo de forma geral a mesma construção feita em [63]. Eles encontram distribuição do grau em lei de potência para as redes espaço- L e exponencial para as redes espaço- P . Embora a definição de nó como estações seja clara nos trabalhos precedentes, a definição das arestas não era consensual. Sendo que a definição da aresta da rede como sendo a existência de pelo menos uma rota de ônibus, trem ou outro modal, descrito pela primeira vez naquele trabalho.

Um trabalho comparando vários PTNs de várias cidades do mundo mostrou que a distribuição do grau poderia ser em Lei de Potência ou em Exponencial, no entanto, o modelo usado não diferencia a topologia de vários tipos de transporte. De modo que para algumas cidades são usados dados de transporte de ônibus somente, para outras, os dados de transporte de metrô, ônibus e trem, por exemplo [48].

Estudos de redes de fluxo de tráfego e análises topológicas de sistemas de transporte também foram bastante estudados em [33], onde Kurant e Thiran desenvolvem a análise usando tabela de horários para gerar três diferentes espaços: espaço de transferências, espaço de paradas e espaço de estações. O espaço de transferências se assemelha com o espaço- P e o espaço de estações, com o espaço- L . O modelo define para o espaço de transferências que todo trem é um subgrafo e o número de subgrafos é definido como um peso da aresta. É gerado uma rede em multicamadas [28], onde a estrutura física e o fluxo são camadas distintas.

Nosso trabalho se restringe ao estudo empírico de redes de transporte urbano, que aceitamos como sendo um subgrupo de redes de transporte, inserido no contexto de redes espaciais.

3.1 Redes de Transporte Público

Redes Transporte público são sistemas em que as rotas e os horários de viagens são fixos, disponíveis para serem usados por qualquer pessoa que pagar a tarifa estabelecida. Os modelos mais comuns são realizados por trem, ônibus e metrô.

Uma das redes reais estudadas pela Teoria de Redes Complexas é a Rede de Transporte Público, ou *PTN*, do inglês, *Public Transport Network*. A Figura (3.1) mostra

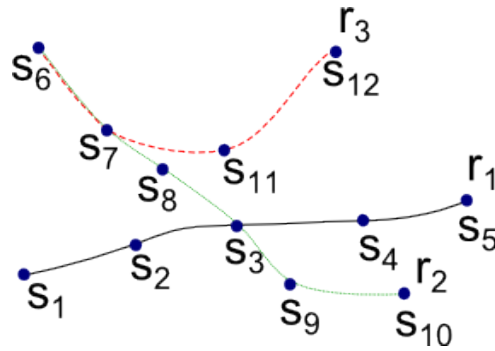


Figura 3.1 - Rede de Transporte Público formada por 3 rotas (r_1, r_2, r_3) e 12 estações/pontos de paradas ($s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}$), sendo o s_3 a estação de transferência entre as rotas r_1 e r_2 e as estações s_6 e s_7 , entre das rotas r_1 e r_3 .

um exemplo simples de um sistema de transporte constituído por 3 rotas (r_1, r_2, r_3) e 12 estações/pontos de paradas ($s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}$), sendo o s_3 a estação de transferência entre as rotas r_1 e r_2 e as estações s_6 e s_7 , entre das rotas r_1 e r_3 .

Em geral são realizadas análises empíricas, simulações e abordagens teóricas para descrever essas redes [62, 63, 64, 33, 65, 66, 67, 68, 48, 70, 71]. Em grande parte, o objetivo é encontrar alguma característica de eficiência e a robustez por meio de análises das propriedades de redes complexas, tais como: grau de centralidade, coeficiente de aglomeração, comprimento do menor caminho médio, coeficiente de intermediação e outras.

São abordados vários tipos de sistemas de transportes: de ônibus [65, 67, 66, 68, 64], trem [64], metrô [20, 62] etc. Para estudos de PTNs são geralmente construídos quatro tipos de redes: espaço- L , espaço- P , espaço- C e espaço- B [63, 64, 65, 66, 48]. A Figura (3.2) demonstra como essas redes podem ser construídas a partir do sistema de transporte da ilustrada na Figura (3.1).

- espaço- P : Nessa rede as estações representam os nós e as arestas são representadas pela existência de rotas que passam em duas estações [63, 64, 33, 65]. Nesse caso, um nó é vizinho de primeira ordem de outro se esse outro for atendido por pelo menos uma rota que passa pelo primeiro nó. Ou seja, tanto faz se a estação é a seguinte na via ou se é a última de uma rota (um terminal, por exemplo), ambos são considerados primeiros vizinhos do nó analisado. Assim, todas as estações que são atendidas por uma determinada rota serão nós vizinhos de primeira ordem um do outro.
- espaço- L : Nesse tipo de rede, as estações são nós e as rotas são as arestas.

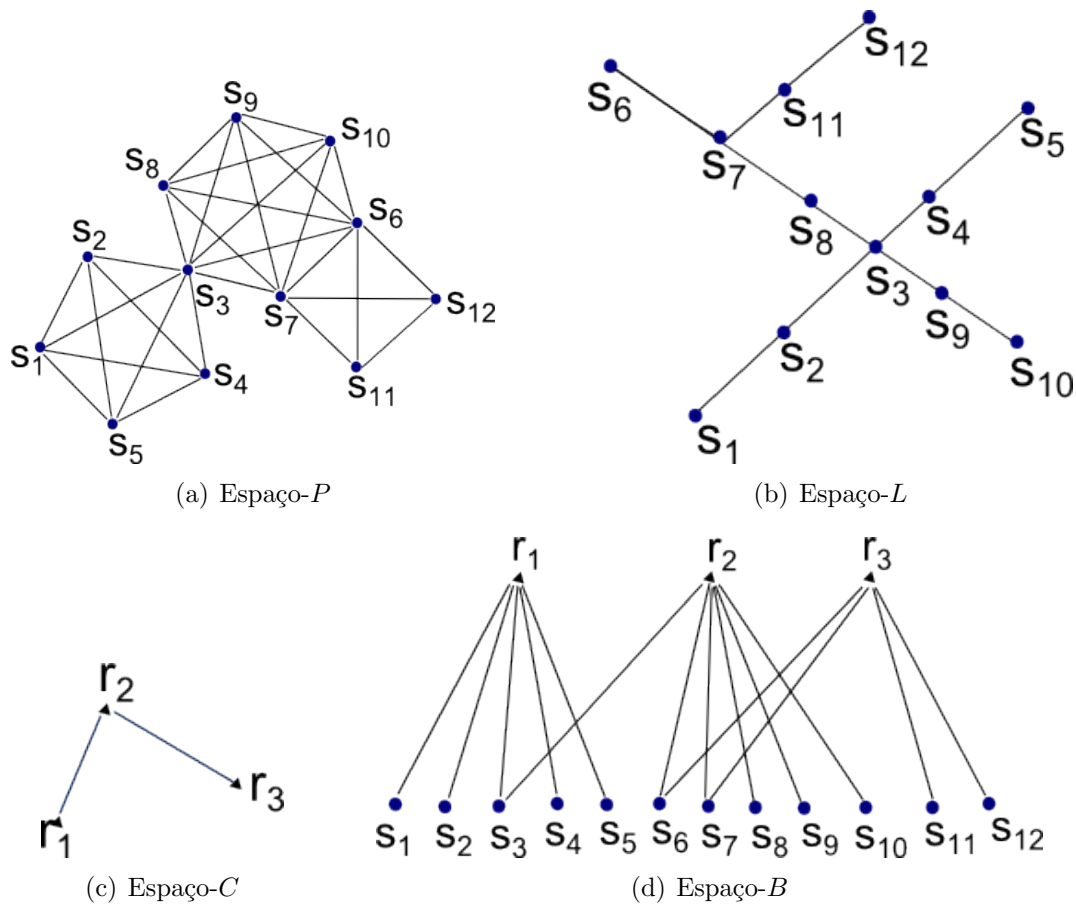


Figura 3.2 - Ilustração dos tipos de redes PTN – *Public Transport Network* que são geradas a partir de um sistema de transporte: espaço- P , espaço- L , espaço- C e espaço- B .

Contudo, um nó j só será vizinho do nó i se o j for sucessivo ao nó i na rota [64, 33, 65].

- espaço- C : Para essa rede as definições são opostas das definidas para o espaço- P . Nesse caso, as rotas representam os nós. Se duas rotas passam por pelo menos uma estação em comum, então existe aresta entre esse par de rotas [65, 68].
- espaço- B : essa é uma rede bipartida que foi concebida na análise de redes do sistema de transportes da China [65, 68]. Tanto rotas quanto estações são nós da rede. Um nó tipo rota só terá aresta com um nó tipo estação, se a rota atende a essa estação.

Tal como qualquer outro tipo de redes, é usada a matriz adjacente para análise de cada uma dessas redes.

3.2 Grau do nó em PTN

O grau do nó leva a definição padrão, no entanto, muitas vezes é chamado de grau de conectividade do nó, iremos usar o termo conectividade a partir daqui por ter mais significado prático em redes de transportes. De modo que, valor de k_i é o número de nós conectados ao nó i .

Para cada tipo de rede, o grau de conectividade do nó possui uma interpretação particular em redes de transportes. No caso do espaço- P , o grau de conectividade do nó é a quantidade de estações que um passageiro pode acessar a partir da estação i usando uma única rota sem necessidade de transbordo. Essa rede é importante para a análise das estações centrais de um PTN , por meio da análise da distribuição do grau de conectividade do nó podemos verificar se é uma rede sem escala ou de pequeno mundo e, a partir disso, verificar a robustez do sistema. Na prática, o grau de conectividade podem auxiliar na escolha da melhor estação para um sistema de transporte integrado. Podemos escolher estações com maior grau para servir de estação central de uma cidade.

No caso da rede espaço- C , é a quantidade de rotas com as quais uma rota pode conectar, como definido anteriormente, duas rotas se conectam, se e somente se, houver pelo menos uma estação/ponto de parada comum às duas rotas. A ideia é descobrir se existem rotas isoladas, se o $k_i = 0$ significa que não há nenhuma estação pertencente a i que é atendida por outra rota. Assim, os passageiros dessa rota não poderão acessar outra estação senão aquelas que pertencem à rota. Além disso, a medida de grau de conectividade do nó de uma rede espaço- C informa se possível acessar uma determinada região usando duas rotas, r_1 e depois r_2 . O grau do nó, portanto, é o número de rotas com os quais a rota analisada pode cruzar. Procedendo-se com a análise do grau de conectividade do nó dessa rede, podemos observar se uma rota é muito parecida com outra, se o nó dessa rede espaço- C tem um grau grande, sugere que essa rota conecta bastante com as outras rotas existente. Assim, um passageiro que tenha que se decidir se utiliza uma rota r_i ou uma rota r_j para tomar outra rota r_h , ele pode se decidir em tomar aquela rota com o maior grau, visto que a probabilidade desse passageiro conseguir tomar outra rota é maior para o caso de rotas com maior grau.

Em redes do tipo espaço- L , k_i representa o número de estações sequencialmente anteriores e posteriores à estação i . Em geral, uma estação tem grau $k_i = 2$, pois i é conectado com a estação imediatamente anterior e imediatamente posterior. Contudo, se tratando de uma estação central que possui acesso a várias rotas, o grau de conectividade será $k_i > 2$. Essa quantidade, pode auxiliar na decisão de se escolher as estações *hub* (que são as estações conectadas com muitas outras [31]), juntamente com o grau de

conectividade do espaço- C . A partir dessa medida é possível se verificar qual estação está próximo a uma bifurcação de rotas. Muitas vezes duas ou três rotas atendem um conjunto de estações em comum por algum pedaço de trecho e se bifurcam em algum ponto. Se o grau de conectividade for $k_i > 2$, certamente i será uma bifurcação, a última estação para transbordo entre duas rotas.

O cálculo do grau é realizado tal como exposto subseção 2.4.1. Incluindo os conceitos grau para redes direcionadas e ponderadas.

3.3 Comprimento do menor caminho em PTN

O significado do comprimento médio do menor caminho em redes reais de PTN é diferente para cada tipo de rede. Para a rede espaço- P , o \mathcal{L} conta o número médio de transferências que tem de ser feitas numa viagem entre duas estações quaisquer, que chamamos de estações de transbordo. Na rede espaço- P todas as estações que pertencem a uma rota são vizinhos de primeira ordem um do outro. Significa dizer que não há necessidade de transbordo para se acessar qualquer nó j pertencente a rota a partir do nó i também pertencente a mesma rota. Assim, a distância d_{ij} para todos os nós da rota é $d_{ij} = 1$, ou seja, um transbordo. No entanto, quando dois nós i e j quaisquer do espaço- P não se conectam por meio de uma única rota, devemos descobrir se por meio de um nó intermediário t esta conexão é possível. Ou seja, se existem as arestas (i, t) e (t, j) , então i e j são vizinhos de segunda ordem um do outro e teremos a necessidade de um transbordo.

Usando a rede espaço- L , o comprimento médio do menor caminho \mathcal{L} revela o número de estações pelas quais é necessário passar antes de chegar ao nó de destino. Ou seja, representa o tamanho das rotas em número de estações que uma PTN possui.

Já na rede espaço- C , se o comprimento $d_{ij} = 1$ significa que duas rotas são conectadas por ao menos uma estação, se $d_{ij} = 2$ significa que é necessário se utilizar uma outra rota t que conecte a rota i e j .

3.3.1 Topologia de PTN integrada

Usamos as características de redes de pequenos mundos para encontrarmos respostas sobre a tarifação única do sistema em ocasião de um sistema integrado física e tecnologicamente. A questão era sobre a quantidade máxima de viagens que seria permitida para uma única tarifa, caso não houvesse uma rota direta entre dois pontos de paradas. Criamos um algoritmo para o cálculo da matriz de transbordo usando as ideias de comprimento do menor caminho das redes de pequenos mundo se calculamos o número máximo de utilização de rotas, ou o número máximo de transbordos que um passageiro

necessita realizar para acessar qualquer estação do Distrito Federal.

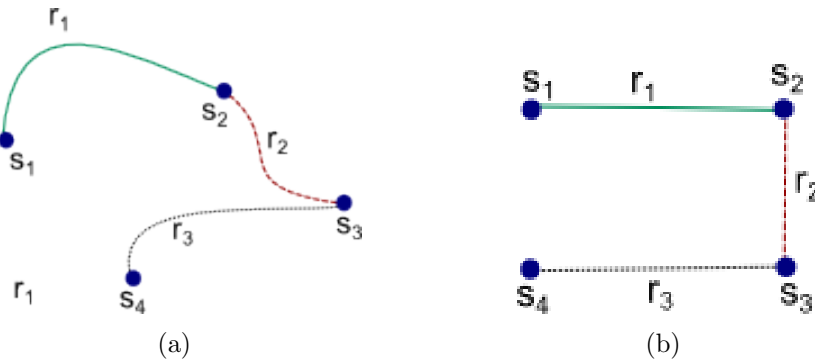


Figura 3.3 - 3.3(a): Rede de ônibus formada por 3 rotas (r_1, r_2, r_3) e 4 estações (s_1, s_2, s_3, s_4); 3.3(b): Rede espaço- P ; Esquema ilustrativo de uma rede de ônibus espaço- P para o cálculo do menor caminho.

Usamos a matriz adjacente do espaço- P para encontrar a matriz de transbordo \mathcal{T} do PTN. A rede espaço- P é uma rede simples, ou seja, não há autoarestas e nem multiarestas. Além disso, representamos como uma rede não direcional, de modo que a matriz adjacente seja simétrica. Assim, a matriz adjacente \mathbf{A} é formada por elementos $\{a_{ij}\}$ com valor $a_{ij} = 1$ quando existir uma rota entre os nós i e j e $a_{ij} = 0$ para os elementos da diagonal e quando não existir rota entre os nós. Portanto, para o exemplo da Figura (3.3), temos:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Queremos saber se existe um caminho entre i e j . Assumimos a partir de agora que existirão autoarestas, assim, somamos a matriz identidade \mathbf{I} , que é $\mathbf{A}^T \mathbf{A}$.

$$\mathbf{A}^1 = \mathbf{A} + \mathbf{I} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} 0 & 1 & ? & ? \\ 1 & 0 & 1 & ? \\ ? & 1 & 0 & 1 \\ ? & ? & 1 & 0 \end{pmatrix}$$

Se $a_{ij}^1 = 1$ significa que há um caminho entre i e j e preenchemos a matriz de transbordo \mathcal{T} com valor $t_{ij} = 1$. Se $a_{ij}^1 = 0$, teremos que encontrar um nó intermediários u em que exista um caminho, tal que $a_{iu} = 1$ e $a_{uj} = 1$, isto é, que a_{iu} seja igual a um e a_{uj} também seja igual a um, e que a estação u esteja presente em alguma rota na qual pertence a estação i ao passo que u esteja presente também em alguma rota na qual pertence a estação j .

$$\mathbf{A}^2 = \mathbf{A}^1 \times \mathbf{A}^1 = \begin{pmatrix} 2 & 2 & 1 & 0 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 2 \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} 0 & 1 & 2 & ? \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ ? & 2 & 1 & 0 \end{pmatrix}$$

Na matriz $\mathbf{A}^2 = \mathbf{A}^1 \times \mathbf{A}^1$ observamos apenas os elementos da posição $a_{ij}^2 = 1$ para i e j em que $a_{ij}^1 = 0$. E então preenchemos a matriz \mathcal{T} com o elemento $t_{ij} = 2$. Se ainda restarem $a_{ij}^2 = 0$ teremos que proceder novamente com a busca de um novo nó intermediário tal que a_{iu} , a_{uv} e a_{vj} existam.

$$\mathbf{A}^3 = \mathbf{A}^2 \times \mathbf{A}^1 = \begin{pmatrix} 4 & 5 & 3 & 1 \\ 5 & 7 & 6 & 3 \\ 3 & 6 & 7 & 5 \\ 1 & 3 & 5 & 4 \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix}$$

E feito isso, preenchemos a matriz \mathbf{L} com o valor 3 nos elementos da posição i e j em que $a_{ij}^3 = 1$. E seguimos a implementação dessa rotina até que não sobrem mais nenhum elemento da matriz \mathbf{A}^n com valor igual a zero. Percebemos que a multiplicação inicial pela matriz identidade é fundamental para o algoritmo pois evita o aparecimento de novos valores de zero na matriz \mathbf{A}^n , devido ao valores iguais a zero da diagonal principal da matriz de adjacências.

Uma vez escrita a matriz \mathbf{L} , subtraímos $\mathbf{1}$ para obter a matriz de transbordos de um PTN:

$$\mathcal{T} = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{pmatrix}$$

tal que t_{ij} é o número de transbordos entre as estações i e j , ou seja é o número de transferências entre rotas. A matriz de menor caminho, significa o número de rotas necessárias para se ir de i a j .

3.4 Coeficiente de aglomeração em PTN

Em uma rede de transporte público é frequente a necessidade de se alterar as rotas da rede para atender a novas demandas seja devido ao crescimento da cidade, seja devido à construção de um novo shopping center. Da mesma forma, com o passar do tempo, a demanda de uma rota já existente pode diminuir, seja por causa da desativação de uma fábrica, ou ainda, por construção de uma nova via de acesso mais curta. Essa contínua necessidade de alterar o PTN passa por um processo de desativação de rotas e/ou a criação de nova rotas. Além de se buscar a contínua melhora no sistema para se diminuir os custos da operação do sistema, isso quer dizer que, duas rotas podem ser acopladas em uma única rota se isso for economicamente mais viável ou ainda uma rota pode ser separada em outras duas.

Nesse contexto, como o estudo de propriedades de centralidade em redes complexas poderia auxiliar na tomada de decisão quanto a ativação ou desativação de uma rota? O coeficiente de aglomeração da rede bipartida espaço- B pode auxiliar a avaliar se provocará ou não prejuízo ao PTN a desativação de uma rota específica. Verificando-se o coeficiente de aglomeração dos nós tipo rota, se o coeficiente for zero, significa que a rota analisada é a única que atende ao conjunto de estações da rota. O resultado revela que não há nenhuma outra rota que os conecte senão a própria rota analisada.

Da mesma forma, se o coeficiente de aglomeração for igual a 1 (um) ou próximo de 1 (um), certamente aquela rota analisada não fará falta para conectar as estações que atende a rota.

3.5 Coeficiente de intermediação em PTN

Para redes PTN, o coeficiente de intermediação (*betweenness*) \mathcal{B}_i mensura a importância de um nó com respeito à conectividade entre os outros nós da rede [48]. Para a rede espaço- P , o coeficiente de intermediação indicará quão importante é o nó (estação) analisado para se manter um sistema de transporte integrado. Quanto maior o coeficiente, maior é sua importância para a integração. Se o coeficiente de intermediação desse nó for alto significa que é um nó muito servido pelo transporte quando há necessidade de transbordo. Pois nesse caso, a conexão direta entre os nós j e k não são contados, mas apenas aquelas que necessitam de uma integração, ou seja, os nós k e j só se conectam se

necessariamente passar pelo nó i . A medida do coeficiente de intermediação de i verificará a sua importância para a conexão de todos os nós da rede. É possível, por essa medida decidir qual nó (estação) poderia, portanto, ser melhor para a construção de um terminal de integração, pois será uma estação com muito embarque e desembarque.

3.6 Robustez em PTN

Em redes PTN é necessário prever falhas na operação do sistema e tentar mitigar os possíveis danos. Uma via congestionada provocada por algum acidente de trânsito, um bloqueio de via por causa de alguma manifestação ou ainda por motivos de forças da natureza, como uma enchente, são fatores que podem impedir o fluxo do transporte de passageiros. No modelo de redes complexas, podemos estudar esses casos como sendo uma retirada de vértices ou arestas. Na rede espaço- P a retirada de um nó significa que aquela estação não está mais acessível. Teremos então uma nova rede com $N - 1$ nós. A análise da robustez indica quanto uma rede pode sofrer falhas e ainda assim manter um mínimo de eficiência.

Da mesma forma, quando uma linha não estiver mais em operação, por motivos diversos, seja por não possuir demanda que justifique os custos da operação, seja por motivos de quebra de veículo, é possível analisar qual o impacto que esses tipos de eventualidades pode prejudicar toda a rede de transporte. Nesse caso, podemos analisar a robustez da rede espaço- C , cujas as rotas do sistema PTN são representadas por vértices.

Assim, a análise da robustez das redes espaço- P e espaço- C podem informar o nível de eficiência do fluxo da rede, mesmo que ela sofra algum tipo de falha. O quanto essas falhas podem provocar o isolamento de uma ou mais regiões faz parte do que pode ser investigado por esse método.

No entanto, características geográficas podem provocar alterações na robustez da rede, tal que a investigação da rede espaço- L torna-se necessária.

3.7 Estudo de caso: Transporte Urbano do Distrito Federal

Em 2010 foi realizado um estudo da rede de ônibus urbano do Distrito Federal, Figura (3.4). Mapeamos 3.438 estações, Figura (3.4(a)), dos quais usamos apenas 3.272 estações do componente gigante, e 856 rotas, Figura (3.4(b)). Construímos as três redes PTN: espaço- P , espaço- C e espaço- B . Na época a falta de dados do sistema de transporte impossibilitou a construção da rede espaço- L , cujos nós são estações sequenciais das rotas [40]. Foram verificadas as propriedades de centralidade, tais como o grau de conectividade do nó e analisadas as distribuições da conectividade das redes.

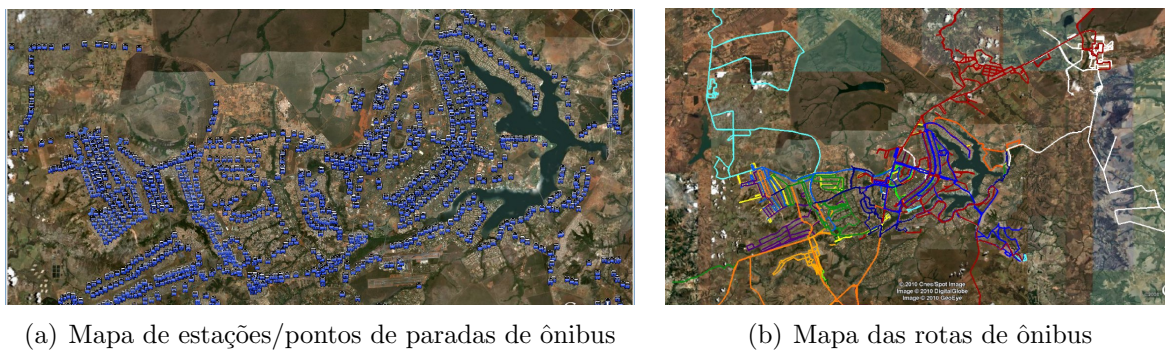


Figura 3.4 - Dados mapeados no Google Earth para geração das redes espaço- P , espaço- C e espaço- B . (Fonte: [40])

Em 2009 discutia-se a implantação do sistema de tarifa única para o sistema de transporte urbano do DF. O Plano Diretor de Transporte Urbano - PDTU [72] previa a necessidade de implantação de um sistema integrado, física e tecnologicamente. A integração física possibilitaria que um passageiro pudesse utilizar duas ou mais rotas como se fosse uma única viagem. Para isso, seria necessária também uma integração tecnológica, cujo intuito seria uma integração tarifária, ou seja, a exemplo de grandes cidades pelo mundo, com um único bilhete/cartão, o passageiro poderia realizar transbordos/transferências afim de chegar ao seu destino.

A dúvida consistia em determinar qual seria a quantidade máxima de transbordos para o DF, ou seja, quantas rotas diferentes seriam necessárias para um passageiro conseguir sair de qualquer lugar do DF e chegar a qualquer outro dentro do DF, de modo que todo o sistema de ônibus pudesse estar conectado. Uma questão muito simples de se responder com o modelo de redes complexas a partir da análise da topologia de uma rede PTN, usando o método derivado do cálculo do menor caminho da rede espaço- P , descrito na Subseção 3.3.1.

Assim, usando a matriz de Transbordo, \mathcal{T} , aplicada ao sistema inteiro, obtivemos o valor de máximo de 3 transbordos para todo o DF, E, portanto, para acessar qualquer estação a partir de outro seriam necessários no máximo o uso de 4 rotas diferentes para o sistema PTN de ônibus do DF [40].

Todavia, é importante salientar que a efetiva possibilidade de transferência entre rotas só é possível levando-se em conta os horários das viagens de cada rota. Se um passageiro necessita tomar duas rotas para chegar a seu destino, saindo de i para j , realizando um transbordo em t , necessariamente. O horário da viagem da segunda rota no ponto de paradas t não deve ser anterior ao horário que a primeira rota, vindo de i a t

chegue a t . Caso contrário, o passageiro não poderia realizar o transbordo. Esse estudo de transbordo não possui informações sobre viagens, tão somente de rotas e estações/pontos de paradas.

Além disso, realizamos um estudo de redes de pequenos mundos, com o propósito de averiguar se a rede de transporte urbano do DF poderia se enquadrar nessas classes de redes. Para isso usamos a rede espaço- P , calculamos o comprimento médio do menor caminho e o coeficiente de aglomeração, propriedades de redes de mundo pequenos. Obtivemos $\mathcal{L}/\mathcal{L}_0 \cong 0.23$ e $\mathcal{C}/\mathcal{C}_0 \cong 0.96$, sugerindo que a rede espaço- P do PTN de ônibus do DF seja uma rede de pequenos mundos[40].

Em resumo, vimos que comprimento de menor caminho da rede espaço- P significa na prática o número de transferências ou transbordos que um passageiro deve fazer para ir de um ponto de paradas a outro. Isso porque nessa representação todos os pontos de paradas de uma rota são nós vizinhos de primeira ordem um do outro. E vimos que a matriz de adjacências para esse caso é binário, 1 para nós conectados e 0 para nós não conectados. Nesse caso, nós vizinhos de segunda ordem representam pontos de paradas que não estão conectados diretamente, mas necessitam de um outro nó intermediário para a conexão, que na prática significa que será necessária a utilização de duas rotas que tenham pelo menos um ponto de paradas em comum.

3.7.1 Grau do nó da rede espaço- B

Após a obtenção do grau do nó de cada nó dessas redes, observamos a a frequência com o qual cada valor de $k_i = k$ se repete, ou seja, contamos quantos nós têm o mesmo grau k . Assim, obtivemos a seguinte distribuição do grau para o Sistema de Transporte Urbano do DF, Figura (3.5(a)).

Foi obtida uma distribuição do grau do nó do espaço- B , Figura (3.5), que parece seguir um decaimento de Lei de potência. Que é um comportamento observado em muitas redes reais. Tendo, portanto, indícios de que seja uma rede sem escala. Assim, obtivemos uma curva de ajuste que obedece a equação:

$$P(k_B) = a(k_B)^{-\gamma_B}, \quad (3.1)$$

onde $a = 4.921, 94 \pm 4.674$ e $\gamma_B = 1, 14 \pm 0, 08$. O erro de γ_B é de 7,438% nesse ajuste. Todavia, o erro de a é bastante significativo, chegando a ser de 94,96%. Dessa forma, maior esclarecimento, fizemos o gráfico da distribuição em escala log-log e, portanto, sugere se tratar de uma curva lei de potência.

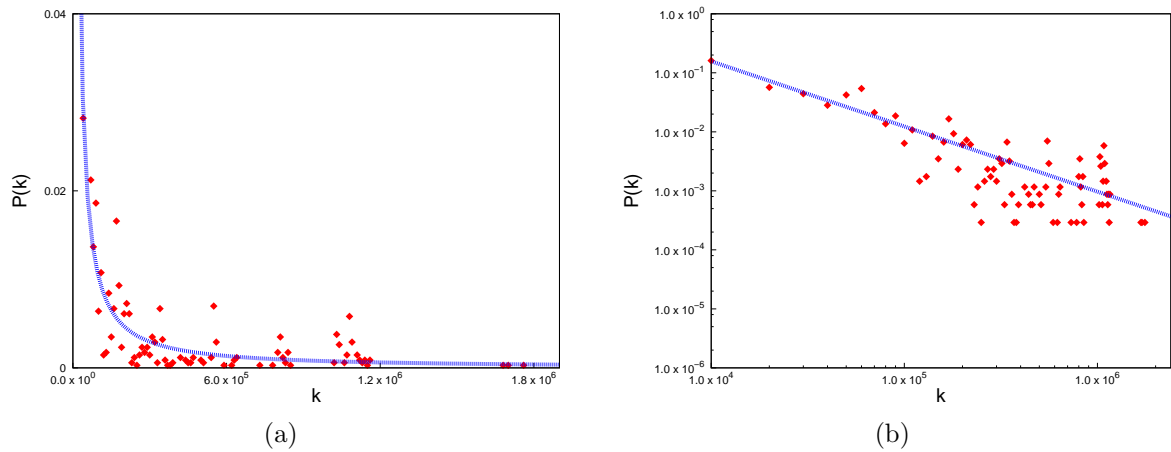


Figura 3.5 - Distribuição do grau do nó da rede espaço- B .; 3.5(a): Distribuição do Grau do nó-ponto no espaço- B . Cada ponto representa uma média ponderada da quantidade de k_i contidos num intervalo $\Delta k = 10.000$ valores de k_i ; 3.5(b): Gráfico dilog da distribuição do grau no espaço- B por intervalo $\Delta k = 10.000$. (Fonte: [40])

3.7.2 PTN com peso

Vimos que as redes PTN são geralmente representadas como redes simples. Isso significa que não existem auto-arestas nem multiarestas. Contudo, redes PTN possuem muitas rotas que compartilham as mesmas estações. A Figura (3.6) apresenta uma rede de transporte urbano formada por duas rotas, r_1 e r_3 . A rota r_1 atende as estações (s_1, s_2, s_3, s_4 e s_5) e a rota r_2 atende as estações (s_1, s_2, s_3, s_{11} e s_{12}). Para a construção da rede espaço- P , que têm definidas as estações como nós e as rotas como arestas.

Vemos que existem duas arestas entre os nós s_1 e s_2 e entre s_2 e s_3 , isto é, ambos os pares de nós possuem multiarestas, que são associadas as rotas r_1 e r_3 . Por outro lado, os outros pares de nós possuem apenas uma rota que os atende, as estações s_4 e s_5 são atendidas apenas pela rota r_1 e as estações s_{11} e s_{12} são atendidas apenas pela rota r_3 , Figura (3.6(a)). Podemos representar as multiarestas como um peso das arestas. De forma que podemos dizer que seria mais provável ir da estação s_1 para s_2 ou s_3 , do que ir da estação s_4 para s_5 ou da estação s_{11} para estação s_{12} , pois existem mais rotas servindo s_1, s_2 e s_3 que servindo o restante dos nós, como mostrado na Figura (3.6(b)).

Então, a partir do peso das arestas, podemos calcular o grau ponderado ou força do nó de cada nó da rede espaço- P ponderada, onde o número de rotas distintas que servem os pontos de paradas atribuem um peso para a conexão, é calculado pela equação (2.29). Consideramos o peso da conexão como o número de multiarestas.

Fizemos o estudo da distribuição do grau do nó ponderado ou força do nó

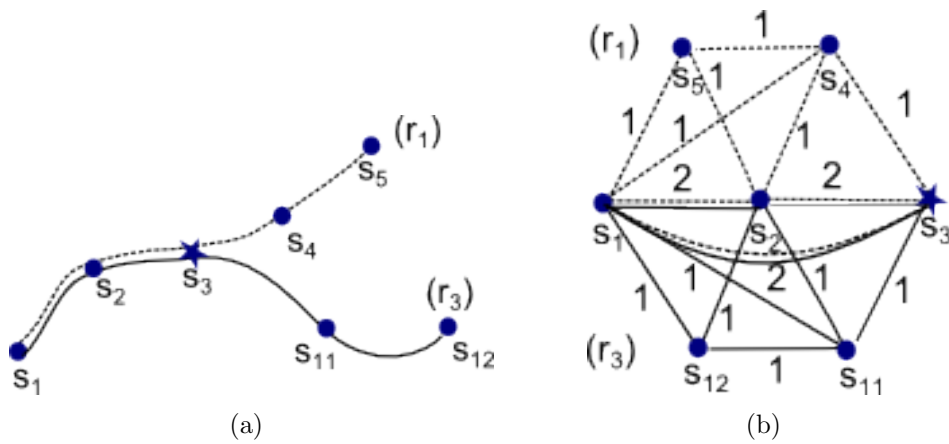


Figura 3.6 - Rede de transporte de ônibus formada por duas rotas (r_1, r_3); 3.6(a): Rede de transporte de ônibus com duas rotas; 3.6(b): Rede espaço- P ponderada.

das redes ponderadas do espaço- P e espaço- C .

A análise da distribuição de s da rede espaço- P , entretanto, apresentou uma curva de distribuição ilustrada na Figura (3.7).

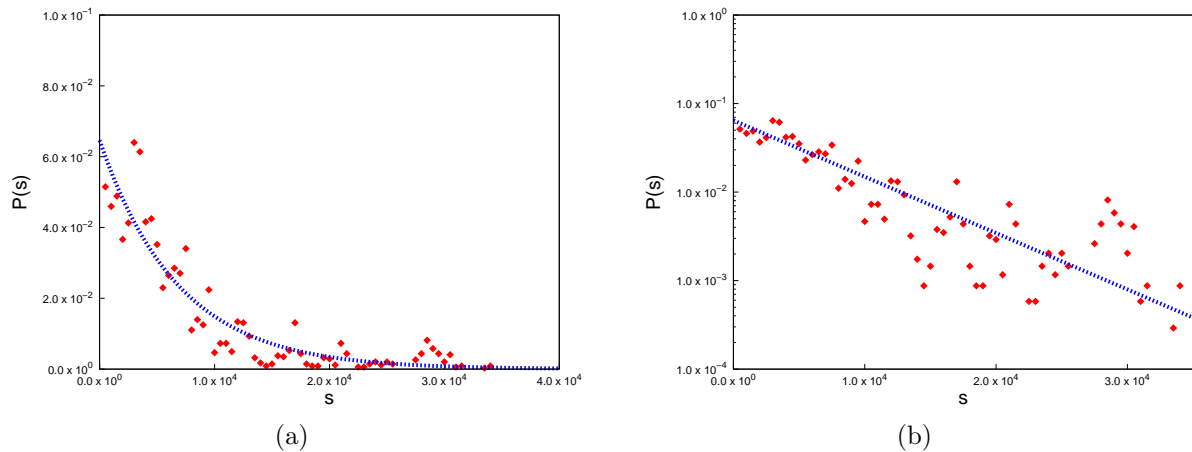


Figura 3.7 - Distribuição do grau do nó da rede espaço- P ; 3.7(a): Distribuição do grau do nó do espaço- P ponderada. Foi utilizado para cada valor no gráfico, o total de Δs compreendido em intervalos de 500 em 500; 3.7(b): Distribuição do grau do nó em escala monolog. Percebe-se que a curva de ajuste atende bem os valores obtidos do sistema. (Fonte: [40])

O melhor ajuste para esse caso foi feito através de uma curva de decaimento exponencial,

$$P(s) = be^{-\beta s}, \quad (3.2)$$

onde $b = 6,50 \times 10^{-2} \pm 3,80 \times 10^{-3}$, com erro de 5,93%, e $\beta = 1,47 \times 10^{-4} \pm 1,19 \times 10^{-5}$, com erro de 8,14%. Pela análise da curva de juste na escala monolog, vemos que a aproximação é razoavelmente boa, Figura (3.7(b)). Embora tenhamos tentado um ajuste com uma função lei de potência na esperança de averiguar se esta rede poderia se enquadrar nas classes de redes sem escalo, a tentativa não foi boa. Sendo o melhor ajuste com feita pela função de decaimento exponencial.

Distribuição do grau do nó sem escala são do tipo que possuem uma cauda em decaimento de lei de potência. Vimos na literatura que redes sem escala dependem de duas propriedades: preferência de conexão e crescimento. Vimos, ainda, que Barabasi afirma que essas duas propriedades são simultaneamente necessárias. Ele demonstra essa tese no modelo A e no modelo B, descrito na Subseção 2.2.2.

No modelo A demonstram a formação de redes com a presença de crescimento, mas não apresentam preferência de conexão. Ou seja, os nós mais conectados não possuem maior probabilidade de serem conectados, de modo que as conexões são feitas de forma aleatória. No entanto, por apresentar crescimento, significa que no processo de formação das redes dessa categoria ouve a inserção de novos nós, portanto, não sendo redes fechadas. Esses tipos de redes apresentam curvas da distribuição do grau do nó em decaimento exponencial.

O resultado da distribuição da força do nó da rede espaço- P do transporte do DF pode indicar que no processo de formação houve a inclusão de novas estações (nós) ao sistema, pois curvas exponenciais para a distribuição do grau sugerem que o fator crescimento está presente, mas o fator preferência de conexão não. Mas a inclusão desses novos nós não obedeceu a propriedade de preferência de conexão, ou seja, quando uma nova estação é incluída no sistema, não necessariamente se faz numa rota que já tenha muitas estações, podendo ser por meio de criação de estações em rotas curtas. Ou ainda, que a criação de novas rotas de ônibus não obedecem à propriedade de preferência de conexão. A criação de uma nova rota, obedece às necessidades de transporte público da população, dessa forma, novas rotas são criadas para novas áreas de povoamento, que muitas vezes não existia nenhuma conexão.

No entanto, quando se trata de local de destino, vemos que o destino de interesse não varia muito, sendo sempre para o centro de oportunidades, geralmente, a região central de Brasília. Por esse ponto de vista poderíamos concluir que a rede deveria obedecer a uma distribuição do grau do nó com preferência de conexão. Obedecendo, dessa forma, uma distribuição do tipo decaimento de lei de potência. Todavia, o melhor ajuste da curva obtido foi de uma decaimento exponencial, como é observado na Figura

(3.7). O que nos leva a questionar se a investigação de propriedades estruturais de grau da rede topológica é suficiente para indicar a classe de rede a qual pertence.

