



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21c/2018/04.17.13.57-TDI

MOBILIDADE URBANA EM UM DIA TÍPICO - (GEO)GRAFOS VARIANTES NO TEMPO

Jéssica Domingues Santos

Dissertação de Mestrado do
Curso de Pós-Graduação em
Computação Aplicada, orientada
pelos Drs. Leonardo Bacelar Lima
Santos, e Marcos Gonçalves Quiles,
aprovada em 27 de março de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3QTL422>>

INPE
São José dos Campos
2018

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GBDIR)

Serviço de Informação e Documentação (SESID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

E-mail: pubtc@inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação (CPG)

Membros:

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (COCST)

Dr. André de Castro Milone - Coordenação-Geral de Ciências Espaciais e Atmosféricas (CGCEA)

Dra. Carina de Barros Melo - Coordenação de Laboratórios Associados (COCTE)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia e Tecnologia Espacial (CGETE)

Dr. Hermann Johann Heinrich Kux - Coordenação-Geral de Observação da Terra (CGOBT)

Dr. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CGCPT)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SESID)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SESID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Duca Barbedo - Serviço de Informação e Documentação (SESID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SESID)

EDITORAÇÃO ELETRÔNICA:

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SESID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SESID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21c/2018/04.17.13.57-TDI

MOBILIDADE URBANA EM UM DIA TÍPICO - (GEO)GRAFOS VARIANTES NO TEMPO

Jéssica Domingues Santos

Dissertação de Mestrado do
Curso de Pós-Graduação em
Computação Aplicada, orientada
pelos Drs. Leonardo Bacelar Lima
Santos, e Marcos Gonçalves Quiles,
aprovada em 27 de março de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3QTL422>>

INPE
São José dos Campos
2018

Dados Internacionais de Catalogação na Publicação (CIP)

Santos, Jéssica Domingues.
S59m Mobilidade urbana em um dia típico - (Geo)grafos variantes no tempo / Jéssica Domingues Santos. – São José dos Campos : INPE, 2018.
xxi + 62 p. ; (sid.inpe.br/mtc-m21c/2018/04.17.13.57-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2018.

Orientadores : Drs. Leonardo Bacelar Lima Santos, e Marcos Gonçalves Quiles.

1. Redes complexas. 2. Mobilidade urbana. 3. (Geo)grafos.
I.Título.

CDU 004.72:711.7



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

Aluno (a): **Jéssica Domingues Santos**

Título: "MOBILIDADE URBANA EM UM DIA TÍPICO - (GEO) GRAFOS VARIANTES NO TEMPO"

Aprovado (a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título de **Mestre** em **Computação Aplicada**

Dr. Alan James Peixoto Calheiros



Presidente / INPE / Cachoeira Paulista - SP

Participação por Vídeo - Conferência

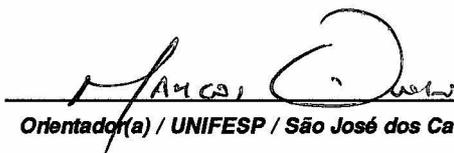
Dr. Leonardo Bacelar Lima Santos



Orientador(a) / INPE / São José dos Campos - SP

Participação por Vídeo - Conferência

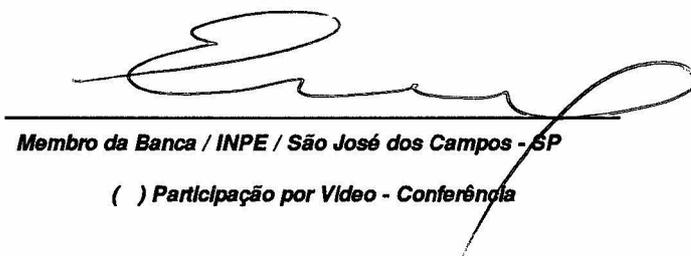
Dr. Marcos Gonçalves Quiles



Orientador(a) / UNIFESP / São José dos Campos - SP

Participação por Vídeo - Conferência

Dr. Elbert Einstein Nehrer Macau



Membro da Banca / INPE / São José dos Campos - SP

Participação por Vídeo - Conferência

Dra. Luciana de Resende Londe



Convidado(a) / CEMADEN / São José dos Campos - SP

Participação por Vídeo - Conferência

Dr. Tiago José de Carvalho



Convidado(a) / IFSP / Campinas - SP

Participação por Vídeo - Conferência

Este trabalho foi aprovado por:

maioria simples

unanimidade

“Não tenha medo de sonhar e lutar!”.

PROF. MSC. DANIEL MERLI LAMOSA
em “*Universidade de Taubaté - Unitau*”, 2013

A meus pais, Geni e Adilson.

Dedicatoria

AGRADECIMENTOS

Agradeço, primeiramente, aos meus pais que sempre me apoiaram, especialmente minha amada mãe Geni Domingues que com todo amor, carinho e ensinamentos sempre me deu um grande abraço, o melhor olhar, as melhores palavras e “os piores puxões de orelhas”, e meu querido pai Adilson Moreira dos Santos, que com seu jeito brincalhão e alegre me mostra como encarar os problemas sorrindo, para eles meu coração é cheio de agradecimentos, amor e ternura.

Ao meu amor Daniel por estar sempre do meu lado, não deixando eu desistir dos meus sonhos e me apoiando incondicionalmente, agradeço principalmente por ser meu melhor professor, me ensinando a viver a vida. Ele em todos momentos difíceis afastou qualquer dor ou angústia do meu coração com todo carinho e compreensão, me deu forças para continuar e me deu motivos para sonhar.

As crianças que são minha luz: Luiza, Vinícius e Ana Júlia. Eles transformaram o meu dia com brincadeiras, com risadas, com cada palavra nova que aprendem e aquecem meu coração com um amor puro.

Aos meus amigos no qual sempre encontrei carinho e estímulos.

Aos meus orientadores e professores Dr. Leonardo Bacelar Lima Santos e Dr. Marcos Gonçalves Quiles. Ao Prof. Dr. Leonardo por todo suporte, debates, paciência e pelos desafios propostos que contribuíram para o engrandecimento desse trabalho, agradeço toda a disponibilidade e incentivo na condução da pesquisa, sempre com interesse e seriedade. Ao Prof. Dr Marcos pelas importantes sugestões apresentadas a este trabalho.

A todos os professores que tive durante minha vida.

As funcionárias do programa de pós-graduação em computação aplicada do INPE, sempre eficientes e disponíveis.

Aos amigos que fiz durante as disciplinas, em especial ao Rolf Ezequiel Simões, que me ensinou a pegar ônibus em São José dos Campos, foi meu parceiro em trabalhos de matemática, computação, de café da manhã na ADC para encarar as disciplinas e dar risadas sinceras.

Ao João Eliakin por me estender a mão e me ajudar nos primeiros passos no desenvolvimento com a biblioteca *igraph*.

A Alice N. Marteli por me ensinar a criar mapas no *QuantumGIS*

Enfim, gostaria de agradecer ao CNPq pelo suporte financeiro e aos projetos CNPq 454267/2014-2, FAPESP 2015/50122-0 e DFG-IRTK 1740/2 por proporcionar suporte para publicações e eventos.

RESUMO

Um sistema complexo pode ser modelado através do conceito de rede complexa que, por meio de grafos, retrata um conjunto de conexões com características topológicas não completamente regulares nem mesmo completamente aleatórias. No clima, na biologia, na sociologia e em muitas outras áreas, essas características apresentam semelhanças permitindo a reutilização das soluções a partir da generalização dos modelos. Este trabalho utiliza Redes Complexas para estudar mobilidade urbana. O campo de estudo é a cidade de São José dos Campos, Brasil. Os dados reais de deslocamento consistem em uma pesquisa Origem-Destino: a cidade foi dividida em 55 zonas de tráfego e mais de 20 mil pessoas foram questionadas sobre o horário de partida e chegada de cada viagem. O desenvolvimento foi dividido em 3 etapas: pré-processamento, processamento e pós-processamento. No pré-processamento, um grafo de origem e destino com representação matricial em 3 dimensões, em linguagem C/C++, foi criado, no qual cada vértice representa uma zona de tráfego, e as arestas são ponderadas pelo fluxo de pessoas, com variação no tempo, totalizando 24 matrizes, uma para cada hora do dia. No processamento, em C, foi utilizado a biblioteca *igraph* para calcular as medidas topológicas tais como grau (número de conexões), coeficiente de aglomeração (redundância de conexões entre vizinhos) e diâmetro (maior menor distância) de uma rede de mobilidade ao longo de um dia típico e também aplicamos o algoritmo *walktrap* para detecção de comunidades. No pós-processamento, utilizando o conceito de (geo)grafos, grafos representados com geolocalização, foi desenvolvido a GeoCNet, ferramenta que permite a criação de um *shapefile* com as propriedades topológicas do grafo. A estrutura modular mostrou que no momento de alto fluxo a rede se torna muito conectada e a estrutura modular desaparece. Por outro lado, nos momentos de baixo fluxo, a rede apresenta estrutura modular bem definida e as sub-regiões da cidade são representadas pelas comunidades detectadas.

Palavras-chave: redes complexas. mobilidade urbana. (geo)grafos.

URBAN MOBILITY IN A TYPICAL DAY - TIME SERIES USING COMPLEX NETWORK

ABSTRACT

A complex system can be modeled through the complex network concept that, by using graphs, represents a connection set with topological characteristics which are neither completely regular nor completely random. In sociology, biology, urban mobility and many other areas those characteristics show similarities which are useful for the reuse of the solutions from the models generalization. This work focused in the problem of Urban mobility, which in an unplanned urban growth scenario might generate negative impacts, like traffic jams, air pollution and infrastructure flaws. Based on real data for the city of São José dos Campos, the mobility of a typical day was represented. These data consist of an Origin-Destination survey: the city was divided into 55 traffic zones and more than 20 thousand people were asked about the time of departure and arrival of each trip. The development was divided in 3 steps, pre-processing, processing and post-processing. In preprocessing, an origin destination graph was generated with a 3-dimension matrix representation, in language C++, in which each vertex represents a traffic zone and the edges are weighted by the flux of people, with 24 time variations, one for each hour of the day. In the processing, in C, the igraph library was used to calculate the topological properties such as degree (number of connections), clustering coefficient (neighbors redundancy) and diameter (longest distance) of a network of mobility over a typical day and we also applied the textit walktrap algorithm for community detection. In the post-processing, using the concept of (geo) graphs, graphs represented with geolocation, the GeoCNet was developed. It is a tool that allows the creation of a textit shapefile with the topological properties of the graph. The results of the dynamic behavior of urban mobility showed cohesive values of the calculated topological properties, highlighting areas with more concentrated flow and the analysis revealed that some communities persist throughout the day.