Contudo, esse resultado é semelhante ao obtido na distribuição do grau do nó da rede espaço- P de algumas cidades, tais como o no sistema de Berlim, Düsseldorf e Hong Kong [48]. Das quais apenas o de Hong Kong é um sistema de ônibus, das demais cidades são sistemas completos incluindo trem, metrô e outros modais de transporte de passageiros além da modal ônibus.

Para a distribuição do grau do nó do espaço- C , Figura (3.8), calculamos o valor do $\langle s \rangle = 11.616,55$ e o desvio padrão igual a $\sigma = 6.391,59$. Encontramos uma distribuição muito diferente do encontrado para a rede espaço- P . Vemos que no caso do espaço- P , a distribuição não apresenta um pico, ou seja, não temos um valor de s característico no sistema. Todavia, no caso da espaço- C , encontramos um pico em torno de $\langle s \rangle$. Ajustamos os dados com uma curva gaussiana:

$$P(s_C) = ce^{-\frac{(s_C-d)^2}{r^2}}, \quad (3.3)$$

onde $c = \frac{1}{\sigma\sqrt{2}} = 0,0571686 \pm 0,003016$, $d = \langle s \rangle = 9.617,38 \pm 524,5$ e $r = \sqrt{2}\sigma = 10.960,9 \pm 838$. Onde σ é o desvio padrão.

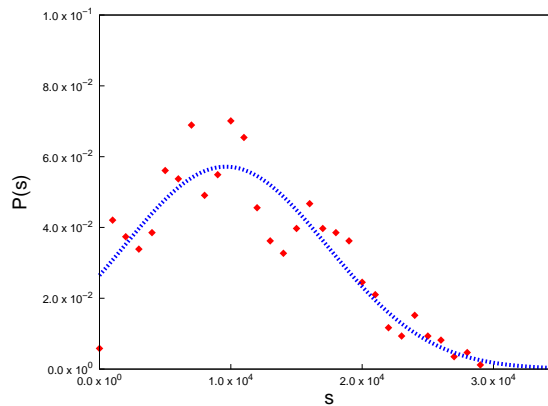


Figura 3.8 - Distribuição do grau do nó da espaço- C ponderada. Onde o intervalo $\Delta s = 300$. (Fonte: [40])

Isso sugere que na rede espaço- C , comparando o resultado ao modelo B, da Subseção 2.2.2, que a quantidade de nós (rotas) no sistema permanece quase estática, muitas vezes, o nome da rota permanece o mesmo, mas o itinerário sofre alteração. O pico sugere que a maioria dos nós possuem o grau do nó com valor próximo ao valor da

média. Isso sugere que com o passar do tempo, as rotas de ônibus chegaram a possuir o mesmo tamanho, em termos de quantidades de estações atendidas. Essa distribuição é diferente dos obtidos em outros estudos de PTN, da cidade de Londres, Los Angeles e Paris, sugerem uma distribuição do tipo decaimento exponencial [48]. Dentre as quais, apenas Los Angeles é PTN apenas de ônibus.

Pela construção das 3 redes do sistema de transporte urbano do DF, espaço- P , espaço- C e espaço- B , podemos observar diferentes distribuições do grau do nó. Encontramos que na rede- B essa distribuição seguia uma lei de potências, que evidencia a ausência de um grau característico. A espaço- P ponderada, por sua vez, apresentou uma distribuição do tipo decaimento exponencial e a espaço- C ponderada, uma distribuição gaussiana, sugerindo elementos de aleatoriedade nas conexões entre os nós.

CAPÍTULO 4

Modelo alternativo para PTN

Se por um lado as redes de transporte podem ter a topologia bem representada pelas redes espaço- P , espaço- C , espaço- L e espaço- B , por outro lado, contudo, se perde informações importantes do sistema de transporte real quando não levamos em conta o fluxo ou o tráfego. Vimos que sistema de transporte ferroviário (metrô, trem) e rodoviário (ônibus) estão sendo representados da mesma forma (Capítulo 3). Contudo, são diferentes na operação e na topologia, isto é, são diferentes quanto ao fluxo de viagens e de passageiros e quanto a conexão entre os estações ou pontos de paradas.

Em um sistema ferroviário, mesmo que existam dois trilhos, um de s_1 para s_2 e outro de s_2 para s_1 , as estações permitem a troca de sentido da viagem sem que se saia da estação, portanto, sem que pague por outra viagem. Assim, é aceitável que as redes PTN ferroviárias sejam redes não-direcionadas. E ir de s_1 a s_2 é tão provável quanto ir de s_2 a s_1 . E a rede espaço- P como rede de transferência/transbordo se mostra correta.

Porém, devemos considerar a rede rodoviário de transporte por ônibus algo diferente da rede ferroviária. Num sistema rodoviário por ônibus, o pagamento ou a validação do bilhete é feita no próprio veículo. Em caso de mudança de sentido na mesma rota será necessário desembarcar e embarcar em um outro veículo, tal como numa rede ferroviária. Contudo, mesmo que um passageiro permaneça usando a mesma rota, se ele não mudar de estação, que geralmente se encontra do outro lado da rua, não conseguirá voltar. Assim, a aresta (s_1, s_2) não é igual a aresta (s_2, s_1) , pois a conexão entre s_2 e s_1 não existe, pois o ônibus não volta pela mesma via. Então, haverá um outro par de nós, tal como (s_{12}, s_{11}) pertencente ao sentido oposto da mesma rota, para fazer a viagem de volta, resultando em topologias distintas para as redes de trem e de ônibus, Figura (4.1). Esta troca de estação obrigatória tem consequências para o passageiro. Em uma estação ferroviária, não é necessário que se pague uma nova passagem, mas em estações de ônibus, será cobrado um novo valor.

A Figura (4.2) apresenta outro exemplo para ilustrar a diferença entre uma rede PTN ferroviária e rodoviária. A Figura (4.2(a)) ilustra uma rede não direcionada, representando uma rede ferroviária. Já na Figura (4.2(b)) temos uma rede direcionada, que deve ser a mais apropriada para representar uma rede PTN de ônibus, pois geralmente, em redes de ônibus, para cada localidade da cidade, existem dois pontos de paradas, uma para cada sentido da via. Em redes ferroviárias, que são representadas por redes não direcionadas, tanto faz ir da estação s_i para s_j como ir da estação s_j para s_i . De forma

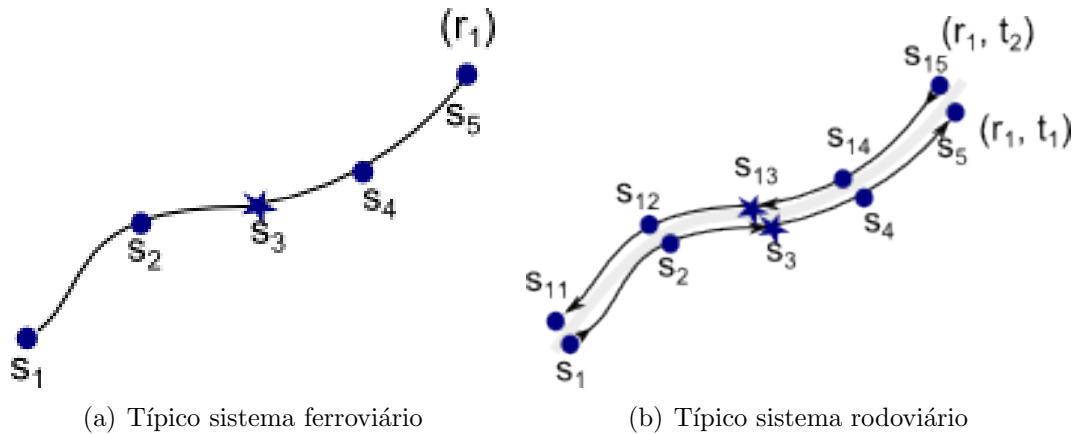


Figura 4.1 - Comparação entre sistemas de trem e de ônibus formado por uma rota r_1 e dois sentidos, que justifica a necessidade de se construir redes direcionadas para o tratamento de redes de ônibus.

que, para fazermos uma viagem da estação s_4 , pertencente a rota r_1 até a estação s_{11} , pertencente a rota r_3 , seria necessário fazer uma transferência de rota na estação s_3 . No caso da rede não direcionada, essa viagem de s_4 até s_{11} , fazendo uma transferência de rota de r_1 para r_3 , seria permitida.

No entanto, redes de ônibus são redes direcionadas. A Figura (4.2(b)) ilustra as rotas r_1 , com sentido de s_1 para s_5 , e outra rota, r_3 no sentido de s_1 para s_{12} . Suponhamos que exista um passageiro que deseja realizar uma viagem de da estação s_4 para s_{11} . Nesse exemplo não é possível realizar essa viagem, visto que não temos nenhuma rota no sistema que realize a viagem saindo da estação s_4 em direção a s_{11} . Poderíamos pensar novamente em uma viagem com integração, tal que o passageiro pudesse realizar uma transferência de rotas, usando 2 ou mais rotas para a viagem. Mesmo que tenhamos estação pertencentes a ambas as rotas, não há viagens no sentido da estação s_4 para s_3 .

Desse modo, a representação da direção da viagem em redes de ônibus, podem trazer mais informações quanto a possibilidade de transferência entre rotas de um sistema.

Ao analisarmos o uma rede de transporte em busca de estações tipo *hub* para a integração do sistema, fica clara a diferença entre os sistemas ferroviário e rodoviário. Ao observar o PTN rodoviário por ônibus percebemos que o conjunto de estações das rota em um sentido são diferentes do conjunto das estações da mesma rota, porém no sentido oposto.

As PTN ferroviárias também podem ser representadas por meio de redes direcionadas, para isso, seria necessário dados de estações sequencias ordenadas no sentido

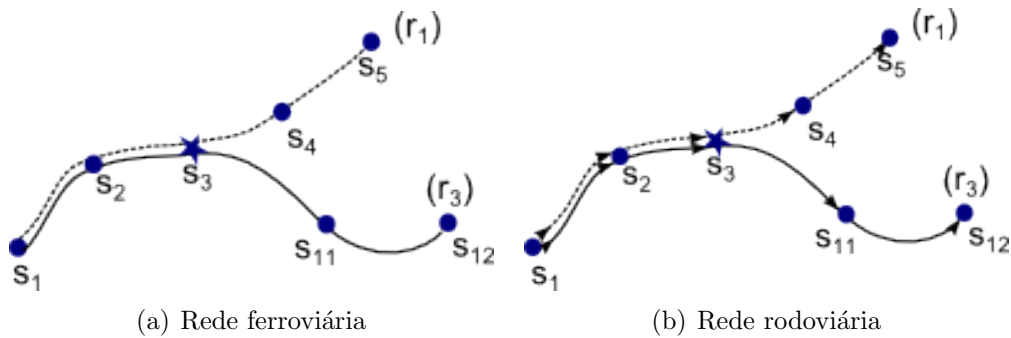


Figura 4.2 - Rede de transporte de ônibus formada por duas rotas (r_1, r_3) ; 4.2(a): Rede não-direcionada; 4.2(b): Rede direcionada.

da rota. Mostramos no Capítulo 5 a importância da construção de redes direcionadas para o cálculo da Pendularidade, uma medida atribuída a distribuição de viagens no tempo. Inclusive, para se estudar a pendularidade de PTN ferroviário, seria necessário a transformação de redes não direcionadas em redes direcionadas.

A mesma importância à direção das arestas é observada para se calcular a distribuição espacial das viagens de passageiros, que é abordado no Capítulo 6. Pois é necessário saber de onde e para onde vão as viagens, ou seja, qual a direção das viagens.

4.1 Dados do Transporte Urbano do Distrito Federal

Os dados do sistema de transporte de ônibus do Distrito Federal estão caracterizados pelo sentido da viagem. Para uma determinada rota, temos dois sentidos, ida e volta. Isso significa que a ida tem um conjunto de pontos de paradas e a volta tem outro conjunto de pontos de paradas referente aqueles posicionado na via de sentido contrário. A Figura (4.3) ilustra duas rotas r_1 e r_2 com suas respectivas viagens, uma em cada sentido. Podemos então usaremos a representação das redes PTN direcionadas.

4.2 Rede PTN direcional

Em algumas redes complexas, tais como *www* [8] ou redes de citações [3], é bastante comum a aplicação de rede direcionada. O grau de conectividade se diferencia em grau de entrada (k^{in}) e grau de saída (k^{out}). Em uma rede direcional, um mesmo nó i pode ser contado para o grau do nó de entrada para uns nós ou para o grau do nó de saída para outros nós.

Podemos escrever a matriz adjacente de uma rede não-direcionada, que raramente será uma matriz simétrica para redes reais. O cálculo do grau de conectividade

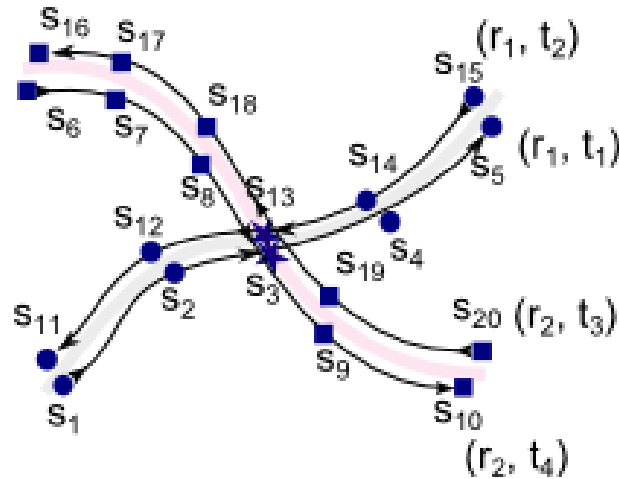


Figura 4.3 - Ilustração de uma rede de ônibus, com duas rotas (r_1, r_2) cada uma com duas viagens, uma de ida e outra de volta, configurando viagens duas viagens diferentes para cada rota, $r_1 = (t_1, t_2)$ e $r_2 = (t_3, t_4)$.

do nó é simples:

$$k_j^{in} = \sum_{i=1}^n a_{ij} \quad (4.1)$$

$$k_i^{out} = \sum_{j=1}^n a_{ij}. \quad (4.2)$$

Como exposto na Subseção 2.4.1, o grau médio é dado por $\langle k \rangle = m/n$, equação (2.28).

Redes de transporte de ônibus em geral possuem conjuntos diferentes de pontos de paradas para as viagens da ida e para as viagens da volta de uma mesma rota.

A partir dos dados descritos na seção 4.1, geramos uma rede espaço- L e outra espaço- P , ambas direcionadas, onde as estações estão em sequência numa rota e são distintas para cada sentido da viagem, tal como ilustrado na Figura (4.4). Assim, trataremos o sistema de transporte de ônibus do Distrito Federal como uma rede direcionada, uma mesma rota terá dois sentidos. Sentido de ida de uma via e de volta com ambas as direções compostas por estações distintas.

4.3 Oferta e demanda em PTN

Vimos na subseção 3.7.2 que podemos definir o peso da aresta como o número de rotas distintas que conectam dois nós. Em estudos de redes ponderadas de aeroportos o peso é definido como o número de tráfego entre dois aeroportos [44]. Assim,

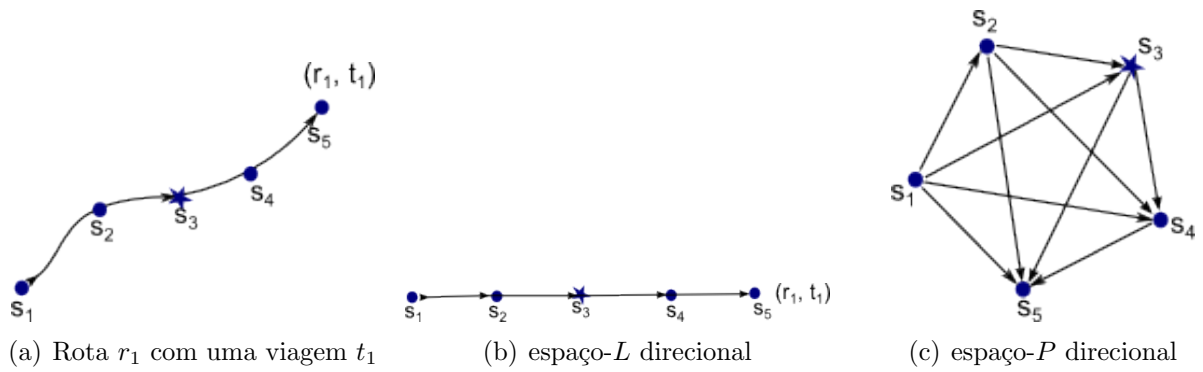


Figura 4.4 - Esquema ilustrativo de construção das redes espaço- L e espaço- P de uma rota de ônibus.

para o sistema de transporte público de ônibus do DF, uma vez estabelecidas as rotas e as estações, podemos verificar como é a operação do sistema.

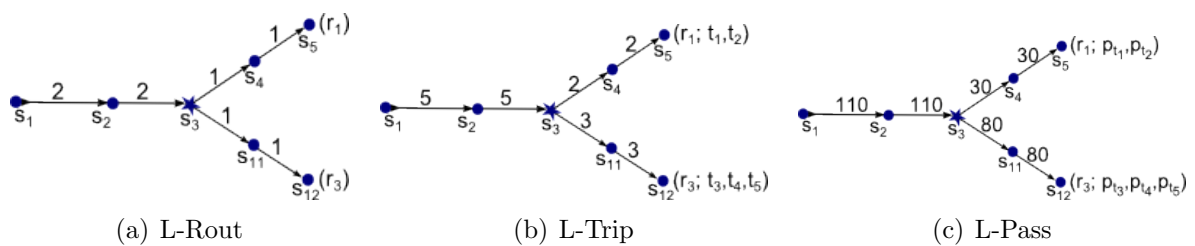


Figura 4.5 - Rede de transporte de ônibus 4.1(b) com rota r_1 com duas viagens t_1 e t_2 e rota r_2 com três viagens t_1 , t_2 e t_3 . Cada viagem com o número de passageiros diferente, $t_1 = 10$, $t_2 = 20$, $t_3 = 50$, $t_4 = 10$ e $t_5 = 20$; 4.5(a) ilustra a Rede espaço- L ponderada pelo número de rotas; 4.5(b) Rede espaço- L operacional ponderada com número de viagens; 4.5(c) Rede espaço- L operacional ponderada com número de passageiros.

Se existem passageiros que queiram ir de uma região A para uma região B , significa que existe demanda pelos serviços de transporte. Se existe uma viagem entre A e B , significa que há oferta de serviços para atender a necessidade dos passageiros.

Assim, definimos o número de passageiros do PTN como a demanda e o número de viagens como oferta. Contudo, a grande dificuldade que temos é de encontrar os dados de operação disponíveis para a pesquisa. Muitos dados topológicos de redes PTN estão disponíveis abertamente na internet, uma boa fonte desses dados é o *Google Transit*. Nesse site estão disponíveis dados geo-referenciados das rotas e estações de PTN de várias cidades pelo mundo.

A partir das redes topológicas PTN direcionada introduzimos os fatores

relacionados ao número de viagens (oferta) e de número de passageiros (demanda) como um peso para a aresta.

Redefinimos então o espaço- P em classes de redes: P -Topo (topologia), P -Rout (rotas), P -Trip (viagens) e P -Pass (passageiros). Igualmente fizemos para a rede espaço- L : L -Topo (topologia), L -Rout (rotas), L -Trip (viagens) e L -Pass (passageiros).

Ilustramos na Figura (4.5) um exemplo de como geramos as redes PTN usando as informações de viagens e de passageiros de uma rede de transporte de ônibus direcionada, Figura (4.1(b)), com rota r_1 com duas viagens t_1 e t_2 e rota r_2 com três viagens t_1, t_2 e t_3 . Cada viagem com o número de passageiros diferente, $t_1 = 10$ passageiros, $t_2 = 20$ passageiros, $t_3 = 50$ passageiros, $t_4 = 10$ passageiros e $t_5 = 20$ passageiros. Essas medidas de viagens e de passageiros podem ser descritas pelo pelo das arestas. Assim, a Figura (4.5(a)) representa a rede espaço- L ponderado com o número de rotas diferentes que geram pesos diferentes às arestas. A rede ponderada por rotas é chamada de rede L -Rout. Da mesma forma, podemos atribuir o valor das multiarestas que representam as várias viagens entre um par de nós como pesos. A Figura (4.5(c)) representa a rede no espaço- L ponderada pelo número de viagens, que chamamos de L -Trip. A Figura (4.5(b)) representa a rede espaço- L ponderada com número de passageiros, que chamamos de L -Pass. Além dessas novas redes ponderadas para o espaço- L , também geramos as redes ponderadas para rede duas outras redes espaço- P , que são P -Rout, P -Trip e P -Pass.

4.4 Resultados

Calculamos as propriedades de centralidade e de caminho das redes espaço- L e espaço- P . Na rede espaço- L temos $n = 3,476$ nós $em = 5,093$ arestas e no espaço- P temos $n = 3,476$ nós e $m = 743,104$ arestas:

Tabela 4.1 - Propriedade do espaço- L , valores médios de: $\langle k \rangle$ – grau do nó, $\langle s \rangle$ – força do nó; $\langle c \rangle$ – coeficiente de aglomeração; $\langle l \rangle$ – comprimento do menor caminho e $\langle \mathcal{B} \rangle$ – coeficiente de intermediação (*betweenness*).

	$\langle k \rangle$	$\langle s \rangle$	$\langle c \rangle$	$\langle l \rangle$	$\langle \mathcal{B} \rangle$
L -Topo	2.93	2.93	0.06	24.21	$6.70e - 3$
L -Rout	2.93	28.48	0.06	24.21	$6.70e - 3$
L -Trip	2.93	623.79	0.06	24.21	$6.70e - 3$
L -Pass	2.93	25,537.26	0.06	24.21	$6.70e - 3$

Usamos a definição de grau ponderado, ou força, dada pela equação (2.29). A ferramenta usada para os cálculos dessas propriedades foi o NetworkX. Como se verifica,

Tabela 4.2 - Propriedade do espaço- P , valores médio de: $\langle k \rangle$ – grau do nó, $\langle s \rangle$ – força do nó; $\langle c \rangle$ – coeficiente de aglomeração; $\langle l \rangle$ – comprimento do menor caminho e $\langle \mathcal{B} \rangle$ – coeficiente de intermediação (*betweenness*).

	$\langle k \rangle$	$\langle s \rangle$	$\langle c \rangle$	$\langle l \rangle$	$\langle \mathcal{B} \rangle$
P -Topo	427.56	427.56	0.62	2.21	$3.48e - 4$
P -Rout	427.56	1,014.69	0.62	2.21	$3.48e - 4$
P -Trip	427.56	21,139.44	0.62	2.21	$3.48e - 4$
P -Pass	427.56	874,608.42	0.62	2.21	$3.48e - 4$

não foi possível diferenciar as propriedades de coeficiente de aglomeração ($\langle c \rangle$), comprimento do menor caminho ($\langle l \rangle$) e coeficiente de intermediação ($\langle \mathcal{B} \rangle$) para as redes com diferentes pesos. Assim, para a redes do espaço- L temos os resultados da Tabela (4.1) e para a rede do espaço- P temos os resultados da Tabela (4.2).

Focamos na distribuição da força s das redes espaço- L e espaço- P .

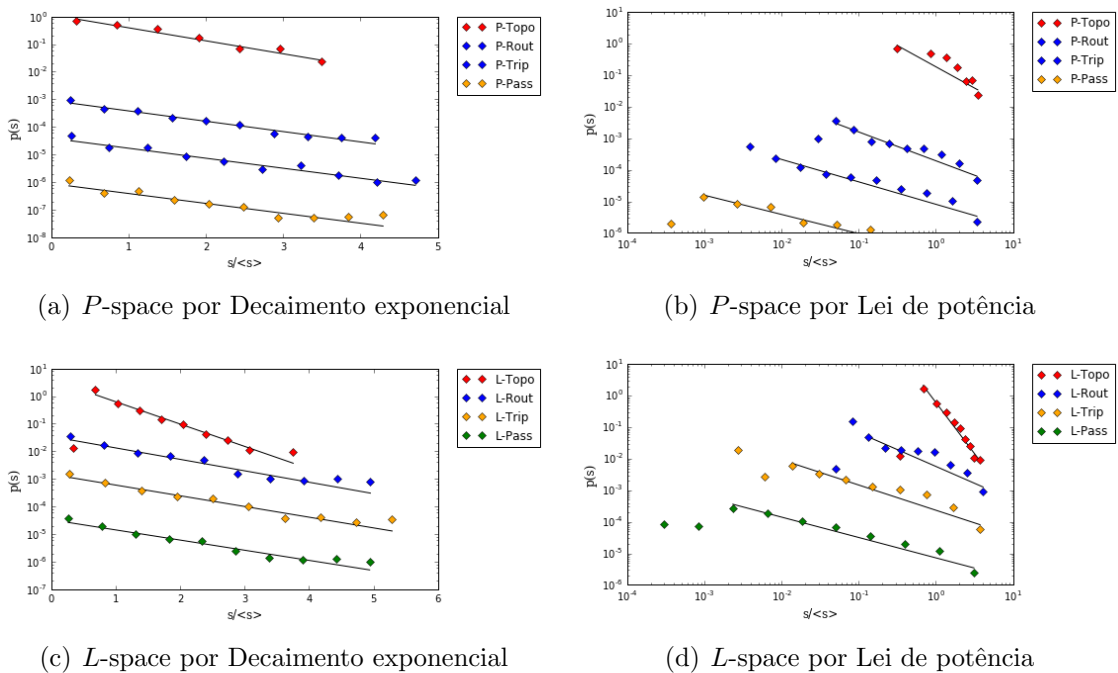


Figura 4.6 - Distribuição do grau das redes (4.6(a) e 4.6(b)) P -Topo (topologia), P -Rout (rotas), P -Trip (viagens) e P -Pass (passageiros); Distribuição do grau das redes (4.6(c) e 4.6(d)) L -Topo (topologia), L -Rout (rotas), L -Trip (viagens) e L -Pass (passageiros).

Calculamos a força do nó nas redes P -Topo, P -Rout, P -Trip e P -Pass. Obtivemos um ajuste de curva para os quatro tipos de redes em Decaimento Exponencial,

Figura (4.6(a)), melhor que Decaimento em Lei de Potência, Figura (4.6(b)).

Apesar do primeiro estudo com os modelos de rede L e P [63] sugerirem que o decaimento exponencial é observado para as redes L , se esperava que a rede P seguisse um decaimento em Lei de Potências. Verificamos que as redes ponderadas também parecem obedecer a mesma distribuição da rede simples (L - e P -Topo) cuja a cauda da distribuição do grau, ou força, para o caso ponderado, comportam-se como um decaimento exponencial. Comparando com resultados de transporte de ônibus de outras cidades do mundo obtidos em [48], verificamos que das 14 cidades daquele trabalho, os dados das cidades de Dallas, Hong Kong, Los Angeles, São Paulo, Sidney e Taipei são apenas de transporte de ônibus, cujos conjuntos de pontos de paradas são distintos para as viagens de ida e de volta de uma mesma rota. Os autores concluem que não há um padrão para a distribuição do grau para os PTNs das cidades analisadas. No entanto, Hong Kong, Sidney e Taipei tiveram ajustes mais confiáveis para decaimento exponencial, como encontramos para o caso do PTN do Distrito Federal.

Em tese, podemos então sugerir que o comportamento de decaimento exponencial do grau independente da escala, pois observamos este comportamento não importa o tamanho da rede, ou não importa a quantidade de multiarestas que a rede possui. No entanto, todas as redes ponderadas tem a mesma topologia, o que difere entre elas é o valor das multiarestas.

Acreditamos que os resultados deste trabalho possam trazer algum acréscimo ao estudo de redes PTN, pois partimos de um método um pouco mais elaborado, inserindo informações como número de rotas, viagens e passageiros às redes e diferenciamos uma rede não direcionada e uma rede direcionada e encontramos Decaimento exponencial para a distribuição da força s , tanto para as redes do espaço- P , quanto para as redes do espaço- L . No entanto, seria interessante calcular algum parâmetro entre as redes L -Topo, L -Rout, L -Trip e L -Pass, usando algum método de redes em multicamadas. O mesmo seria feito para as redes P -Topo, P -Rout, P -Trip e P -Pass.

No entanto, não sabemos quanto a simplificação feita nas viagens tenha afetado este resultado, uma vez que consideramos todas as viagens de uma rota como indistinguíveis, não valorando por horário ou algo do tipo que desse pesos distintos de acordo com cada viagem. Por exemplo, uma viagem realizada às 8h da manhã ser mais importante que uma viagem às 10h, quando a demanda é bem pequena. Ou o contrário, uma viagem das 10h ser mais importante, pois há menos ofertas nesse horário. Da mesma forma fizemos para os passageiros, considerando-os apenas como uma quantidade. Há estudos que discutem a distinguibilidade dos viajantes em redes complexas de mobilidade

humana [73]. Não usamos aqui, todavia.

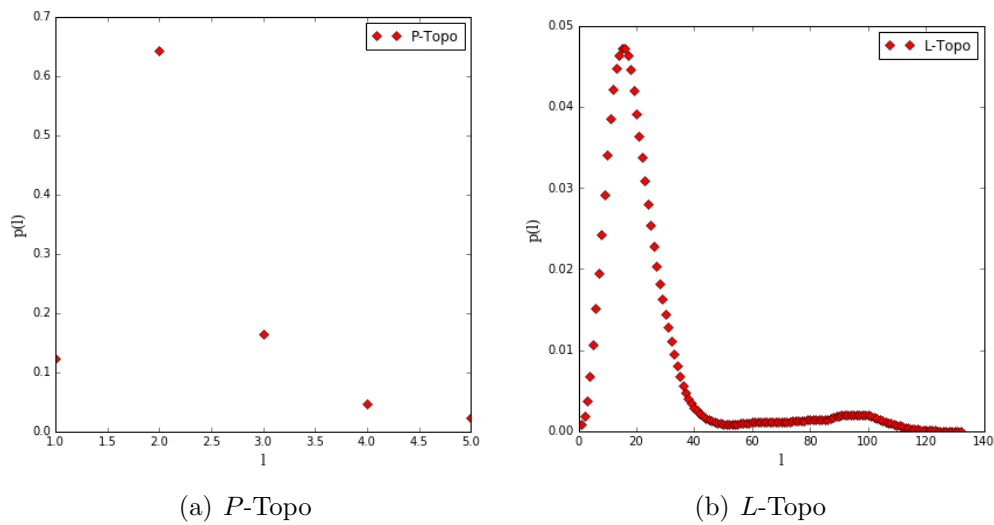


Figura 4.7 - Distribuição do comprimento do menor caminho pra as redes topológicas P -Topo e L -Topo.

Além disso, fizemos o gráfico da distribuição do comprimento de menor caminho das redes P -Topo e L -Topo para ilustrar a diferença de tamanhos das redes PTN, Figura (4.7). Na prática, o comprimento do menor caminho tem diferentes significados para as redes L -Topo e P -Topo.

Para o caso da rede P -Topo, significa o número de transferências necessárias para se ligar duas estações quaisquer. Vemos que grande parte necessita duas rotas (arestas) para uma conexão, maior que 60% dos casos, esse resultado corrobora com os obtidos anteriormente (Subseção 3.3.1). Já o comprimento de menor caminho da rede L -Topo significa o tamanho da rota em termos de números de estações. Na sua grande maioria está concentrada em torno de 20 estações.

Optamos por calcular o comprimento médio de menor caminho apenas para as redes topológicas, pois o valor do peso quando colocado no algoritmo do networkX interpreta como tamanho do caminho, que não é o caso. Mas fica aberta a possibilidade de se inserir novos dados ao sistema.

CAPÍTULO 5

Pendularidade em PTN

Neste capítulo apresentamos um método de análise de redes PTN usando informações da operação do sistema de transporte, tais como horários, viagens e passageiros, que possam ajudar a compreender melhor a dinâmica de uma rede complexa.

Apresentamos uma motivação relacionada ao movimento diário das pessoas em um território, ou seja, a pendularidade da mobilidade urbana, inserida no contexto das redes complexas.

Trabalhamos a partir da distribuição dessas medidas de viagens e de passageiros no tempo para definir coeficientes, que denominamos de R_{CM} e h_{CM} , cujos valores podem elucidar mais detalhes da dinâmica das redes PTN. Os coeficientes são análogo a qualquer sistema de massas, que no nosso estudo foram usados como massa os dados de viagens e os dados de passageiros de cada rota.

Definimos os coeficientes harmônicos construídos a partir da ideia de movimentos harmônicos e periódicos.

Calculamos para a rede L -Trip e L -Pass os valores de R_{CM} e h_{CM} , usando o conceito de força das arestas de um rede complexa, que podem trazer novas informações das redes PTN, que vão além da sua topologia.

5.1 Conceituação Demográfica

Em uma situação com recursos infinitos, comprar uma quantidade de veículos igual à razão passageiro/lotação do veículo seria muito simples. Em uma cidade que existam 100 pessoas que querem ir do ponto i ao ponto j , o número de veículos necessário seria dois veículos com capacidade de 50 passageiros cada. Ou, se das 100 pessoas 50 necessitam viajar às 7h da manhã e outras 50 apenas às 8h, e se a viagem de ida e volta, de i a j e retornando a i , tem duração menor que uma hora, a necessidade de veículos cai para metade, ou seja, será necessária a compra de apenas um veículo com capacidade de 50 passageiros para atender a demanda de transporte de 100 pessoas. Se os desejos de viagens de i a j fosse homogênea ao longo do dia, tal que, a cada hora, 50 pessoas quisessem viajar de i a j , um único veículo seria suficiente para atender a demanda da população. E ao fim do dia, o veículo teria realizado 24 viagens. Ou seja, a capacidade de transporte desse sistema formada por um único veículo seria o tempo de percurso vezes a lotação do veículo dividido por 24h.

Contudo, o sistema real não é tão simples assim e a distribuição de demanda de viagens não é uniforme ao longo do dia. Na maioria das vezes o movimento diário da população está vinculado ao trabalho, e por causa dos horários de início e fim de jornada de trabalho da grande maioria serem coincidentes, a maior parte da população está se deslocando ao mesmo tempo. Em uma cidade de distribuição populacional uniforme isso não seria problema. No entanto, muitas cidades apresentam monocentros de interesse da população. Isso significa que todos tenderão a ir para o centro pela manhã para trabalhar e voltar de lá pelo fim do dia para suas casas. Assim, muitas pessoas estão se deslocando ao mesmo tempo e para a mesma direção e sentido. Em cidades com distribuição de postos de trabalho mais homogêneos no território, com descentralização espacial de postos de trabalho, não se verificaria congestionamentos no trânsito. Se a cidade tem poucos habitantes também não terá problemas com congestionamentos. Em metrópoles, todavia, não se observa essa tranquilidade e diariamente é observado congestionamentos cada vez maiores.

Muitas vezes saber definir quando começa o horário de pico e quando termina pode melhorar a alocação de veículos de um sistema de transportes coletivo. Mesmo sendo heterogêneo durante um dia, o movimento de pessoas ao dia parece obedecer a um padrão.

Em mobilidade o conceito de movimento pendular refere-se ao processo de deslocamento da população no território, ou seja, é o vai-e-vem de pessoas num território. Do inglês, *commuting* refere-se a troca de pessoas entre cidades a partir do vai-e-vem diário, seja por motivo de consumo, trabalho, negócios, etc. Esse movimento rítmico cotidiano se assemelha ao movimento de um pêndulo, por isso o nome comumente usado em geografia de pendularidade para o fenômeno de vai-e-vem diário entre municípios [74].

Ainda segundo [74], *commuting*, entretanto, expressa a relação que as cidades estabelecem entre si em relação a bacias de empregos, já pendularidade ressalta o papel que as escolhas individuais e a escolha dos estilos de vida desempenham na estruturação da vida diária. Observamos claramente a analogia empregada entre movimentos de sistemas de osciladores com o desejo de viagens de pessoas num território.

Nos Estados Unidos, os dados de movimentos pendulares são usados para redefinir as áreas metropolitanas, enquanto que no Brasil não são considerados como elementos importantes para a definição de uma região metropolitana, assim, aqui, uma metrópole se confunde com a região metropolitana, não se distinguindo o que seria a delimitação regional por lei de criação da mesma ou a importância sócioeconômica [75]

De forma geral, dados de amostra populacional mostram que a principal motivação para os deslocamentos diários é ainda o trabalho. Para o caso do Distrito Federal, são realizadas as Pesquisas de Amostra Domiciliar que apresentam as características de movimento cotidiano da população por motivo de trabalho, de estudo e outros [76]. A partir desse pesquisa, o estado tem gerado a matriz de Origem-Destino, cujos dados se referem ao movimento da população no território. Esses dados são importantes para que se defina a demanda de serviços públicos, tais como a demanda de transporte público de passageiros. Além disso, esses dados de Origem-Destino são importantes para o planejamento de estruturas rodoviárias, construção de escolas e hospitais públicos e outro.

Para o presente trabalho, analisaremos apenas a dinâmica do movimento, sem abordar a relação individual com profundidade. Apenas serão analisados os aspectos do movimento de vai-e-vem da população em termos de distribuição espacial e temporal.

5.2 Movimento Pendular

Na Física, os movimentos de um objeto em torno de sua posição de equilíbrio, cujos desvios das condições de equilíbrio estável do sistema são suficientemente pequenos, podem ser geralmente descritos como um sistema de osciladores harmônicos lineares acoplados.

O movimento de ida-e-volta da casa-trabalho-casa de um indivíduo em um certo território pode ser considerado como um sistema conservativo quanto ao número de indivíduos. Suponha que, nos movimentos diários de uma região a outra por qualquer que seja o motivo *à priori*, o indivíduo regressa a casa no fim do dia ou noite tal que o número de habitantes de cada região seja uma constante num ciclo, mas que varie ao longo do dia. Assim, considerar-se-á a casa como o ponto de partida e de chegada e o local de trabalho como o ponto em torno do qual os movimentos de oscilação serão descritos. Após um ciclo o número de indivíduos de uma região se conserva, significando que todos os indivíduos regressarão a casa ao fim do dia.

Assim, pela noite, o indivíduo estaria em repouso, literalmente, no seu horário de descanso. Portanto, por simplificação, a cada dia o sistema de transporte coletivo de passageiros do DF será um sistema fechado.

A analogia dos deslocamentos diários da população com sistemas de osciladores tem suas vantagens. A presença de características físicas tais como o período, frequência e fase são algumas delas. A análise desses dados podem produzir resultados quanto as características do movimento em questão. O período de oscilação será de um

dia, 24 horas, uma vez que as atividades serão repetidas no dia seguinte (dia útil).

O sistema de transporte público coletivo de passageiros surge da necessidade de transportar os trabalhadores das suas casas ao trabalho e vice-versa. Da mesma forma, para cada viagem de ônibus, com uma identidade distinta, a viagem só se repetirá no dia seguinte. Ou ainda, um trabalhador sairá da sua casa rumo ao trabalho novamente no dia seguinte. Se seu horário é regular, espera-se que ele sairá de casa no mesmo horário e local todos os dias da semana.

Assim, analisar a distribuição de viagens de ônibus ao longo do dia ou a distribuição de passageiros pode trazer luz à demanda de transporte coletivo para cada região calculando-se as propriedades físicas dos movimentos periódicos do sistema. Inclusive, respostas mais precisas quanto aos horários de pico para cada local.