Keywords: Complex network. Urban Mobility. (geo)graph. GeoCNet.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Grafo Ponderado	5
2.2 Diâmetro de um grafo	6
2.3 a) Grafo com todos os vértices; b) Grafo com vértice c retirado; c) Grafo com o vértice d retirado.	6
2.4 a) Grafo com todas as arestas; b) Grafo com aresta eliminada desco- nectando os vértice b e c	7
2.5 Grau de um grafo	8
2.6 Estrutura de comunidade	9
3.1 55 zonas de tráfego	15
3.2 Vazios urbanos	16
3.3 Representação do último maior diâmetro	17
3.4 Representação matricial em 3 dimensões	18
3.5 A Matriz A representa a matriz inicial de fluxo; A matriz B representa a soma entre os vértices i e j : $A(i, j) + A(j, i)$ e a diagonal zerada	18
3.6 Ferramentas utilizadas no desenvolvimento da GeoCNet	20
3.7 (geo)grafo da cidade do Rio de Janeiro - RJ	21
3.8 Dados de chuva por Radar	22
4.1 Grafo da rede de um dia típico (24 horas)	23
4.2 (geo)grafo da rede de um dia típico (24 horas)	24
4.3 Histograma dos Graus da rede de 24 horas - O eixo Y representa a frequência de vértices e o eixo x representa o grau encontrado	25
4.4 Grau de cada vértice da rede com 24 horas: o eixo Y representa o grau e o eixo x representa o vértice	26
4.5 (geo)grafo - 24h	27
4.6 Tabela de atributo topológico	28
4.7 (geo)grafo com limiar crítico- 24h com grau de propriedade topológica	29
4.8 Gráficos gerados com intervalo de tempo igual a 15, 60 e 180 minutos, respectivamente.	30
4.9 Gráfico do Grau médio da rede com janela temporal de 60min.	31
4.10 Gráfico do Coeficiente de aglomeração médio com janela temporal de 60min.	32
4.11 Gráfico do diâmetro da rede com janela temporal de 60min.	33
4.12 Gráficos gerados com os dados normalizados da medidas	34

4.13	Limiares crítico	35
4.14	Gráficos gerados com os dados normalizados da medidas	36
4.15	(geo)grafo do acumulado do dia da cidade de São José dos Campos	38
4.16	(geo)grafo do acumulado do dia da cidade de São José dos Campos com limiar crítico	39
4.17	Modularidade para limiar 1	40
4.18	(geo)grafo para 0h da cidade de São José dos Campos	41
4.19	(geo)grafo para às 7h da cidade de São José dos Campos	42
4.20	(geo)grafo para às 12h da cidade de São José dos Campos	43
4.21	(geo)grafo para às 17h da cidade de São José dos Campos	44
4.22	(geo)grafo para às 22h da cidade de São José dos Campos	45
4.23	Modularidade para limiar crítico	46
4.24	(geo)grafo para 0h da cidade de São José dos Campos.	47
4.25	(geo)grafo para 7h da cidade de São José dos Campos	48
4.26	(geo)grafo para 12h da cidade de São José dos Campos.	49
4.27	(geo)grafo para 17h da cidade de São José dos Campos.	50
4.28	(geo)grafo para 22h da cidade de São José dos Campos.	51
4.29	Região sul com diferentes limiares e horários	52
4.30	Padrão da região sul	52

LISTA DE TABELAS

	<u>Pág.</u>
3.1 55 zonas de tráfego	13
3.2 55 zonas de tráfego	14
3.3 Igraph em C e as medidas da rede	19
4.1 Fator de normalização (valor máximo) das medidas topológicas da rede .	34
4.2 Fator de normalização (valor máximo) das medidas topológicas da rede .	36

LISTA DE ABREVIATURAS E SIGLAS

CBTT	– Congresso Brasileiro de Transporte e Trânsito.
CNMAC	– Congresso Nacional de Matemática Aplicada e Computacional.
CNPq	– Conselho Nacional de Desenvolvimento Científico e Tecnológico.
FAPESP	– Fundação de Amparo à Pesquisa do Estado de São Paulo.
GeoCNET	– Geographical Complex Network.
INPE	– Instituto Nacional de Pesquisas Espaciais.
INPI	– Instituto Nacional da Propriedade Industrial.
IPPLAN	– Instituto de Pesquisa, Administração e Planejamento.
SBRH	– Simpósio Brasileiro de Recursos Hídricos.
SIG	– Sistema de Informação Geográfica.

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 REVISÃO DA LITERATURA	5
2.1 Definições básicas de Teoria de Grafos	5
2.2 Aplicando teoria de grafos em mobilidade urbana	10
3 METODOLOGIA	13
3.1 Descrição dos dados	13
3.2 Pré-processamento dos dados	18
3.3 Desenvolvimento	19
3.3.1 GeoCNet	19
4 RESULTADOS OBTIDOS	23
4.1 Rede estática	23
4.1.1 Fluxo completo	23
4.1.2 Limiar crítico	28
4.2 Rede dinâmica	30
4.2.1 Fluxo completo	31
4.2.2 Limiar crítico	35
4.3 Comunidades	37
4.3.1 Rede Estática	38
4.3.2 Rede Dinâmica	40
4.4 Planejamento urbano	52
5 CONCLUSÕES	55
5.1 Principais contribuições	55
5.2 Perguntas científicas	56
5.3 Trabalhos futuros	57
5.4 Publicações	57
REFERÊNCIAS BIBLIOGRÁFICAS	59

1 INTRODUÇÃO

A interação e a circulação de pessoas em regiões urbanas ajudam a moldar a evolução das cidades ao longo do tempo. Essa evolução envolve questões culturais, socioeconômicas e ambientais. Os impactos dessa evolução se agravam com o crescimento dos grandes centros urbanos (KINDERSLEY, 2000; FERNANDES et al., 2015; TOMINAGA et al., 2009), quando emergem problemas como inundações e deslizamentos com impactos à vida humana e infraestruturas.

A complexidade das atividades que acontecem simultaneamente em uma cidade com milhares de residências, comércios, indústrias e intenso deslocamento entre regiões, torna necessário um planejamento. Entretanto, esse planejamento não é trivial e demanda estudos multidisciplinares envolvendo diversas áreas, como o planejamento urbano, engenharia de trânsito, infraestrutura, gestão de riscos de desastres, entre outros (TUCCI; BERTONI C, 2003; LOPES S. B.; PFAFFENBICHLER, 2012).

Existem várias formas de abordar as características das cidades em seus diferentes aspectos apresentados. Atualmente, tem crescido o uso da teoria de grafos no contexto de redes complexas. Define-se um grafo através de dois elementos fundamentais: vértices que representam os objetos, e arestas que representam as relações entre eles (COSTA et al., 2007; COSTA et al., 2011). Duas características em especial são de interesse nesse estudo: a componente espacial cuja posição geográfica é utilizada para modelar o sistema, e a dinâmica temporal que modifica as relações entre objetos ao longo do tempo (BARTHELEMY, 2010; HOLME; SARAMÄKI, 2013). Com a componente espacial e a dinâmica temporal é proposta uma análise de (geo)grafos variantes no tempo - *time-varying (geo)graphs*. Um grafo geográfico, ou (geo)grafo, é uma estrutura na qual cada objeto (vértice) é uma representação do espaço geográfico e a conexão (aresta) reflete uma relação espacial entre os objetos (SANTOS et al., 2017).

Conhecer padrões de mobilidade urbana é vital para áreas como epidemiologia, planejamento urbano, redes de comunicação e estratégias de eventos. Porém, possuir os dados relativos à mobilidade urbana das atividades humanas diárias, trajetórias de veículos, entre outros, é um grande desafio, pois muitas vezes os dados são controlados por grandes empresas de redes sociais, GPS e aplicativos mobile, que não disponibilizam abertamente os dados. Uma alternativa para a obtenção de dados são as pesquisas de origem e destino (MONTIS et al., 2010). Na cidade de São José dos Campos há um conjunto de dados coletados pelo IPPLAN (Instituto

de Pesquisa, Planejamento e Administração), relativos a uma pesquisa de origem-destino (IPPLAN, 2014). Estes dados são um retrato dos deslocamentos cotidianos das pessoas na cidade ao longo do dia. Para representar a dinâmica da cidade, os dados são aqui agrupados em uma sequência ordenada de intervalos temporais, onde a série temporal é representada por um grafo que se altera a cada instante, caracterizando um grafo dinâmico.

O objetivo deste trabalho é apoiar tomadores de decisões das demandas do desenvolvimento urbano, via análise das componentes espacial e temporal do fluxo dos deslocamentos na cidade, modelado por redes complexas.

O presente trabalho tem as seguintes perguntas científicas:

- a) Como a rede de deslocamentos se comporta ao longo do dia? Ou seja, qual a dinâmica espaço-temporal das propriedades topológicas da rede?
- b) Existem padrões espaço-temporais? Ou seja, há uma estrutura espacial nos deslocamentos?
- c) Se existirem: há persistência ao longo do dia? Ou seja, a estrutura é a mesma para diferentes instantes da amostragem dos dados de mobilidade urbana?

Do ponto de vista institucional, este estudo se alinha ao objetivo estratégico do Instituto Nacional de Pesquisas Espaciais (INPE) de prover produtos e serviços inovadores com geoinformática para o monitoramento e apoio à gestão territorial e ambiental, com base na meta de desenvolver uma metodologia para análise de tendência de crescimento urbano para apoiar o Sistema Nacional de Monitoramento da Ocupação Urbana do Ministério das Cidades (INPE, 2016).

O restante desse documento está organizado da seguinte maneira:

- a) Capítulo 2. Nesse capítulo, apresenta-se os conceitos de definições básicas de Teoria de Grafos e aplicações em mobilidade urbana.
- b) Capítulo 3. Esse capítulo apresenta a descrição dos dados, o pré-processamento e o desenvolvimento realizado.
- c) Capítulo 4. Os resultados obtidos, juntamente com as considerações finais são apresentadas nesse capítulo.

d) Capítulo 5. As conclusões e propostas futuras do trabalho são apresentadas.

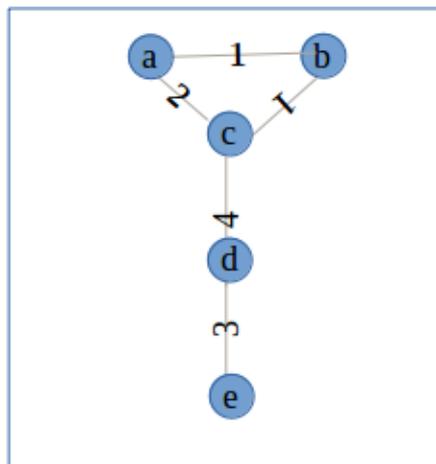
2 REVISÃO DA LITERATURA

2.1 Definições básicas de Teoria de Grafos

Um grafo pode ser definido como $G(V, A)$, com V um conjunto finito não vazio de vértices, e A um conjunto de arestas. O vértice (V) é a representação de uma entidade do problema a ser estudada no grafo, podendo ser pessoas, ruas, moléculas, etc. A aresta (A) é a representação da relação entre os objetos. A ordem de um grafo é dada pelo seu número de vértices, e o tamanho é dado pelo número de arestas.

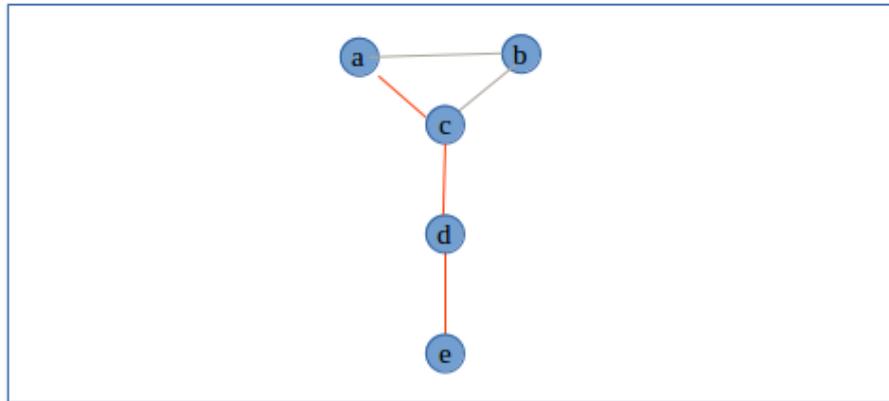
Caminho em um grafo é definido por uma sequência de vértices conectados por arestas e o comprimento é dado pelo número de arestas necessárias para completar o percurso. As arestas podem assumir pesos, tornando o grafo ponderado e sua soma resulta no custo do caminho. Essa medida permite uma análise do caminho não apenas pelo seu comprimento, mas também pelo seu custo final. A Figura 2.1 ilustra um grafo ponderado.

Figura 2.1 - Grafo Ponderado



A Figura 2.2 exemplifica o diâmetro, outra importante medida em redes complexas, relativo ao comprimento do maior menor caminho entre um par de vértices.