5.3 Estudo do centro de massa

Em ciências demográficas é bem comum a ideia de pendularidade, um sistema formado por pessoas interessadas em ir e vir de/para uma região. Sejam em movimentos diários, semanais, mensais em busca de realizações pessoais tais como emprego, lazer, estudo.

A grandeza física característica de sistemas pendulares é o período. Se um sistema de transporte público de passageiros, de ônibus, por exemplo, tem uma certa oferta de viagens ao longo da semana, que se repete de segunda a sexta-feira, podemos supor que se trata de um movimento periódico diário.

Queremos observar a distribuição de ofertas de viagens e de demanda de passageiros.

Supondo que num ponto qualquer identificamos três viagens ao longo do dia, às 9h, 12h e 15h, com carregamento de 500, 100 e 200, respectivamente. Na Figura (5.1(a)) vemos a distribuição dessas viagens e o R_{CM} e θ_{CM} indicando o centro de massa. E na Figura (5.1(b)), a distribuição de viagens em que se leva em conta o número de passageiros. Assim, temos dois parâmetros: o raio que indica a concentração, se $R_{CM} = 1$ significa que todas as viagens se encontram num mesmo horário, se $R_{CM} = 0$, a distribuição será uniforme ao longo do dia. O ângulo θ_{CM} se refere a horário típico da movimentação observada. O valor de R_{CM} expressa o grau de concentração de movimentação em torno do valor definido por θ_{CM} .

Assim, a distribuição de ofertas de transporte por ponto de parada por dia

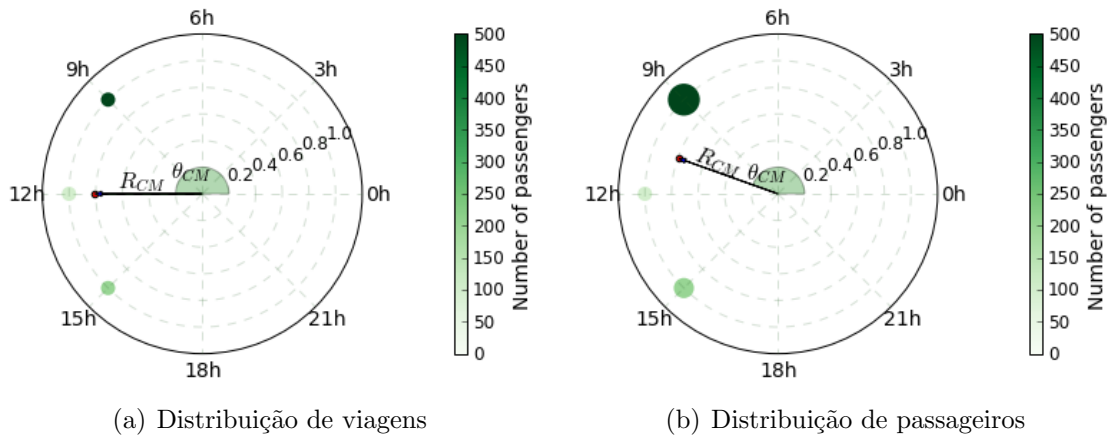


Figura 5.1 - Representação das medidas R_{CM} e θ_{CM} : (a) Distribuição de viagens num ponto de parada ao longo do dia; (b) Distribuição de passageiros num ponto de parada ao longo do dia.

pode ser caracterizado a partir de duas grandezas: raio R_{CM} e ângulo θ_{CM} . E portanto, faremos uma analogia a posição do Centro de Massa de um sistema de partículas. Sabemos que

$$\mathbf{R}_{CM} = \frac{1}{M} \sum_i^N m_i \mathbf{r}_i \quad (5.1)$$

$$M = \sum_i^N m_i \quad (5.2)$$

onde m_i é a massa de cada partícula e \mathbf{r}_i sua posição, é a equação do raio do centro de massa. Como estamos interessados apenas no horário de cada evento, definimos $r_i = 1$, tal que, $x_i = \cos \theta_i$ e $y_i = \sin \theta_i$. De forma que, podemos representar o centro de massa por X_{CM} e Y_{CM} :

$$\sum_i m_i \mathbf{r}_i = \sum_i m_i (x_i \hat{\mathbf{x}} + y_i \hat{\mathbf{y}}), \quad (5.3)$$

tal que, as componentes são:

$$X_{CM} = \frac{1}{M} \sum_{i=1}^N m_i \cos \theta_i \quad (5.4)$$

e

$$Y_{CM} = \frac{1}{M} \sum_{i=1}^N m_i \sin \theta_i. \quad (5.5)$$

Podemos considerar a 'massa' m_i de uma viagem i que serve um certo ponto

de parada ao longo do dia, $\theta_i = 2\pi h_i/T$ como uma grandeza relacionada ao horário h_i da viagem i . Para um período de $T = 24$ horas, temos:

$$X_{CM} = \frac{1}{M} \sum_{i=1}^N m_i \cos\left(\frac{2\pi h_i}{24}\right) \quad (5.6)$$

e

$$Y_{CM} = \frac{1}{M} \sum_{i=1}^N m_i \sin\left(\frac{2\pi h_i}{24}\right). \quad (5.7)$$

Assim,

$$R_{CM} = \sqrt{X_{CM}^2 + Y_{CM}^2} \quad (5.8)$$

$$\theta_{CM} = \arctan(Y_{CM}/X_{CM}) \quad (5.9)$$

E a partir disso podemos calcular também o horário do CM:

$$h_{CM} = \frac{24}{2\pi} \arctan \frac{Y_{CM}}{X_{CM}} + \begin{cases} 24, & \text{se } X_{CM} > 0 \text{ e } Y_{CM} < 0 \\ 0, & \text{se } X_{CM} > 0 \text{ e } Y_{CM} > 0 \\ 12, & \text{se } X_{CM} < 0 \end{cases} \quad (5.10)$$

Assim o tamanho do raio R_{CM} representará a concentração de viagens em algum horário do dia. Esse resultado não indica o horário de pico das viagens, mas o horário típico das viagens. R_{CM} será tal que $0 \leq R_{CM} \leq 1$.

Quanto menor for o R_{CM} , mais homogênea será a distribuição ao longo do dia; e quanto mais próximo de 1, mais heterogênea será a distribuição, significando que há uma maior concentração de viagens em um certo intervalo de tempo do dia. Assim, podemos verificar a partir de R_{CM} o horário de maior concentração para um determinado ponto de parada. Para o caso de termos uma grande quantidade de dados, podemos escrever 5.6 e 5.7 para um caso contínuo:

$$X_{CM} = \frac{\int_{\theta_1}^{\theta_2} \cos(\theta) d\theta}{\int_{\theta_1}^{\theta_2} d\theta}, \quad (5.11)$$

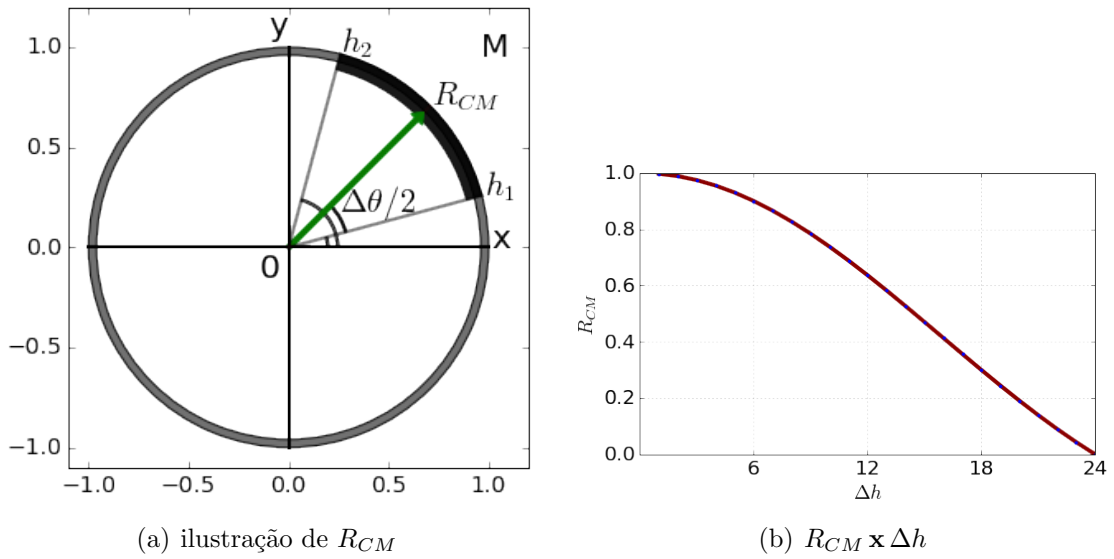


Figura 5.2 - Comportamento do R_{CM} em função do $\Delta h = h_2 - h_1$.

e

$$Y_{CM} = \frac{\int_{\theta_1}^{\theta_2} \sin(\theta) d\theta}{\int_{\theta_1}^{\theta_2} d\theta}, \quad (5.12)$$

E portanto, podemos escrever R_{CM} em função de $\Delta\theta$ (ver apêndice A):

$$R_{CM} = \frac{2}{\Delta\theta} \sin\left(\frac{\Delta\theta}{2}\right). \quad (5.13)$$

onde $\Delta\theta = 2\pi(h_2 - h_1)/24$ e R_{CM} dependem apenas dos horários do ônibus num certo ponto de parada. A Figura (5.2) mostra como o raio R_{CM} depende de $\Delta h = (h_2 - h_1)$. Viagens mais concentradas implicam em menor Δh e maior R_{CM} .

Para o tratamento de viagens carregadas de passageiros, uma viagem não é igual a outra e $m_i \neq 1$. Podemos, então, atribuir um valor de peso em cada viagem tal que m_i seja agora o valor do número de passageiros de uma certa viagem i . Tal modificação será dada nas equações 5.4 e 5.5. E a soma $\sum_i m_i$ será o total de passageiros transportados no dia. Vale ressaltar que estamos nos referindo ao número de passageiros transportados e não ao número de pessoas beneficiadas pelos serviços de transporte, uma vez que não foi identificada a pessoa e nem a quantidade de vezes que ela utilizou o serviço ao longo do dia.

Esse estudo só é possível ser realizado quando usamos a rede espaço- L dire-

cionada, pois necessitamos da sequência dos nós de uma viagem. Com a PTN de ônibus as arestas são naturalmente direcionadas, o que facilita identificar o sentido do movimento. Nos estudos realizados com redes PTN não direcionadas, não se sabe de onde e para onde estão indo as viagens. Para o estudo de redes ferroviárias pode ser necessário incluir a direcionalidade explicitamente. O uso das redes tipo espaço- L desenharam o caminho das viagens, que não seria possível com o uso das redes tipo espaço- P .

5.4 Solução para sistemas periódicos

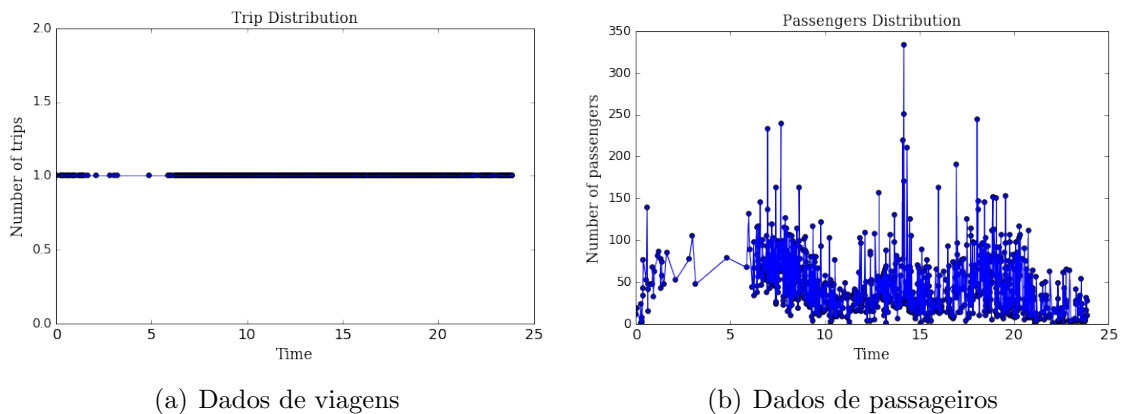


Figura 5.3 - Gráfico do (a) Número de viagens e do (b) Número de passageiros que passam no ponto de parada 72122.

A Figura (5.3) mostra os registros das viagens realizadas (5.3(a)) em um dia, assim como a quantidade de passageiros (5.3(b)) em cada uma das viagens. Na prática, essas informações seriam o equivalente ao resultado de um contador instalado num certo ponto de parada, tal que, toda vez que uma viagem de ônibus passasse no local, essa viagem seria registrada, fornecendo horário e carregamento (quantidade de passageiros total).

Vamos supor que a distribuição de viagens ao longo do dia se repita no dia seguinte e no próximo e assim por diante, tal que queiramos descrever uma função do número de viagens pelo tempo $X(t)$. Essa função deve ser periódica, já que supomos que irá se repetir a cada dia $X(t + T) = X(t)$, onde $T=24$ horas é o período.

Sabemos que $\sin(x)$ e $\cos(x)$ são funções periódicas, e a soma de funções periódicas também é uma função periódica. Dessa forma, podemos propor uma solução

possa ser expressa por uma série de Fourier,

$$X(t) = \sum_{n=0}^{\infty} [a_n \cos(\omega_n t) + b_n \sin(\omega_n t)] , \quad (5.14)$$

onde a_n e b_n dependem da função $X(t)$.

Como consequência das condições de ortonormalidade (ver Apêndice B), podemos encontrar os coeficientes a_n e b_n . Para isso, multiplicamos a equação (B.1) por $\cos(2\pi mt/T)$ e integramos em t de $[-T/2, T/2]$, ou $[0, T]$, mantendo $m > 0$ fixo,

$$\begin{aligned} \int_{-T/2}^{T/2} X(t) \cos\left(\frac{2\pi m}{T}t\right) dt &= \frac{a_0}{2} \int_{-T/2}^{T/2} \cos\left(\frac{2\pi m}{T}t\right) dt \\ &+ \sum_{n=1}^{\infty} a_n \int_{-T/2}^{T/2} \cos\left(\frac{2\pi n}{T}t\right) \cos\left(\frac{2\pi m}{T}t\right) dt \\ &+ \sum_{n=1}^{\infty} b_n \int_{-T/2}^{T/2} \sin\left(\frac{2\pi n}{T}t\right) \cos\left(\frac{2\pi m}{T}t\right) dt. \end{aligned} \quad (5.15)$$

Como o único termo não-nulo à direita do sinal de igualdade é quando o termo $m = n$, temos

$$\int_{-T/2}^{T/2} X(t) \cos\left(\frac{2\pi n}{T}t\right) dt = \frac{T}{2} a_n, \quad n = 1, 2, \dots, \quad (5.16)$$

Para encontrar b_n multiplicamos a equação (B.1) por $\sin(2\pi mt/T)$ e repetimos o processo. E podemos calcular os valores dos coeficientes de Fourier por

$$a_0 = \frac{2}{T} \int_{-T/2}^{T/2} X(t) dt \quad (5.17)$$

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} X(t) \cos\left(\frac{2\pi n}{T}t\right) dt, \quad n = 1, 2, \dots \quad (5.18)$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} X(t) \sin\left(\frac{2\pi n}{T}t\right) dt, \quad n = 1, 2, \dots \quad (5.19)$$

Podemos expressar duas grandezas derivadas dos coeficiente de Fourier para caracterizar

nossa sistema,

$$M = a_0 \quad (5.20)$$

$$R = \sqrt{a_n^2 + b_n^2} \quad (5.21)$$

$$\theta = \arctan\left(\frac{b_n}{a_n}\right). \quad (5.22)$$

Comparando, respectivamente, com as equações (5.2), (5.8) e (5.9), podemos a partir dos dados das viagens, encontrar as variáveis do centro de massa, usando $n = 1$. O cálculo para a distribuição de passageiros é análogo.

5.5 Dados do Transporte Urbano do Distrito Federal

No Distrito Federal existem dois modais de transporte público: o modal rodoviário e o metroviário. O estudo empírico foi realizado com os dados de transporte por ônibus do Distrito Federal. O DFTRANS, autarquia gestora do transporte urbano do DF, forneceu os dados de passageiros que foram extraídos do banco de dados do Sistema de Bilhetagem Automática (SBA) e os dados geo-referenciados em formato GTFS com as previsões de pontos de paradas, rotas e viagens.

Outros dados importantes para a construção da rede de transporte urbano foram fornecidos pela Secretaria de Estado de Mobilidade do Distrito Federal – SEMOB, do qual foram obtidos os dados de polígonos de cada zona do Distrito Federal (geo-referenciado), que deram origem aos dados do PDTU 2009 e pela COODEPLAN, da qual foram obtidos os dados de censo, tais como quantidade de trabalhadores e de habitantes de cada Região Administrativa do Distrito Federal.

Dos dados GTFS referentes aos dados planejados correspondem ao total de 35.463 viagens diferentes, 781 rotas, e 3.673 pontos de paradas cadastrados. Os dados de viagens apresentam o código de identificação, horários estimados em cada ponto de parada, identificação da rota a que pertence a viagem e sequência de pontos de paradas ao longo da trajetória. A esses dados chamaremos de dados das viagens planejadas.

Outro conjunto de dados fornecidos pelo DFTRANS se referem aos dados extraídos do SBA, que correspondem aos dados das viagens operadas. Os dados de viagens operadas apresentam o código de identificação, nome da rota, horário de início da viagem, horário de término da viagem, local de origem, local de destino e o número de passageiros transportados. Esses dados correspondem ao total de 27.320 viagens realizadas no dia 13/10/2014 em 876 rotas distintas, com o total de 1.104.908 passageiros transportados.

As identificações das viagens das duas bases de dados não são compatíveis, isso gerou a necessidade de eliminar as viagens operadas que não fazia parte da base de dados de viagens planejadas, pois necessitamos dos dados dos pontos de paradas de cada rota e de cada viagem. A única correspondência entre os dois banco de dados era o nome da rota.

Entre as 876 rotas operadas, apenas 574 pertenciam a base de dados de rotas planejadas. Esse é o total de rotas que puderam ser identificadas em ambas as bases de dados. A redução foi necessária porque os dados das viagens operadas não possuem informações de pontos de paradas, tais como, localização, sequencia e horário. E porque os dados das viagens planejadas não possuem dados de carregamento de passageiros e horários de atendimento em cada ponto de parada.

Para a correspondência entre os dados das viagens planejadas e das viagens operadas, utilizamos algumas aproximações: 1 - eliminamos as viagens operadas que o nome da rota não pertencia à base de dados das viagens planejadas; 2 - eliminamos as viagens operadas que o local de origem da rota não correspondia ao local de origem da correspondente rota pertencente à base de dados das viagens planejadas.

Dos dados de segunda-feira dia 13/10/2014 houve a redução de 876 rotas para 574 rotas, equivalente a 66% do total; de 27.320 viagens para 17.479 viagens, equivalente a 64% do total de viagens; e de 1.104.908 passageiros transportados para 710.765 passageiros, que equivale a 64% do total transportado no dia.

Posteriormente, complementamos as viagens operadas com dados dos pontos de paradas da rota e o horário estimado em cada ponto de parada, a partir do horário de início e local de início da viagem operada, complementamos com os pontos de paradas ao longo da rota e horários estimados em cada ponto de parada, a partir dos dados de intervalo de tempo entre dois pontos de paradas consecutivos extraídos do banco de dados das viagens planejadas.

Ao final conseguimos uma tabela com os seguintes dados de cada viagem: nome da rota, identificação da viagem, local de início da viagem, local de término da viagem, pontos de paradas das rota em sequência, horário em cada ponto de parada, latitude dos pontos de paradas da rota, longitude dos pontos de paradas da rota e número de passageiros da viagem. Usamos um pouco de mineração de dados para tratar os dados brutos para se obter dados para a construção da rede complexa (Apêndice D).

5.6 Análise da Distribuição dos horários em Redes espaço- L

Vimos no Capítulo (3) os tipos de redes que podem ser geradas a partir de um sistema de transporte público levando-se em conta a sua topologia, quais sejam: espaço- P , espaço- L , espaço- C e espaço- B , ilustrados na Figura (3.2).

Percebemos que muitos autores tratam da mesma forma um sistema ferroviário e um sistema rodoviário, representando-os como redes não direcionadas. Sugerimos que o modelo pode não ser o mais adequado, pois as estações/pontos de paradas de ônibus possuem geralmente duas estações distintas, uma em cada sentido da via, logo, uma em cada sentido da rota. Dessa forma, para a mudança de sentido de viagem em uma rota de transporte de ônibus, é necessário que o passageiro troque de estação, em geral. Isso é diferente de um sistema ferroviário, em que é possível mudar o sentido da viagem sem que mude de estação.

Isso permite dizer que as redes ferroviárias podem ser representadas pelas redes não direcionadas PTNs, contudo, as redes rodoviárias, por serem naturalmente direcionadas, devem ser representadas por redes direcionadas, como discutido no Capítulo 4. Uma grande diferença é que teremos viagens de i a j , mas não o contrário em redes rodoviárias. No entanto, se desejarmos aplicar a metodologia proposta nesse trabalho às rotas ferroviárias, será necessário utilizarmos redes direcionadas, de forma que as arestas possuam sentido de fluxo matinal e vespertino diferentes, caso contrário, não será possível medir a pendularidade.

Vimos no Capítulo 4 que o estudo de topologia realizado em redes de transporte não é suficiente para entender como é o real funcionamento delas. Propusemos estudos com redes ponderadas para verificar alguma propriedade da rede que pudesse elucidar sua dinâmica. Sugerimos o estudo de redes ponderadas para se evitar maior complexidade no modelo, tal que os pesos sejam algum elemento do funcionamento operacional do sistema de transportes, tais como o número de rotas, viagens e passageiros. Assim, mostramos na Figura (4.6) que o peso das redes ponderadas, com número de rotas, viagens ou passageiros, não interfere na forma de distribuição do grau do nó. E, talvez possamos usar melhor uma rede ponderada para entender a dinâmica de um PTN.

Admitindo a direção e peso das arestas, fizemos o estudo da pendularidade para cada ponto de paradas do DF, cujo conceito está relacionado ao movimento de ir e vir em uma rede. A partir disso, podemos verificar as tendências de deslocamento entre diferentes nós e arestas ao longo de um ciclo. Adotamos o ciclo com o período de $T = 24$ horas.

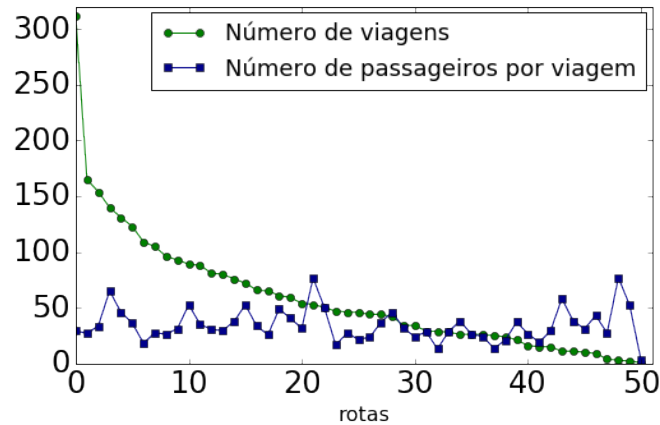


Figura 5.4 - Distribuição de número de viagens e de passageiros por viagem das rota de ônibus do DF com maior número de viagens.

Percebemos que para uma determinada rota, a proporção do número de passageiros e do número de viagens são diferentes. A Figura (5.4) mostra que ora o número de passageiros é maior, ora o número de viagens é maior. Isso significa que ora o veículo viaja lotado, ora viaja vazio, respectivamente. Uma forma de se inserir essas informações à rede PTN é através dos pesos e direção.

Outro fato importante, se refere a distribuição de viagens e de passageiros em uma determinada rota ao longo do dia. Essa distribuição no tempo não é uniforme, isso tem a ver com a dinâmica de uma região onde está inserida a rede PTN.

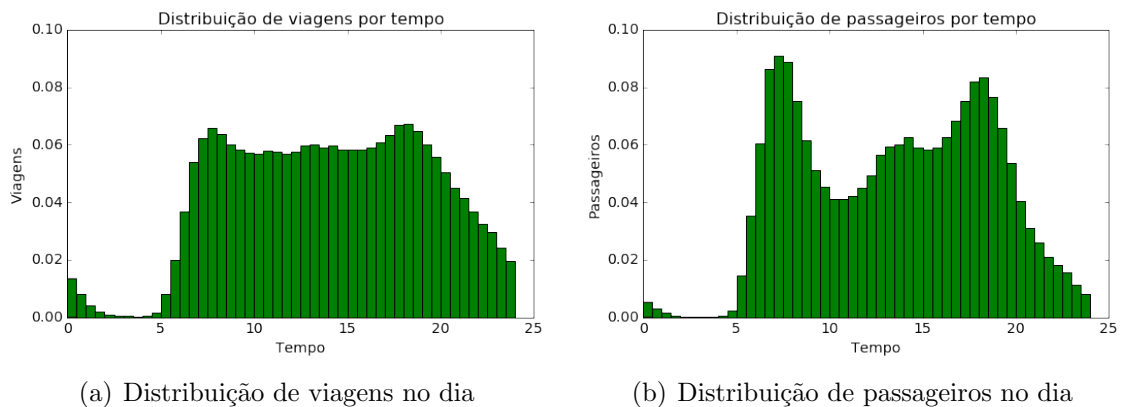


Figura 5.5 - Distribuição do número de viagens e do número de passageiros do DF ao longo do dia, considerando todos os pontos de paradas.

Observando o total de viagens e de passageiros ao longo do dia, percebemos que existem horários em que a atividade é mais frequente. São os chamados horários de

pico. A Figura (5.5) mostra a distribuição das viagens e de passageiros ao longo do dia. Vemos maior atividade de transporte em alguns horários e menor em outros. Justamente o que corrobora com a dinâmica da região, o ir e vir das pessoas.

Assim, podemos observar a distribuição de viagens ofertadas ao longo do dia no Figura (5.5(a)) e vemos que a maior parte dos pontos de parada estão sendo servidos por alguma viagem em dois picos: um primeiro por volta das 7h a 8h da manhã e outra por volta das 18h da tarde. E entre esses dois picos, observa-se que a oferta de serviço continua de forma quase constante.

No entanto, ao analisarmos o gráfico da Figura (5.5(b)), cuja curva representa a distribuição de passageiros transportados, vemos que os picos se acentuam mais nos intervalos de horários já observados no gráfico de distribuição de viagens como horários de pico, porém, percebe-se que a demanda não segue uniforme no intervalo entre picos, tendo uma diminuição bastante acentuada no horário entre picos.

Os resultados são esperados, visto que para o DF a demanda diária é concentrada no pico da manhã, definido no intervalo entre 6h às 8h30min e no pico da tarde, definido no intervalo entre 17h às 19h. Fora desses horários, principalmente entre 8h30min a 17h, é considerado horário de entre picos, cuja demanda cai significativamente, ocasionando a paralisação de cerca de 40% da frota operante na área central [72].

Isso nos motiva a investigar melhor as redes de transporte observando os elementos relacionados ao horário e ao peso. Assim, construímos medidas de pendularidade R_{CM} e h_{CM} , cujas equações são 5.8 e 5.10, respectivamente. Para isso, primeiramente construímos a rede espaço- L , onde os nós são os pontos de paradas e a onde a existência de uma rota entre dois nós consecutivos é a aresta. Posteriormente construímos as redes ponderadas L -Trips e L -Pass, cujos pesos das arestas são, respectivamente, número de viagens e número de passageiros.

A primeira aproximação que faremos para o estudo de redes ponderadas será a quantidade fixa de passageiros em uma viagem ao longo do percurso, ou seja, consideraremos que o veículo inicia e termina a viagem com carregamento fixo, não havendo portanto embarque e desembarques. Essa simplificação é feita porque não possuímos dados de sobe-e-desce em cada ponto de parada ao longo da rota. O sistema de bilhetagem automática conta apenas o número de passageiros transportados numa viagem de uma determinada rota, mas não trás nenhuma informação de onde o passageiro embarcou ou onde ele desembarcou.

Calculamos o R_{CM} e o h_{CM} de cada nó. Fizemos alguns exemplos de alguns ponto de parada específicos do DF. Observamos que o horário de maior interesse dos passageiros difere ligeiramente do horário de maior oferta de viagens, Figuras (5.6(c)) e (5.6(a)).

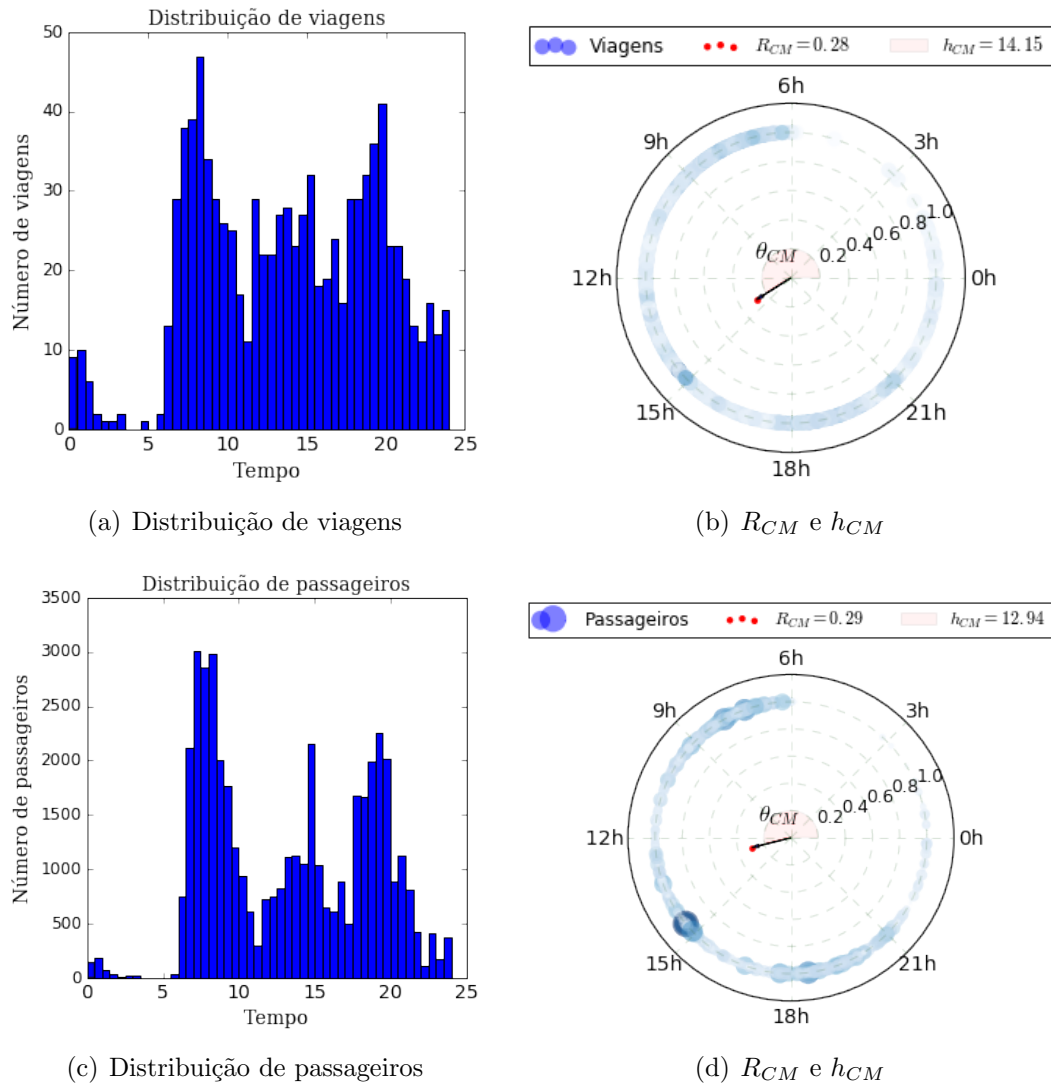


Figura 5.6 - Analisadas 928 viagens e 44,424 passageiros do ponto de parada 72122 - centro de Brasília, próximo ao Banco Central: (a) Distribuição de viagens por intervalos de meia hora; (b) $R_{CM} = 0,28$, $h_{CM} = 14,15$ das viagens; (c) Distribuição de passageiros por intervalo de meia hora; (d) $R_{CM} = 0,29$, $h_{CM} = 12,94$ dos passageiros.

A Figura (5.6(b)) ilustra os valores do $R_{CM} = 0,28$, $h_{CM} = 14,15$ das viagens que passam pelo ponto de parada 72122 - centro de Brasília, próximo ao Banco Central, que é próxima à região de maior concentração de postos de trabalho e atende o fluxo matinal proveniente das outras regiões. Mas observando a Figura (5.6(d)), que ilustra

os valores do $R_{CM} = 0,29$, $h_{CM} = 12,94$ dos passageiros, vemos uma leve diferença nos valores. Comparando as duas distribuições, Figuras (5.6(a)) e (5.6(c)), vemos que há mais flutuação na distribuição de passageiros que na distribuição de viagens, a proporção de passageiros em cada horário é muito mais heterogêneo que a proporção de viagens, isso reflete na medida de R_{CM} . Assim, o R_{CM} de passageiros é maior que o R_{CM} de viagens. Também vemos que ao atribuímos o peso das viagens por passageiros, o horário médio de maior concentração é também deslocado do horário de concentração das viagens.

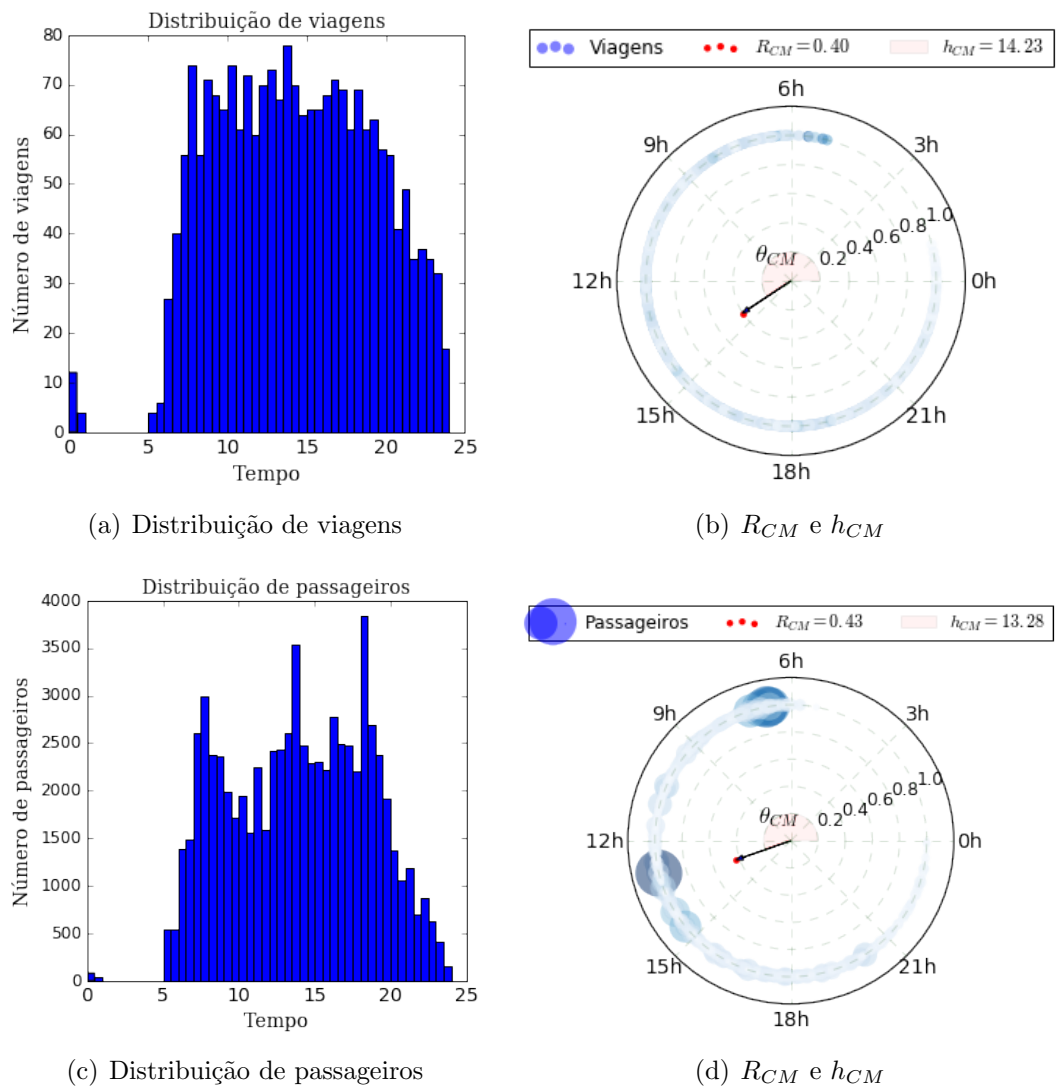


Figura 5.7 - Analisadas 1717 viagens e 78.186 passageiros que passam pelo ponto de parada 72447 - W3 sul, Brasília: (a) Distribuição das viagens a cada intervalo de tempo de meia hora; (b) $R_{CM} = 0,40$, $h_{CM} = 14,23$ das viagens; (c) Distribuição dos passageiros a cada intervalo de tempo de meia hora; (d) $R_{CM} = 0,43$, $h_{CM} = 13,28$ dos passageiros.

A Figura (5.7) mostra o estudo de um ponto de parada 72447 - W3 sul,

Brasília, localizada numa via próxima a escolas, faculdades e comércio, incluindo atividades noturnas como bares e restaurantes e faculdades e cursos noturnos. A distribuição das viagens a cada intervalo de tempo de meia hora parece bastante uniforme entre 8h da manhã até 18h da tarde, mas é intensa a quantidade de viagens nos horários até 23h. Quando observamos a proporção dos passageiros que utilizam ônibus que passam nesse nó, vemos que a demanda não é tão uniforme quanto a oferta de viagens, no entanto, comparando-se com o outro nó (72122), vemos que, tanto o número de viagens quanto o número de passageiros no período noturno não podem ser ignorados. Assim, para o caso de viagens que passam nesse ponto de parada temos o valor de $R_{CM} = 0,40$ e $h_{CM} = 14,23$ e no caso da distribuição de passageiros, temos $R_{CM} = 0,43$ e $h_{CM} = 13,28$.

5.7 Resultados

Fizemos o cálculo do R_{CM} e h_{CM} para todos os nós, que representam pontos de paradas do DF, Figura (5.8). A partir disso, investigamos quais os valores mais expressivos de cada medida. Encontramos desde $R_{CM} = 1$ a R_{CM} muito próximo de zero, mas são em quantidade muito pequena comparada à grande maioria que está entre 0.2 e 0.6 para L -Trip e L -Pass. Assim, para extrair os melhores parâmetros para o mapa de cores da Figura (5.11), fizemos a distribuição de h_{CM} e R_{CM} , calculamos os valores médios e o desvio padrão de cada medida, Figuras (5.9) e (5.10).