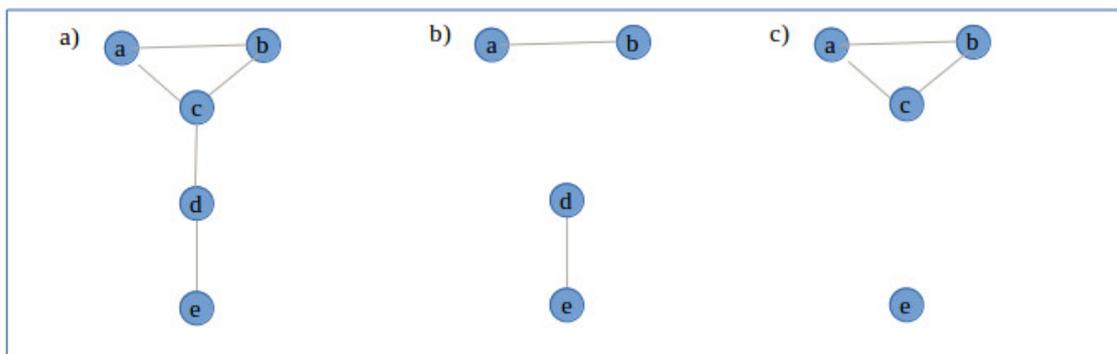
Figura 2.2 - Diâmetro de um grafo



Na Figura 2.2 o valor do diâmetro é 3, sendo a maior menor distância é entre o V_a ou V_b até o V_e . Para obter a menor distância e tendo como ponto de partida o V_a , o próximo passo é para o V_c , depois para o V_d e finalmente no V_e .

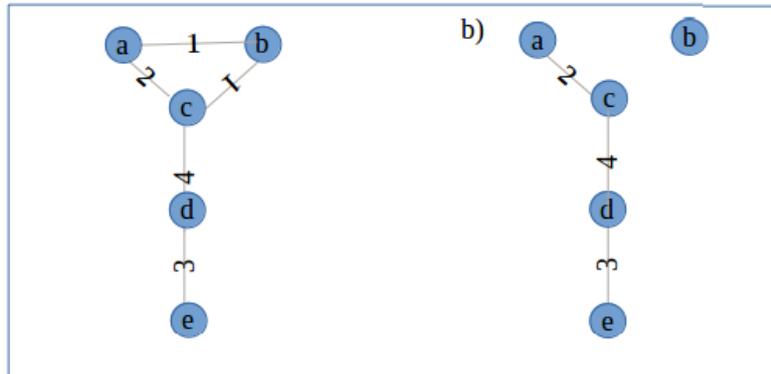
Um grafo pode ser particionado pela eliminação de vértices ou de aresta, o que resulta em alterações de caminhos (BORODIN et al., 2010), como na Figura 2.3.

Figura 2.3 - a) Grafo com todos os vértices; b) Grafo com vértice **c** retirado; c) Grafo com o vértice **d** retirado.



Ao perder arestas o grafo também tem sua estrutura alterada, o que pode ocorrer ao se mudar a regra de criação da aresta: em um primeiro instante o grafo tem uma regra para conectar os vértices e depois pode ter outra regra, Figura 2.4.

Figura 2.4 - a) Grafo com todas as arestas; b) Grafo com aresta eliminada desconectando os vértice b e c.



No grafo **a** a regra de criação é para todo peso ≥ 1 , enquanto no grafo **b** a regra de criação da aresta é para todo peso ≥ 2 . Técnicas de alterar a estrutura inicial da rede podem ser aplicadas para extrair informações da rede a partir de uma nova perspectiva (DENSAR et al., 2008).

A seguir são apresentadas algumas das mais fundamentais medidas em Redes Complexas:

- **Grau:** O grau de um vértice i , denotado por k_i , é o número de arestas que estão conectadas ao vértice V (COSTA et al., 2007). Um vértice V de grau 0 é chamado de vértice isolado. Um vértice V de grau 1 é chamado vértice pendente. Um grafo que possui todos os vértices com o mesmo grau é chamado de Grafo regular. Para redes não direcionadas, o grau pode ser calculado a partir da Equação 2.1:

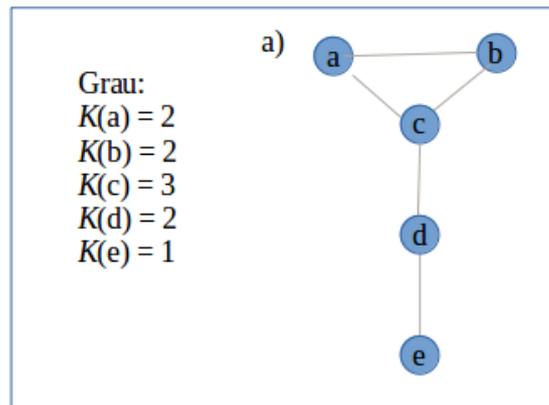
$$k_i = \sum_j a_{ij} = \sum_i a_{ji} \quad (2.1)$$

O grau médio de uma rede é a média de k para todos os vértices na rede, Equação 2.2:

$$\langle k \rangle = \frac{1}{N} \sum_i k_i = \frac{1}{N} \sum_{ij} a_{ij} \quad (2.2)$$

Figura 2.5 exemplifica o calculo do grau de cada vértice.

Figura 2.5 - Grau de um grafo



Na Figura 2.5 no Grafo cada vértice apresenta o seu grau e o grau médio é $\langle k \rangle = 2$.

Para estudar a estrutura cíclica da rede (sua tendência para formar conjuntos de vértices fortemente conectados), uma medida topológica utilizada é o coeficiente de aglomeração, que permite quantificar como o vértice e seus vizinhos estão agrupados. Esta medida foi introduzida pela primeira vez por Watts e Strogatz (1998).

- Coeficiente de aglomeração: A medida de coeficiente de aglomeração, denotada por C , também conhecido como transitividade, (SOFFER; VÁZQUEZ, 2005), é contextualizada pelo número de ligações entre os vizinhos mais próximos de um vértice. O Coeficiente mede a extensão da interconectividade entre os vizinhos de um vértice que possivelmente poderiam estar conectados entre eles, assim se um vértice i está conectado ao vértice j , e o j está conectado ao vértice k , então é calculado a probabilidade de i também estar conectado com k , (BARABÁSI; ALBERT, 2002). O calculo é feito a partir da Equação 2.3 (BESSA et al., 2013).

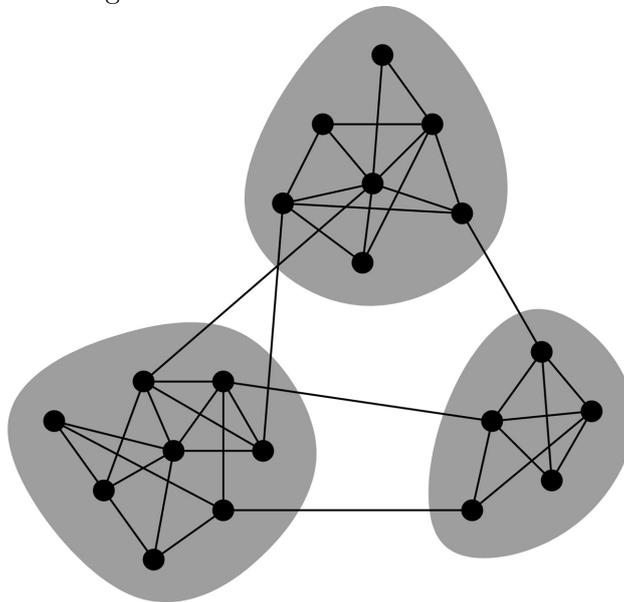
$$C_i = \frac{2n_i}{k_i(k_i - 1)} \quad (2.3)$$

, onde n_i é o número de arestas entre os vértices incidentes a uma mesma aresta e k_i o grau do vértice. Com base na Equação 2.3, que calcula o coeficiente de aglomeração de cada vértice é possível calcular o o coeficiente de aglomeração médio, que é a média dos coeficientes de aglomeração de todos os vértices da rede Equação 2.4

$$\langle C \rangle = \frac{1}{N} \sum_i C_i \quad (2.4)$$

A rede também pode apresentar uma estrutura homogênea ou não. Redes reais normalmente apresentam estruturas não homogêneas, definidas pela presença de grupos de vértices fortemente conectados com conexões dispersas entre os grupos. A Figura 2.6 apresenta um exemplo uma rede com uma estrutura de comunidade bem definida, dividida em 3 comunidades.

Figura 2.6 - Estrutura de comunidade



Fonte: Newman (2006)

Cada grupo apresentado na Figura 2.6 representa uma comunidade e em cada uma os vértices estão densamente conectados dentro da própria comunidade. Essa abordagem tem sido utilizada em diversas áreas, como reconhecimento de organização terrorista, gerenciamento de organização, análise de rede biológica, mineração na Web, clusterização de documentos, etc (FORTUNATO, 2010).

Para avaliar a qualidade de uma divisão da rede em comunidades específicas, Newman e Girvan (2004) propuseram uma medida denominada modularidade, que é tipicamente representada por Q , capaz de medir a qualidade da divisão da rede. Essa medida pode ser positiva ou negativa, sendo que os valores positivos indicam

uma provável comunidade. Assim, pode-se procurar a estrutura de comunidade observando as divisões que tenham, de preferência, altos valores positivos (NEWMAN, 2006).

Diversos métodos foram propostos para detectar comunidades. DONGEN (2000) propuseram um algoritmo Markov Cluster, baseado em caminhadas aleatórias na rede e com operações algébricas estocásticas associadas a sua matriz. (PONS; LATAPY, 2005) propuseram uma medida de semelhanças entre os vértices da rede baseada em caminhadas aleatórias e desenvolveram um algoritmo para computar a estrutura de comunidade, sendo utilizado para redes ponderadas ou não. Rovall and Bergstrom (2008) ROVALL; BERGSTROM (2008) utilizaram o fluxo de probabilidade de caminhadas aleatórias em uma rede como um representante para fluxos de informação no sistema real e decomposeram a rede em módulos equivalendo-se que a descrição sucinta do fluxo de informação é um problema de codificação ou compressão. WEINAN et al. (2008) propuseram uma estratégia para encontrar a melhor partição de uma grande rede complexa em um pequeno número de *cluster*, propondo a utilização de cadeias de Markov associadas à dinâmica dessas redes. Para detecção de comunidades em redes com grafos em que cada aresta tem um sinal positivo ou negativo. YANG et al. (2007) propuseram um algoritmo com heurística de agrupamento de redes com base no modelo de caminhada aleatória de Markov. Muitos outros algoritmos de detecção de comunidades foram desenvolvidos, seja com método exato, com heurísticas, com compressão de dados, para redes ponderadas ou não, para redes grandes ou pequenas, todos buscando descobrir as propriedades entre o micro-mundo e o macro-mundo de redes complexas (YANG et al., 2016).

2.2 Aplicando teoria de grafos em mobilidade urbana

Densar et al. (2008) fizeram um estudo de uma área metropolitana na Finlândia criando um grafo não ponderado e sua base de dados foi o arruamento da cidade com sistema de informação geográfico, sendo as intersecções dos cruzamentos os vértices e a conexão entre eles representado por arestas. As medidas estudadas foram centralidade por grau, coeficiente de aglomeração e cortes de vértices. O objetivo era a identificação de locais onde uma interdição causaria danos e perturbações (DENSAR et al., 2008).

Montis et al. (2010) fizeram um estudo comparativo entre duas ilhas da Itália: Sardenha e Sicília. A base de dados foi uma pesquisa estática de origem-destino. Em ambos os estudos os sistemas foram modelados com grafos ponderados e sime-

trizados. As técnicas de análise de redes complexas aplicadas foram grau e coeficiente de aglomeração. O trabalho realiza as comparações e destaca que sistemas com características geográficas similares podem ter o mesmo comportamento.

Soh et al. (2010) utilizaram grafo ponderado de uma base de dados de rotas de viagens nos sistemas de transporte ferroviário e rodoviário de Cingapura, criando uma matriz de adjacência e matriz de pesos representando o número de passageiros viajando entre os locais V_i e V_j em um único dia. As medidas topológicas estudadas foram grau, coeficiente de aglomeração, diâmetro e menor caminho. Com o objetivo de realizar uma análise temporal dos sistemas (dias de semana e final de semana) os fluxos de tráfego foram observados e mostraram um contraste com vértices com pouca conectividade. O trabalho destacou uma diferença significativa nos fluxos de tráfego para ambas as redes durante os dias da semana e fins de semana, sugerindo a importância de adicionar uma perspectiva temporal ausente de muitos estudos anteriores. Como conclusão afirmam que um estudo dinâmico contribui para as informações obtidas pela análise topológica tradicional e que a análise de redes complexas ponderadas é útil para compreender um grande sistema complexo.

Guo et al. (2012) investigaram a mobilidade em Shenzhen, China, a partir de dados de origem-destino baseado em dados de taxi. Duas técnicas são exploradas: a) agrupamento espacial de vizinhos mais próximos; b) extração e mapeamento das medidas de fluxo. O objetivo do trabalho foi detectar padrões na distribuição espacial e entender as tendências temporais dos movimentos.

Gao et al. (2013) consideraram dados de celular para detecção de comunidade e encontraram padrões influenciados pelo contexto geográfico. Louail et al. (2015) também extraiu dados de telefones celulares, criando matrizes de origem e destino para analisar os fluxos entre áreas residenciais e empregatícias na cidade.

Emmerich et al. (2013) relacionaram custos de viagem para destinos individuais permitindo determinar trajetórias entre locais distintos e verificar viagens que são propensas a grande fluxo, mostrando os padrões de tráfegos em diferentes partes da cidade.

Os trabalhos apresentados com diferentes bases de dados utilizaram redes complexas ponderadas pelo fluxo para análise espacial ou temporal dos dados. Os estudos com dinâmica temporal dos dados destacaram que a técnica enriquece o estudo diante análises obtidas por dados estáticos. Os resultados enfatizaram que os estu-

dos são capazes de auxiliar decisões de planejamento urbano.

3 METODOLOGIA

A seguir é descrito o conjunto de métodos que compõe a metodologia desenvolvida. Primeiramente é apresentada a base de dados, seguida da descrição do pré-processamento e do desenvolvimento com as medidas de redes complexas.

3.1 Descrição dos dados

Nesta dissertação são usados dados de mobilidade urbana relativos à Pesquisa Origem-Destino de São José dos Campos/SP, cidade com aproximadamente 700 mil habitantes, situada no Vale do Rio Paraíba. Entre abril e setembro de 2011, a pedido da Secretaria de Transportes da Prefeitura Municipal de São José dos Campos, foi realizado um estudo em 4.185 domicílios, no qual todos os moradores presentes foram entrevistados. Os dados referentes ao fluxo de pessoas entre as zonas contêm cerca de 24 mil relatos. Este estudo retrata os deslocamentos e as atividades diárias dos moradores, questionando a trajetória, origem e destino, o motivo, meio de transporte, e o tempo das viagens. As entrevistas foram realizadas entre terça e sexta e os entrevistados responderam sobre as viagens realizadas no dia anterior, representando um dia típico

Os domicílios estão distribuídos nos perímetros urbano e rural da cidade. A cidade foi dividida em 55 zonas de tráfego, apresentadas na Figura 3.1. O nome de cada zona é descrito na Tabela 3.1.

Tabela 3.1 - 55 zonas de tráfego

Zonas	
1 = Centro	8 = Santana /Vila Rossi
2 = Vila Maria/Vila Santa Luzia	9 = Vila Industrial/VilaTatetuba
3 = Jardim Paulista/Monte Castelo	10 = Jardim da Granja/Jardim Souto
4 = Vila São Pedro/Jardim Jussara	11 = CTA/INPE
5 = Vila Piratininga/Jardim Augusta	12 = Vila Nova Conceição/Vila São Bento
6 = Vila Adyanna/Vila Betânia	13 = Jardim Satélite/Floradas de São José
7 = Vila Ema/Jardim Esplanada	14 = Jardim Aquarius/Jardim Colinas

(continua)

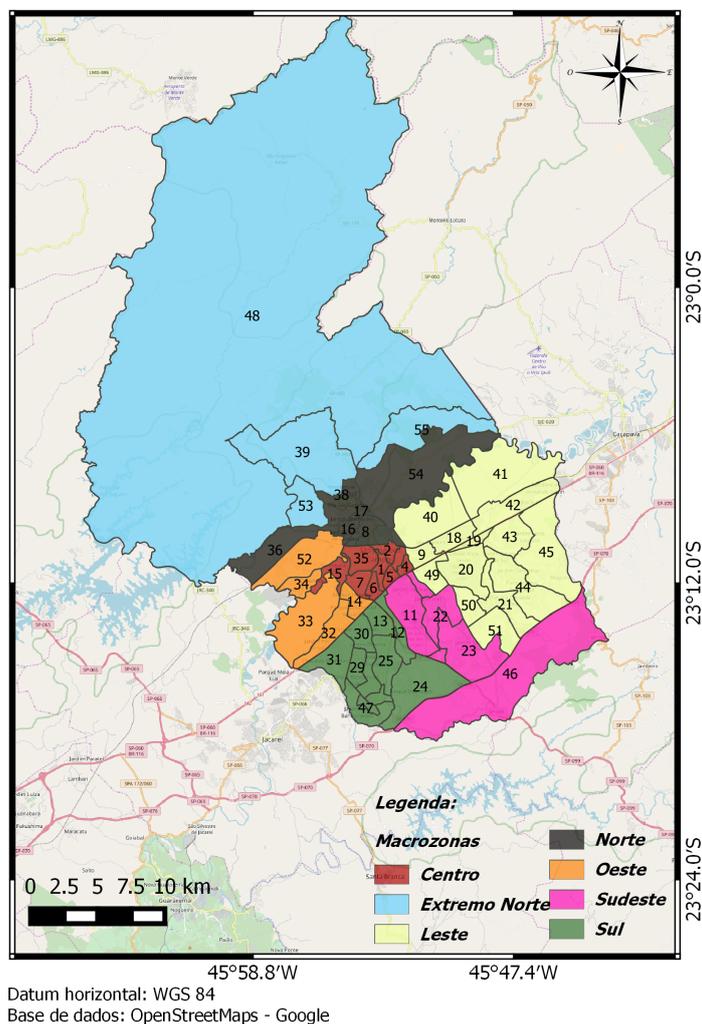
Fonte: Adaptada de (Instituto de Pesquisa, Administração e Planejamento, 2014)

Tabela 3.2 - 55 zonas de tráfego

Zonas	<i>(conclusão)</i>
15 = Condomínio Esplanada do Sol/- Serimbura	36 = Vazio urbano
16 = Jardim Telespark/Jardim Altos de Santana	37 = Vila São Geraldo/Jd. Minas Ge- rais
17 = Alto da Ponte/Vila Veneziani	38 = Vila Paiva/Jardim Boa Vista
18 = Vila Tesouro/Jardim Universo	39 = Chác. dos Freitas/Chác. do Bon- sucesso
19 = Jardim Motorama/Jardim Nova Detroit	40 = Res. Vista Linda
20 = Jardim Americano/Vista Verde	41 = Vazio urbano
21 = Chác. Pousada do Vale/Campos de São José	42 = Eugênio de Melo/Galo Branco
22 = Residencial Flamboyant/Jardim Uirá	43 = Jd Santa Inês/Res. Frei Galvão
23 = Jardim do Lago/Vila Iracema	44 = Novo Horizonte/Res.Dom Bosco
24 = Parque Interlagos/Altos da Quinta	45 = Jd. Ebenezer/Chác. Sítio Jataí
25 = Bosque dos Eucaliptos/Quinta das Flores	46 = Cajuru
26 = Conjunto Habitacional Dom Pe- dro	47 = Jd. República/Vila das Flores
27 = Jardim Imperial/Jardim Colo- nial	48 = São Francisco Xavier/área rural
28 = Bosque dos Ipês/Campo dos Alemães	49 = Jd. Diamante/REVAP
29 = Cidade Morumbi/Res. União	50 = Vazio urbano
30 = Parque Industrial/Jd Oriente	51 = Santa Cecília/Jardim Helena
31 = Palmeiras de São José/Rio Com- prido	52 = Paratehy/Residencial Jaguarly
32 = Jardim Alvorada/Jd. das Indús- trias	53 = Chácara Fazenda Caeté
33 = Limoeiro	54 = Águas de Canindú/Chác. Havaí
34 = Urbanova	55 = Chác. Pedra D'água II
35 = Banhado	

Fonte: Adaptada de Instituto de Pesquisa, Administração e Planejamento (2014)

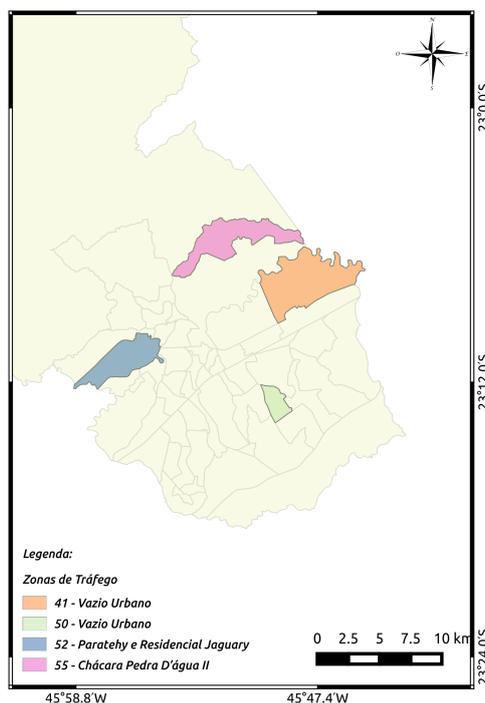
Figura 3.1 - 55 zonas de tráfego



A macrozona Sul, que é a mais populosa, concentra mais de 37% dos residentes. A macrozona Centro é a que tem o maior número de empregos em comércio e serviços, enquanto as macrozonas Sudeste e Leste oferecem o maior número de empregos na indústria. A macrozona Sul se distingue pela grande oferta de empregos nas áreas de comércio, serviços e educação. A macrozona Sul se destaca em ensino infantil e fundamental. A macrozona Oeste, com a maior renda per capita, conta com duas grandes universidades particulares e apresenta a mais elevada relação de matrículas por habitante da cidade. O índice de matrículas escolares por habitante na região leste é inferior à média municipal.

A Figura 3.2 expõe as 4 zonas que não são zonas de tráfego.

Figura 3.2 - Vazios urbanos



Os critérios de uma zona de tráfego são estabelecidos segundo diversos aspectos, como setores censitários do IBGE, sistemas viários, barreiras físicas naturais e características homogêneas no uso e ocupação do solo (Instituto de Pesquisa, Administração e Planejamento, 2014). Os dados se tornaram o ponto de partida para uma série de estudos, planos e projetos, como em (SILVA, 2016; SILVA et al., ; SANTOS et al., 2017) que estudaram a representação da rede em Sistemas de Informação Geográfica (SIG).

A partir dos dados coletados (de origem, destino, instante de partida, instante de chegada e a medida que representa o fluxo entre as zonas, denominada fator de expansão, representando o peso da aresta), foi gerada uma rede de fluxo da cidade de São José dos Campos. Assim, os vértices representam as zonas e as arestas as conexões entre elas. Para relacionar zonas e representar a cidade geograficamente todas as zonas da cidade foram consideradas, totalizando 55 vértices.