A Figura (5.9) mostra os valores de R_{CM} e h_{CM} das redes L -Trip e L -Pass. Mas na Figura (5.10), para cada valor de R_{CM} e h_{CM} de cada ponto de parada, nós multiplicamos pelo valor de M , número de viagens, se L -Trip e número de passageiros, se L -Pass. E fizemos novamente a distribuição, dos valores de R_{CM} e h_{CM} . Dessa forma,

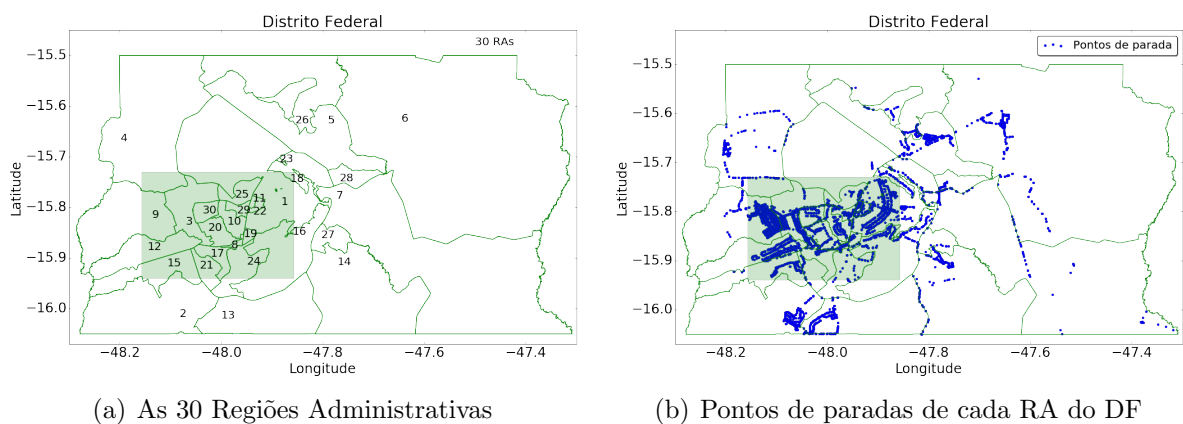


Figura 5.8 - O Distrito Federal é dividida em 30 Regiões Administrativas (RAs). A área sombreada é mostrada nas Figuras (5.11) e (5.14).

obtivemos uma ponderação no histograma, tal que os valores de R_{CM} e h_{CM} mais importantes fossem evidenciados. A importância de ponderar o histograma, inserindo um pesos na distribuição na Figura (5.10) está relacionada com os valores que iremos mostrar nos mapas da Figura (5.11), pois queremos evidenciar o intervalo de R_{CM} mais importante em vez de todo o intervalo $0 < R_{CM} < 1$, uma vez que há quantidades desprezíveis de alguns valores de R_{CM} , por exemplo, são poucos os nós que possuem valor de $R_{CM} = 0.1$. Seguimos o mesmo método para definir o melhor intervalo de h_{CM} , ou seja, em vez de usarmos todo o intervalo de $0 < h_{CM} < 24$, evidenciamos o intervalo de valores de h_{CM} que possuem as maiores barra no histograma.

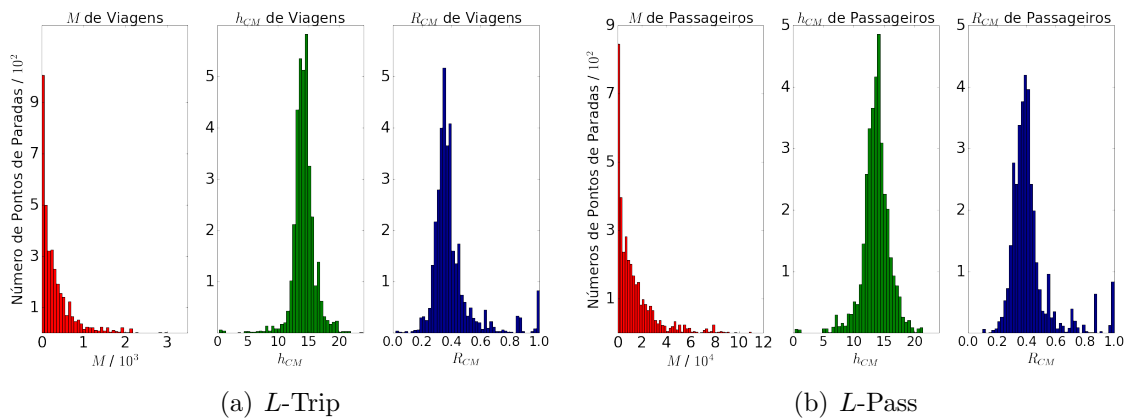


Figura 5.9 - Distribuição de M , R_{CM} e h_{CM} das viagens das redes de nós da L -Trips e L -Pass. Os gráficos 5.9(a) são referentes a rede de viagens e os gráficos de 5.9(b) são referentes a rede de passageiros.

Assim, observamos que alguns valores de R_{CM} e h_{CM} desaparecem quando utilizamos pesos no histograma. Isso sugere que atribuir novas propriedades as redes topológicas PTN seja necessário para melhor analisar as redes de transporte. Ou seja, o número de viagens e o número de passageiros, modificam a topologia da rede PTN.

A partir da distribuição da Figura (5.10) calculamos o valor médio (μ) e o desvio padrão (σ) de R_{CM} e de h_{CM} . E deles definimos o intervalo para os mapas da Figura (5.11). Para o mapa de R_{CM} usamos $[\mu_R - 2\sigma_R, \mu_R + 2\sigma_R]$. Para o mapa de h_{CM} , fizemos uma divisão em dois períodos: matutino, cujo intervalo é $[\mu_h - 2\sigma_h, \mu_h]$; e vespertino, cujo intervalo é $[\mu_h, \mu_h + 2\sigma_h]$.

Os mapas da Figura (5.11) mostram cada ponto de parada (nós da rede espaço- L) geograficamente localizados. Combinamos as três medidas, chamados de R_{CM} , h_{CM} matutino e h_{CM} vespertino para as redes L -Trip e L -Pass. Nos seis mapas, o tamanho

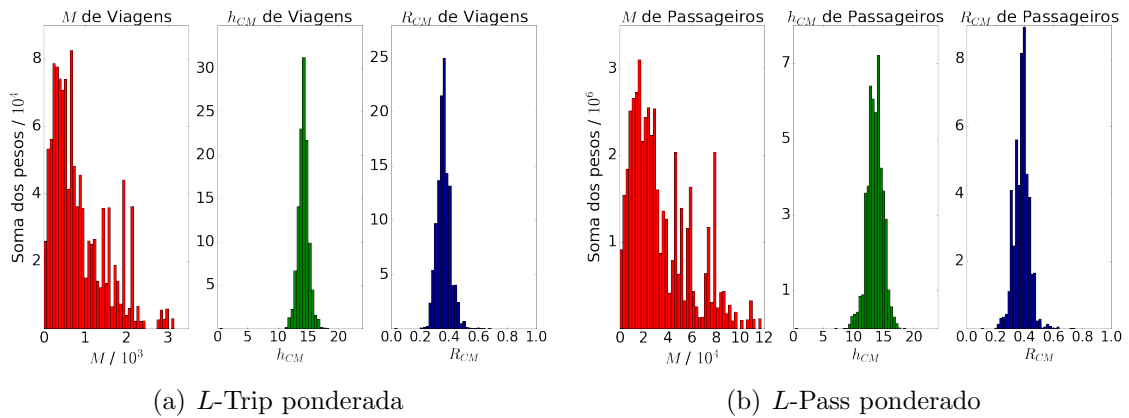


Figura 5.10 - Distribuição de M , R_{CM} e h_{CM} das viagens das redes de nós da L -Trip e da L -Pass ponderados. Os gráficos da Figura 5.10(b) são referentes a rede de viagens ponderada, já os gráficos de 5.10(a) são referentes a rede de passageiros ponderada.

de cada nó se refere ao valor de M , que para o L -Trip equivale ao número de viagens que servem cada nó, e para o L -Pass, equivale ao número de passageiros que passam em cada nó. Significa que quanto maior o nó, mais viagens e passageiros passam pelo nó, respectivamente. A barra de cores se refere ao intervalo de cada medida usando-se os critérios definidos na Seção 5.6.

Um nó pendular é definido como aquele cuja demanda existe para uma direção ao longo do dia. E o não-pendular é definido como o nó demandado para ambas as direções durante o pico matutino e o pico vespertino. Por essa razão, nós não-pendulares possuem h_{CM} próximo ao valor médio μ_h . E pendular, possuem valores mais extremos, mais distantes do valor médio. Essa análise só é possível pois trabalhamos com redes direcionadas, tal que as arestas conectando dois nós estão sempre em um único sentido.

A Figura (5.5) mostra que a demanda de passageiros é maior em dois picos bem distintos, enquanto que a oferta de viagens tem diferenças bem mais suaves. Isso implica na diferenciação de cores dos mapas. Tal que percebemos que os mapas da rede L -Pass possuem nós com h_{CM} mais distantes do valor médio. E portanto, a cor dos nós é mais para o magenta, comparado com o mapa da rede L -Trip.

Os mapas 5.11 e 5.14 se referem à região sombreada dos mapas da Figura (5.8), que mostram as RAs do Distrito Federal. A região foi escolhida como destaque por conter a maior parte da população do DF, englobando a região mais populosa do DF, Ceilândia (9), com mais de 15% dos trabalhadores do DF, seguida por Brasília (1), Taguatinga (3) e Samambaia (12) e quase 50% dos empregos do DF se encontram em Brasília (1), seguido por Taguatinga e Ceilândia. (ver a Figura (6.4)).

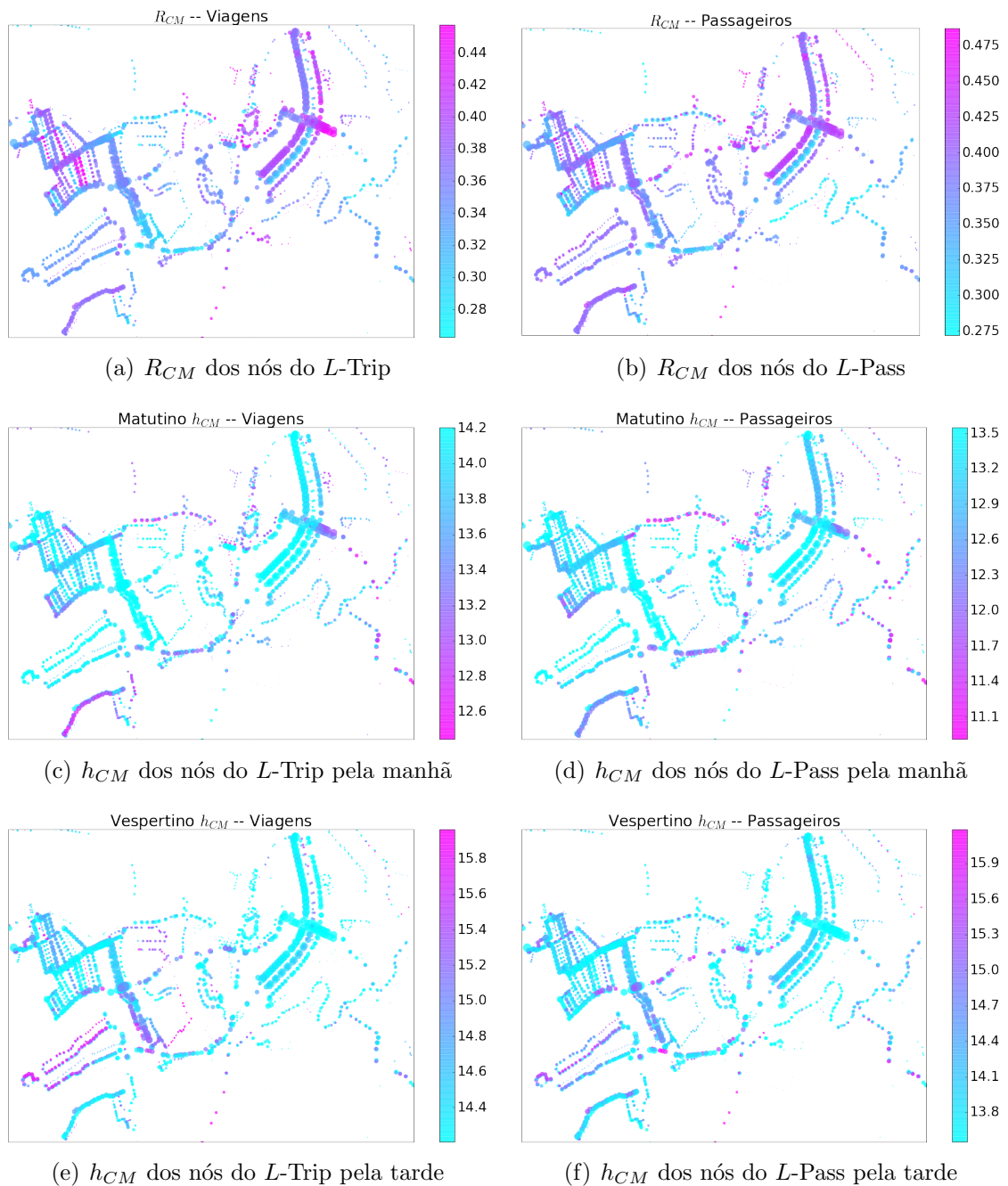


Figura 5.11 - Mapas das propriedades R_{CM} e h_{CM} dos nós das redes L -Trip e L -Pass.

Observamos que o horário de maior concentração difere em cada ponto. Além de serem diferente quanto ao número de viagens que fornecem o serviço em cada ponto, o número de passageiros passando por esses pontos também são distintos. Para melhor análise, vamos trabalhar com as arestas.

A pendularidade também foi calculada para as arestas. Nesse caso o horá-

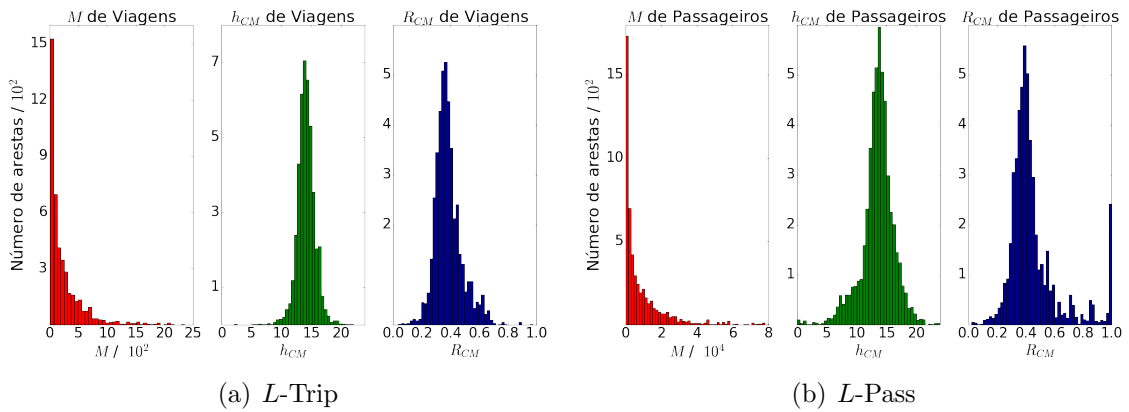


Figura 5.12 - Distribuição de M , R_{CM} e h_{CM} das viagens das redes de arestas da L -Trip e da L -Pass ponderadas. Os gráficos 5.12(a) são referentes a rede de viagens, já os gráficos de 5.12(b) são referentes a rede de passageiros ponderada.

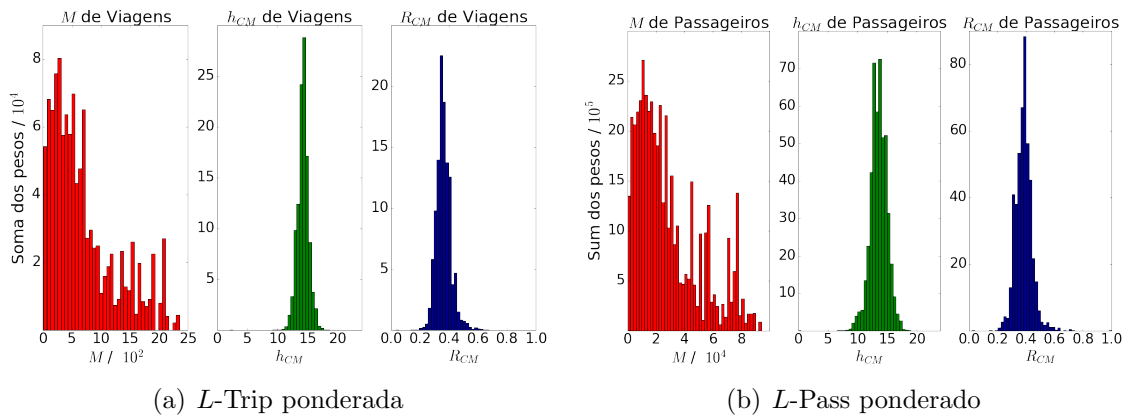


Figura 5.13 - Distribuição de M , R_{CM} e h_{CM} das viagens das redes de arestas da L -Trip e da L -Pass ponderadas. Os gráficos 5.13(a) são referentes a rede de viagens ponderada, já os gráficos de 5.13(b) são referentes a rede de passageiros ponderada.

rio das viagens de cada aresta foi calculado como a média dos horários entre dois nós consecutivos, $h = (h_2 + h_1)/2$. E obtivemos o resultado apresentado na Figura (5.12) e (5.13).

Usando as distribuições das Figuras (5.12) e (5.13), fizemos os mesmos cálculos para os valores médios e desvios padrão definido para a rede de nós.

Além disso, os mapas do L -Trip fornecem informações sobre o tráfego de veículos e os mapas do L -Pass fornecem informações sobre a utilização do serviço de transporte público.

O mapa do R_{CM} de viagens, Figura (5.14(a)), mostra que as regiões de

coloração magenta são aquelas rotas em que há menor Δh , significando uma maior concentração de viagens em um determinado período do dia. Observamos que a região das Esplanadas dos Ministérios, cuja coloração magenta se destaca do lado direito em um formato retangular de Brasília, tem a concentração das viagens maior que a média do DF. Uma explicação para isso é que a região é composta basicamente por Ministérios e outros órgãos públicos, onde o horário de expediente da maioria dos funcionários é de 8h às 18h. Outro local importante se situa nas arestas que passam pela via W3 sul (terceira via do lado esquerdo da figura que forma um pássaro, na parte superior-direita do mapa), vemos que as viagens são mais espalhadas ao longo do dia, tendo ofertas também durante a noite, já que é uma região de escolas e faculdades além de ser uma região com muitos comércios. Outras arestas interessantes estão situadas no eixo-L sul, onde vemos que a pendularidade é mais baixa que a média do DF, é uma região em que existe demanda de serviços em casas, onde o funcionário chega por volta das 7h da manhã, assim como pertence uma região com bastante bares tanto acima da via quanto abaixo dela. Isso leva uma necessidade de transporte nos horários noturnos, isso reflete no valor de R_{CM}

Quando atribuímos peso nas arestas com informações de passageiros, Figura (5.14(c)), vemos que o valor de R_{CM} se modifica para cada aresta, comparado com os valores de contagem de viagens. Isso porque os pesos dão maior relevância para aquelas viagens que carregam mais passageiros do que aquelas que carregam poucos passageiros.

Da mesma forma, comparando os horários de maior concentração das viagens e de passageiros ao longo do dia, percebemos que a atribuição de pesos pode esclarecer melhor como é o mapa das ofertas de viagens e o mapa das demandas de passageiros, Figuras (5.14(e)), (5.14(b)), (5.14(d)) e (5.14(f)).

Percebemos que os padrões de tráfego são distintos para os períodos da manhã e da tarde. O tráfego matutino concentra-se nas horas iniciais da manhã e o tráfego vespertino é mais espalhado ao longo do período, isso se deve pelo fato de existirem atividades após o horário de expediente comercial. Uma dessas atividades está relacionada ao comércio de grandes shoppings e bares e casas noturnas que se encontram em algumas regiões, além das zonas de faculdades e universidades em outras regiões que possuem também horários noturnos.

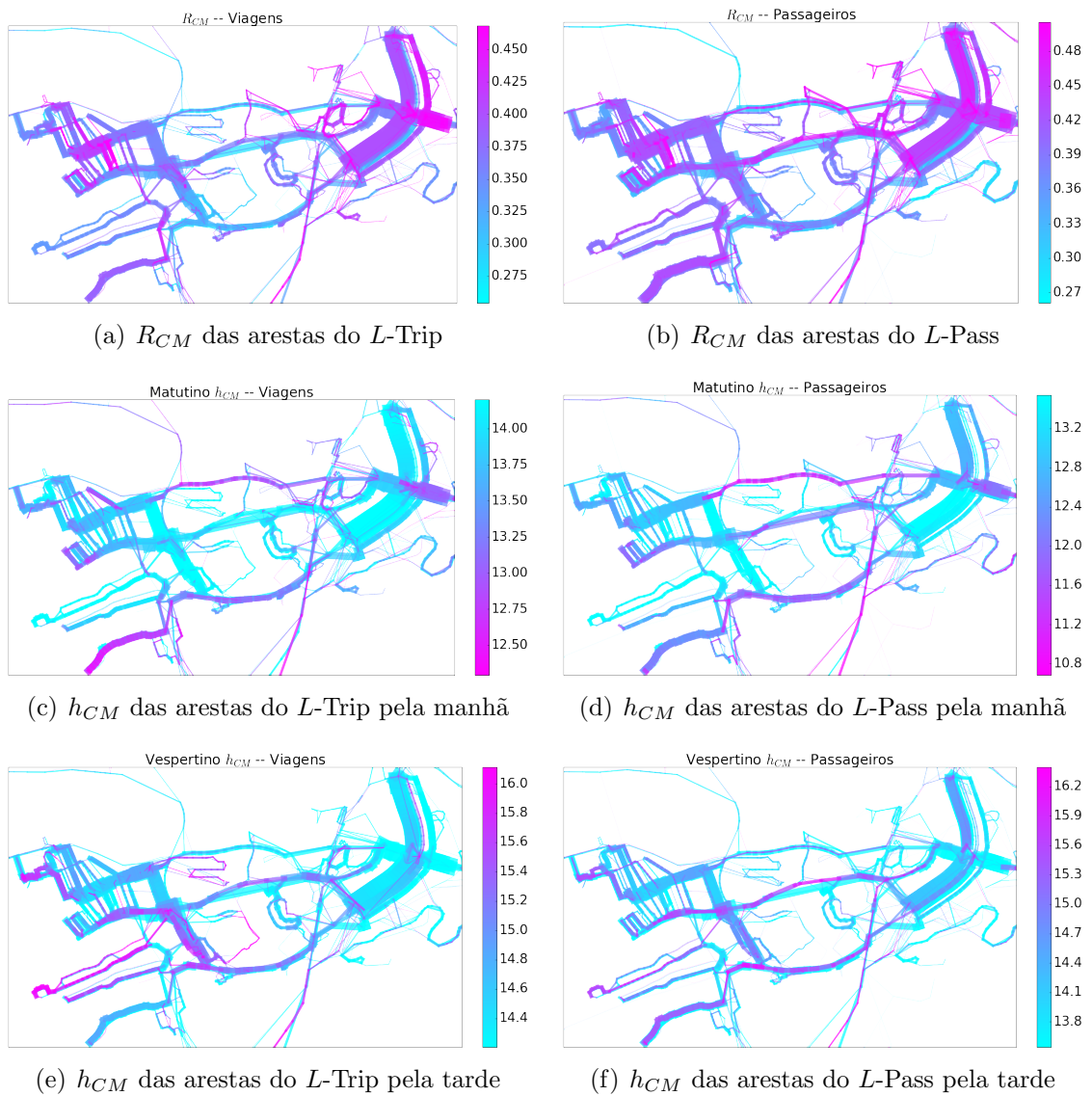


Figura 5.14 - Mapas das propriedades R_{CM} e h_{CM} das arestas das redes *L-Trip* e *L-Pass*.

CAPÍTULO 6

Distribuição de Viagens

Tratamos a estrutura e dinâmica da rede de transporte a partir da teoria de redes complexas. Usamos as propriedades topológicas de redes complexas para propor um modelo de distribuição de viagens que atenda às necessidades dos passageiros de uma cidade. Fizemos um caso especial com os dados de transporte urbano de ônibus do Distrito Federal.

Neste capítulo apresentamos a importância dos modelos de distribuição de viagens para o planejamento de transportes de uma região. Discutimos as características das viagens de ônibus do DF em termos de embarques e desembarques de passageiros em uma viagem. E a partir disso, propomos um modelo de distribuição de viagens de passageiros usando a medida de distância em redes complexas da rede espaço- L .

No tradicional modelo gravitacional a equação de distribuição de viagens é inversamente proporcional à distância, entretanto, sugerimos um modelo gravitacional alterado em que na equação as viagens entre i e j são mais prováveis quanto maior for a distância d_{ij} entre os nós i e j da rede L -Trip.

Essa alteração é motivada pela constatação, a partir de dados empíricos, de que os embarques tendem a se concentrar no início das viagens ao passo que os desembarques apresentam uma tendência de se concentrarem nos últimos pontos de parada. Isso sugere que os passageiros parecem percorrer maior distância e não o contrário.

6.1 Modelo de Planejamento de Transportes

Os objetivos no planejamento de transportes podem ser agrupados em três grandes categorias, relativas aos requerimentos das três partes interessadas do sistema: passageiros, operador e comunidade [77]. E esses objetivos correspondem, respectivamente, a:

- Realizar trabalho máximo de transporte: Expresso por número de passageiro-viagem ou passageiro-km, esta categoria implica em fornecer viagens altamente velozes, conveniência ao passageiro e outros elementos que os atraiam.
- Alcançar um impacto operacional máximo: Este objetivo pode ser expresso ao final por custo total mínimo do sistema para um dado nível de desempenho.
- Criar impactos positivos: Geralmente inclui desde efeitos de curto alcance, tal

como reduzir o congestionamento em rodovias, bem como metas de longo alcance, tal como alcançar um nível mais alto de mobilidade populacional e maior qualidade de vida.

Compreender a distribuição das viagens de uma cidade pode auxiliar muito na otimização dos serviços de transporte oferecidos para sua população, com o intuito de alcançar um impacto operacional máximo. Isso significa que minimizar viagens que não são necessárias pode trazer benefícios para reduzir os custos operacionais.

Nesta seção apresentaremos brevemente como o modelo gravitacional está inserido no contexto do planejamento urbano e no planejamento de transportes. O primeiro são considerados o comportamento e as características populacionais da região, com o objetivo de qualificar e quantificar as viagens realizadas entre as regiões. São considerados fatores que podem gerar a necessidade de ir e vir das pessoas, tais como a localização de um shopping ou oportunidades de empregos. Por outro lado, são consideradas também as características da população, tais como renda e tamanho da família. O segundo considera um método de planejamento do sistema de transporte urbano que atenda à demanda da cidade ou região. Nesse sentido, descrevemos a relevância da utilização do modelo gravitacional e discutiremos sua importância para a geração da própria rede de transporte.

6.1.1 Modelo Comportamental

As decisões de planejamento urbano dependem fortemente do conhecimento da relação entre local de atividade e o comportamento das viagens dos usuários dessas atividades. Para relacionar a taxa de viagens da população que exerce certa atividade em um certo local é necessário uma técnica de análise, que pode levar em conta a distância entre a população usuária e os locais de atividade. Isso tem sido desenvolvido e implementado empiricamente, sendo a técnica mais usada o modelo gravitacional e suas variantes. Estudos empíricos realizados nos anos de 1940 por George Kingsley Zipf e John Q. Stewart mostraram que vários fenômenos de atividades humanas envolvendo interação espacial na escala macroscópica podiam ser modelados usando a Lei de Newton [78],

$$F_{AB} = K \frac{P_A P_B}{d^2}, \quad (6.1)$$

onde P_A é a massa de A , P_B é a massa de B , d é a distância entre A e B e K é uma constante. O modelo basicamente substitui os valores das massas de corpos celestes pela população das cidades A e B , onde d é a distância entre as duas cidades. Uma simples interação de dois corpos com um fator de atratividade. Com esse modelo foi possível prever diversos exemplos de interações humanas sobre o espaço, tal como o número de

peessoas que viajam de ônibus, trem ou avião entre duas cidades, o número de chamadas telefônicas, o volume em toneladas de embarques em trem expresso entre duas cidade e o número de estudantes de cada estado inscritos na Universidade de Princeton [78].

Uma variante do modelo gravitacional puro é o número de viagens entre i e j , que depende do número de moradias (H) de i e de oportunidades disponíveis (S) em j , além do custo da viagem D entre i e j , expresso por,

$$T_{ij} = K \frac{H_i^\alpha S_j^\beta}{D_{ij}^\lambda}, \quad (6.2)$$

onde K é uma constante e α , β e λ são constantes determinadas empiricamente. Além disso, esses parâmetros podem ser estimados estatisticamente. As oportunidades S podem ser uma biblioteca, uma loja, um hospital, ou qualquer outra estrutura que atraia pessoas.

Um caso especial com $\alpha = \beta = 1$, $\lambda = 2$ se reduz ao modelo gravitacional, que é o modelo gravitacional sem vínculos [78].

Embora os estudos de viagens levem em conta as características de atratividade entre as regiões, nem sempre esse fator é suficiente para prever as viagens na vida real. Um exemplo disso é o estudo do consumo de uma população, que nesse caso, depende não apenas do local do mercado ou do shopping, mas também das condições externas, tais como a renda e o tamanho da família.

Isso significa que será necessário inserir um vínculo que combine o número de viagens feitas para cada centro j a partir de uma dada origem i , não importando a atratividade dos centros, mas mantendo constante no nível fixo \bar{T}_i , de modo que a equação de vínculo seja,

$$\bar{T}_i = \sum_j T_{ij}, \quad (6.3)$$

e deva ser inserida no modelo. De forma que,

$$\bar{T}_i = \sum_j K \frac{H_i^\alpha S_j^\beta}{D_{ij}^\lambda}. \quad (6.4)$$

E dessa equação, é definida a constante K ,

$$K = \frac{\bar{T}_i}{H_i^\alpha \sum_j \frac{S_j^\beta}{D_{ij}^\lambda}}. \quad (6.5)$$

Substituindo na equação (6.2) e desde que possamos cancelar H_i^α , temos,

$$T_{ij} = \bar{T}_i \left(\frac{S_j^\beta}{D_{ij}^\lambda} / \sum_j \frac{S_j^\beta}{D_{ij}^\lambda} \right), \quad (6.6)$$

que é a equação do modelo gravitacional com competição de Wilson, de 1967 [78], o qual veremos com mais detalhes depois. Essa equação pode ser reescrita como,

$$\frac{T_{ij}}{\bar{T}_i} = \frac{S_j^\beta}{D_{ij}^\lambda} / \sum_j \frac{S_j^\beta}{D_{ij}^\lambda}, \quad (6.7)$$

desde que os termos de cada lado da igualdade estejam entre zero a um, e desde que a soma dos termos sobre todos os centros j 's seja igual a um, eles são ambos probabilidades. Ou seja, se T_{ij} e \bar{T}_i já tenha ocorrido e sendo observações em vez de previsões, então o termo à esquerda da igualdade é uma probabilidade empírica. Em todo caso, o termo pode ser escrito como uma probabilidade p_{ij} , que é a probabilidade que uma viagem seja feita para o j desde que o valor das viagens realizadas de i seja dada. Assim,

$$p_{ij} = \frac{S_j^\beta}{D_{ij}^\lambda} / \sum_j \frac{S_j^\beta}{D_{ij}^\lambda}, \quad (6.8)$$

tal que a equação (6.6) possa ser escrita como,

$$T_{ij} = p_{ij} \bar{T}_i. \quad (6.9)$$

Assim, o vínculo ao modelo gravitacional, equação (6.9), prediz o número de viagens de i a j pela alocação total de um dado centro i , \bar{T}_i , para cada centro j . Por outro lado, vemos pela equação (6.8) que a probabilidade depende das oportunidades oferecidos em um dado centro (S_j), distante da residência i ao centro j por D_{ij} , e a atração realizada por todos os centros concorrentes sobre os residentes em i , $\sum_j S_j^\beta / D_{ij}^\lambda$ [78].

Para calibrar o modelo gravitacional com vínculos a partir da estimação empírica, podemos cancelar o H_i , como foi feito na equação (6.6), podemos escrever a equação (6.2), como,

$$T_{ij} = K_i \frac{S_j^{\beta_i}}{D_{ij}^{\lambda_i}} \quad (6.10)$$

$$= K_i S_j^{\beta_i} D_{ij}^{-\lambda_i}. \quad (6.11)$$

onde a contante K_i e os expoentes β_i e λ_i são específicos de cada local de residência i . Vemos que a partir do dados empíricos esses três parâmetros para um certo local de residências podem ter a forma,

$$\log T_{ij} = \log(K_i) + \beta_i \log(S_j) - \lambda_i \log(D_{ij}). \quad (6.12)$$

Que pode ser obtido ajustando-se os parâmetros por,

$$\log T_{ij} = a + b_1 \log(S_j) + b_2 \log(D_{ij}). \quad (6.13)$$

com K_i estimado por,

$$K_i = \frac{\sum_j \hat{T}_{ij}}{\sum_j \hat{S}_j^{\beta_i} \hat{D}_{ij}^{-\lambda_i}}. \quad (6.14)$$

Onde os símbolos com $(\hat{\cdot})$ são os valores observados. Outra forma de fazer a distribuição de viagens um pouco mais atenuante quanto ao fator de distância é escrever,

$$T_{ij} = K S_j^\beta e^{(-\lambda D_{ij})}, \quad (6.15)$$

tal que K_i é estimado como,

$$K_i = \frac{T_i}{\sum_j S_j^\beta e^{(-\lambda D_{ij})}} \quad (6.16)$$

Portanto, empiricamente podemos as viagens T_{ij} de uma região podem ser estimadas pela equação (6.15), quando são fornecidos os dados característicos de cada região e a distância D_{ij} .

6.1.2 Modelo 4 Etapas

O modelo de planejamento de transporte tradicional é o chamado Modelo 4 Etapas: Geração de viagens, Distribuição de viagens, Divisão/Repartição Modal e Alocação de viagens [78, 79].

- Geração de viagens: A primeira etapa gera, a partir de dados censitários, a previsão de demanda de viagens para a região. Nessa etapa são usados os dados referentes à residência, ao local de trabalho, ao local de estudo, e a outros que podem motivar o movimento de pessoas em um certo território.
- Distribuição de viagens: Na segunda etapa é gerada a matriz de origem-destino,

cujos dados são relacionados ao ir e vir da população entre zonas ou áreas bem delimitadas dentro de um certo território. Aqui são usadas alguma propriedade da região como atratores de viagens.

- Divisão Modal: Uma vez conhecido a demanda de viagens, decide-se qual a melhor opção de modal de transporte para atender a demanda de fluxo. Nessa etapa são decididos qual meio de transporte utilizados.
- Alocação de viagens: A última etapa serve para equilibrar e ajustar a necessidade de viagens com o custo do serviço de forma a alocar os veículos em cada rota para servir as viagens.

O modelo gravitacional é usado para gerar a matriz Origem-Destino, ou seja, esse modelo está inserido na etapa de distribuição de viagens no processo de planejamento de transportes.

6.2 Modelo Gravitacional

Ao Planejar um sistema de transporte é preciso conhecer a demanda, quantas viagens diárias são necessárias, em quais locais e em quais horários devem passar o ônibus. Atualmente muitos planejadores fazem uso do modelo Gravitacional de Transportes e seus derivados para estimar as viagens de transporte para uma cidade ou região incluindo-se dados demográficos de cada região e algum dado de interação entre as regiões, tais como distância, tempo de viagem, ou algum tipo de impedância. Inclusive, são usados para estimar a demanda de transporte em tempos futuros, desde que se tenha algum tipo de dado estimado das regiões, tais como o crescimento populacional ou construção de algum shopping.

6.2.1 Modelo de Wilson

Nesse modelo é expressa a função de distribuição de viagens entre duas zonas, i e j , como uma equação diretamente proporcional ao número de viagens que se originam em i e ao número de viagens que se destinam a j e inversamente proporcional a distância entre i e j , ou seja, análogo ao Modelo Gravitacional de Newton para a atração entre dois corpos massivos distantes um do outro por uma medida de distância d_{ij} tal como,

$$f_{ij} = \gamma \frac{m_i m_j}{d_{ij}^2}, \quad (6.17)$$

onde γ é a constante gravitacional. Assim, para a distribuição de viagens, Wilson propõe a equação

$$T_{ij} = k \frac{O_i D_j}{d_{ij}^2}, \quad (6.18)$$

onde k é uma constante. T_{ij} é o número de viagens entre i e j , O_i é o número de viagens que se originam em i , D_j , o número de viagens que se destinam a j e d_{ij} , a distância entre i e j . Ou seja, uma analogia a um sistema de dois corpos (duas regiões) que possui um fator de atração. Assim, é necessário que os vínculos,

$$\sum_j T_{ij} = O_i \quad (6.19)$$

e

$$\sum_i T_{ij} = D_j, \quad (6.20)$$

estejam definidos para que a expressão (6.18) seja válida. Sendo T_{ij} uma matriz, O_i é soma das linhas e D_j , a soma das colunas dessa matriz.

Essas equações de vínculo podem satisfazer um conjunto de constantes A_i e B_j associados às zonas de produção e de atração, respectivamente. Além disso, ainda segundo o modelo de Wilson, T_{ij} depende de uma função de distância. E a equação (6.18) pode ser modificada para

$$T_{ij} = A_i B_j O_i D_j f(d_{ij}), \quad (6.21)$$

onde

$$A_i = \left[\sum_j B_j D_j f(d_{ij}) \right]^{-1} \quad (6.22)$$

e

$$B_j = \left[\sum_i A_i O_i f(d_{ij}) \right]^{-1}. \quad (6.23)$$

As equações de vínculo podem ser satisfeitas se o conjunto de constantes A_i e B_j associado com a zona de produção de viagens e zona de atração de viagens, respectivamente, forem introduzidas. A_i e B_j são muitas vezes chamados de fatores de balanço. Desde que as equações de vínculo (6.19) e (6.20) sejam satisfeitas, as equações (6.21-6.23) descrevem um modelo gravitacional análogo ao modelo de Newton. Nesse caso, d_{ij} é uma medida de impedância generalizada, que pode ser medida de distância entre i e j , medida de tempo, medida de custo, ou alguma medida que atribua alguma combinação de pesos que resulte em um fator de custo generalizado [80].

Para se empregar o método acima, é necessário definir um conjunto de

variáveis que especificam o sistema e enumerar alguns vínculos dessas variáveis. Wilson supõe que a probabilidade de ocorrer uma distribuição $\{T_{ij}\}$ é proporcional ao número de estados do sistema que dão origem à distribuição $\{T_{ij}\}$. E que, se $w(T_{ij})$ é o número de maneiras pelas quais os indivíduos podem se organizar para produzir a distribuição total $\{T_{ij}\}$, então a probabilidade de ocorrer $\{T_{ij}\}$ é proporcional a $w(T_{ij})$. Isso assumindo que um estado é formado por microestados, estes últimos definidos como formas individuais de viagens, por exemplo, a viagem de cada trabalhador de sua residência ao trabalho. Portanto, o número total de tais combinações é,

$$\sum w(T_{ij}), \quad (6.24)$$

onde a soma é sobre toda a distribuição que satisfaça os vínculos do problema [80].