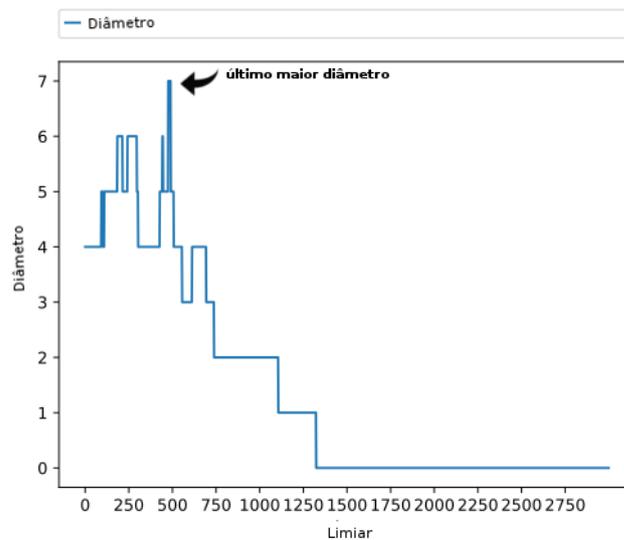
Nesse contexto, este trabalho investiga, via análise de redes complexas, o fluxo de pessoas entre as regiões de uma cidade ao longo do dia. Duas análises são exploradas: rede estática e rede dinâmica. Especificamente, para cada rede, há duas abordagens para a construção do grafo geográfico: fluxo completo e limiarizado.

Para a construção a partir do fluxo completo todas as viagens são consideradas, ou seja, em todo fluxo entre zonas foi criada uma aresta conectando os vértices. Para o fluxo limiarizado, a conexão entre as zonas ocorre a partir de um determinado número de viagens. Por exemplo: a partir de 10 viagens entre as zonas ocorrem as conexões e todos valores inferiores não são considerados.

O valor para a limiarização é determinado de acordo com um valor crítico. Este valor tem por objetivo manter as conexões com alto fluxo, ou seja: as principais informações da rede são mantidas eliminando pequenos valores.

A técnica utilizada para a determinar o limiar crítico foi considerar todos os seus valores possíveis e, a partir de cada um, calcular o diâmetro do grafo. Escolhe-se o limiar cujo valor do diâmetro seja o último maior. A Figura 3.3 apresenta, em destaque, o último maior diâmetro e, conseqüentemente, o valor do limiar crítico.

Figura 3.3 - Representação do último maior diâmetro

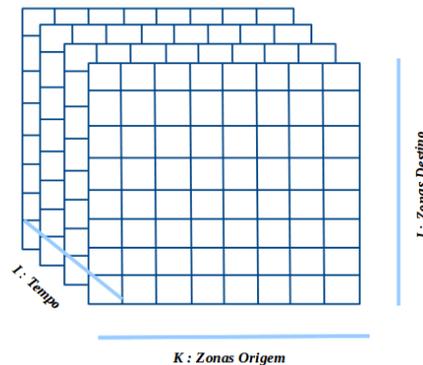


No exemplo da Figura 3.3 o instante que apresenta o último maior diâmetro representa o instante crítico da rede. Ao eliminar arestas considerando limiares maiores, a quantidade de caminhos tende a diminuir e, como consequência, a distância aumenta. Contudo, após um certo valor de limiar, a rede se divide e caminhos previamente existentes são eliminados. Neste trabalho, o limiar crítico é adotado como sendo o valor associado à rede com maior diâmetro, ou seja, o limiar observado antes da quebra do grafo.

3.2 Pré-processamento dos dados

Para uma análise temporal foi necessário realizar um pré-processamento e o desenvolvimento em Linguagem C++. O tempo, a origem e o destino foram representados de forma matricial em 3 dimensões. A Figura 3.4 exemplifica a representação matricial.

Figura 3.4 - Representação matricial em 3 dimensões



As dimensões, i, j, k são definidas por :

- i : Representa o tempo.
- j, k : Representam o fluxo entre a zona de origem e a de destino, que são ponderadas pelo fluxo de pessoas.

A construção da matriz de fluxo foi feita a partir da desconsideração do fluxo dentro da própria zona e da técnica de simetrização, Figura 3.5.

Figura 3.5 - A Matriz **A** representa a matriz inicial de fluxo; A matriz **B** representa a soma entre os vértices i e j : $A(i, j) + A(j, i)$ e a diagonal zerada

Matriz A				Matriz B			
0	1	2	0	0	3	2	0
2	0	0	1	3	0	0	2
0	0	1	0	2	0	0	0
0	1	0	0	0	2	0	0

A partir da matriz inicial, a simetrização é realizada pela soma do fluxo da zona 1 para a zona 2 com o fluxo da zona 2 para a zona 1, representando o fluxo total e não direcional entre zonas. Os valores da diagonal, que representam o fluxo intra-zona, são desconsiderados neste trabalho.

Para a escolha da janela temporal, os dados foram agrupados com 3 variações de tempo: 15min, 60min e 180min, detalhados posteriormente.

3.3 Desenvolvimento

A partir do pré-processamento foi criada a rede ponderada não direcionada com variação no tempo, representada com sequências ordenadas de hora em hora, sobre o conjunto fixo de zonas, os vértices. Neste contexto, medidas como coeficiente de aglomeração, grau e diâmetro foram calculadas utilizando a biblioteca *igraph*, em linguagem C. Foram mensuradas grandezas fundamentais para o estudo dos dados da série temporal como tamanho, grau de um vértice, grau médio, coeficiente de aglomeração de um vértice, coeficiente de aglomeração médio e diâmetro de um grafo.

A Tabela 3.3 lista as principais funções utilizadas para o desenvolvimento do trabalho.

Tabela 3.3 - Igraph em C e as medidas da rede

Medidas	igraph
Inicialização do grafo	<code>igraph_weighted_adjacency()</code>
Tamanho	<code>igraph_ecount()</code>
Grau de um vértice	<code>igraph_degree()</code>
Coeficiente médio de aglomeração	<code>igraph_transitivity_avglocal_undirected()</code>
Diâmetro	<code>igraph_diameter()</code>
Detecção de comunidade	<code>igraph_community_walktrap()</code>

3.3.1 GeoCNet

Para analisar as redes complexas geograficamente, foi desenvolvida a ferramenta GeoCNET (Geographical Complex Networks). A GeoCNet permite agregar a componente geográfica com propriedades topológicas de redes complexas em um

(geo)grafo, o que possibilita a visualização da rede em um SIG (Sistema de Informações Geográficas) e permite a manipulação dos dados com *shapefile* de diferentes tipos de dados, como, por exemplo, dados climáticos, arruamentos, bacias hidrográficas, etc. Um *shapefile* é um formato de armazenamento de dados de vetor para armazenar a posição, forma e atributos de feições geográficas. Nas redes complexas de mobilidade urbana desenvolvidas (estáticas e dinâmicas), em um *shapefile* o (geo)grafo é composto pelo centroide de cada zona da cidade (vértice) e pela conexão que reflete a relação espacial entre eles (aresta).

A aplicação foi desenvolvida com um conjunto de ferramentas, apresentados na Figura 3.6.

Figura 3.6 - Ferramentas utilizadas no desenvolvimento da GeoCNet

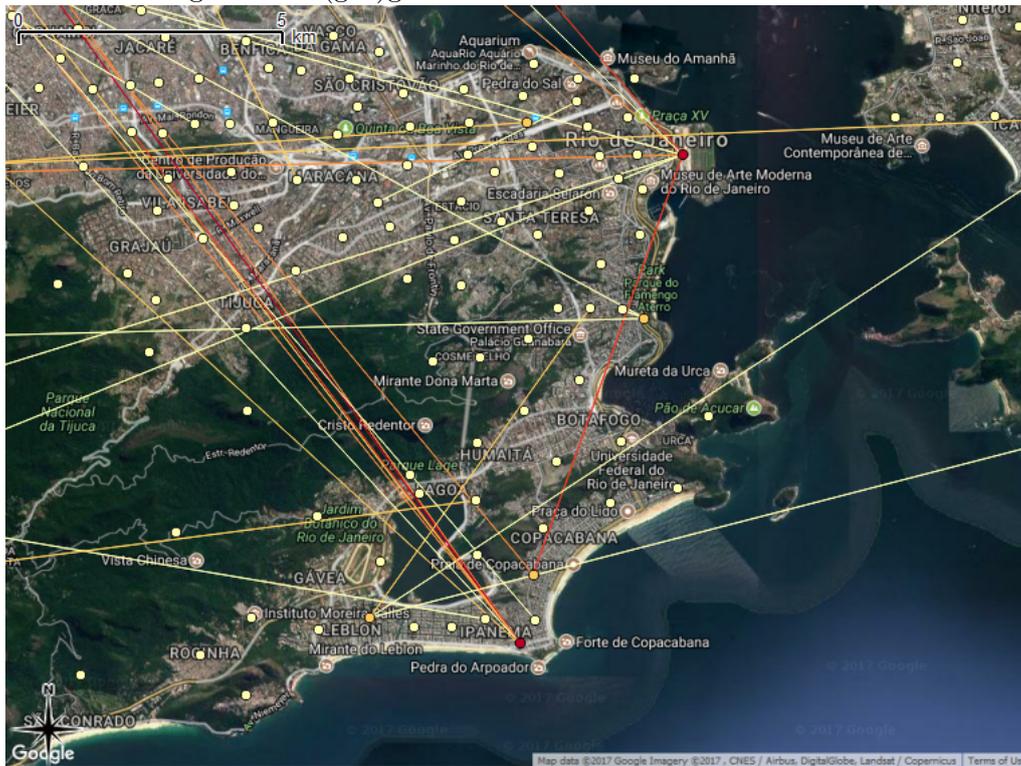


A Figura 3.6 apresenta o conjunto de ferramentas utilizadas. A GeoCNet foi desenvolvida em C com a biblioteca *igraph* para o processamento dos dados e em Python3 para o pós-processamento. Em python é feita a comunicação com o banco de dados PostgreSQL e extensão PostGIS, que permite a criação de um *shapefile* para visualização de um (geo)grafos, o software utilizado para a visualização foi o QuantumGIS. Assim, as dependências são: biblioteca Igraph em C, PostgreSQL 9.5+, PostGis 2.2 e Python 3. Para o funcionamento da aplicação são necessários dois arquivos de entrada: Um *shapefile* dos vértices com uma coluna obrigatória do tipo inteiro chamado id, representando o ID dos nós e matriz de adjacência ponderada que representa as conexões entre esses nós. Como resultado um arquivo *shapefile* de linha é criado representando as arestas da rede. O outro arquivo criado é um *shapefile* de pontos. Ambos terão atributos topológicos de vértices e arestas, respectivamente.

Assim, com a ferramenta é possível visualizar geograficamente a rede complexa com propriedades topológicas em um *shapefile* de vértices e um *shapefile* de arestas em um SIG.

A ferramenta permite a importação de diversos tipos de dados. A Figura 3.7 apresenta os dados de mobilidade urbana de um estudo para a cidade do Rio de Janeiro - RJ.

Figura 3.7 - (geo)grafo da cidade do Rio de Janeiro - RJ

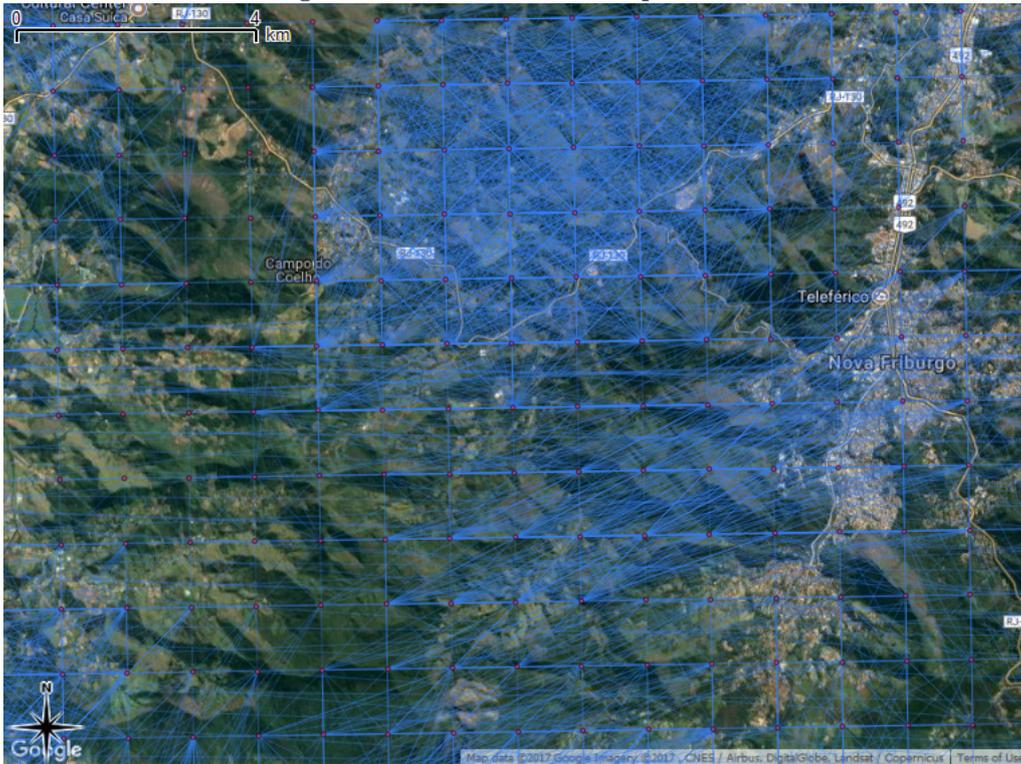


Fonte: Santos et al. (2017)

Na Figura 3.7, a escala de cor para o vértice está relacionada com o seu grau e a aresta com o peso. Nos vértices, quanto mais intenso o tom da cor, maior é o grau e para a aresta maior o peso.