Para estudos de viagens de origem-destino dos trabalhadores da residência ao trabalho, Wilson utiliza o modelo gravitacional convencional, cuja matriz de viagens, \mathbf{T} , deve satisfazer aos dois vínculos expressos pelas equações (6.19) e (6.20) e a equação de vínculo adicional,

$$\sum_i \sum_j T_{ij} c_{ij} = C, \quad (6.25)$$

onde c_{ij} é uma impedância, ou custo generalizado, de viajar entre i e j , que substitui d_{ij} da equação original, portanto, a medida de impedância não precisa ser exatamente a distância. Esse vínculo implica que o total gasto nas viagens entre as regiões, num determinado tempo, é fixado em uma quantidade C .

Como a suposição básica do método é que a distribuição $\{T_{ij}\}$ seja proporcional ao número de estados do sistema que dá origem a essa distribuição, e que satisfaça os vínculos, supões-se, então, que

$$T = \sum_i O_i = \sum_j D_j \quad (6.26)$$

é o número total de viagens, e o número de arranjos individuais distintos que dão origem à distribuição T_{ij} é

$$w(T_{ij}) = \frac{T!}{\prod_{ij} T_{ij}!}, \quad (6.27)$$

desde que não exista nenhum arranjo de interesse particular de viagens. Ou seja, para um total de T viagens, queremos escolher T_{11} de T possíveis viagens, e T_{12} , não mais de T ,

mas de $(T - T_{11})$ possibilidades, e assim por diante, tal que,

$$\begin{aligned}
 w(T_{ij}) &= \frac{T!}{T_{11}!(T - T_{11})!} \frac{(T - T_{11})!}{T_{12}!(T - T_{11} - T_{12})!} \cdots \frac{(T - T_{11} - \dots - T_{n(n-1)})!}{T_{nn}!(T - T_{11} - \dots - T_{nn})!} \\
 &= \frac{T!}{T_{11}!T_{12}!\dots T_{nn}!} \\
 &= \frac{T!}{\prod_{ij} T_{ij}!}.
 \end{aligned} \tag{6.28}$$

O número total de estados possíveis é portanto,

$$W = \sum w(T_{ij}), \tag{6.29}$$

onde a soma sobre T_{ij} satisfaz as equações de vínculo (6.19) e (6.20). Contudo, o valor máximo de $w(T_{ij})$ domina os outros termos da soma tal que a extensão da distribuição T_{ij} , que dá origem esse máximo é a distribuição mais provável [80]. Para obter o conjunto de T_{ij} 's que maximizem $w(T_{ij})$ como definido em (6.27) satisfazendo os vínculos (6.19), (6.20) e (6.25), é preciso encontrar o máximo da função \mathcal{M} ,

$$\mathcal{M} = \log w + \sum_i \lambda_i^{(1)} \left(O_i - \sum_j T_{ij} \right) + \sum_j \lambda_j^{(2)} \left(D_j - \sum_i T_{ij} \right) + \beta \left(C - \sum_i \sum_j T_{ij} c_{ij} \right), \tag{6.30}$$

sendo $\lambda_i^{(1)}, \lambda_j^{(2)}$ e β os multiplicadores de Lagrange. Então, maximizando $\log w$, em vez de w , é possível usar a aproximação de Stirling,

$$\log N! = N \log N - N, \tag{6.31}$$

para estimar os termos fatoriais. O T_{ij} que maximiza \mathcal{M} , e que portanto, constitui a distribuição de viagens mais provável, é a solução de

$$\frac{\partial \mathcal{M}}{\partial T_{ij}} = 0 \tag{6.32}$$

e as equações de vínculo (6.19), (6.20) e (6.25). Usando a aproximação de Stirling, equação (6.31), nota-se que

$$\frac{\partial \log N!}{\partial N} = \log N \tag{6.33}$$

e então

$$\frac{\partial \mathcal{M}}{\partial T_{ij}} = -\log T_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}, \tag{6.34}$$

que se anula quando

$$T_{ij} = e^{(-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij})}. \quad (6.35)$$

Então, substituindo em (6.19) e (6.20) para obter $\lambda_i^{(1)}$ e $\lambda_j^{(2)}$:

$$e^{[-\lambda_i^{(1)}]} = O_i / \left[\sum_j e^{(-\lambda_j^{(2)} - \beta c_{ij})} \right] \quad (6.36)$$

$$e^{[-\lambda_j^{(2)}]} = D_j / \left[\sum_i e^{(-\lambda_i^{(1)} - \beta c_{ij})} \right]. \quad (6.37)$$

Escrevendo

$$A_i = e^{(-\lambda_i^{(1)})} / O_i \quad (6.38)$$

e

$$B_j = e^{(-\lambda_j^{(2)})} / D_j \quad (6.39)$$

e então

$$T_{ij} = A_i B_j O_i D_j e^{(-\beta c_{ij})}, \quad (6.40)$$

onde, usando as equações (6.36) e (6.37),

$$A_i = \left[\sum_j B_j D_j e^{(-\beta c_{ij})} \right]^{-1} \quad (6.41)$$

e

$$B_j = \left[\sum_i A_i O_i e^{(-\beta c_{ij})} \right]^{-1}. \quad (6.42)$$

Assim, a distribuição de viagens mais provável é o mesmo que a distribuição do modelo gravitacional, definido nas equações (6.21-6.23). A estatística derivada constitui uma nova teoria do modelo gravitacional. C , vínculo de custo da equação (6.25) não precisa ser conhecido, como esta equação não é resolvida na prática para β . Este parâmetro será encontrado por um método de normalização. Contudo, se C é conhecido, a equação (6.25) pode ser resolvido numericamente para β .

Assim, para calcular a distribuição mais provável, O_i , D_j e c_{ij} são definidos previamente. A expressão $\exp(-\beta c_{ij})$ é interpretado como a forma preferencial da função de dissuasão à distância e o parâmetro β é determinado na teoria pela equação de vínculo do custo, mas ainda é comum que se interprete como um tipo de distância média percorrida, em que maior o β , menor é a distância média percorrida.

Portanto, pela teoria estatística de distribuição de viagens [80] é mostrado que, dado um número total de viagens de origem e destino para cada zona para uma categoria proposta como viagem-pessoa homogêneo, dado o custo das viagens entre cada zona, e dado que existe um número fixo total gasto no transporte da região, então existe uma distribuição mais provável entre as viagens entre as zonas, e a distribuição é a mesma que uma normalmente descrita como a distribuição do modelo gravitacional [80].

6.2.2 Modelo Gravitacional com termo de competição simples

Uma variação simples para o modelo gravitacional convencional que pode ser feito. Podemos escolher todos os B_j s iguais a 1, de modo que não haja um vínculo no destino, equação (6.20). Então este modelo é

$$T_{ij} = A_i O_i D_j f(c_{ij}), \quad (6.43)$$

onde

$$A_i = \left[\sum_j D_j f(c_{ij}) \right]^{-1}, \quad (6.44)$$

para assegurar que,

$$\sum_j T_{ij} = O_i, \quad (6.45)$$

mas $D_j^* = \sum T_{ij}$, onde D_j^* não é necessariamente igual a D^j . Nesse caso especial, quando D_j é alguma medida de atratividade, D_j^* é definido como o resultado do número de viagens produzido no modelo. E seguindo a mesma técnica de maximização, pode se chegar a expressão,

$$T_{ij} = \exp(-\lambda_i^{(1)} - \beta c_{ij}), \quad (6.46)$$

onde o $\lambda_i^{(1)}$ e β são os multiplicadores de Lagrange. Podemos encontrar o valor de $\lambda_i^{(1)}$ substituindo a equação (6.46) na equação (6.45) diretamente, resultando em,

$$T_{ij} = \frac{O_i \exp(-\beta c_{ij})}{\sum_j \exp(-\beta c_{ij})}. \quad (6.47)$$

Se D_j for uma tipo de medida de atratividade, e não o número de viagens que terminam em j , e assumindo que o viajante recebe algum tipo de benefício W_j para ir para uma região que outra, então a equação (6.47) pode ser reescrito como,

$$T_{ij} = \frac{O_i \exp(\beta W_j - \beta c_{ij})}{\sum_j \exp(\beta W_j - \beta c_{ij})}. \quad (6.48)$$

que pode ser comparado às equações (6.43) e (6.44), tomando $\exp(\beta W_j)$ como uma medida de atratividade para D_j .

Assim, a equação (6.48) é a equação do modelo gravitacional com competição.

6.3 Redes Multiarestas

O conceito de redes multiarestas é dado por Sagarra [73] em cujo trabalho desenvolve a mecânica estatísticas de redes multiarestas. Ele aponta que a distinguibilidade de unidades de peso são cruciais em redes de transporte, nas quais humanos fluem entre duas localidades. Em planejamento urbano, se constrói a matriz origem-destino com base em informações diversas relacionadas ao uso e ocupação do solo e ao comportamento de viagens para se fazer uma previsão de transporte urbano, ou ainda, usam o modelo 4 etapas para isso. Em todo caso, o método mais empregado é ainda o modelo gravitacional modificado com os fatores de calibre em conformidade com as propriedades urbanas incluídas na previsão [78].

Essa abordagem é baseada na mecânica estatística para redes complexas geradas por eventos simples distinguíveis que podem ser agrupados em arestas em analogia à alocação de partículas em níveis de energia na mecânica estatística clássica. Nessas tais redes, o peso das arestas codifica informações acerca de processos independentes por diferentes agentes em diferentes tempos que são representados por um número inteiro e quantizado. Ou seja, formado por processos distinguíveis.

6.4 Dados do Transporte Urbano do Distrito Federal

Utilizamos as mesmas fontes de dados citados no Capítulo Pendularidade em PTN, Seção 5.5, para a construção do modelo de distribuição de passageiros na rede de transporte urbano do DF.

À Secretaria de Estado de Mobilidade Urbana do Distrito Federal - SEMOB-DF solicitamos os dados das delimitações de zonas censitárias utilizadas para a PDAD-2011 e para o PDTU-2009. O Distrito Federal e Entorno são divididos em 419 zonas, das quais 381 são dentro do DF. Utilizamos apenas as zonas do DF neste trabalho. Os dados de zonas são também geo-referenciadas, descrevendo cada área como um polígono. Esses polígonos foram importantes para definirmos os pontos de paradas pertencentes a cada área e, portanto, para definirmos as rotas que atendem cada área. Posteriormente separamos as 381 zonas em 30 RAs, tendo em vista que os dados demográficos adquiridos para a pesquisa são por RAs.

O primeiro desafio em trabalhar com dados reais, foi a grande dificuldade na integração dos dados, os quais foram obtidos de várias fontes e por isso não eram interoperáveis. Precisamos desenvolver algoritmos próprios para a integração deles, pois nem todos eram geo-referenciados. Os dados operacionais fornecidos pelo DFTRANS informavam o número de passageiros por viagem, cujo código identificador era completamente diferente do código dos dados GTFS dos dados de rotas, viagens, pontos de paradas.

6.5 Características do Transporte Urbano do Distrito Federal

Os estudos empíricos de carregamento de algumas rotas de ônibus do Distrito Federal mostram o quantitativo de passageiros que embarcam e desembarcam em cada ponto de paradas ao decorrer de uma viagem [81]. O relatório [81] se refere a uma pesquisa em campo de sobe-e-desce, ou seja, de embarque e desembarques, de algumas rotas específicas. A pesquisa amostral foi realizada registrando-se os dados do quantitativo de embarques e de desembarques em cada ponto de paradas. Algumas dificuldades encontradas aqui se referem à não padronização da coleta. Percebemos que alguns colaboradores daquela pesquisa informaram apenas o embarque e desembarque nos pontos de paradas/estações em que o ônibus parava, ignorando os pontos de paradas em que não havia solicitação de passageiros. Outros tiveram o cuidado de registrar o ponto de parada mesmo não tendo solicitação. Utilizamos estes últimos para estimar o perfil de embarques e desembarques por distância da origem em número de pontos de paradas.

Assim, a Figura (6.1) mostra o número de passageiros embarcado em cada ponto de parada dividido pelo número total de passageiros transportado na viagem. Essa figura apresenta o perfil da lotação do veículo ao longo da viagem. Traçamos a média (linha tracejada) de passageiros em cada ponto de paradas e a partir dela, ajustamos ($f(x)$) a curva (linha contínua).

Ajustamos o valor médio do número de passageiros embarcados ponto a ponto por $f(x)$, cuja curva é uma gaussiana. O que percebemos é que em boa parte da viagem o veículo permanece bem cheio, com mais de 60% do carregamento da viagem. Quanto mais próximo do zero é o pico, mais rápido satura a viagem, sugerindo que os passageiros embarcam nos primeiros pontos de parada da rota.

As amostras de embarque e desembarque sugerem que o veículo geralmente tem muitos embarques no início da viagem e quase não há desembarques nesses pontos. E os desembarques costumam ocorrer no fim da rota. Isso significa que a maior parte dos passageiros embarca no início da rota e vai até o fim da rota, tendo poucos embarques e desembarques no meio do caminho.

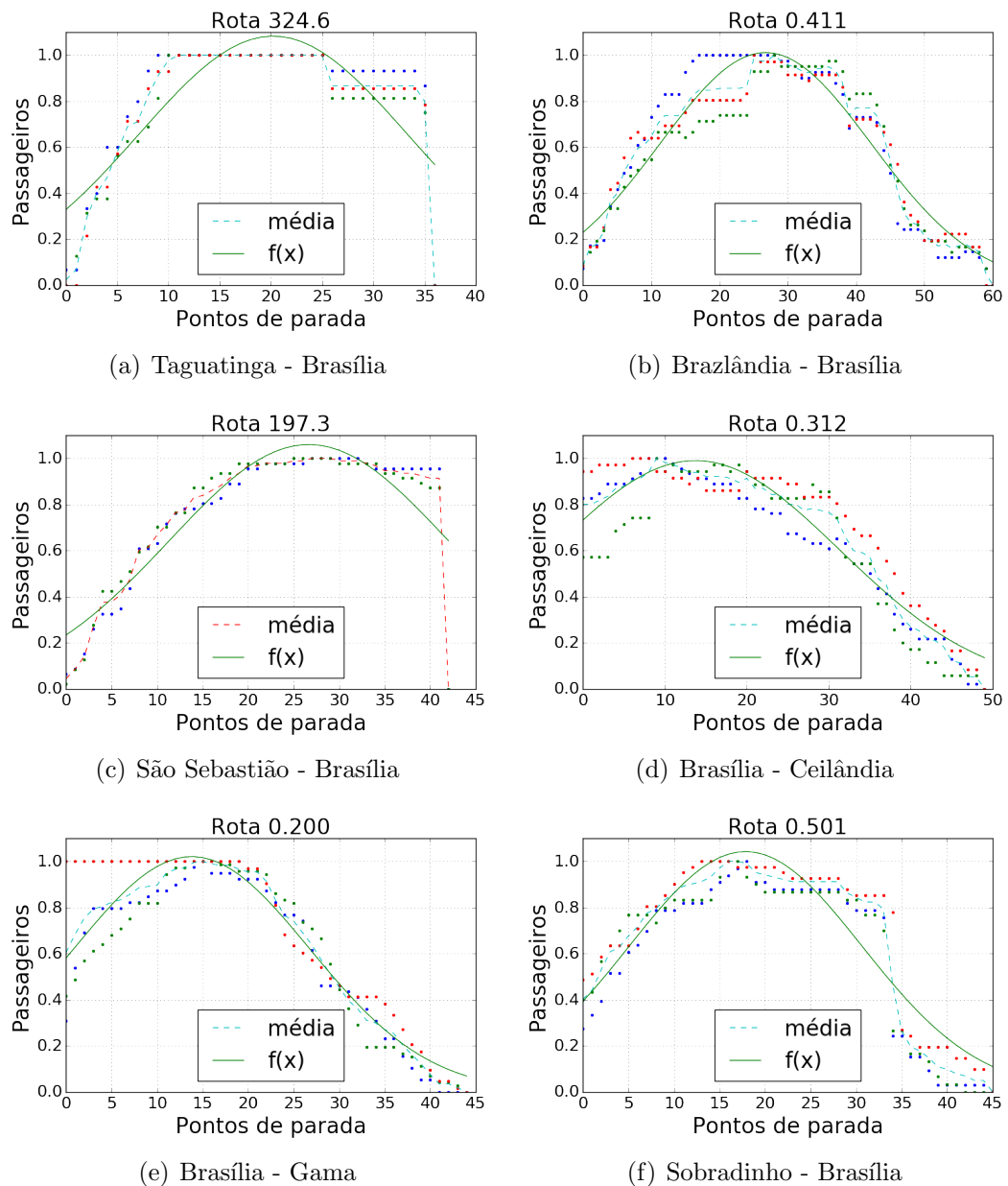


Figura 6.1 - Amostra do carregamento de viagens em rotas ônibus do DF. Esta figura apresenta o perfil da lotação do veículo ao longo da viagem. Traçamos a média de passageiros em cada ponto de paradas (linha tracejada) e a partir dela, ajustamos a curva (linha contínua). Dados retirados do relatório [81].

Esse comportamento pode ser visto na Figura (6.2), onde o número de entradas, X_i^E , é grande no início e decresce ao final, e no número de

A curva X_i^E mostra como o número de entradas acumuladas de passageiros ao passar pelo ponto de parada i , dividido pelo número total transportado. Ao contrário do que é observado na curva X_i^D , definido como número de passageiros total transportado

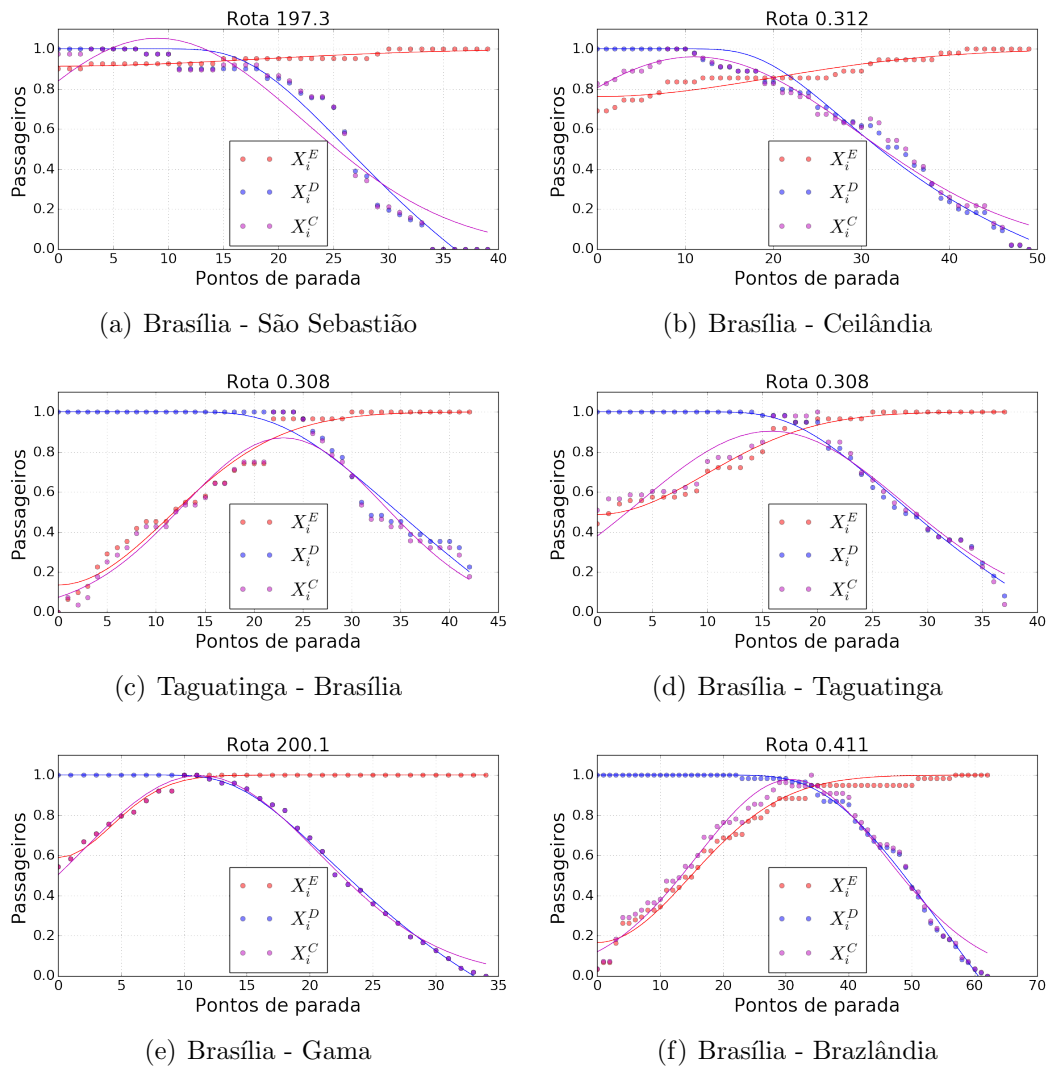


Figura 6.2 - Amostra do Embarque e Desembarques de viagens em rotas de ônibus do DF. O eixo-x representa os sucessivos pontos de paradas de cada rota. Para cada rota são mostradas as entradas de passageiros X_i^E , definido como número total de passageiros que entraram até o ponto de paradas i , dividido pelo número total transportado e das decidas de passageiros X_i^D , definido como número de passageiros total transportado menos o número de passageiros que desceram até o ponto i dividido pelo número total transportado e a lotação de passageiros dividido pelo número total transportado na viagem X_i^C .

menos o número de passageiros que desceram até o ponto i dividido pelo número total transportado. Já a lotação de passageiros dividida pelo número total transportado na viagem é representado por X_i^C . Quanto mais contante é o valor de X_i^C , mais homogênea seria os embarque e desembarque de uma rota.

Fizemos o ajuste das curvas da distribuição de embarque (X_i^E) e desembarque (X_i^D) da Figura (6.2) por curvas exponenciais. Tal que $X_i^E = a \exp(-b^2 d_i^2) + 1$ e $X_i^D = a \exp(-b^2/d_i^2) + 1$. Onde d_i é a distância em relação a origem da viagem da rota.

A Figura (6.2) juntamente com a Figura (6.1) sugerem que as viagens de passageiros entre os primeiros pontos de paradas de uma rota é nula ou muito improvável, pois não há desembarques, apenas embarques. Da mesma forma, sugerem que para pontos mais próximos do final da rota, também a probabilidade de viagens entre esses pontos é nula ou muito improvável, pois não temos embarques. Isso pode mostrar que para d_{ij} pequeno as chances de uma viagem acontecer é pouco provável, mas para d_{ij} grande, parece ser mais provável. A partir dessas observações podemos implementar um modelo de distribuição de passageiros usando a rede L -Trip de PTNs, cujos pontos de paradas são nós sequenciados e as arestas presentes apenas em nós consecutivos da rota.

6.6 Modelo de distribuição de viagens para o transporte do DF

Nosso trabalho propõe um modelo de distribuição de viagens, tal que o coeficiente de atratividade seja uma propriedade de redes complexas que possa expressar o movimento diário dos trabalhadores para cada região do DF como passageiros de transporte público coletivo do DF.

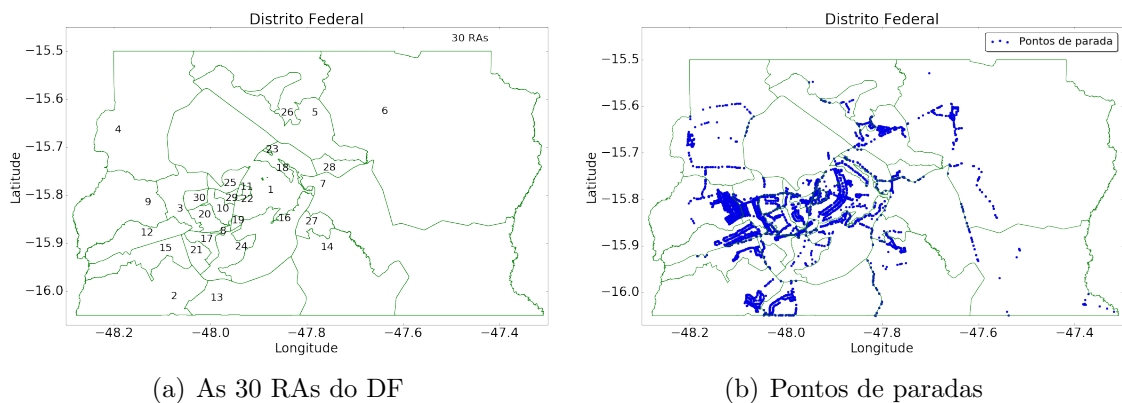


Figura 6.3 - Mapas: 6.3(a) Divisão do Distrito Federal em 30 Regiões Administrativas e 6.3(b) os 3673 pontos de paradas de ônibus.

Distrito Federal (DF) é formado por 30 Regiões Administrativas (RAs), Figura (6.3), com adensamento populacional não-homogêneo e com postos de trabalho bastante centralizados em algumas poucas zonas, em 2011 sua população foi estimada em 2.556.149 habitantes, dos quais 1.078.261 possuem alguma ocupação laboral, segundo os dados censitários da Pesquisa Distrital de Amostra Domiciliar 2011 – PDAD-2011 [76].

O fluxo de migrações diárias por motivo de trabalho concentra-se muito para a RA de Brasília. Somente em Brasília são 477.125 postos de trabalho ocupados,

seguido em segundo lugar por Taguatinga, com apenas 89.585, Figura (6.4). No entanto, 87.736 dos postos de trabalho de Brasília são ocupados por trabalhadores que residem na própria RA, o restante das vagas é preenchido por habitantes de outras RAs do DF.

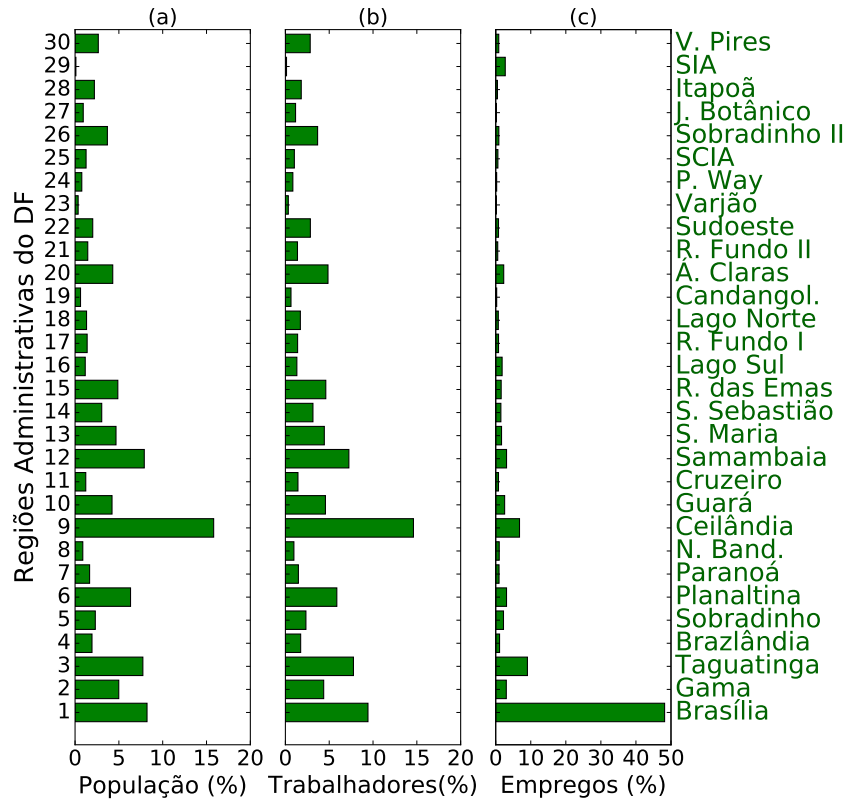


Figura 6.4 - Distribuição Populacional do DF: (a) fornece dados percentuais da população do DF por RA; (b) fornece dados percentuais dos trabalhadores de cada RA e (c) fornece dados dos empregos (postos de trabalho ocupados) de cada RA. Dados retirados da relatório [76].

Os dados de mobilidade diária das pessoas de uma certa região por motivo de trabalho são elementos observados para se definir as necessidades de transporte urbano dessa região. Os dados de moradia e local de trabalho são importantes para se verificar a necessidade de deslocamento diário.

A partir desses dados, é possível construir a Matriz de Origem-Destino (M) e definir de onde e para onde vão as pessoas. Na Figura (6.5) é representada a Matriz OD do Distrito Federal dos trabalhadores da região. O eixo das abscissas representa a RA de origem e o eixo das ordenadas representa a RA de destino. Cada ponto é um elemento da Matriz (m_{ij}), sua coloração está relacionada com a quantidade de pessoas que vão de i a j . Fizemos a barra de cores em escala logarítmica para melhorar a nitidez. Percebemos

que grande parte da população se destina para a RA 1, Brasília.

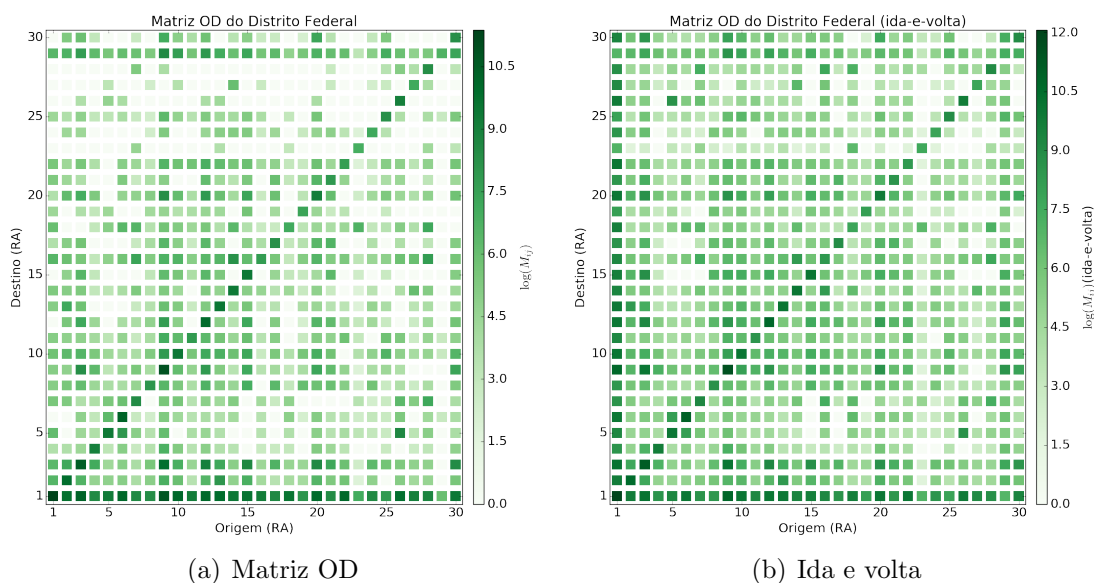


Figura 6.5 - Movimentos diários por motivo de trabalho entre as 30 Regiões Administrativas do DF. A Figura (6.5(b)) é obtida pela simetrização da matriz da Figura (6.5(a)), ou seja, da soma com sua transposta. Os dados foram retirados do relatório [76].

No Plano Diretor de Transporte Urbano do Distrito Federal realizado em 2009 em ocasião à previsão de um novo sistema de transporte urbano, foi usado o Modelo Gravitacional com atratores para se estimar a demanda em 2010 e 2020 [72]. Em geral, os coeficientes de atratividade estão relacionadas aos dados de população, tais como renda, trabalho etc.

O modelo gravitacional de distribuição de viagens em geral é feito por centróides [82]. Um centróide é definido como um ponto que representará a posição de uma região de onde se originam as viagens e para onde se destinam essas viagens. No Distrito Federal, as RAs são consideradas os centróides, com um ponto posicionado na posição de maior densidade da RA. Percebemos que a analogia do centróide para um nó da rede é bem sugestiva. O que possibilita a representação através de uma rede complexa, cujas arestas são pessoas e cujos nós são as RAs, vejamos a Figura (6.6). Na rede de RAs, o tamanho dos nós é proporcional ao grau do nó. Fizemos uma diferenciação entre os grau do nó ponderado de saída (k_i^{in}), onde o peso é o número de trabalhadores que saem do nó i e de entrada (k_i^{out}), onde o peso é o número de trabalhadores que entram do nó i . Definimos a intensidade das cores dos nós tal que os mais escuros são aqueles que possuem o maior grau do nó na rede de RAs. Descrevemos na tabela 6.1 as RAs e seus índices, que

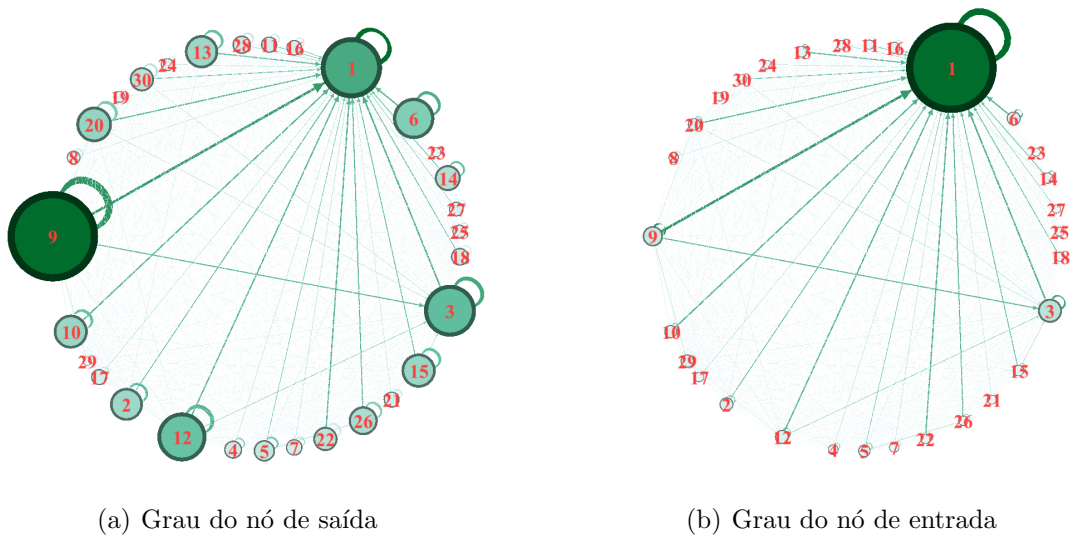


Figura 6.6 - Distribuição espacial dos movimentos diários por motivo de trabalho entre as 30 Regiões Administrativas do DF. Em (6.6(a)) destacam-se as regiões com maior número de trabalhadores ativos que residem em cada RA, pelo tamanho e coloração do nó, quanto maior e mais escuro é o círculo, maior é o grau de saída do nó. Em (6.6(b)) representa-se as regiões quanto ao número de postos de trabalho, ou seja, quanto maior e mais escuro é o nó, maior é o grau de entrada do nó.

foram usados na representação dos nós da Figura (6.6).

Tabela 6.1 - Regiões Administrativas (RAs) do Distrito Federal.

id	nome	id	nome	id	nome
1	Brasília	11	Cruzeiro	21	Riacho Fundo II
2	Gama	12	Samambaia	22	Sudoeste
3	Taguatinga	13	Santa Maria	23	Varjão
4	Brazlândia	14	São Sebastião	24	Park Way
5	Sobradinho	15	Recanto das Emas	25	SCIA(Estrutural)
6	Planaltina	16	Lago Sul	26	Sobradinho II
7	Paranoá	17	Riacho Fundo I	27	Jardim Botânico
8	Núcleo Bandeirante	18	Lago Norte	28	Itapoã
9	Ceilândia	19	Candangolândia	29	SIA
10	Guará	20	Águas Claras	30	Vicente Pires

Retornando às PTNs, definimos o sistema de transporte de uma região formada por S pontos de paradas e T viagens entre os locais i e j . Geramos uma rede do tipo espaço- L , onde os nós são pontos de paradas sequenciais e as arestas são as rotas

entre pontos de paradas sequenciais e uma rede tipo espaço- P , onde os nós são os pontos de paradas e as arestas são as rotas [48]. No entanto, para o estudo de distribuição de viagens, as arestas serão as viagens T . Assim, usaremos a rede L -Trip e P -Trip, previamente definidos na Seção (4.3), cuja nomenclatura passamos a usar para distinguir as PTNs ponderadas do espaço- P e espaço- L . De forma que dividimos o espaço- P em classes de redes: P -Topo (topologia), P -Rout (rotas), P -Trip (viagens) e P -Pass (passageiros). Igualmente fizemos para a rede espaço- L : L -Topo (topologia), L -Rout (rotas), L -Trip (viagens) e L -Pass (passageiros).

Definimos as viagens entre os nós s_i e s_j por

$$T_{ij} = \sum_r T_{ij}^{(r)}, \quad (6.49)$$

onde r é cada rota (linha de ônibus) do sistema e T^r é a viagem da rota r . Assim, $S = \{s_1, s_2, s_3, \dots, s_N\}$ são todos os N pontos de paradas e $T = t_1, t_2, \dots, t_M$ serão todas as viagens de ônibus. Se entre s_i e s_j existe uma viagem, de qualquer que seja a rota, então existirá uma aresta. Podemos definir o grau do nó como sendo o número de viagens entre dois nós s_i e s_j de toda a rede de transportes,

$$k_i = \sum_j T_{ij}. \quad (6.50)$$

Se quisermos representar uma rede ponderada do sistema, o peso da aresta entre os nós s_i e s_j será o número de multiarestas representadas pelas viagens entre s_i e s_j de qualquer das rotas que atendem esses dois nós. No entanto, cada viagem tem uma importância diferente para os passageiros, por exemplo, uma viagem no horário de pico é mais importante do que uma viagem no horário entre-pico, que consideramos igual ao número de passageiros no estudo de pendularidade (Capítulo (5)). A distinguibilidade pode ser relacionada ao horário da viagem de cada rota, propriedade que não faremos distinção. Simplificamos o modelo para um sistema com viagens indistinguíveis de tal forma que possamos usar uma rede ponderada com multiarestas iguais ao peso da aresta.

O que observamos é que isso sugere a relação entre o grau k_i definido para o P -Trip, tal que,

$$k_i = \sum_r \kappa_i^{(r)}, \quad (6.51)$$

onde κ_r é o grau do nó do P -Rout, contado como número de viagens da rota r que atendem o nó i .

Assim, se quisermos saber se existe uma forma de se alcançar o ponto de paradas representado pelo nó s_j a partir de s_i , basta verificar se existe pelo menos uma aresta entre eles e a rede P -Trip fornece essa informação com muita facilidade. Podemos ainda descrever se há ou não conexões com dois ou mais nós intermediários obrigatórios para conectar s_i e s_j a partir da distância de rede.

Como apresentado no Capítulo 4, usaremos a definição de distância $d_{i,j}$ como número de arestas para sair de s_i e chegar em s_j tal como é feito usualmente em redes complexas. Portanto, numa rede espaço- P , a distância $d_{i,j}$ se refere ao número de transbordos, ou transferências, entre as rotas ou viagens. Como nesse caso todos nós s_i de uma rota ou de uma viagem são vizinhos de grau 1, isso significa que eles têm uma distância $d_{i,j} = 1$. No entanto, quando não há caminho direto e $d_{i,j} = 0$, podemos verificar se inserindo um nó intermediário s_k existe o caminho partindo de s_i para s_k e de s_k para s_j . Se isso é verdadeiro, significa que a distância será $d_{i,j} = 2$ e s_i e s_j estão conectados, desde que passe pelo nó intermediário s_k . Na prática isso significa que o passageiro terá de realizar 1 transbordo para trocar de rota em s_k .

No caso da rede espaço- L , se entre s_i e s_j houver uma aresta, ou seja, se existir pelo menos uma rota ou viagem entre s_i e s_j , então existirá um caminho com distância $d_{i,j} = 1$, se, no entanto, para ir de s_i a s_j não houver um caminho direto, mas sendo necessário passar por s_k , então, a distância será $d_{i,j} = 2$. Como a rede é direcionada e sequencial, o valor máximo da distância em uma viagem com N pontos de parada da rota ou viagem será $d_{i,j}^{max} = N - 1$ e será o valor do comprimento total do caminho da rota ou viagem.

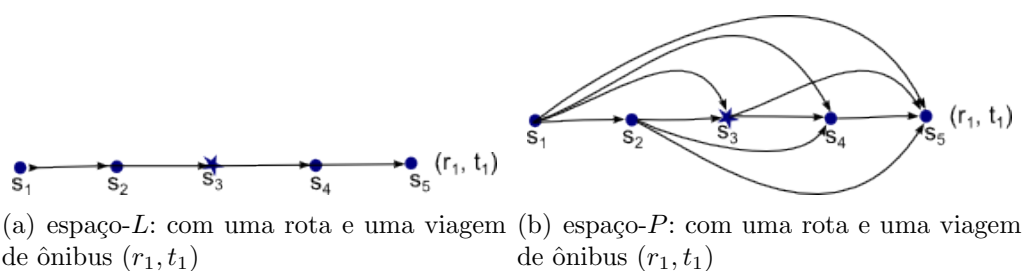


Figura 6.7 - Esquema ilustrativo de viagem com $N=5$ nós, $S = \{s_1, s_2, s_3, s_4, s_5\}$ pontos de paradas. 6.7(a) representa tanto um L -Topo, L -Rout de uma rota, L -Trip de uma viagem ou L -Pass de um passageiro. 6.7(b) representa tanto um P -Topo, P -Rout de uma rota, P -Trip de uma viagem ou P -Pass de um passageiro.

Para analisar a distribuição das conexões, tomemos como exemplo a rede da Figura (6.7(a)). Considerando que a viagem de uma rota qualquer é direcionada, os pontos

de paradas serão sequenciais, tais que $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4 \rightarrow s_5$, como podemos ver na Figura (6.7(a)). A probabilidade de conexão entre os nós s_i e s_j das redes L -Topo e P -Topo, referentes às redes de topologia, será igual para qualquer que seja o s_i e o s_j desde que $j > i$, de modo que teremos uma distribuição homogênea,

$$P_{ij}^P = \frac{1}{\frac{N(N-1)}{2}} \quad (6.52)$$

onde N é o número de nós da rota ou da viagem. Dessa forma, a probabilidade P_{ij} será equiprovável para qualquer par de nós. Igualmente podemos definir para as rede L -Topo e podemos expressar a equação de $P_{d_{ij}}$ por

$$P_{d_{ij}}^L = \frac{1}{\frac{L(L-1)}{2}} \quad (6.53)$$

onde d_{ij} é o menor caminho entre s_i e s_j e L é o número de menores caminhos possíveis entre todos os nós da rede. Ou seja, podemos ter os caminhos, $L = \{l_{12}, l_{13}, l_{14}, l_{23}, l_{24}, l_{34}\}$, tais que, $l_{12} : (s_1 \rightarrow s_2)$, $l_{13} : (s_1 \rightarrow s_2 \rightarrow s_3)$, $l_{14} : (s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4)$, $l_{23} : (s_2 \rightarrow s_3)$, $l_{24} : (s_2 \rightarrow s_3 \rightarrow s_4)$ e $l_{34} : (s_3 \rightarrow s_4)$.

Daí temos a mesma distribuição de viagens entres os nós i e j como uniformemente distribuídos para ambas as redes L -Topo e P -Topo. Isso é plausível, uma vez que necessariamente a viagem irá ocorrer entre o ponto inicial e o ponto final passando por todos os outro pontos de paradas intermediários no caso do L -Topo.

E podemos escrever as condições para P^P e P^L dadas por

$$\sum_{i,j>i} P_{ij}^P = 1 \quad (6.54)$$

e

$$\sum_{i,j>i} P_{ij}^L = 1. \quad (6.55)$$

Podemos generalizar tanto P^P quanto P^L pela equação

$$P_{ij} = \frac{1}{n} \quad (6.56)$$

onde n é número de combinações de possíveis de embarques em i e desembarques j é tal que

$$n = C_2^N = \frac{N(N-1)}{2}, \quad (6.57)$$

tal que se $j = i$ consideramos $P_{ij} = 0$, não teremos autoarestas. Esta equação vale para a rede L -Pass também, pois contamos apenas as combinações entre os nós iniciais e finais. Além disso, para $j < i$ $P_{j<i} = 0$, pois a rede é direcionada. Isso se deve ao fato de que um passageiro não pode desembarcar em um ponto anterior ao qual embarcou, pois o ônibus segue em uma direção do trajeto. Salvo o caso das linhas circulares, que, se o passageiro permanecer no veículo, poderá fazer um trajeto fechado, mas não maior que o comprimento total da rota.

Para o caso das redes ponderadas L -Rout e P -Rout, podemos considerar o peso das arestas como um fator probabilidade, tal que maior o peso, maior a probabilidade de conexão entre os nós. Da mesma forma será para o L -Trip e P -Trip. Ou ainda para o L -Pass e P -Pass, que definimos que o número de passageiros estaria relacionada ao carregamento da viagem, independente de onde embarcou ou desembarcou o passageiro pertencente àquela viagem.

Considerando um sistema formado por várias rotas e várias viagens, podemos representar a probabilidade de viagens entre s_i qualquer e s_j qualquer da mesma forma. Todavia, percebemos que para o caso de redes ponderadas por rotas e viagens, quanto maior o número de multiarestas, ou quanto maior o peso da aresta, maiores são as chances de que a viagem ocorra entre dois nós quaisquer. Assim, para as redes L -Rout e P -Rout, a probabilidade de ir de s_i a s_j será,

$$P_{ij}^P = \frac{1}{\frac{N(N-1)}{2}} \sum_{ij} p_{ij}, \quad (6.58)$$

onde p_{ij} é o peso da aresta ij ou o número da multiplicidade de arestas. E ao questionarmos qual a probabilidade de um passageiro qualquer ir de s_i a s_j usando o sistema de transportes urbano, podemos responder a partir das rotas que o sistema oferece. Isso apenas contando as rotas do sistema.

Da mesma forma será entre as redes L -Trip e P -Trip, já que a aresta nas duas redes é a existência de viagens. E uma viagem passa necessariamente por todos os pontos de paradas pertencentes a rota, salvo exceções relacionadas com acidentes ou desvio de rota. No entanto, em geral todos os nós de uma viagem serão igualmente acessíveis. Assim, a equação de distribuição de viagens entre os nós da rota do L -Trip e P -Trip pode ser expresso pela equação (6.58). Para o cálculo dos resultados de pendularidade, consideramos o peso do L -Pass e do P -Pass como o número total de passageiros que utilizaram uma viagem específica. Não questionamos quais eram os pontos iniciais e finais de cada passageiro.

Para que possamos saber precisamente quais os nós de início e de fim de viagem cada passageiro percorreu, seria necessário que os dados de passageiros das viagens de ônibus estivessem discriminados por ponto de embarque e de desembarque. De modo que teríamos a distinguibilidade entre os passageiros. No entanto, não existem esses dados, tão somente os dados quantitativos da viagem inteira.

O questionamento sobre os pontos de embarque e de desembarque de cada passageiro é importante para verificar o desejo de viagens dos passageiros. Isso geraria uma rede cujos nós teriam peso de importância a partir do interesse dos passageiros. Vimos que estatisticamente um passageiro que embarca no nó s_2 , pode desembarcar em s_3 ou s_4 ou s_5 , Figura (6.7), com probabilidades iguais, pois não sabemos qual a importância de cada nó para esse passageiro. Ou ainda, considerando que pontos de paradas onde passam mais rotas e mais viagens tenham pesos diferentes, podemos atribuir uma probabilidade a partir da oferta de serviços de transporte.

Mas podemos mudar um pouco o ponto de vista, de forma que passamos a verificar a importância do nó a partir da quantidade de embarques e desembarques que são realizados. Essa questão é relevante para o planejamento do sistema de transporte urbano, pois saber qual a demanda de viagens pode facilitar muito a vida de quem deve oferecer as viagens. A partir do questionamento de onde e para onde um certo passageiro do transporte urbano quer ir, verificando os desejos de viagens dele, podemos tentar estimar como seria um novo sistema que atenda favoravelmente a esses desejos. Vimos que existem alguns métodos a partir do estudo do uso do solo e de teorias comportamentais que possibilitam prever a demanda. E vimos que uma ferramenta bastante útil é a matriz origem-destino feita por cetróides, a qual é produzida pelo modelo gravitacional.

Embora um pouco mais complicado, pois raramente temos os dados de passageiros ponto a ponto, podemos tentar inferir a partir da observação das viagens dos passageiros do sistema de transporte, qual a característica do ir e vir dos usuários. E se isso está de acordo com a demanda da população. Dessa forma, uma vez obtendo os dados de embarque e desembarque ponto a ponto, poderíamos estatisticamente inferir qual o desejo de viagem dos passageiros. Essa abordagem poderia ser realizada para o *L-Pass* e o *P-Pass* quando se sabe precisamente onde cada passageiro embarcou e onde desembarcou, de modo que teríamos noção da importância de cada ponto de parada, e portanto, de cada região, ou vice-versa.

Quando observamos a distribuição de passageiros de uma viagem em uma rota qualquer do sistema de transporte urbano, percebemos que podem ter diferentes tipos de distribuição de probabilidade. Se a distribuição é uniforme, ou seja, se todos os pontos

de paradas tiverem a mesma relevância, podemos estimar novamente a probabilidade de uma viagem entre s_i e s_j pela equação anterior (6.58). Mas sabemos a partir dos dados empíricos que essa distribuição não é uniforme, que existem regiões mais importantes que outras, como podemos observar na Figura (6.6) e, portanto, que existem pontos de paradas mais importantes que outros, como exemplo, podemos retornar à Figura (6.2) e verificar que existem pontos de paradas que não há embarque e nem desembarque. A distribuição de passageiros ao longo de uma rota não é uniforme.

Assim, usando a rede L -Trip direcionada, podemos ver que um passageiro que embarca no ponto de parada s_1 , tem a possibilidade de desembarcar em s_2 , s_3 , s_4 ou s_5 , se um passageiro embarca em s_2 , ele poderá desembarcar em s_3 , s_4 ou s_5 , jamais em s_1 , Figura (6.7). Generalizando temos que se um passageiro que embarca numa viagem no ponto s_i , ele só poderá desembarcar em qualquer s_j , com $j > i$.

Em redes reais, portanto, nem sempre os embarques e desembarques são uniformes entre os pontos de paradas da rota, como foi mostrado nos gráficos da Figura (6.2) na Seção 6.5. Vemos pelos dados empíricos de embarque e desembarque ponto a ponto que há uma tendência de haver mais embarque no início da viagem que no final e mais desembarques no final da viagem que no seu início.

Daí observamos que o transporte de ônibus do Distrito Federal não tem muito sobe-e-desce de passageiros nos pontos intermediários de uma rota. Sugerindo que um veículo de uma certa viagem já inicia com um número alto de passageiros e percorre quase toda a rota sem alteração no saldo de passageiros dentro do veículo até passar pelos pontos finais da rota. Na Figura (6.1) podemos ver como é o comportamento do saldo de passageiros em algumas viagens. Isso significa que para o passageiro que embarcou em s_1 , a probabilidade de desembarcar em s_5 é maior que desembarcar em s_2 .

No entanto, mesmo observando os dados empíricos de algumas rotas presentes em [81], não temos os passageiros distinguíveis. Podemos tentar estimar a partir dos dados empíricos da Figura (6.2) como seria a probabilidade P_{ij} de embarcar em s_i e desembarcar em s_j . A única opção aqui seria usar a simplificação e considerar os eventos de embarque independentes dos eventos de desembarque. Tal que $P_{ij} = P_i^E P_j^D$, onde P_i^E seria a probabilidade de embarcar em s_i e P_j^D , a probabilidade de desembarcar em s_j . Utilizando essa distribuição teremos os resultados apresentados na Figura (6.8). Contudo, observamos que esse caso não é o melhor, pois permite que um passageiro embarque e desembarque no mesmo ponto de parada, que é algo que não acontece no sistema real.

Em geral para o planejamento e previsão de demanda de viagens de um

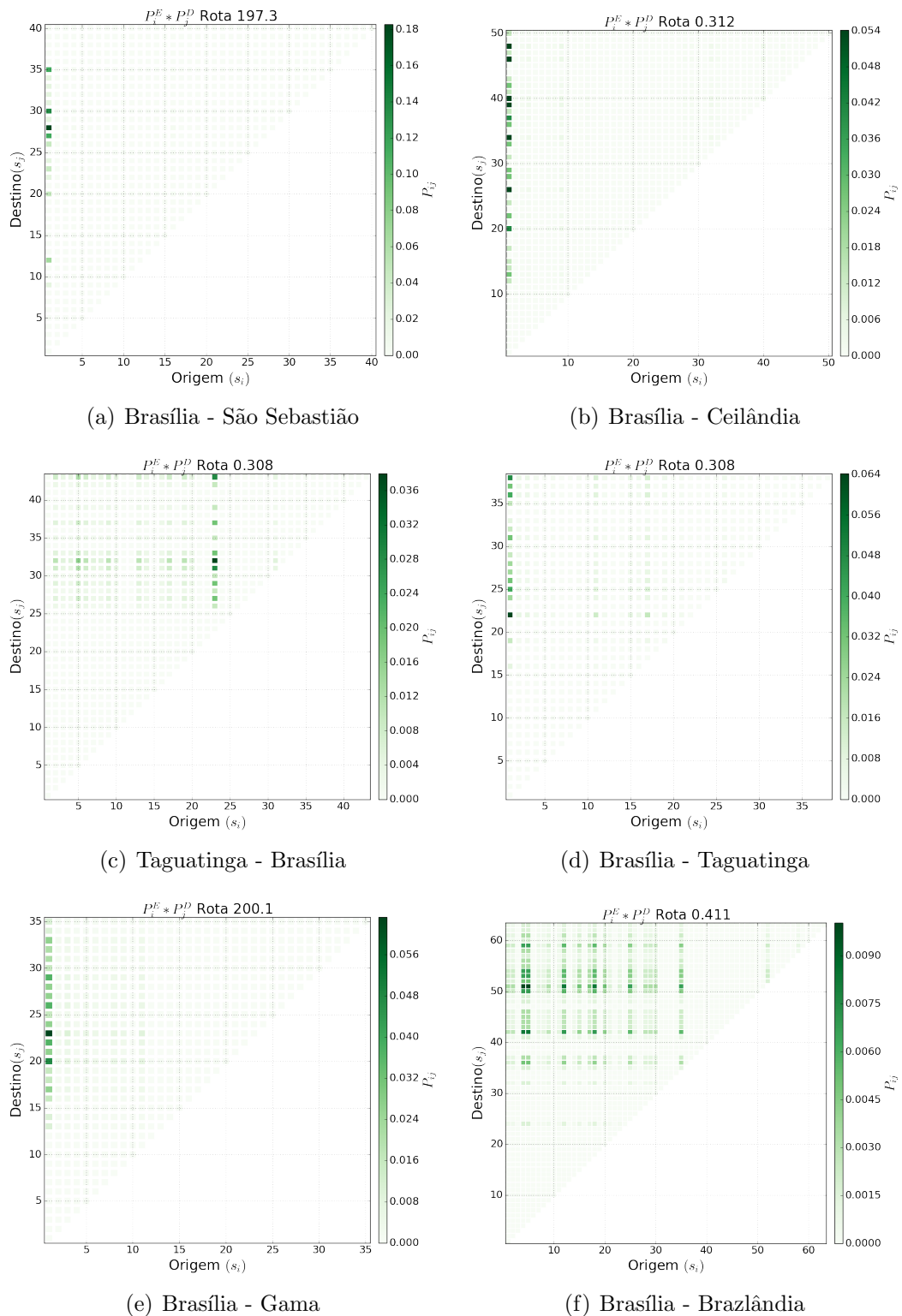


Figura 6.8 - Amostra do Embarque e Desembarques de viagens em rotas ônibus do DF ponto a ponto usando a distribuição de probabilidade P_{ij} como probabilidade de embarcar e probabilidade de desembarcar sejam eventos independentes.

sistema de transporte urbano é usado o Modelo gravitacional, desde que se tenha os dados de custos e população. Entre os vários elementos de custo, pode ser usada a distância como um deles. Contudo, a equação é dada por $T_{ij} \sim \exp(-b c_{ij})$, onde c_{ij} é o fator de custo da viagem, que pode ser um fator de tempo, distância, etc. Se definirmos o custo como distância da viagem, teremos o resultado em que maior a distância d_{ij} , menor as chances de ocorrer a viagem T_{ij} . Isso é plausível, imaginemos que existam 3 padarias, a chance de se ir à padaria mais próxima é maior, que aquela mais distante.

Os dados das rotas de ônibus do DF trazem uma característica interessante. Percebemos que os embarques ocorrem no início da viagem, em geral o número de passageiros dentro do ônibus vai aumentando, até que começam os desembarques. Isso sugere que os passageiros tendem a percorrer viagens com maior distância em vez de menor distância.

Então, propomos um modelo matemático em que se possa representar as características de embarques e desembarque do sistema de transporte urbano do DF, em que a probabilidade de viagem entre uma origem e um destino seja tal que a situação mais provável seja aquela em que um passageiro desembarcar em um ponto mais distante do ponto embarcado. Ou seja, teremos um sistema em que maior será a probabilidade de embarque e desembarque, quanto maior for a distância entre nós, d_{ij} , contado em número de arestas. Simulamos o modelo para os pontos de paradas e depois para RAs. Assim, faremos uma modificação na equação do modelo gravitacional usando apenas a distância característica de redes d_{ij} como vínculo do sistema, tal que a probabilidade de ocorrer uma viagem entre s_i e s_j seja proporcional a

$$P_{ij} \sim e^{(-b^2/d_{ij})}, \quad (6.59)$$

onde d_{ij} é a distância entre os nós i e j em termos de número de arestas da rede espaço- L . Como

$$P = \sum_i \sum_j P_{ij} = 1, \quad (6.60)$$

temos

$$P_{ij} = \frac{e^{(-b^2/d_{ij})}}{\sum_{i,j>i} e^{(-b^2/d_{ij})}}. \quad (6.61)$$

Dessa forma, a equação de distribuição de viagens é.

$$T_{ij} \sim P_{ij}, \quad (6.62)$$

onde a equação (6.61) é o vínculo do sistema. E finalmente definimos a distribuição de viagens por,

$$T_{ij} = \frac{2}{N(N-1)} \frac{e^{(-b^2/d_{ij})}}{\sum_{i,j>i} e^{(-b^2/d_{ij})}}, \quad (6.63)$$

onde $N(N-1)/2$ é o número de total de possíveis pares de nós da rede.

Podemos regular o fator b para apresentar maior ou menor homogeneidade na distribuição. Ao aumentarmos b , aumenta-se a probabilidade de viagens a grandes distâncias e quanto menor for b , mais homogênea será a distribuição com d_{ij} .

A Figura (6.9) mostra o comportamento do fator b . Quando $b = 0$, teremos o caso de distribuição uniforme, de modo que não importa a distância d_{ij} dos nós i e j , pois o vínculo se anula. Percebemos que quanto mais b se aproxima de zero, menos relevante será a distância d_{ij} da rede. E se b , maior será a relevância para d_{ij} . De modo que a diferença do valor de d_{ij} acarreta em maior diferença na probabilidade de ocorrer uma viagem entre viagens com d_{ij} pequeno e grande. Ou seja, mais relevante é o vínculo.

A Figura (6.10) representa a distribuição por distância d_{ij} , que permite analisar o comportamento da função (6.63), considerando uma viagem de uma rota com $N = 70$ pontos de paradas. Observamos que, para b próximos de zero a distribuição é mais homogênea, sendo zero em $j = i$, e portanto, depende pouco da distância d_{ij} . Ao contrário, quanto maior o valor de b , maior será a diferença na distribuição, tal que para uma distância $d_{ij} = 2$ a probabilidade de embarque/desembarque em dois pontos de uma viagem é menos provável que para uma distância $d_{ij} = 50$.

Vemos que no caso da distribuição uniforme, a maior probabilidade é de se viajar a uma distância pequena, pois temos no caso 69 $d_{ij} = 1$, 68 $d_{ij} = 2$, até 1 $d_{ij} = 69$ e o histograma por d_{ij} obedece ao fator de número de combinações (s_i, s_j) da rede. Ao crescermos o valor de b , embora se mantenha o número de pares de nós com o mesmo d_{ij} , o valor de T_{ij} mostra maior relevância para aqueles pares de nós (s_i, s_j) que possuam maior d_{ij} . Isso significa que o peso estatístico dado pelo P_{ij} se torna mais relevante, quanto maior é o b .

A distribuição de probabilidade por distância d_{ij} é apresentada na Figura (6.10). Não consideramos uma diferença entre os nós, tão somente a distância entre eles, ou seja, não atribuímos qualquer característica demográfica ou algum peso aos pontos de paradas, relacionado à importância da região onde o ponto de paradas está inserido, como, por exemplo, a proximidade de uma escola, shopping ou hospital.

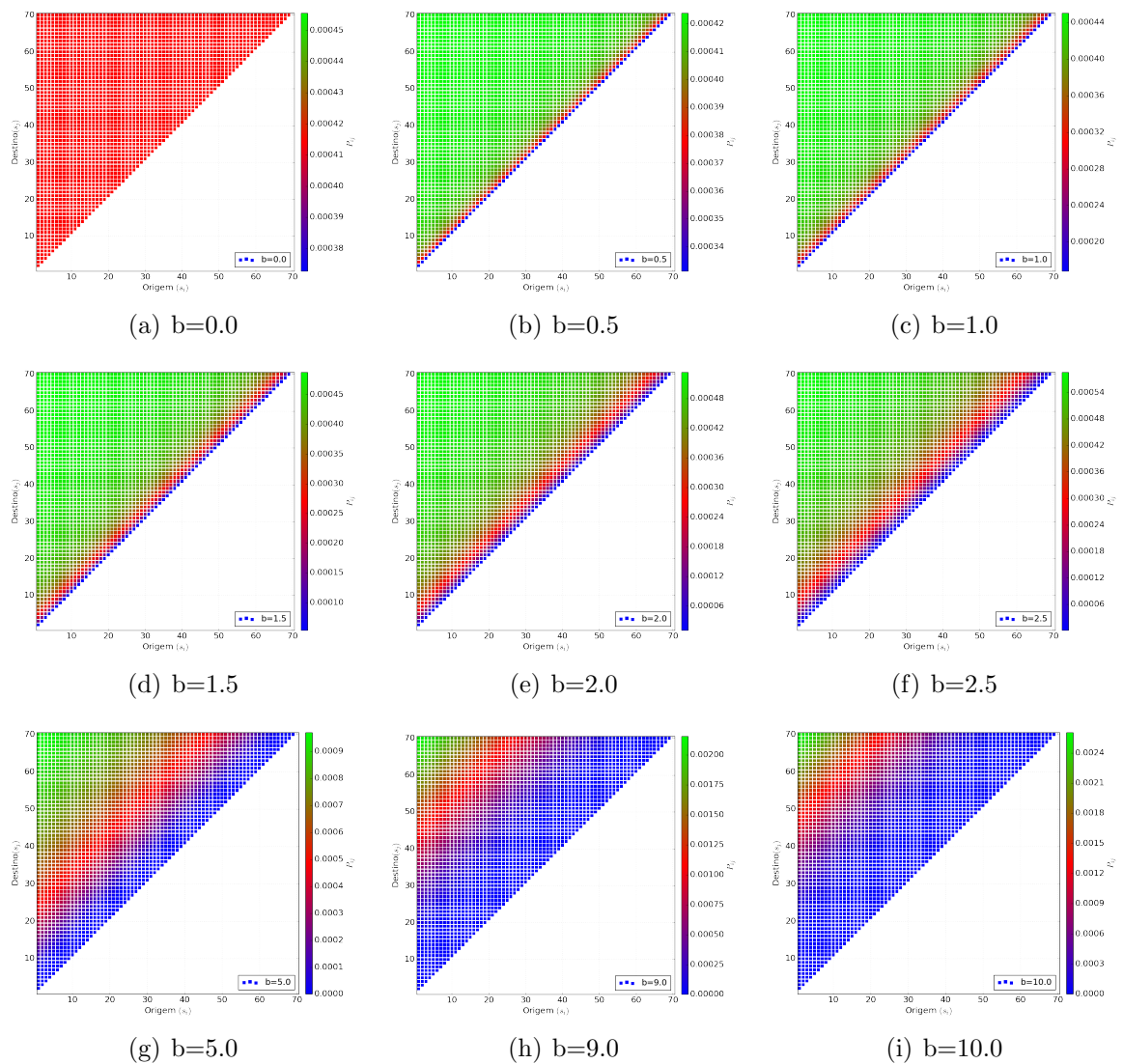


Figura 6.9 - Modelo de Distribuição de passageiros simulado para uma viagem com $N=70$ pontos de paradas. A cor representa a probabilidade de uma viagem originada em s_i (eixo-x) e terminada em s_j (eixo-y). A escala de cores se refere à probabilidade de viagem, quanto mais para o verde, mais provável é a viagem e quanto mais para o azul, menor é a probabilidade.

Após calcular a probabilidade de distribuição usando o modelo, expresso pela equação (6.61), para todas as viagens de todas as rotas, fizemos a soma por RAs, somando o valor dos P_{ij} de todos os pontos que originam em uma RA determinada e se destina a todos os pontos de paradas de outra RA.

A Figura (6.11) ilustra como foi feita a distribuição por RAs. Uma rota de ônibus passa pelos pontos de paradas $\{s_1, s_2, s_3, s_4, s_5, s_6\}$, que estão distribuídas ao longo das regiões $\{A, B, C, D\}$. Para encontrarmos a probabilidade de viagens $T_{i'j}$, onde i' é a região de origem e j' é a região de destino, fazemos a soma dos T_{ij} que pertencem a cada

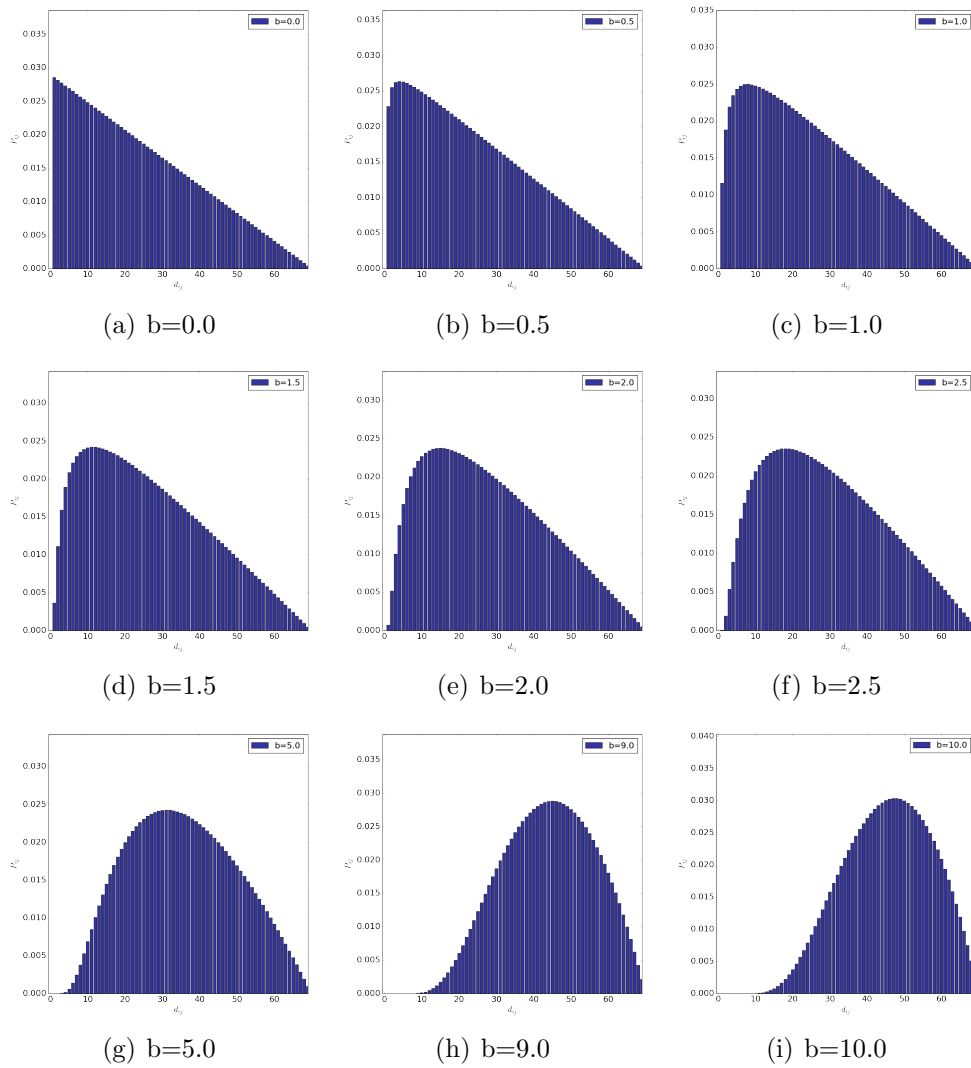


Figura 6.10 - Distribuição de passageiros por d_{ij} simulado para uma viagem com $N=70$ pontos de paradas. No eixo-x estão os d_{ij} e no eixo-y estão os T_{ij} .

região. Por exemplo, se queremos calcular as viagens entre as regiões A e B , calculamos $T_{AB} = T_{s_1s_2} + T_{s_1s_3}$. Da mesma forma, se queremos calcular as viagens entre as regiões B e D , calculamos $T_{BD} = T_{s_2s_5} + T_{s_2s_6} + T_{s_3s_5} + T_{s_3s_6}$.

Podemos expressar a distribuição de viagens por região como,

$$T_{i'j'} = \sum_{i \in i'} \sum_{j \in j'} T_{ij}. \quad (6.64)$$

Para exemplificar o uso do modelo em um sistema de transporte real, vamos analisar a distribuição de viagens em algumas rotas de ônibus do DF. A rota/linha 163.3,

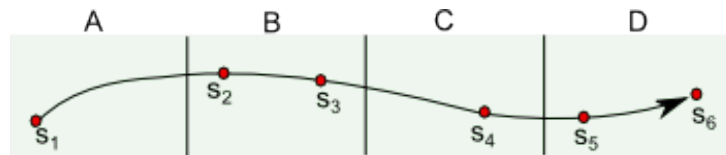


Figura 6.11 - Esquema ilustrativo da distribuição de viagens agrupados por RA. Uma rota de ônibus passa pelos pontos de paradas $\{s_1, s_2, s_3, s_4, s_5, s_6\}$, que estão distribuídas ao longo das regiões $\{A, B, C, D\}$. Para encontrarmos a probabilidade de viagens $T_{i'j'}$, onde i' é a região de origem e j' é a região de destino, fazemos a soma dos P_{ij} que pertencem a cada região. Por exemplo, $T_{AB} = P_{s_1s_2} + P_{s_1s_3}$.

possui $N = 70$ pontos de paradas, é uma linha circular originada em Núcleo Bandeirante e que passa pela Candangolândia, com destino à Brasília, retornando pelo Sudoeste e voltando ao Núcleo Bandeirante. Calculamos a probabilidade de viagens P_{ij} usando o modelo de distribuição para diferentes valores de b , para $b = 0.5$, $b = 1.0$ e $b = 2.0$, e somamos para cada Ra, Figura (6.12).

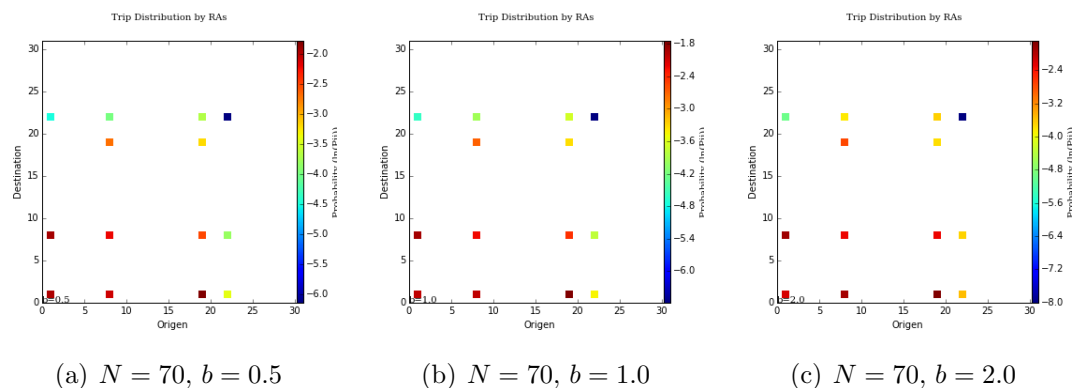


Figura 6.12 - Cálculo da Distribuição de viagens pelo modelo alterado calculado para a rota 163.3. O gráfico representa a distribuição ponto a ponto em escala logarítmica. No eixo-x encontram-se as RAs de origem e no eixo-y, as Ras de destino. A escala de cores informa a probabilidade de ocorrer $T_{i'j'}$, quanto mais para o vermelho, maior é a probabilidade, quanto mais para o azul, menor a probabilidade.

Fizemos o cálculo das probabilidades de viagens para todas as viagens de todas as rotas e multiplicamos pelo número de passageiros de cada viagem para estimar a distribuição de viagens de passageiros em todo o território do Distrito Federal para obtermos um L -Pass real. Lembrando que no Capítulo (4) o peso das arestas da L -Pass era definido como o número total de passageiros transportados na viagem e iguais para todas as arestas geradas pela viagem, ou seja, o peso de todas as arestas da viagem teria o mesmo peso. Ao somarmos o peso produzido por uma viagem a todos os outros pesos produzidos pelas outras viagens existentes entre i e j conseguíamos definir o peso

das arestas na rede L -Pass ou P -Pass. Usando o modelo de distribuição proposto neste capítulo, podemos obter uma rede L -Pass em que cada aresta gerada pela viagem terá um peso diferente, dependendo do valor do número de passageiros que vão do ponto de parada i até o ponto de parada j . E podemos obter uma rede L -Pass em conformidade com as reais viagens entre os pontos de paradas do DF.

6.7 Resultados

Calculamos a probabilidade de viagem entre todos os pontos de paradas atendidos por todas as viagens de todas as rotas de ônibus do DF, usando a equação (6.63), e obtivemos a distribuição de probabilidade mostrada na Figura (6.13), onde origem e destino são pontos de paradas localizados no DF.

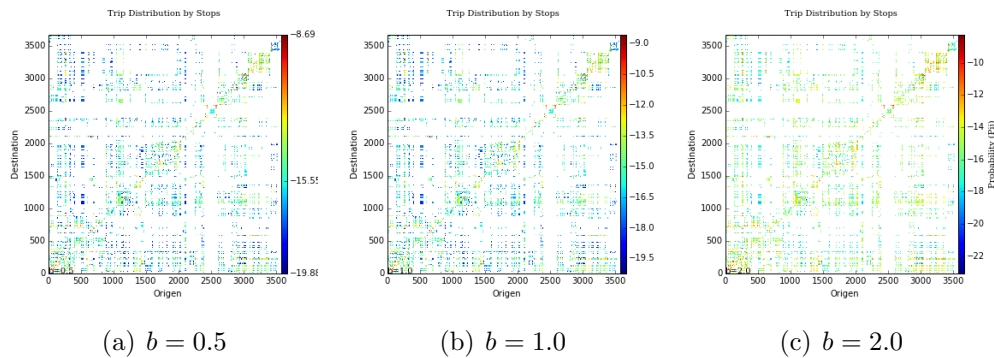


Figura 6.13 - Distribuição da probabilidade de viagens para todos os pontos de paradas de ônibus, com as cores em escala logarítmica. No eixo-x estão posicionados os pontos de paradas de origem (s_i 's) e no eixo-y, os pontos de paradas de destino (s_j 's). A escala de cores corresponde à probabilidade P_{ij} de se ir de s_i a s_j , a cor para o vermelho representa maior probabilidade e a cor mais para o azul representa menor probabilidade.

A nomenclatura dos 3674 pontos de paradas da Figura (6.13) é aleatória, ou seja, o ponto s_1 não é mais próximo de s_2 que de s_3 , apenas são pontos diferentes da rede. Esse gráfico representa os elementos da matriz T_{ij} , onde quanto mais a cor está para o vermelho, maior será a probabilidade de viagens e quanto mais a cor está para o azul, menor será a probabilidade de viagens entre os pontos de paradas i e j . O gráfico apresenta espaços em branco, pois existem pontos de paradas de ônibus que não são conectados por nenhuma viagem diretamente.

Posteriormente, fizemos a distribuição por RA, onde temos o gráfico da Figura (6.14). Utilizamos a equação (6.64), com valores de $b = 0.5$, $b = 1.0$ e $b = 2.0$ por ponto de paradas e agrupamos por RA. A distribuição é calculada para todas as viagens de

ônibus no DF, por isso percebemos que a maior probabilidade se encontra dentro da mesma RA e nas viagens para a RA=1, Brasília, nas viagens de volta de Brasília para as outras RAs. O que sugere que as viagens de ônibus do DF estão de acordo com as necessidades de migração diária com motivo de trabalho. A partir desse gráfico percebemos que nem todos as 30 RAs estão conectadas por rotas de ônibus. Isso é bastante razoável, se lembrarmos da distribuição de fluxo de trabalhadores da Figura (6.5).

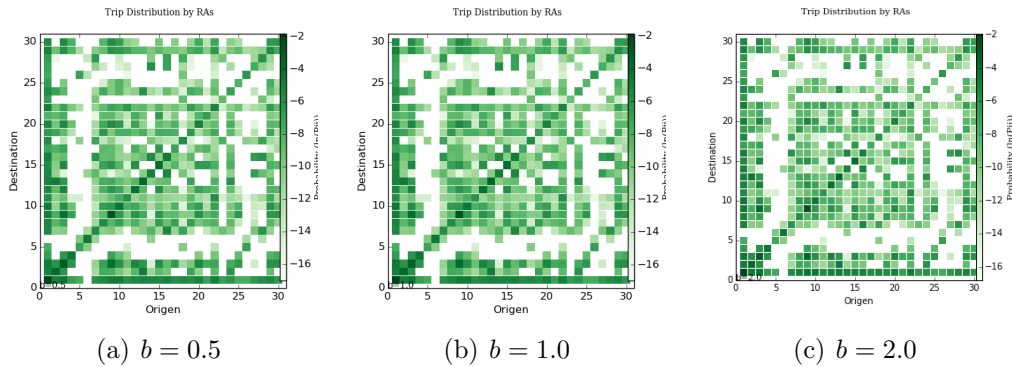


Figura 6.14 - Distribuição de passageiros por RA a partir do modelo de expresso pela equação (6.61) As cores estão em escala logarítmica. No eixo-x estão representadas as RAs de origem (i') e no eixo-y, as RAs de destino (j'). A escala de cores representa a probabilidade $T_{i'j'}$, quanto mais escura é a cor, maior é a probabilidade e quanto mais clara é a cor, menor é a probabilidade.