Figura 3.8 apresenta os dados meteorológicos.

Figura 3.8 - Dados de chuva por Radar



Fonte: Santos et al. (2017)

Na Figura 3.8 os (geo)grafos de correlações para a série temporal de chuvas da região serrana do Rio de Janeiro são apresentados junto à imagem de satélite.

A GeoCNet pode ser encontrada no repositório do github: <https://github.com/jessicadominguess/geocnet>. No repositório há tutoriais em português e em inglês de como configurar o ambiente e o código-fonte da ferramenta.

A ferramenta será cadastrada no INPI (Instituto Nacional da Propriedade Industrial).

4 RESULTADOS OBTIDOS

Os resultados obtidos são divididos em rede estática e rede dinâmica, com fluxo completo e com conexões de fluxo a partir de um limiar crítico, assim a estrutura desse capítulo será dividida em:

- Rede Estática para limiar =1, rede do acumulado de viagens para 24h (agrupamento diário)
- Rede Estática para o limiar crítico, rede de acumulado de viagens para 24h (agrupamento diário)
- Rede dinâmica para limiar =1, cada janela temporal da rede dinâmica é construída por todas viagens iniciadas em cada hora.
- Rede dinâmica para limiar crítico, cada janela temporal da rede é construída por viagens a partir de um limiar crítico.

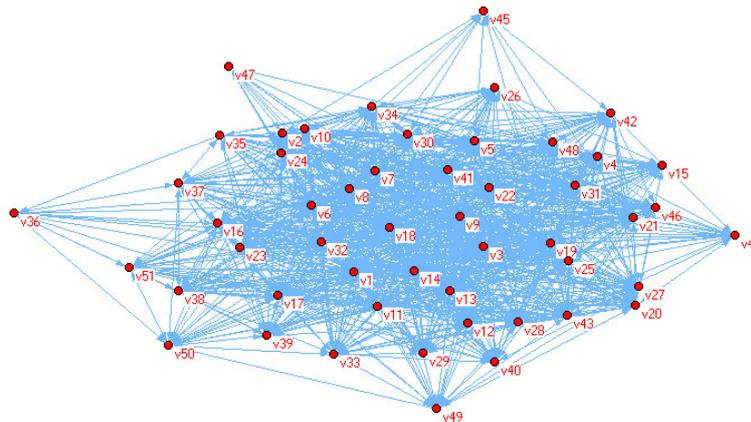
4.1 Rede estática

A rede estática representa o agrupamento de todos os dados para as 24 horas do dia.

4.1.1 Fluxo completo

A Figura 4.1 representa o grafo da rede estática para o fluxo completo

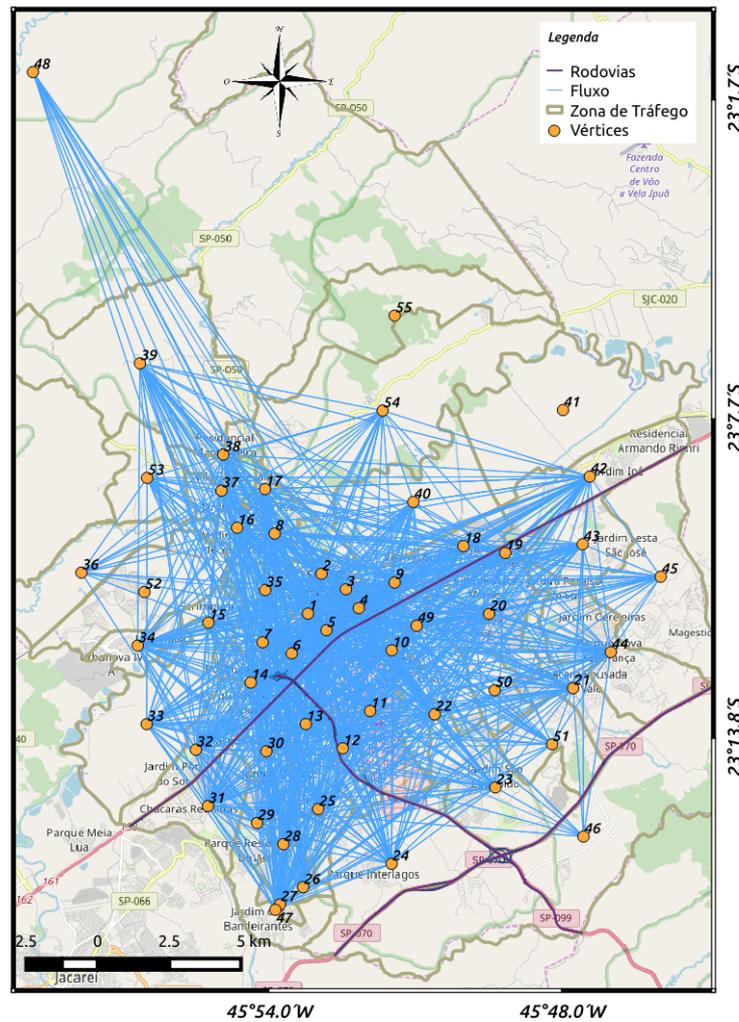
Figura 4.1 - Grafo da rede de um dia típico (24 horas)



O grafo representado na Figura 4.1 é um grafo não direcionado com 55 vértices, porém não está relacionado com o sistema geográfico. O software utilizado para representar a rede do acumulado do dia foi o PAJEK.

A Figura 4.2 apresenta a rede como um (geo)grafo:

Figura 4.2 - (geo)grafo da rede de um dia típico (24 horas)



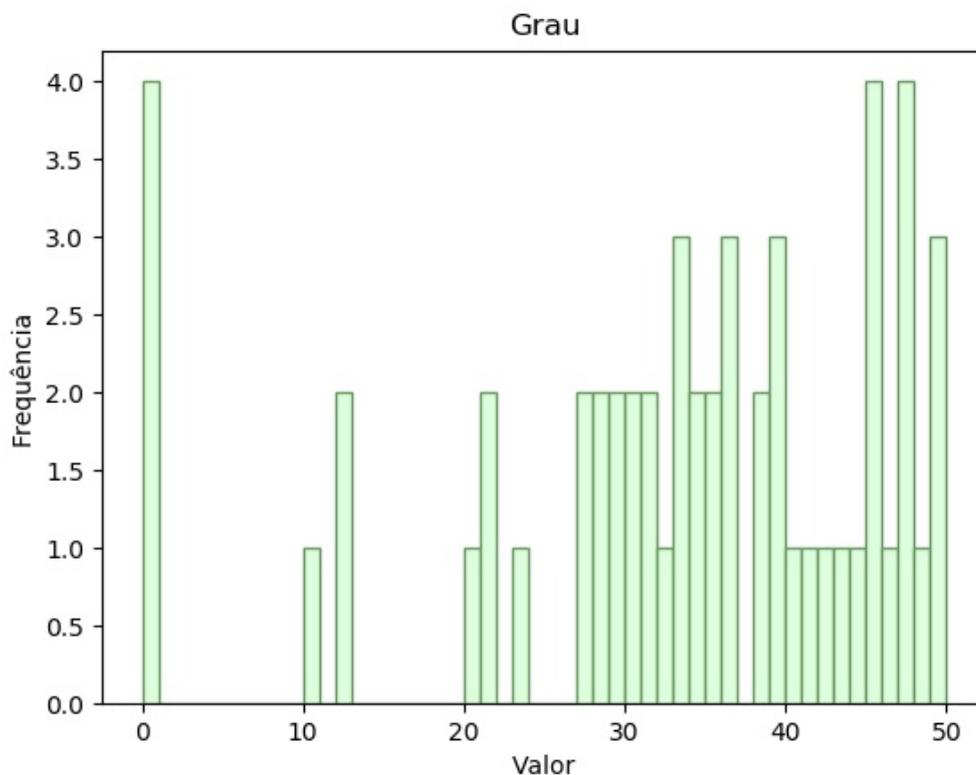
Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Goole

O grafo representado na Figura 4.2 apresenta as características do grafo anterior da Figura 4.1, porém com geolocalização. Os dados foram processados pela GeoC-Net e o software utilizado para representar a rede de 24h foi o QGIS.

Gráficos e tabelas foram criados a partir das métricas obtidas para analisar o comportamento da rede ao longo do dia e para valores acumulados com representação da medida do fluxo total e fluxo crítico.

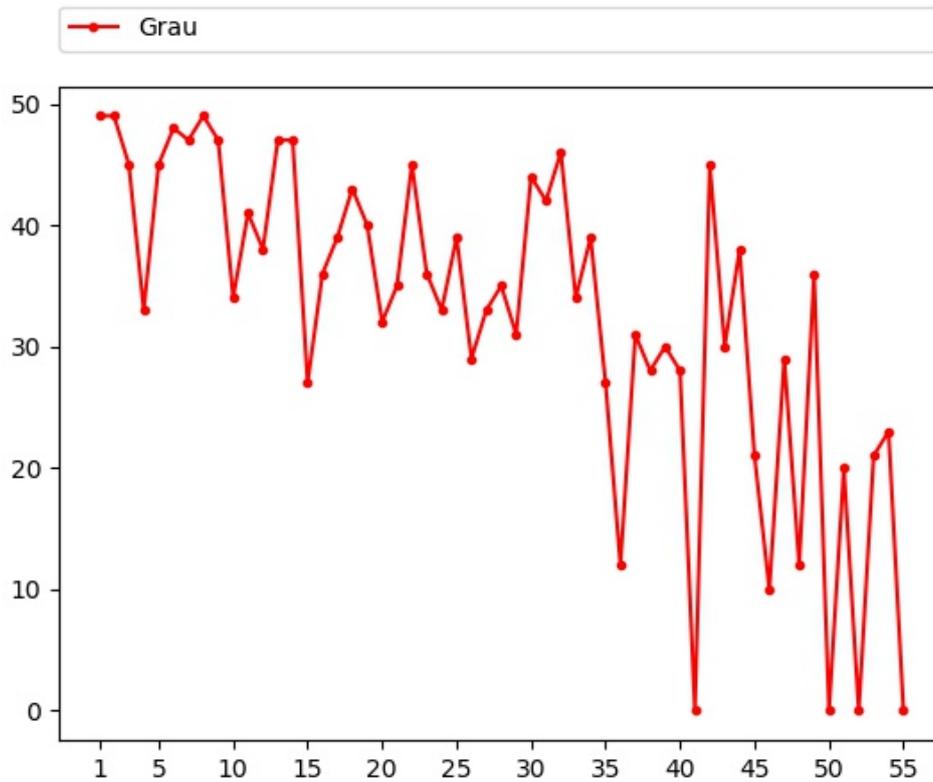
Um histograma da frequência de grau é apresentado na Figura 4.3.