A partir disso, usamos os dados censitários de número de trabalhadores que vivem na RA de origem i' e o número de postos de trabalhos ocupados na RA de destino j' e multiplicamos por $P_{i'j'}$ e pelo número de viagens. A esse procedimento denominamos Modelo Gravitacional Alterado, com o termo $f(d_{ij})$ dado pela equação (6.61).

$$T_{i'j'} = O_{i'} D_{j'} P_{i'j'}, \quad (6.65)$$

onde $O_{i'}$ é o número de trabalhadores de que residem em i' , $D_{j'}$, o número de postos de trabalho em j' , e $P_{i'j'}$ dada pela equação (6.61). Definimos alguns vínculos, tal que:

$$O_{i'} = \sum_{j'} D_{j'} \quad (6.66)$$

$$D_{j'} = \sum_{i'} O_{i'} \quad (6.67)$$

e

$$\sum_{i'} O_{i'} = \sum_{j'} D_{j'} \quad (6.68)$$

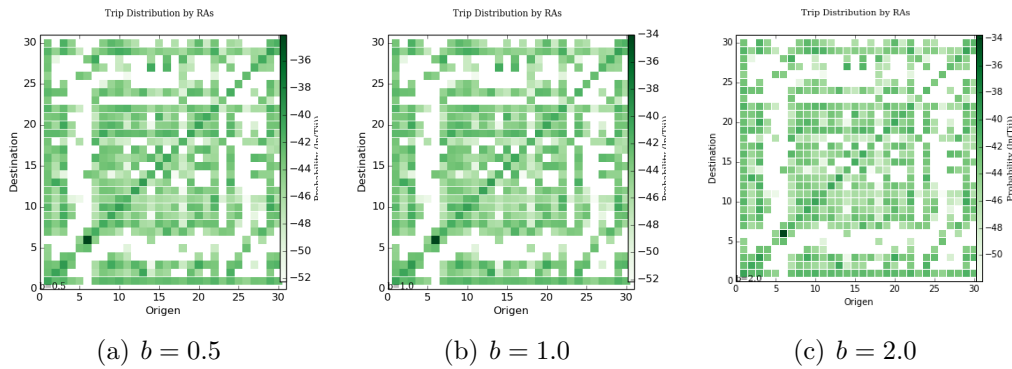


Figura 6.15 - Distribuição de passageiros por RA com o valores censitários da população do DF.

A Figura (6.15) apresenta esse resultado. No entanto, com a inclusão desses vínculos, percebe-se que a distribuição foge da distribuição de trabalhadores mostrada no gráfico da figura (6.5). Inclusive, o ponto origem 6 e destino 6, que representa viagens internas na RA de Planaltina, curiosamente apresenta a maior probabilidade de ocorrência. Percebemos no decorrer do trabalho que os dados de viagens da RA=6 não estavam completas. A base de dados das viagens apresentavam poucas rotas da RA=6. De modo que, necessitamos de dados para melhorar a precisão dos resultados.

A simulação para os dados reais de passageiros da rede *L-Trip* só foi possível usando as RAs como centróides, um vez que para essa unidade de área os valores demográficos foram adquiridos.

Por outro lado, testamos o Modelo Gravitacional Alterado para o caso de algumas rotas reais, que apresentavam os valores de embarque e desembarque em cada ponto de parada da rota. Embora não tenhamos simulado para todos os dados da rede, acreditamos que o resultado deste estudo possa ser útil para o entendimento do modelo e a real aplicação para todo o sistema de transporte, desde que tenhamos os valores de embarque e desembarque ponto a ponto.

Para esse teste, consideramos apenas os valores de embarque E_i , de desembarque D_j e a distância em arestas de s_i a s_j na rede. Simulamos para o caso em que a probabilidade P_{ij} de embarcar em s_i e desembarcar em s_j são eventos independentes, cuja equação é $P_{ij} = P_i^E P_j^D$, onde P_i^E é a probabilidade de embarque e P_j^D , de desembarque.

Simulamos para o Modelo Gravitacional definindo o valor de $O_i = E'_i$, ou seja, o número de pessoas nos pontos de origem equivalem ao número de embarques em s_i e o valor de $D_j = D'_j$, ou seja, o número de pessoas nos pontos de destino equivalem

ao número de passageiros que desembarcam em s_j , tal que a equação se torna $P_{ij} = E'_i D'_j \exp(-b^2 c_{ij})$, com $c_{ij} = d_{ij}$.

E comparamos com o Modelo Gravitacional Alterado, onde modificamos a função de distância, tal que $P_{ij} = E'_i D'_j \exp(-b^2/d_{ij})$.

A Figura (6.16) mostra simulações para uma rotas, considerando que o embarque em s_i e desembarque em s_j são eventos independentes. Usamos o número de embarques em s_i como valor de origens $O_i = E'_i$ e o número de desembarque em s_j como valores de destinos $D_j = D'_j$ para simular o modelo gravitacional e o modelo gravitacional alterado. Simulamos para algumas rotas reais para testar o modelo, que se encontram no Apêndice C.

Observando a diagonal principal da Figura (6.16), percebemos que o vínculo da distribuição de viagens por d_{ij} do modelo proposto neste trabalho evita as viagens com embarque e desembarque no mesmo ponto de parada. Isso faz sentido, pois nenhum passageiro embarca num ponto e desembarque assim que entra no ônibus.

Todavia, para um modelo de distribuição de viagens de passageiros, ou seja, para uma distribuição de que preveja a real P -Pass, seria necessário a obtenção de dados dos passageiros ponto a ponto. Significa que necessitamos da distinção dos passageiros, conhecer o seu ponto de origem e seu ponto de destino. E os dados de embarque poderia substituir o valor dos viajantes que originam em i e se destinam a j para cada viagem de cada rota.

Um fator importante sobre o comportamento dos passageiros de ônibus do Distrito Federal, que diferem das expectativas dos primeiros modelos de distribuição de viagens usando a equação do modelo gravitacional e derivados, pode ser atribuído ao fato de existir um tipo de incentivo para as pessoas que vivem mais distantes dos centros de oportunidades utilizarem o transporte público. Isso colabora com o resultado de viagens serem mais prováveis para distâncias maiores. Por outro lado, equações de distribuição derivados do modelo gravitacional usam fatores de calibração que vai além da topologia da rede, tais como renda do indivíduo, algum fator relacionado a existência de lojas, hospitais, escolas.

Na verdade, no Distrito Federal o uso do transporte público está bastante relacionado com a renda do indivíduo. A maior parte dos passageiros possuem rendas mais baixas que o restante da população do DF, e são justamente aqueles que vivem mais distantes do centro de oportunidades. Um dos incentivos para o uso do transporte público

é o pagamento do Vale Transporte e o preço da tarifa. Aqui o valor da tarifa de ônibus não é proporcional aos custos da viagem, mas possui um fator social embutido no seu cálculo, de forma que os passageiros que percorrem menores distâncias pagam um valor de tarifa que ajuda no custo e, portanto, no valor da tarifa paga pelos passageiros que percorrem maiores distâncias.

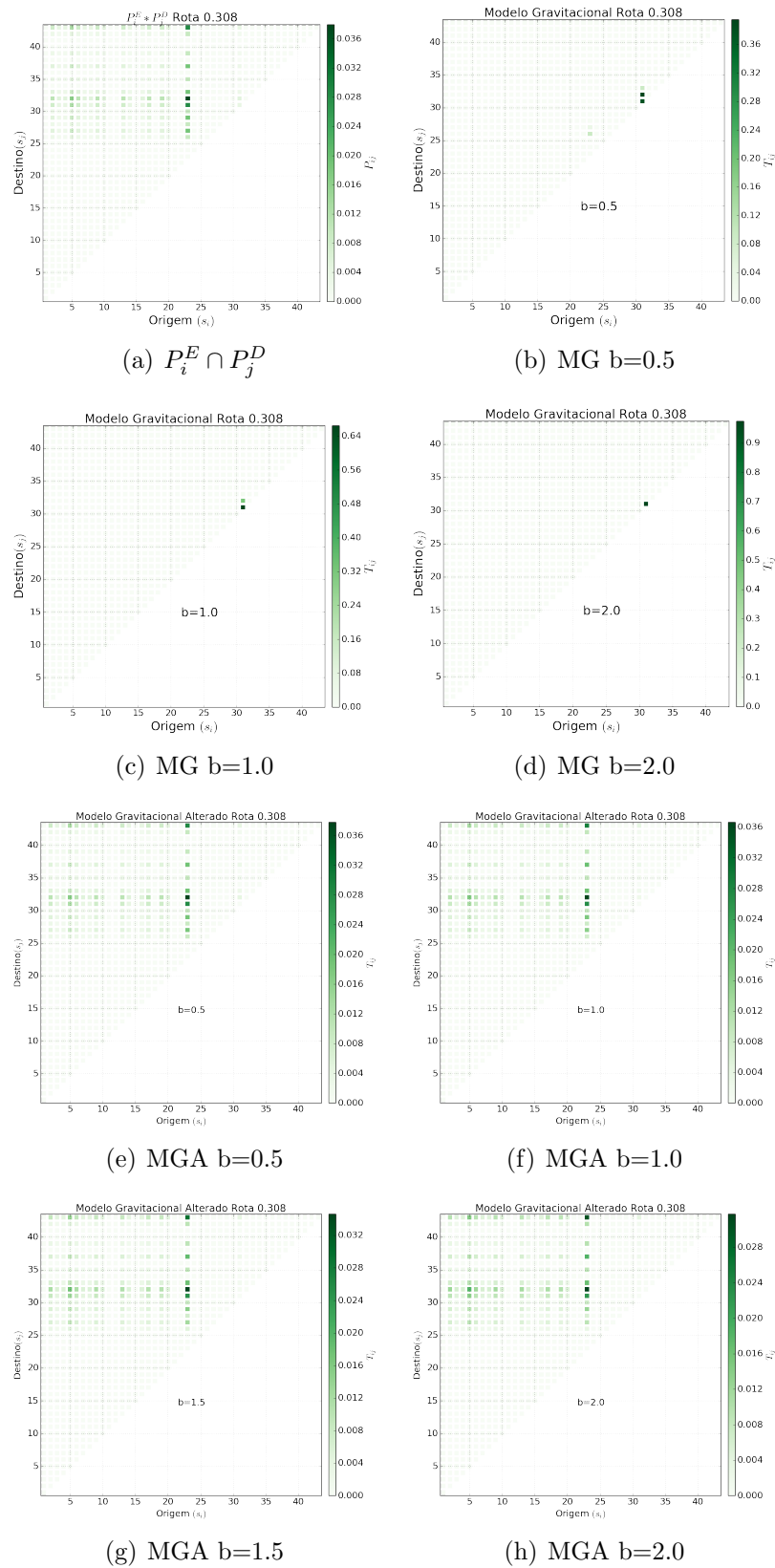


Figura 6.16 - Distribuição de passageiros em uma viagem da rota 0.308. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.

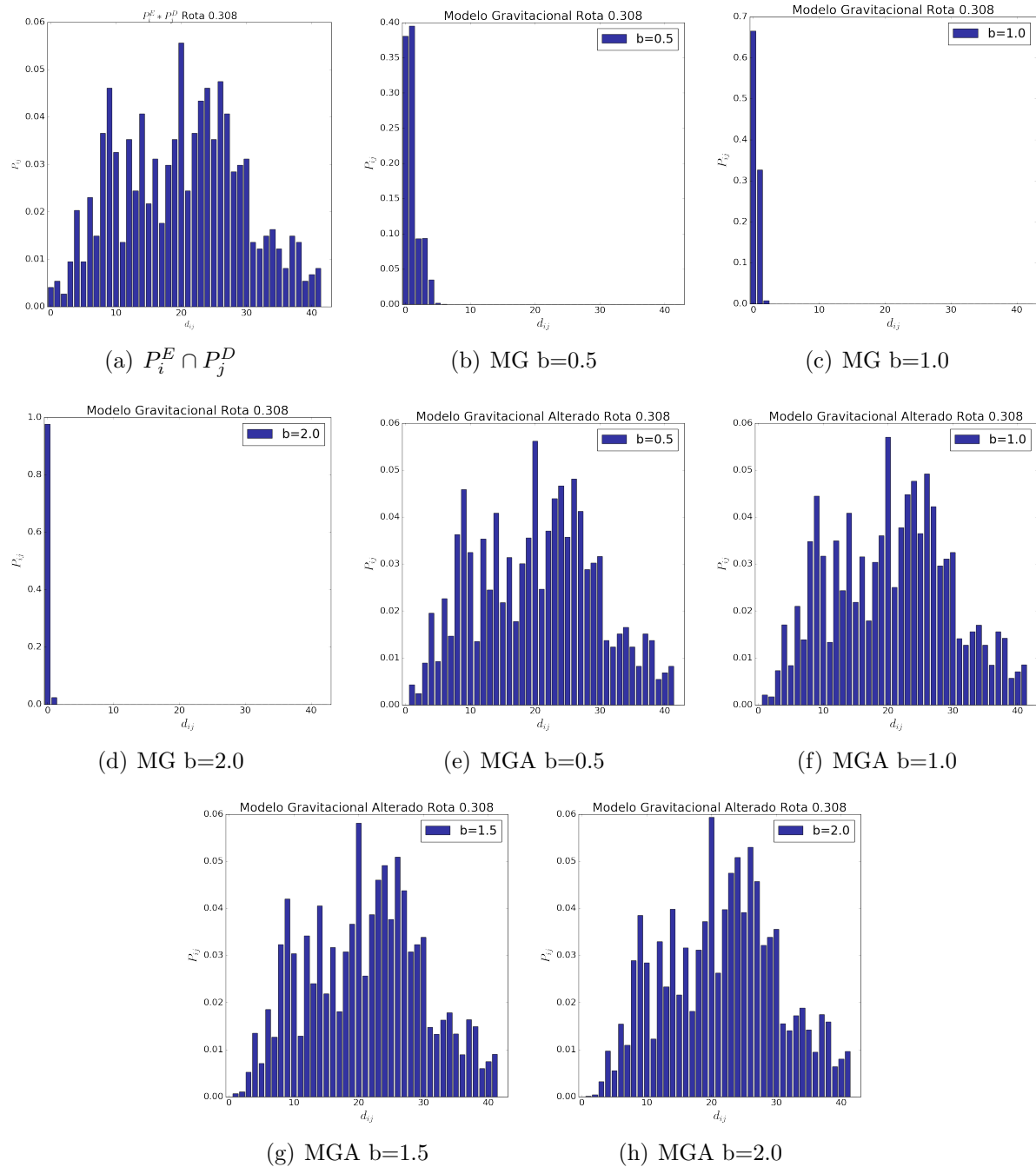


Figura 6.17 - Simulação da distribuição de passageiros simulado para a rota 0.308 usando vários modelos. O eixo-x representa a distância d_{ij} e o eixo-y representa a probabilidade P_{ij} .

CAPÍTULO 7

Conclusão e Perspectivas

Construímos as redes *P-Rout*, *P-Trip* e *P-Pass* como redes ponderadas com número de rotas, número de viagens e número de passageiros, respectivamente, fizemos a distribuição do grau para cada uma delas, bem como para as redes *L-Rout*, *L-Trip* e *L-Pass*, e percebemos que a distribuição do grau do nó para essas redes descreve um curva em decaimento exponencial. Tentamos o ajuste da curva com uma função de decaimento lei de potência, porém, ficou claro que o ajuste com uma função de decaimento exponencial é o mais adequado. Isso, pode trazer informações úteis para o estudo de redes reais *PTN*, pois entendemos que essas redes devem ter a construção diferenciada para redes de trem e redes de ônibus, por causa da propriedade que muitas vezes é negligenciada pelos autores, a direção das conexões.

Mostramos que as grandezas relacionadas a concentração de viagens (R_{CM} e θ_{CM}) pode trazer novas informações para a compreensão da distribuição de viagens num intervalo de tempo em redes ponderadas *L-Rout*, *L-Trip* e *L-Pass*. Essas grandezas podem auxiliar na verificação das viagens em um ponto de parada específico, sendo o valor R_{CM} importante para apontar quanto são concentradas as viagens em torno de um horário específico. Vemos por meio da grandeza θ_{CM} os diferentes horários de concentração de viagens nos pontos de paradas do DF. Assim, com dados mais precisos poderemos verificar o horário de pico ponto a ponto.

Formulamos um modelo de distribuição de viagens baseado no Modelo Gravitacional Tradicional para as viagens de ônibus do Distrito Federal, cujo parâmetro usado para definir a probabilidade de uma viagem ocorrer entre dois lugares diferentes é a distância entre os nós da rede de *L-Trip*. A apresentação da distribuição parece satisfazer as necessidades de demanda de deslocamento diário de trabalhadores no trajeto residência-trabalho, embora não tenhamos dados de população na equação, mas tão somente a distância em número de nós da rede. Sem os elementos de embarque e desembarque ponto a ponto para todos os pontos de paradas de ônibus do DF não podemos usar o modelo gravitacional tradicional para aos prever os embarque e desembarque de uma viagem de rota específica, por isso usamos apenas a distância da rede *L-Trip*. Simulamos com os dados de algumas rotas a título de verificação do modelo.

Propomos uma equação para a distribuição de viagens no P_{ij} em função característica típica dos embarques e desembarques nos ônibus do DF, onde a maior parte dos passageiros embarca no início da rota e somente começa a desembarcar no final dela.

Assim, fizemos uma distribuição em que maior seria a probabilidade de ir de i a j , quanto maior fosse a distância d_{ij} . Simulamos para todas as viagens de todas as rotas do DF com os valores de $b = 0.5$, $b = 1.0$ e $b = 2.0$. No entanto, uma vez tendo mais informações do transporte do DF, poderemos encontrar o b ideal para a distribuição. Pois usando essa equação no modelo gravitacional alterado para a rede L -Trip toda, os resultados não são claros quanto a correspondência da real demanda.

Até agora, fizemos análises das redes do espaço- L , bem como das redes do espaço- P , separadamente, todavia, como perspectivas, podemos modelar as redes ponderadas e direcionadas -Topo, -Rout, -Trip e -Pass em rede de multicamadas, calcular propriedades entre camadas, aplicar um método de ataque ao sistema, retirando nós e/ou arestas e verificar a robustez do sistema a partir do modelo multiplex. E ainda, realizar comparações com diversos modelos de redes, de forma a verificar o mais adequado, usando o modelo de dissimilaridade.

Além disso, podemos aplicar a técnica de Pendularidade a outras cidades cujos dados do sistema de transporte possam ser obtidos, tanto dados da estrutura física (topologia) quanto dados operacionais (processo dinâmico) da rede.

E para o modelo de distribuição de viagens gerar resultados mais precisos, é necessário uma melhor investigação do sistema de transportes com dados reais de embarque e desembarque de passageiros ponto-a-ponto. De modo que, consigamos encontrar um b diferente para cada rota.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. *nature*, Nature Publishing Group, v. 393, n. 6684, p. 440–442, 1998. 3, 10, 11, 12, 26, 27
- 2 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999. 3, 9, 10, 12, 13
- 3 NEWMAN, M. E. J. *Networks: An introduction*. 9. ed. New York: [s.n.], 2015. 772 p. 3, 5, 9, 14, 15, 18, 21, 22, 23, 24, 26, 31, 50
- 4 BOLLOBÁS, B. *Modern graph theory*. [S.l.]: Springer Science & Business Media, 2013. v. 184. 4, 16, 18, 20
- 5 GRIBKOVSKAIA, I.; HALSKAU, Ø.; LAPORTE, G. The bridges of königsberg—a historical perspective. *Networks*, Wiley Online Library, v. 49, n. 3, p. 199–203, 2007. 4
- 6 BÓNA, M. *A walk through combinatorics: an introduction to enumeration and graph theory*. [S.l.]: World Scientific, 2011. 4
- 7 WEST, D. B. et al. *Introduction to graph theory*. [S.l.]: Prentice hall Upper Saddle River, 2001. v. 2. 4
- 8 ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, APS, v. 74, n. 1, p. 47, 2002. 4, 5, 8, 12, 13, 23, 28, 50
- 9 STEEN, M. *Graph Theory and Complex Networks: An Introduction, Maarten van Steen (2010)*. [S.l.]: ISBN. 6
- 10 TRAVERS, J.; MILGRAM, S. An experimental study of the small world problem. *Sociometry*, JSTOR, p. 425–443, 1969. 8
- 11 TRAVERS, J.; MILGRAM, S. The small world problem. *Psychology Today*, JSTOR, v. 1, p. 61–67, 1967. 8
- 12 SIMON, H. A. On a class of skew distribution functions. *Biometrika*, JSTOR, v. 42, n. 3/4, p. 425–440, 1955. 9
- 13 PRICE, D. d. S. Networks of Scientific Papers. *Science*, v. 149, n. 3683, p. 510–515, 1965. 9
- 14 PRICE, D. d. S. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, Wiley Online Library, v. 27, n. 5, p. 292–306, 1976. 9
- 15 BARABÁSI, A.-L.; ALBERT, R.; JEONG, H. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 272, n. 1, p. 173–187, 1999. 10, 12

- 16 NEWMAN, M. E. J. The structure and function of complex networks. *SIAM Review*, SIAM, v. 45, n. 2, p. 167–256, 2003. 10, 13
- 17 BOGUNÁ, M.; KRIOUKOV, D. Navigating ultrasmall worlds in ultrashort time. *Physical Review Letters*, APS, v. 102, n. 5, p. 058701, 2009. 10
- 18 ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Error and attack tolerance of complex networks. *Nature*, Nature Publishing Group, v. 406, n. 6794, p. 378–382, 2000. 10, 11, 12, 28
- 19 ZOU, Z.; XIAO, Y.; GAO, J. Robustness analysis of urban transit network based on complex networks theory. *Kybernetes*, Emerald Group Publishing Limited, v. 42, n. 3, p. 383–399, 2013. 10, 28
- 20 LATORA, V.; MARCHIORI, M. Efficient behavior of small-world networks. *Physical Review Letters*, APS, v. 87, n. 19, p. 198701, 2001. 10, 24, 26, 27, 28, 31, 33
- 21 LATORA, V.; MARCHIORI, M. How the science of complex networks can help developing strategies against terrorism. *Chaos, Solitons & Fractals*, Elsevier, v. 20, n. 1, p. 69–75, 2004. 10, 28
- 22 VRAGOVIĆ, I.; LOUIS, E.; DIAZ-GUILERA, A. Efficiency of informational transfer in regular and complex networks. *Physical Review E*, APS, v. 71, n. 3, p. 036122, 2005. 10
- 23 MELLO, B. et al. Measuring the flow of information among cities using the diffusion power. *Physics Letters A*, Elsevier, v. 374, n. 2, p. 126–130, 2009. 10
- 24 MELLO, B. et al. Network evolution based on minority game with herding behavior. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 76, n. 1, p. 147–156, 2010. 10, 16, 20, 21
- 25 CAJUEIRO, D. O. Optimal navigation in complex networks. *Physical Review E*, APS, v. 79, n. 4, p. 046103, 2009. 10
- 26 CAJUEIRO, D. O. Optimal navigation for characterizing the role of the nodes in complex networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 389, n. 9, p. 1945–1954, 2010. 10
- 27 AMARAL, L. A. N. et al. Classes of small-world networks. *Proceedings of the national academy of sciences*, National Academy of Sciences, v. 97, n. 21, p. 11149–11152, 2000. 10, 31
- 28 KURANT, M.; THIRAN, P. Layered complex networks. *Physical Review Letters*, APS, v. 96, n. 13, p. 138701, 2006. 10, 13, 32
- 29 ARENAS, A. et al. Synchronization in complex networks. *Physics Reports*, Elsevier, v. 469, n. 3, p. 93–153, 2008. 10
- 30 NEWMAN, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, National Academy of Sciences, v. 103, n. 23, p. 8577–8582, 2006. 10

-
- 31 BARABÁSI, A.-L.; BONABEAU, E. Scale-free networks. *Scientific American*, Nature Publishing Group, v. 288, n. 5, p. 50–59, 2003. 12, 31, 35
- 32 WATTS, D. J. The “new” science of networks. *Annu. Rev. Sociol.*, Annual reviews, v. 30, p. 243–270, 2004. 13
- 33 KURANT, M.; THIRAN, P. Extraction and analysis of traffic and topologies of transportation networks. *Physical Review E*, APS, v. 74, n. 3, p. 036114, 2006. 13, 31, 32, 33, 34
- 34 KURANT, M.; THIRAN, P.; HAGMANN, P. Error and attack tolerance of layered complex networks. *Physical Review E*, APS, v. 76, n. 2, p. 026103, 2007. 13, 29
- 35 MUCHA, P. J. et al. Community structure in time-dependent, multiscale, and multiplex networks. *science*, American Association for the Advancement of Science, v. 328, n. 5980, p. 876–878, 2010. 14
- 36 DOMENICO, M. D. et al. Navigability of interconnected networks under random failures. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 111, n. 23, p. 8351–8356, 2014. 14
- 37 GOMEZ, S. et al. Diffusion dynamics on multiplex networks. *Physical Review Letters*, APS, v. 110, n. 2, p. 028701, 2013. 14
- 38 BOCCALETTI, S. et al. Complex networks: Structure and dynamics. *Physics Reports*, Elsevier, v. 424, n. 4, p. 175–308, 2006. 14, 24, 25, 26, 27
- 39 BARTHÉLEMY, M. Spatial networks. *Physics Reports*, Elsevier, v. 499, n. 1, p. 1–101, 2011. 15, 20, 23, 31
- 40 IZAWA, M. *Modelagem do Sistema de Transporte Urbano do Distrito Federal por Redes Complexas*. 64 p. Dissertação (Mestrado) — Universidade de Brasília - UnB, 2010. 16, 17, 22, 25, 40, 41, 42, 43, 44, 46
- 41 NEWMAN, M. E. J. Analysis of weighted networks. *Physical Review E*, APS, v. 70, n. 5, p. 056131, 2004. 17, 21
- 42 NEWMAN, M. E. J.; STROGATZ, S. H.; WATTS, D. J. Random graphs with arbitrary degree distributions and their applications. *Physical review E*, APS, v. 64, n. 2, p. 026118, 2001. 17
- 43 NEWMAN, M. E. J. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical Review E*, APS, v. 64, n. 1, p. 016132, 2001. 17, 26
- 44 BARRAT, A. et al. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, National Academy of Sciences, v. 101, n. 11, p. 3747–3752, 2004. 17, 51
- 45 DIESTEL, R. *Graph theory {graduate texts in mathematics; 173}*. [S.l.]: Springer-Verlag Berlin and Heidelberg GmbH & amp, 2000. 19

-
- 46 STROGATZ, S. H. Exploring complex networks. *Nature*, Nature Publishing Group, v. 410, n. 6825, p. 268–276, 2001. 25
- 47 DOROGOVTSSEV, S. N.; MENDES, J. F. Evolution of networks. *Advances in Physics*, Taylor & Francis, v. 51, n. 4, p. 1079–1187, 2002. 26
- 48 FERBER, C. V. et al. Public transport networks: empirical analysis and modeling. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 68, n. 2, p. 261–275, 2009. 26, 31, 32, 33, 39, 46, 47, 55, 99
- 49 FREEMAN, L. C.; BORGATTI, S. P.; WHITE, D. R. Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, Elsevier, v. 13, n. 2, p. 141–154, 1991. 27
- 50 ESTRADA, E.; HATANO, N.; BENZI, M. The physics of communicability in complex networks. *Physics Reports*, Elsevier, v. 514, n. 3, p. 89–119, 2012. 27
- 51 LI, W.; CAI, X. Statistical analysis of airport network of china. *Physical Review E*, APS, v. 69, n. 4, p. 046106, 2004. 27
- 52 PORTA, S.; CRUCITTI, P.; LATORA, V. The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 369, n. 2, p. 853–866, 2006. 27
- 53 SCHNEIDER, C. M. et al. Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences of the United States of America*, National Academy of Sciences, v. 108, n. 10, p. 3838–3841, 2011. 28, 29
- 54 BERCHE, B. et al. Resilience of public transport networks against attacks. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 71, n. 1, p. 125–137, 2009. 28
- 55 DERRIBLE, S.; KENNEDY, C. The complexity and robustness of metro networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 389, n. 17, p. 3678–3691, 2010. 28
- 56 ZHANG, L. et al. Urban bus transport network optimization from complex network. In: IEEE. *Modelling, Identification & Control (ICMIC), 2012 Proceedings of International Conference on*. [S.l.], 2012. p. 1159–1164. 28
- 57 MOREIRA, A. A. et al. How to make a fragile network robust and vice versa. *Physical Review Letters*, APS, v. 102, n. 1, p. 018701, 2009. 29
- 58 SCHIEBER, T. A. et al. Quantification of network structural dissimilarities. *Nature Communications*, Nature Publishing Group, v. 8, p. 13928, 2017. 30
- 59 SIMÓN, P. Colomer-de; BOGUÑÁ, M. Double percolation phase transition in clustered complex networks. *Physical Review X*, APS, v. 4, n. 4, p. 041020, 2014. 31
- 60 LORDAN, O. et al. Robustness of the air transport network. *Transportation Research Part E: Logistics and Transportation Review*, Elsevier, v. 68, p. 155–163, 2014. 31

- 61 LORDAN, O.; SALLAN, J. M.; SIMO, P. Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda. *Journal of Transport Geography*, Elsevier, v. 37, p. 112–120, 2014. 31
- 62 LATORA, V.; MARCHIORI, M. Is the boston subway a small-world network? *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 314, n. 1, p. 109–113, 2002. 31, 33
- 63 SEN, P. et al. Small-world properties of the indian railway network. *Physical Review E*, APS, v. 67, n. 3, p. 036106, 2003. 31, 32, 33, 55
- 64 SIENKIEWICZ, J.; HOLYST, J. A. Statistical analysis of 22 public transport networks in poland. *Physical Review E*, APS, v. 72, n. 4, p. 046127, 2005. 31, 32, 33, 34
- 65 CHEN, Y.-Z.; LI, N.; HE, D.-R. A study on some urban bus transport networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 376, p. 747–754, 2007. 31, 33, 34
- 66 CHEN, Y.-Z.; LI, N. The randomly organized structure of urban ground bus-transport networks in china. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 386, n. 1, p. 388–396, 2007. 31, 33
- 67 XU, X. et al. Scaling and correlations in three bus-transport networks of china. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 374, n. 1, p. 441–448, 2007. 31, 33
- 68 ZHEN-TAO, Z. et al. An evolutionary model of urban bus transport network based on b-space. *Chinese physics B*, IOP Publishing, v. 17, n. 8, p. 2874, 2008. 31, 33, 34
- 69 KALUZA, P. et al. The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, The Royal Society, v. 7, n. 48, p. 1093–1103, 2010. 31
- 70 BERCHE, B. et al. Transportation network stability: A case study of city transit. *Advances in Complex Systems*, World Scientific, v. 15, n. supp01, p. 1250063, 2012. 33
- 71 DIMITROV, S. D.; CEDER, A. A. A method of examining the structure and topological properties of public-transport networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 451, p. 373–387, 2016. 33
- 72 STDF - Secretaria de Estado de Transporte do Distrito Federal. *Plano Diretor de Transporte Urbano e Mobilidade do Distrito Federal e Entorno*. Brasília-DF: [s.n.], 2009. Accessed in 10/2014. 41, 70, 97
- 73 SAGARRA, O.; VICENTE, C. P.; DÍAZ-GUILERA, A. Statistical mechanics of multiedge networks. *Physical Review E*, APS, v. 88, n. 6, p. 062806, 2013. 56, 91
- 74 JR, E. M.; OJIMA, R. Pendularidade e vulnerabilidade na região metropolitana de campinas: repercussões na estrutura e no habitar urbano. *Revista Brasileira de Estudos Urbanos e Regionais*, v. 16, n. 2, p. 185–204, 2014. 58

- 75 OJIMA, R. Fronteiras metropolitanas: um olhar a partir dos movimentos pendulares. *Revista paranaense de desenvolvimento*, Instituto Paranaense de Desenvolvimento Econômico e Social, n. 121, p. 115–132, 2011. 58
- 76 CODEPLAN - Companhia de Planejamento do Distrito Federal. *Pesquisa distrital por amostra de domicílios (PDAD)–2010/2011*. Brasília-DF: [s.n.], 2011. 59, 95, 96, 97
- 77 VUCHIC, V. R. *Urban transit: operations, planning, and economics*. [S.l.: s.n.], 2005. 80
- 78 KRUECKEBERG, D. A.; SILVERS, A. L. *Urban planning analysis: Methods and models*. [S.l.]: John Wiley & Sons, 1974. 81, 82, 83, 84, 91
- 79 MCNALLY, M. G. The four-step model. In: *Handbook of Transport Modelling: 2nd Edition*. [S.l.]: Emerald Group Publishing Limited, 2007. p. 35–53. 84
- 80 WILSON, A. G. A statistical theory of spatial distribution models. *Transportation research*, Pergamon, v. 1, n. 3, p. 253–269, 1967. 86, 87, 88, 90
- 81 STDF - Secretaria de Estado de Transporte do Distrito Federal. *Relatório da Pesquisa de Quantitativos de Passageiros e de suas formas de pagamento de passagens no Serviço de Transporte Público Alternativo de Condomínios do Distrito Federal - STPAC/DF*. Brasília-DF: [s.n.], 2005. 92, 93, 104
- 82 JENELIUS, E. Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*, Elsevier, v. 17, n. 3, p. 234–244, 2009. 97
- 83 GRUS, J. *Data Science from Scratch: First Principles with Python*. [S.l.]: " O'Reilly Media, Inc.", 2015. 139
- 84 MCFARLAND, D. A.; LEWIS, K.; GOLDBERG, A. Sociology in the era of big data: The ascent of forensic social science. *The American Sociologist*, Springer, v. 47, n. 1, p. 12–35, 2016. 140
- 85 WALLER, M. A.; FAWCETT, S. E. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, Wiley Online Library, v. 34, n. 2, p. 77–84, 2013. 140
- 86 DAVENPORT, T. H.; PATIL, D. Data scientist. *Harvard business review*, v. 90, n. 5, p. 70–76, 2012. 140
- 87 ZAKI, M. J.; JR, W. M.; MEIRA, W. *Data mining and analysis: fundamental concepts and algorithms*. [S.l.]: Cambridge University Press, 2014. 141

Apêndice - Centro de massa

Podemos representa a concentração das viagens ao longo do dia como um caso contínuo, se supormos que temos uma grande quantidade de dados, Figura (5.2):

$$R_{CM} = \sqrt{X_{CM}^2 + Y_{CM}^2}, \quad (\text{A.1})$$

com valores de X_{CM} e Y_{CM} ,

$$\frac{\int_{\theta_1}^{\theta_2} r \cos(\theta) d\theta}{\int_{\theta_1}^{\theta_2} r d\theta} \quad (\text{A.2})$$

$$\frac{\int_{\theta_1}^{\theta_2} r \sin(\theta) d\theta}{\int_{\theta_1}^{\theta_2} r d\theta}. \quad (\text{A.3})$$

Tal que,

$$R_{CM} = \sqrt{\left[\frac{\int_{\theta_1}^{\theta_2} r \cos(\theta) d\theta}{\int_{\theta_1}^{\theta_2} r d\theta} \right]^2 + \left[\frac{\int_{\theta_1}^{\theta_2} r \sin(\theta) d\theta}{\int_{\theta_1}^{\theta_2} r d\theta} \right]^2} \quad (\text{A.4})$$

$$= \sqrt{\frac{[\sin(\theta)|_{\theta_1}^{\theta_2}]^2}{[\theta_2 - \theta_1]^2} + \frac{[\cos(\theta)|_{\theta_1}^{\theta_2}]^2}{[\theta_2 - \theta_1]^2}} \quad (\text{A.5})$$

$$= \sqrt{\frac{[\sin(\theta_2) - \sin(\theta_1)]^2 + [\cos(\theta_1) - \cos(\theta_1)]^2}{[\theta_2 - \theta_1]^2}} \quad (\text{A.6})$$

$$= \sqrt{\frac{\sin^2(\theta_2) - 2 \sin(\theta_2) \sin(\theta_1) + \sin^2(\theta_1) + \cos^2(\theta_1) - 2 \cos(\theta_1) \cos(\theta_2) + \cos^2(\theta_2)}{[\theta_2 - \theta_1]^2}}, \quad (\text{A.7})$$

usando as identidades trigonométricas,

$$\sin^2(\alpha) + \cos^2(\alpha) = 1$$

$$\cos(\alpha \pm \beta) = \sin(\alpha) \sin(\beta) \mp \cos(\alpha) \cos(\beta)$$

$$\cos(2\alpha) = 1 - 2 \sin^2(\alpha),$$

e $\Delta\theta = \theta_2 - \theta_1$, simplificamos a equação,

$$R_{CM} = \sqrt{\frac{1 - 2 \cos(\theta_2 - \theta_1) + 1}{[\theta_2 - \theta_1]^2}} \quad (\text{A.8})$$

$$= \frac{1}{\Delta\theta} \sqrt{2[1 - \cos(\Delta\theta)]} \quad (\text{A.9})$$

$$= \frac{1}{\Delta\theta} \sqrt{2 \left[1 - \left[1 - 2 \sin^2 \left(\frac{\Delta\theta}{2} \right) \right] \right]} \quad (\text{A.10})$$

$$= \frac{1}{\Delta\theta} \sqrt{4 \sin^2 \left(\frac{\Delta\theta}{2} \right)}, \quad (\text{A.11})$$

que resulta na equação,

$$R_{CM} = \frac{2}{\Delta\theta} \sin \left(\frac{\Delta\theta}{2} \right), \quad (\text{A.12})$$

que usamos em 5.13 da seção 5.3, onde $\Delta\theta = 2\pi(h_2 - h_1)/24$.

Apêndice - Série de Fourier

Uma função periódica pode ser representada por somas de senos e cossenos, a mais usada é a série de Fourier,

$$X(t) = \sum_{n=0}^{\infty} [a_n \cos(\omega_n t) + b_n \sin(\omega_n t)] , \quad (\text{B.1})$$

onde a_n e b_n dependem da função $X(t)$. A frequência é dada por $\omega_n = \frac{2\pi n}{T}$, onde T é o período.

Uma função é dita periódica com período $T > 0$ se o domínio de X contém $t + T$ sempre que contiver t e se $X(t + T) = X(t)$ para todo valor de t .