Figura 4.3 - Histograma dos Graus da rede de 24 horas - O eixo Y representa a frequência de vértices e o eixo x representa o grau encontrado



Nesse histograma é possível visualizar a frequência do número de conexões das zonas de tráfego. A soma de todas as frequências é 55, representando cada vértice da rede. Na Figura 4.3 3 vértices apresentam grau 49, o máximo global da rede, ou seja, 3 zonas da cidade estão conectadas com 49 zonas. A Figura 4.4 mostra o número de conexões (grau) de cada zona.

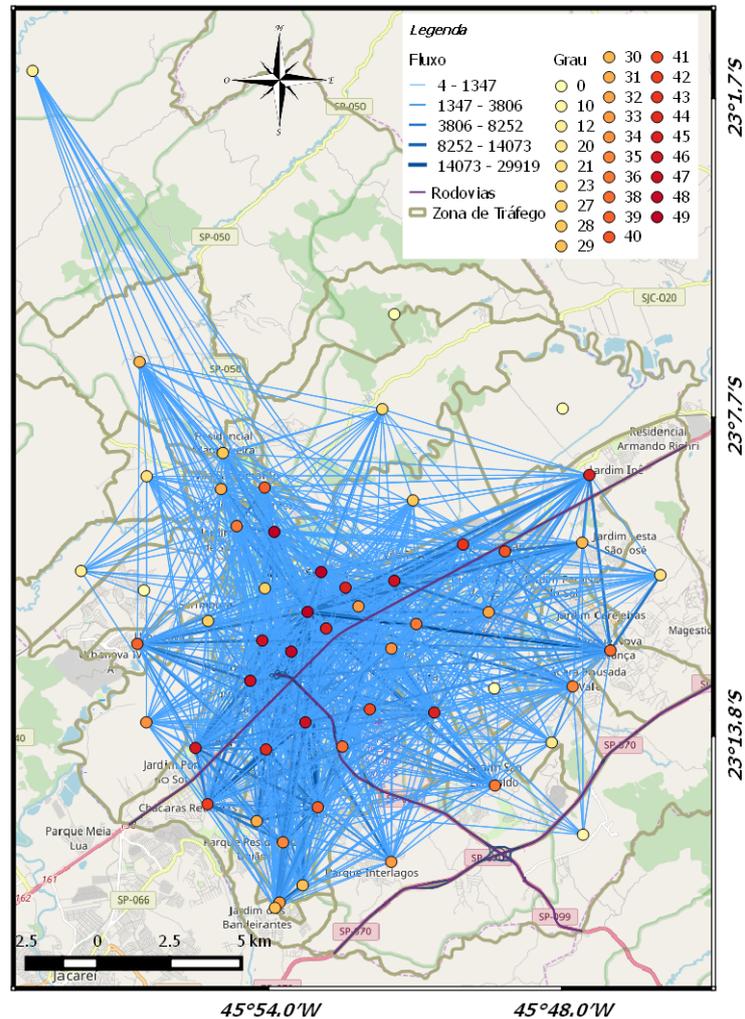
Figura 4.4 - Grau de cada vértice da rede com 24 horas: o eixo Y representa o grau e o eixo x representa o vértice



As zonas que representam o maior valor de conexões (grau) são zona 1 (centro), zona 2 (Vila Maria e Vila Santa Luzia) e zona 8 (Santana e Vila Rossi). As duas primeiras estão na macrozona central da cidade e a última no Norte. Outras zonas apresentam alto número de conexões. Apenas 6 zonas (sem considerar os vértices isolados) estão abaixo de 25 conexões, o que significa que apenas 12% das zonas apresentam menos de 50% de conexões entre zonas.

Com a GeoCNet, um shapefile de linhas e arestas com as propriedades topológicas foram criados. A Figura 4.5 ilustra o (geo)grafo criado.

Figura 4.5 - (geo)grafo - 24h



Na Figura 4.5 o atributo escolhido foi grau (número de conexões) para o vértice e peso (fluxo entre zonas) para as arestas, ou seja, a cor de cada vértice está relacionada com o seu grau: quanto mais intenso o tom do vermelho, maior é o grau. A aresta está relacionada com o seu peso, quanto mais intenso maior o peso.

Todas as medidas estudadas estão armazenadas em uma tabela de atributos no *shapefile*, e a Figura 4.6 é um exemplo. Em razão da existência de muitas arestas é difícil analisar visualmente o (geo)grafo.

Figura 4.6 - Tabela de atributo topológico

	ID	AVG_DEGREE	DEGREE_ALL	CLUSTERING	AVG_CLUSTER	DIAMETER
0	1	32.6900000000	49	0.7100000000	0.8000000000	2
1	10	32.6900000000	34	0.8700000000	0.8000000000	2
2	11	32.6900000000	41	0.7600000000	0.8000000000	2
3	12	32.6900000000	38	0.8500000000	0.8000000000	2
4	13	32.6900000000	47	0.7600000000	0.8000000000	2
5	14	32.6900000000	47	0.7500000000	0.8000000000	2
6	15	32.6900000000	27	0.9100000000	0.8000000000	2
7	16	32.6900000000	36	0.8100000000	0.8000000000	2
8	17	32.6900000000	39	0.7800000000	0.8000000000	2
9	19	32.6900000000	40	0.8200000000	0.8000000000	2
10	2	32.6900000000	49	0.7100000000	0.8000000000	2
11	20	32.6900000000	32	0.8900000000	0.8000000000	2
12	21	32.6900000000	35	0.8200000000	0.8000000000	2
13	22	32.6900000000	45	0.7800000000	0.8000000000	2
14	23	32.6900000000	36	0.8200000000	0.8000000000	2
15	24	32.6900000000	33	0.9000000000	0.8000000000	2
16	25	32.6900000000	39	0.8400000000	0.8000000000	2
17	26	32.6900000000	29	0.8900000000	0.8000000000	2
18	27	32.6900000000	33	0.8700000000	0.8000000000	2
19	28	32.6900000000	35	0.8000000000	0.8000000000	2
20	29	32.6900000000	31	0.9100000000	0.8000000000	2
21	3	32.6900000000	45	0.7500000000	0.8000000000	2
22	30	32.6900000000	44	0.7700000000	0.8000000000	2
23	31	32.6900000000	42	0.8000000000	0.8000000000	2

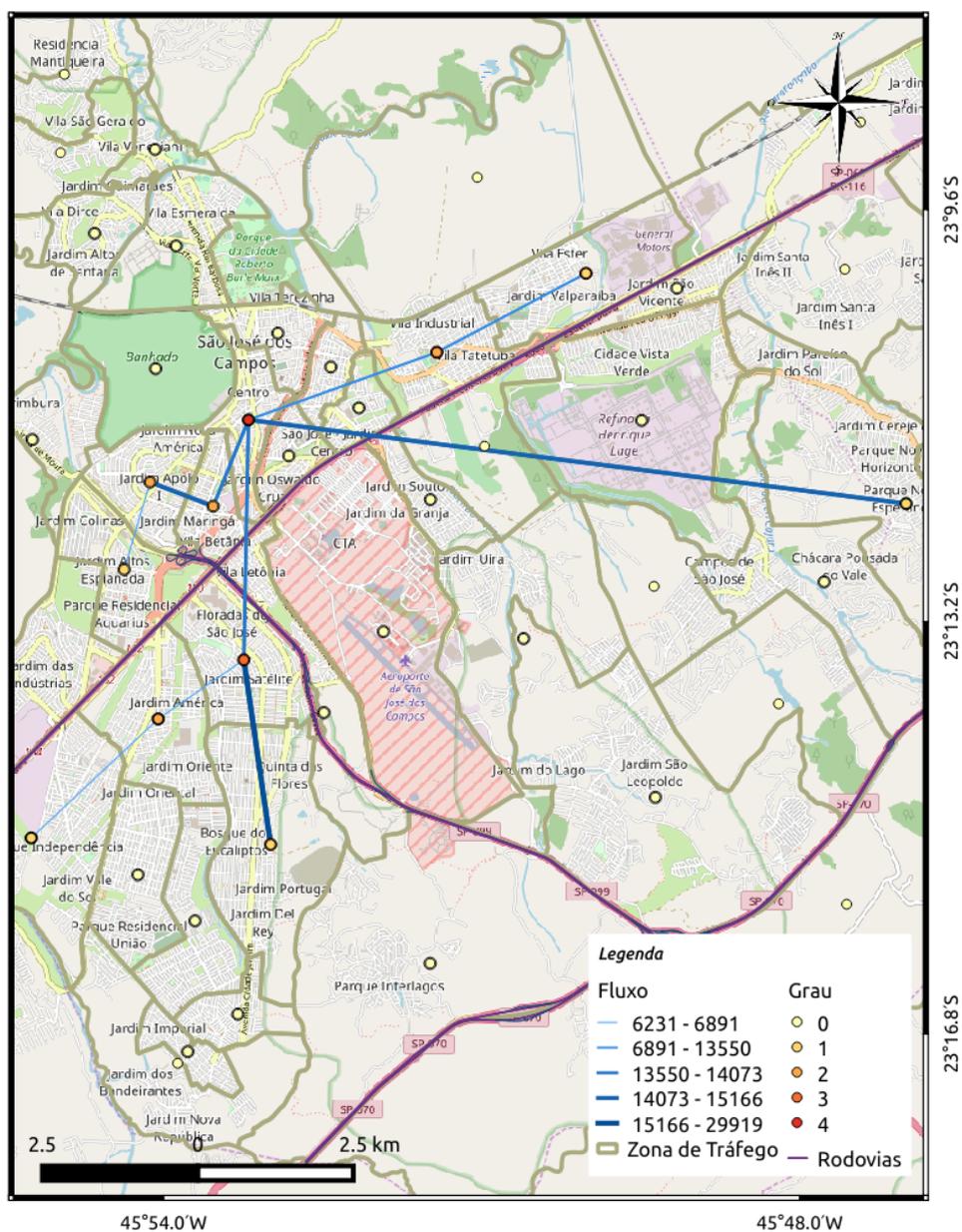
A primeira coluna ID representa o ID de cada zona da tráfego, *avg_degree* o grau médio, *degree_all* soma do grau de entrada com o grau de saída, *clustering* o coeficiente de aglomeração, *avg_cluster*, a média para o coeficiente de aglomeração e *diameter* o diâmetro da rede.

4.1.2 Limiar crítico

Para o calculo das propriedades com a rede limiarizada a regra de conexão foi para todo fluxo maior ou igual a 6231.

A Figura 4.7 representa a rede com as propriedades topológicas com limiar crítico para o acumulado do dia seguindo a mesma regra apresentada anteriormente, vértice relacionado ao grau e a aresta com o fluxo.