Senos e cossenos são funções que obedecem ao critério de ortogonalidade. Se o produto interno (u, v) de duas funções reais u e v no intervalo de $\alpha \leq t \leq \beta$ é definido por

$$(u, v) = \int_{\alpha}^{\beta} u(t)v(t)dt . \quad (\text{B.2})$$

As funções u e v são ditas ortogonais se o seu produto interno for nulo,

$$\int_{\alpha}^{\beta} u(t)v(t)dt = 0 . \quad (\text{B.3})$$

Um conjunto de funções é dito conjunto ortogonal se cada par de funções diferentes pertencentes ao conjunto é ortogonal. Assim, as funções $\sin(2\pi nt/T)$ e $\cos(2\pi nt/T)$, com $n = 1, 2, \dots$ formam um conjunto ortogonal de funções no intervalo $\alpha \leq t \leq \beta$. Pois satisfazem as relações de ortogonalidade:

$$\int_{-T/2}^{T/2} \cos\left(\frac{2\pi mt}{T}\right) \cos\left(\frac{2\pi nt}{T}\right) dt = \begin{cases} 0, & m \neq n; \\ T/2, & m = n; \end{cases} \quad (\text{B.4})$$

$$\int_{-T/2}^{T/2} \cos\left(\frac{2\pi mt}{T}\right) \sin\left(\frac{2\pi nt}{T}\right) dt = 0 \quad \text{para todo } m, n; \quad (\text{B.5})$$

$$\int_{-T/2}^{T/2} \sin\left(\frac{2\pi mt}{T}\right) \sin\left(\frac{2\pi nt}{T}\right) dt = \begin{cases} 0, & m \neq n; \\ T/2, & m = n. \end{cases} \quad (\text{B.6})$$

Esses resultados podem ser obtidos pela integração direta,

$$\int_{-T/2}^{T/2} \sin\left(\frac{2\pi mt}{T}\right) \sin\left(\frac{2\pi nt}{T}\right) dt = \frac{1}{2} \int_{-T/2}^{T/2} \cos\left(\frac{2\pi(m-n)t}{T}\right) - \cos\left(\frac{2\pi(m+n)t}{T}\right) dt \quad (\text{B.7})$$

$$= \frac{1}{2} \frac{T}{2\pi} \left[\frac{\sin\left(\frac{2\pi(m-n)t}{T}\right)}{m-n} - \frac{\sin\left(\frac{2\pi(m+n)t}{T}\right)}{m+n} \right]_{-T/2}^{T/2} \quad (\text{B.8})$$

$$= 0, \quad (\text{B.9})$$

desde que $m+n$ e $m-n$ sejam diferentes de zero. Como m e n são positivos, $m+n \neq 0$. Por outro lado, se $m-n=0$, então $m=n$ e a integral deve ser calculada por

$$\int_{-T/2}^{T/2} \sin\left(\frac{2\pi mt}{T}\right) \sin\left(\frac{2\pi nt}{T}\right) dt = \int_{-T/2}^{T/2} \left(\sin\frac{2\pi mt}{T}\right)^2 dt \quad (\text{B.10})$$

$$= \frac{1}{2} \int_{-T/2}^{T/2} \left[1 - \cos\frac{4\pi mt}{T}\right] dt \quad (\text{B.11})$$

$$= \frac{1}{2} \left[t - \frac{\sin(4\pi mt/T)}{4\pi m/T} \right]_{-T/2}^{T/2} \quad (\text{B.12})$$

$$= T. \quad (\text{B.13})$$

Apêndice - Simulação MGA para viagens reais

Fizemos a simulação do Modelo Gravitacional Alterado usando o número de embarques como o número de pessoas na origem e o número de desembarque como o número de pessoas no destino em cada rota e multiplicamos pela função $P_{ij} = \exp(-b^2/d_{ij})$. Temos os gráficos da distribuição quando P_i^E independente de P_j^D , ou seja, $P_{ij} = P_i^E P_j^D$. Fizemos a simulação usamos o MO - Modelo Gravitacional, com $P_{ij} = E_i D_i \exp(-b^2 d_{ij}^2)$. E finalmente, fizemos a simulação para o MGA - Modelo Gravitacional Alterado, com $P_{ij} = E_i D_i \exp(-b^2/d_{ij})$, onde d_{ij} é a distância em arestas do nó s_i a s_j na viagem de cada rota.

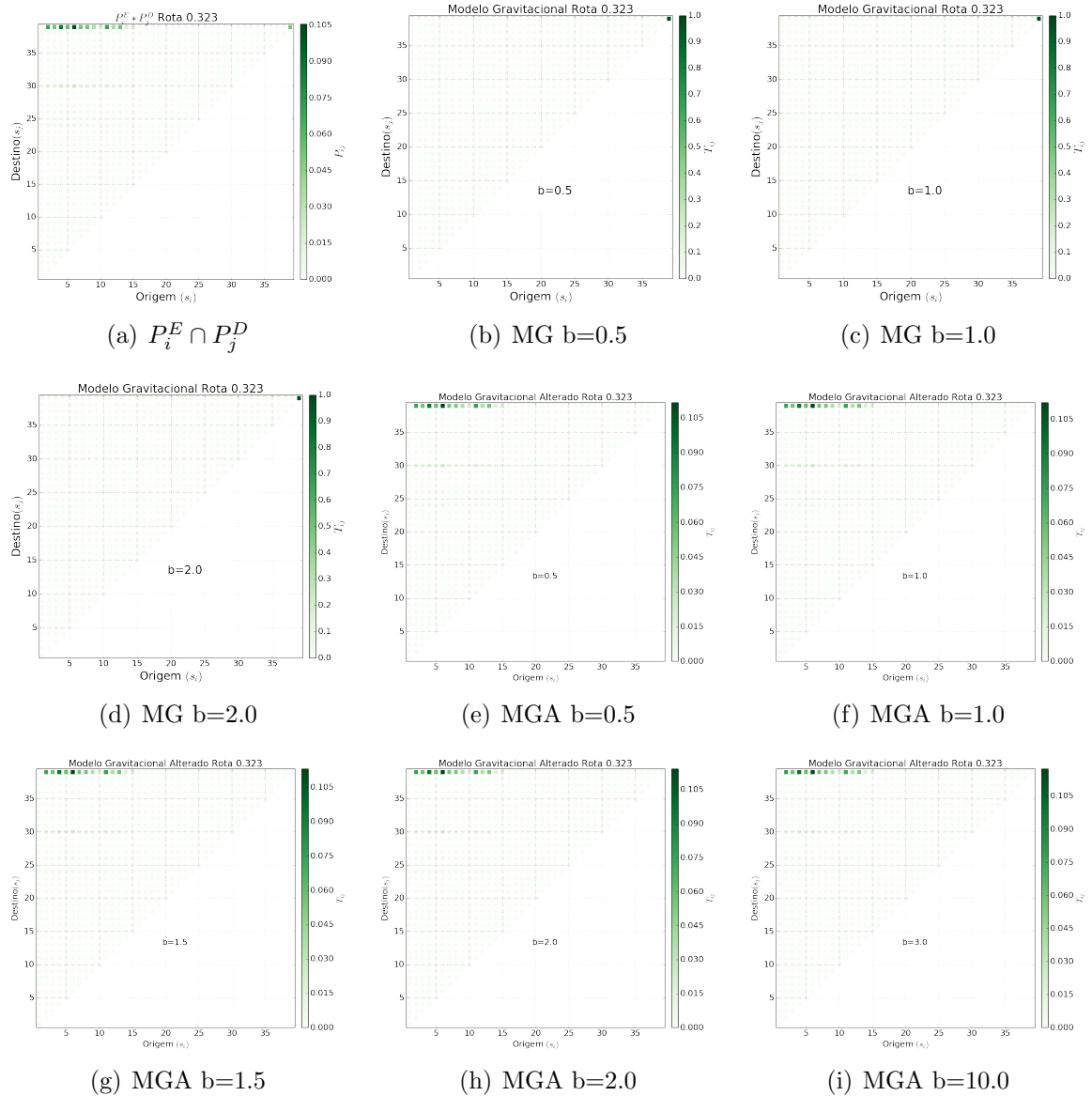


Figura C.1 - Distribuição de passageiros em uma viagem da rota 0.323. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.

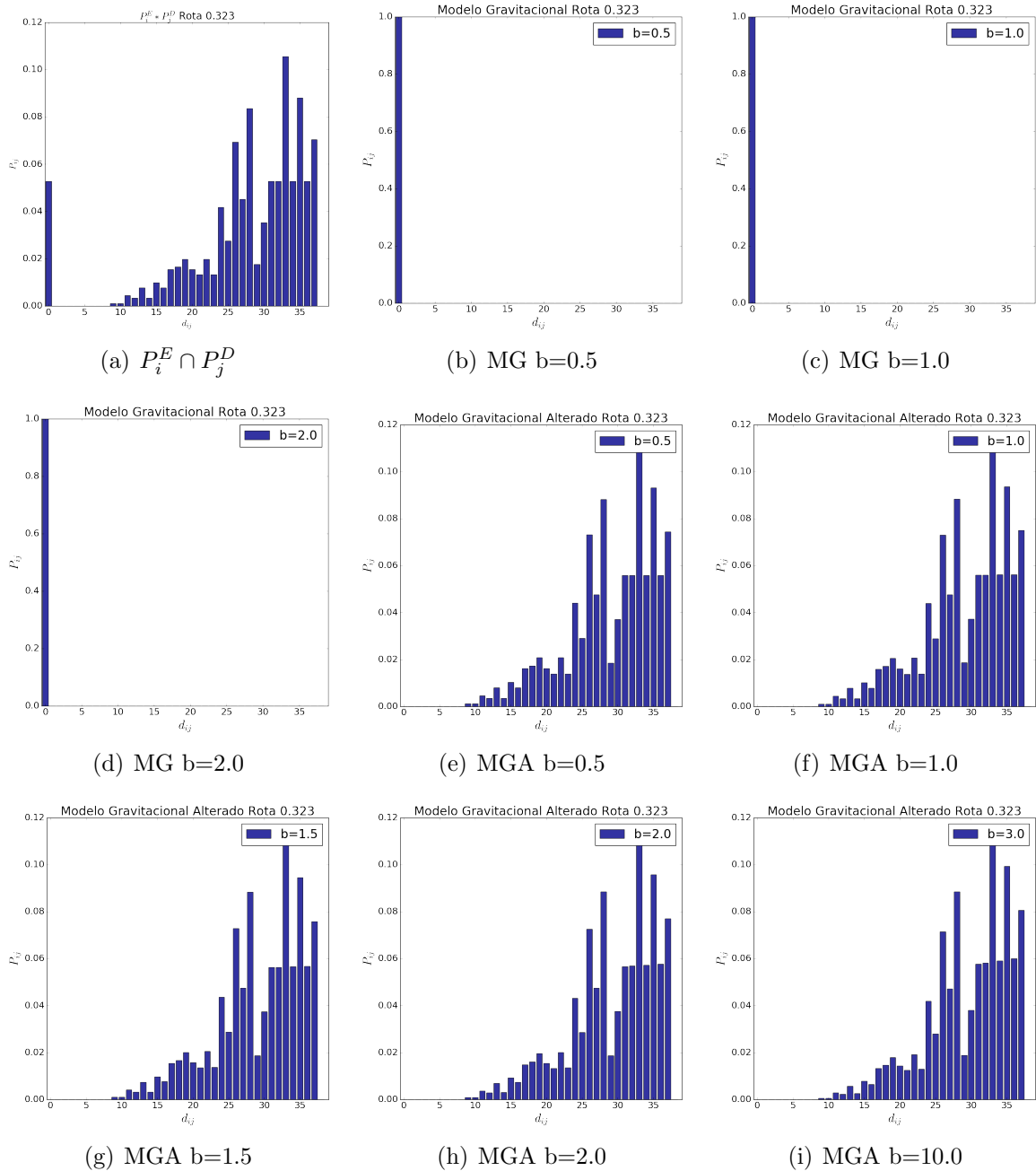


Figura C.2 - Distribuição de passageiros simulado para a rota 0.323.

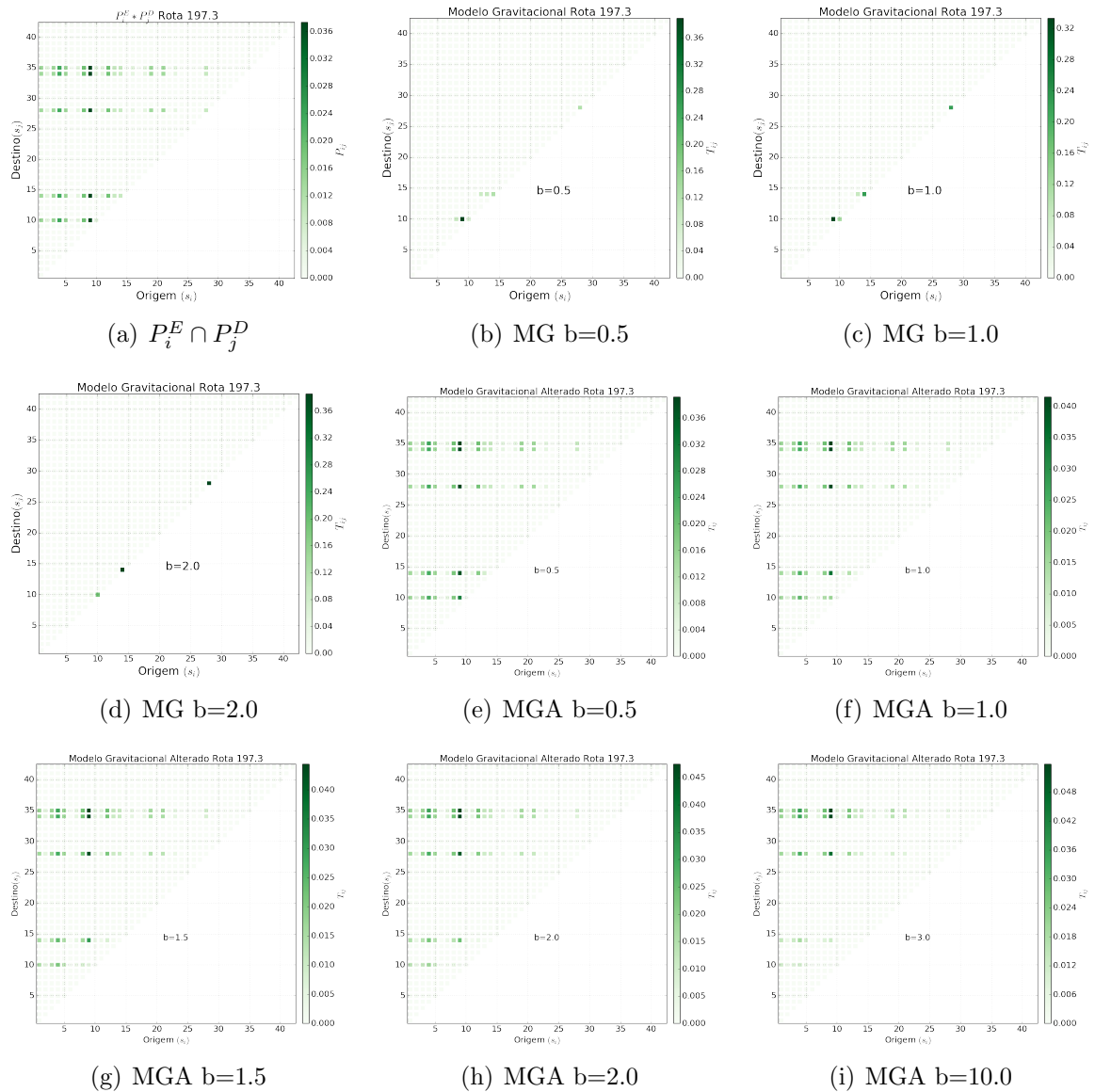


Figura C.3 - Distribuição de passageiros em uma viagem da rota 197.3. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.

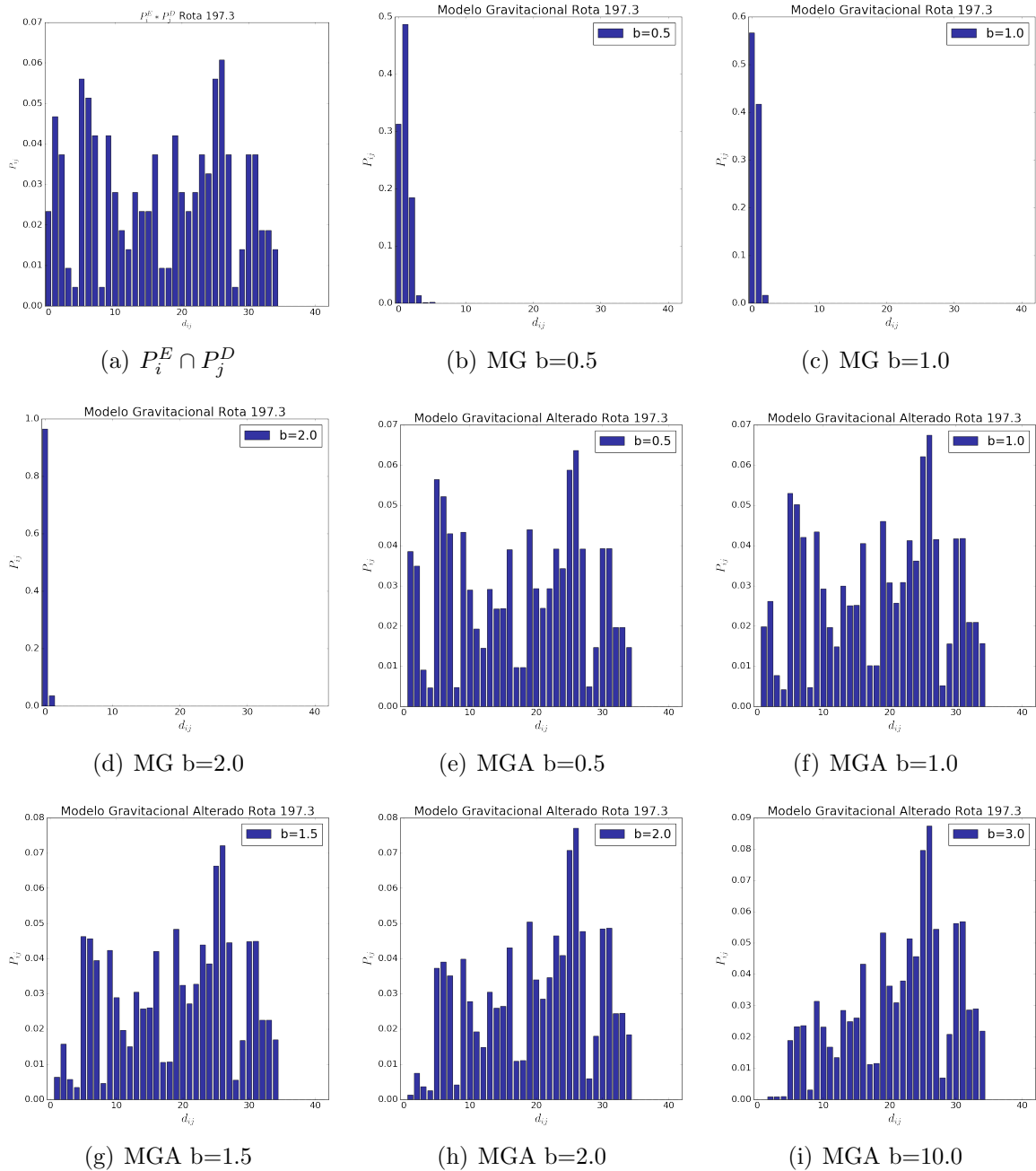


Figura C.4 - Distribuição de passageiros simulado para a rota 197.3.

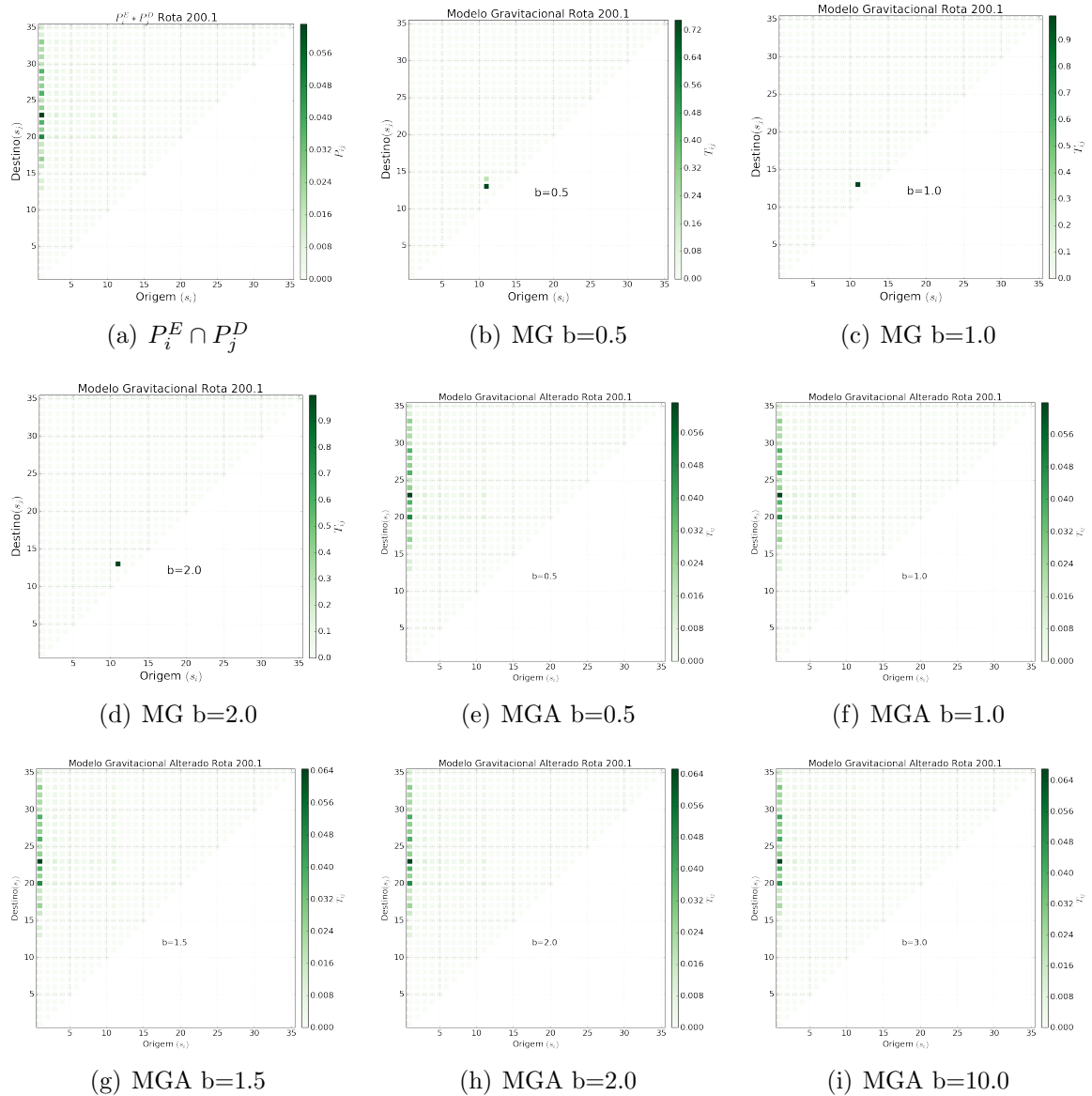


Figura C.5 - Distribuição de passageiros em uma viagem da rota 200.1. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.

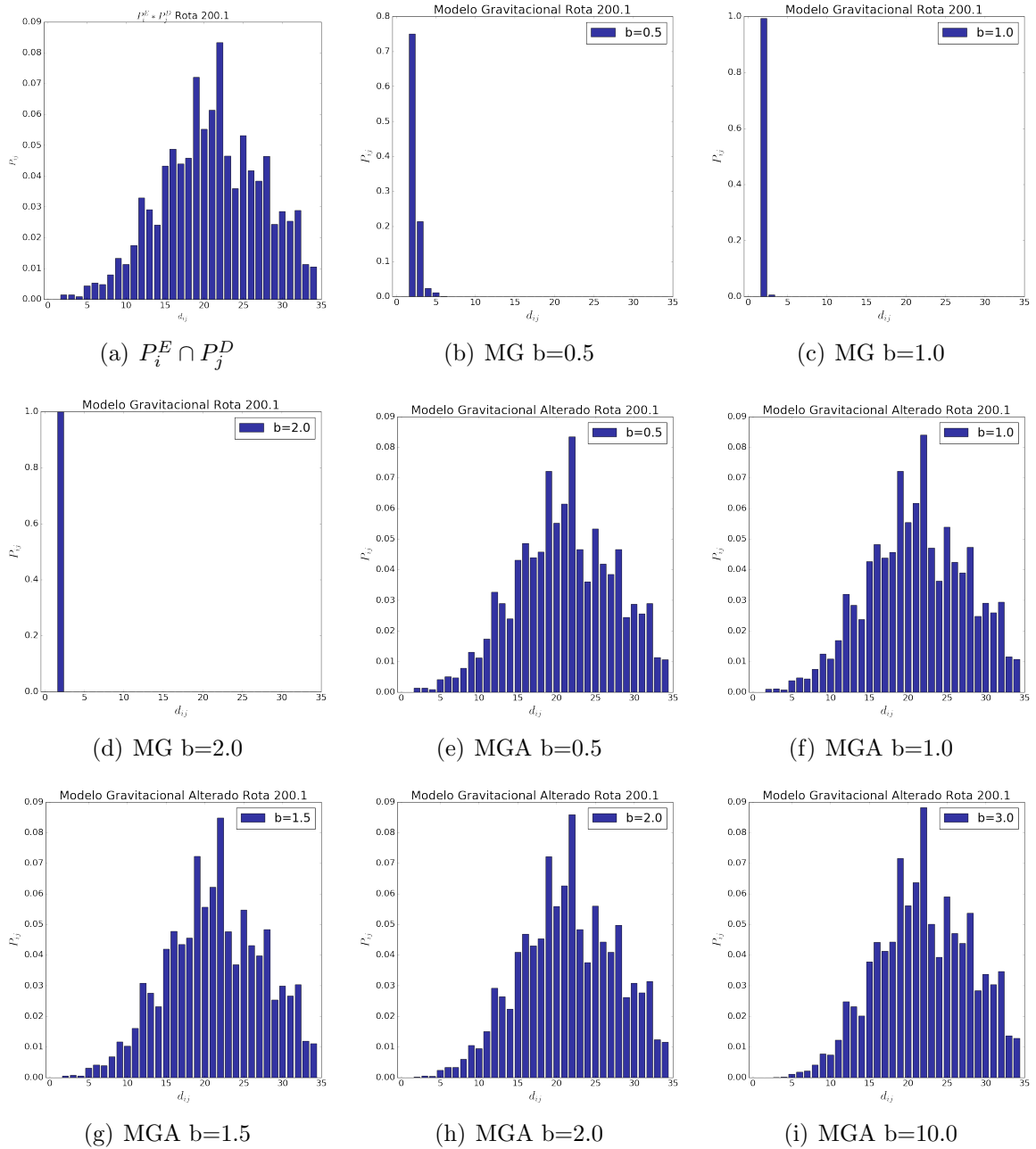


Figura C.6 - Distribuição de passageiros simulado para a rota 200.1.

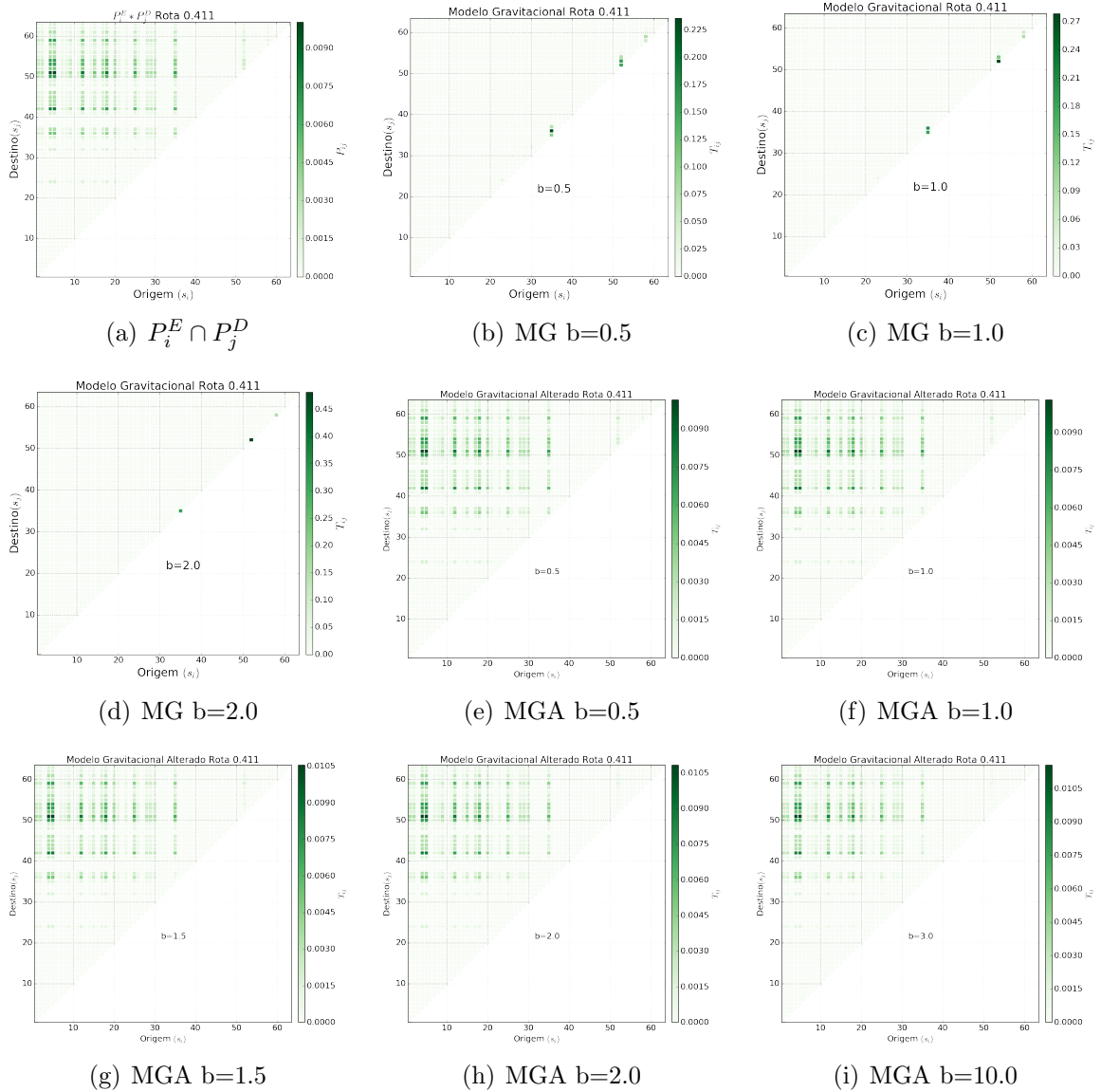


Figura C.7 - Distribuição de passageiros em uma viagem da rota 0.411. Simulado para passageiros indistinguíveis, com o evento de desembarque independente do embarque, para MG - Modelo Gravitacional e MGA - Modelo Gravitacional Alterado.

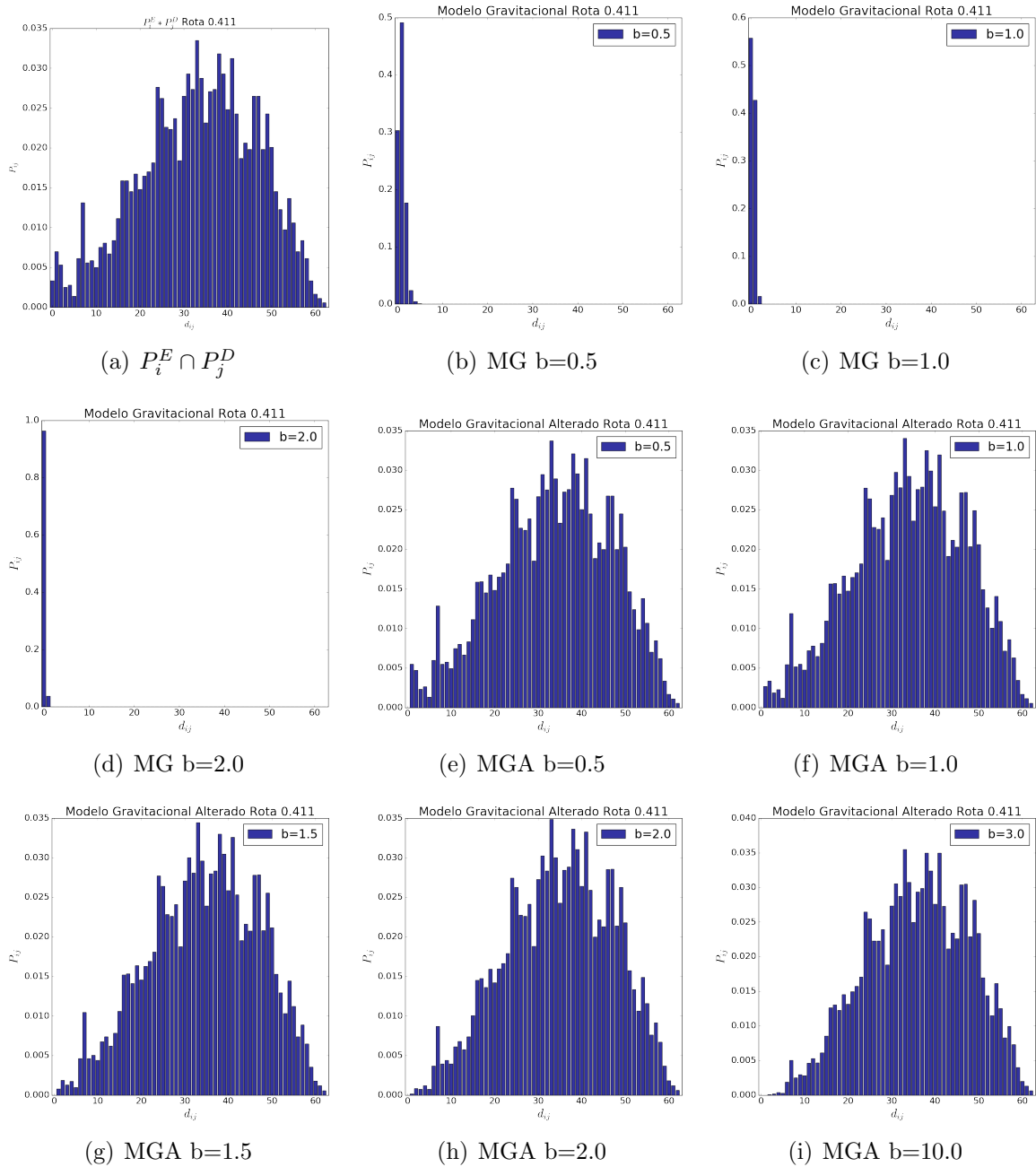


Figura C.8 - Distribuição de passageiros simulado para a rota 0.411.

Apêndice - Mineração de dados

A humanidade passou pelo desafio de poder processar muitos dados em pouquíssimo tempo e de armazenar muitos dados gerados a todo momento. E agora, passa pela dificuldade de poder aproveitar essa imensa quantidade de dados produzidos e armazenados para gerar alguma informação útil. Estamos lidando com a ciência de megadados, que deriva do termo em inglês *Big Data*. O significado de 'mega' pode ser interpretado de três maneiras: o volume de informações que os sistemas devem ser inseridos, processados e disseminados; a variedade e a complexidade dos tipos de informação manipulada; e a velocidade ou taxa em que as informações fluem para dentro ou para fora.

A todo momento sites acompanham cada clique de cada usuário. Telefones inteligentes registram a localidade e velocidade a cada segundo de todo dia. Possuem aplicativos de medidas que gravam todo tipo de dado: batimento cardíaco, hábitos de movimento, dieta, padrões de sono. Os carros inteligentes possuem computadores que coletam hábitos de condução, as casas inteligentes coletam hábitos de vida e os comércios inteligentes coletam hábitos de compra. Ou seja, mas vivemos em um mundo que está se afogando em dados [83].

A própria Internet representa um enorme gráfico de conhecimento que contém uma enorme enciclopédia de referências cruzadas; Bancos de dados específicos de domínio sobre filmes, música, resultados esportivos, máquinas de pinball, memes e cocktails; E muitas estatísticas do governo [83].

O novo desafio consiste em encontrar alguma informação útil que muitas vezes não estão organizados em bancos de dados estruturados, mas espalhados em vários bancos e sistemas de várias entidades, inclusive em nuvens. Dessa forma, a ciência de dados tem sido demandada nos últimos anos justamente pela dificuldade que se integrar dados de distintas fontes e de tratá-los de forma inteligente.

Outra etapa da informação que passamos foi da publicidade de dados governamentais, dados públicos de diferentes órgãos, e a necessidade das informações estatais serem divulgadas com a política de transparência. No entanto, cada órgão ou instituição desenvolveu seu banco de dados e o estruturou de formas não padronizadas.

Embora atendam bem às necessidades dessas instituições em particular, e estejam facilmente disponíveis para qualquer pessoa, um cidadão comum não conseguiria com facilidade cruzar dados de duas fontes distintas para obter uma informação importante, pois não se objetivou qualquer tipo de integração de dados entre vários bancos, mas

apenas o atendimento das necessidades locais de cada instituição ou empresa. Não raros são os casos em que uma única instituição possui dois ou mais sistemas de informação que não são interoperáveis, gerando dificuldades de se levantar informações relevantes para a tomada de decisões.

Quando tratamos de dados de diversas fontes distintas, estamos lidando com dados não padronizados, ou melhor, padronizados individualmente para cada sistema em particular, mas não entre vários sistemas e banco. Nesse contexto, surge a necessidade de um novo tipo de profissional que tenha habilidades de separar dados importantes de dados irrelevantes para obter informações úteis. A nova habilidade exigida para esse fim se chama mineração de dados, que consiste literalmente em 'separar o joio do trigo'. Uma ferramenta particularmente popular para esse fim é o Python, que possui uma sintaxe simples e descomplicada, inclusive para iniciantes em computação.

Essas novas habilidades em busca de informações no cruzamento de dados de diferentes sistemas e bancos de dados, são encontrados muitas vezes em profissionais de áreas que trabalham com formação de padrão em sistemas complexos, tais como físicos, matemáticos, biólogos, que estão acostumados a enxergar padrões. Assim, encontrar uma chave entre duas tabelas de dois bancos distintos é como traçar uma aresta entre dois nós de uma rede. Ou seja, é natural que para a ciência de dados seja empregado métodos de redes complexas, inclusive, para tratar redes reais, como citados em [84].

Os cientistas de dados precisam de um profundo domínio do conhecimento e de um amplo conjunto de habilidades analíticas [85]. Se a sua organização armazena vários petabytes de dados, se as informações mais críticas para o seu negócio residirem em formulários que não sejam linhas e colunas de números, ou se responder a sua maior questão envolveria uma mistura de vários esforços analíticos, então haverá uma oportunidade em ciência de mega dados [86]. Aliar a ciência de dados com análise preditiva e mega dados parece ser o futuro, inclusive para gestão de logística, por exemplo, [85].

A mineração de dados é o processo de descobrir padrões perspicazes, interessantes e inovadores, bem como modelos descritivos, compreensíveis e preditivos a partir de dados em larga escala. É um campo interdisciplinar que mescla conceitos de áreas afins como sistemas de banco de dados, estatísticas, inteligência artificial e reconhecimento de padrões. Faz parte de um processo de descoberta de conhecimento maior, que inclui tarefas de pré-processamento como extração de dados, limpeza de dados, fusão de dados, redução de dados e construção de recursos, bem como etapas de pós-processamento como interpretação de padrões e modelos, hipóteses, confirmação e geração, e assim por diante. Este processo de descoberta de conhecimento e de mineração de dados tende a ser altamente

interativo e iterativo. [87].

Portanto, para encontrarmos grande parte das informações necessárias para a construção das redes PTN do DF, desenvolvemos algoritmos em Python com base no método de mineração de dados para cruzar dados de várias fontes distintas. Assim, para realizar cruzamento de dados pertencentes a duas ou mais base de dados, usamos a ferramenta Pandas. Uma das bibliotecas do Python para a ciência de dados é o Pandas, que contém estrutura de dados de alto nível e ferramentas de manipulação designado para fazer análise de dados rápido e facilmente. Pandas é construído sob Numpy, realiza com muita facilidade operações como *merge* e outras encontradas em ferramentas de base de dados, como aqueles baseados em SQL, desde que os dados estejam estruturas em séries ou tabelas. As bibliotecas de redes complexas como o NetworkX e Gephi foram usadas para o cálculo de propriedades topológicas de redes. E para gerar os gráficos usamos a biblioteca Matplotlib, que permite fazer gráficos na própria linguagem do Python. Essa biblioteca usa bastante recursos do Numpy, pacote do para computação científica do Python.