Figura 4.7 - (geo)grafo com limiar crítico- 24h com grau de propriedade topológica



Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Google

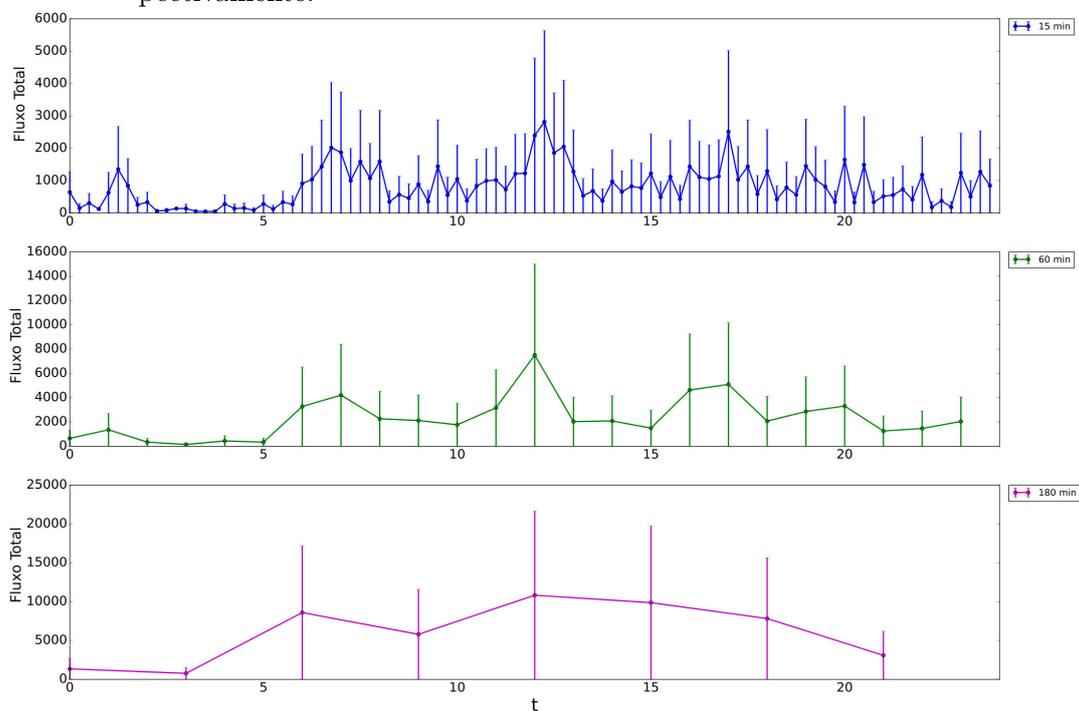
Ao aumentar a regra para criação de conexão (fluxo a partir de 6231) o número de zonas conectadas diminui para 11 vértices, expondo as zonas com fluxo intenso de pessoas ao longo do dia. A zona central se destaca como a zona de maior conexão

(maior grau).

4.2 Rede dinâmica

Para determinar a janela temporal da rede dinâmica, 3 intervalos foram previamente estudados através uma análise estatística. Para isso, 3 gráficos do fluxo foram gerados com intervalo de 15 min, 60min e 180min. A Figura 4.8 representa a janela para cada instante.

Figura 4.8 - Gráficos gerados com intervalo de tempo igual a 15, 60 e 180 minutos, respectivamente.



Todos os gráficos apresentam as mesmas características em sua representação, o eixo das abscissas (eixo t) representa o instante da hora, enquanto o eixo das ordenadas (eixo Fluxo Total) representa a medida de fluxo entre zonas. Cada instante apresenta a soma do fluxo de pessoas de cada horário e a barra na vertical apresenta o valor máximo e mínimo de cada instante.

- **15min:** Revela a alta energia do sistema, com diversos extremos.
- **60min:** Revela valores mais concretos, mostrando que para a janela de tempo de 60 min o maior pico de fluxo ocorre entre 12h e 13h e os valo-

res mínimos no período das 2h às 6h.

- **180min:** Capta praticamente apenas uma grande “onda” de movimento.

Assim, adotou-se para a análise o agrupamento de 60 minutos e a série temporal de propriedades estatísticas mostra que o maior fluxo ocorre entre 12h e 13h, os valores mínimos ocorrem no período das 2h e 6h, e os horários que representam a maior média de fluxo entre zonas de tráfegos ocorrem, respectivamente, entre 7h e 8h, 13h e 14h e 17h e 18h.

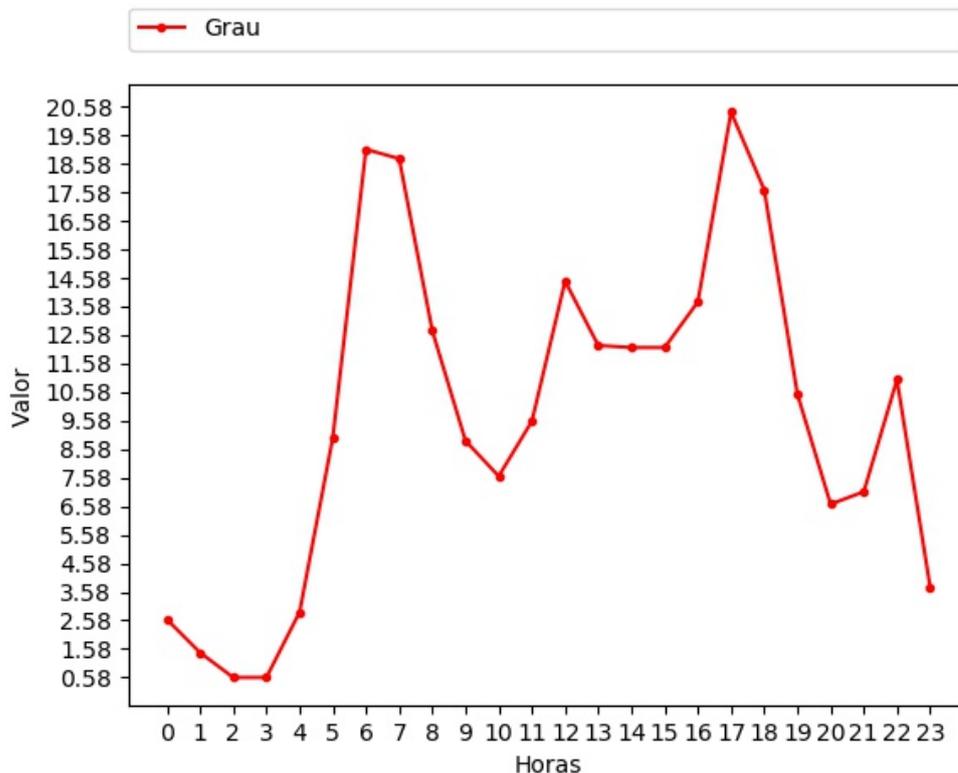
A seguir será apresentado um gráfico para cada medida estudada: grau médio, coeficiente de aglomeração médio e diâmetro da rede com agrupamento de 60min.

4.2.1 Fluxo completo

O fluxo completo considera todo o fluxo entre zonas em cada intervalo de tempo.

- Grau médio para cada hora do dia:

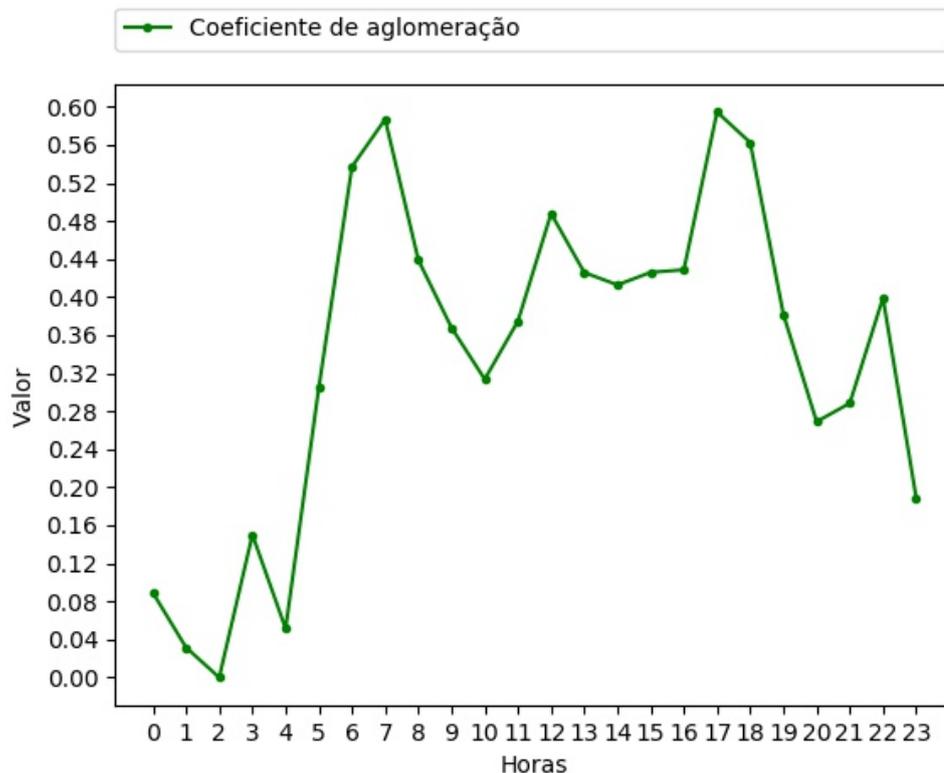
Figura 4.9 - Gráfico do Grau médio da rede com janela temporal de 60min.



Na Figura 4.9 as primeiras horas do dia apresentam uma média de baixos valores o que significa que as conexões entre zonas nesse horário está baixa, ou seja, apesar de haver um fluxo, não há movimentações em várias zonas da cidade. A partir das 4 h a movimentação começa a aumentar, desse modo as pessoas começam a transitar em mais zonas de tráfego, alcançando alguns dos maiores valores às 7h e 8h, o que representa uma conexão média em 35%. No período entre 9 h e 11 h o fluxo de movimentação pelas zonas diminui para uma média de 12,7%. Ao meio-dia a conexão média é de 27% e até às 16 h os valores se mantêm próximos. Às 17 h o valor médio de conexão atinge seu maior valor, representando uma conexão média de aproximadamente 40%. Os valores do gráfico mostram que entre 18 h e 21 h, ao diminuir o número de conexões, o fluxo de pessoas não está espalhado na cidade, porém às 22 h com o aumento da conexão é possível observar que a média ocorre em aproximadamente 21% das zonas.

- Coeficiente de aglomeração médio para cada hora do dia:

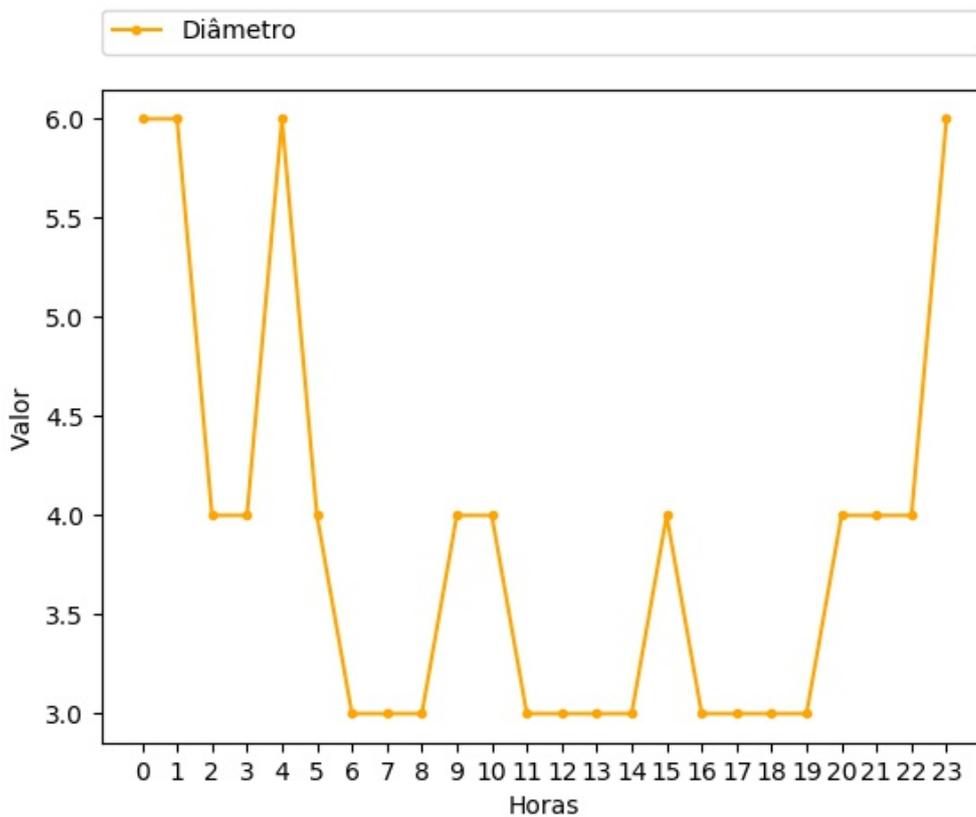
Figura 4.10 - Gráfico do Coeficiente de aglomeração médio com janela temporal de 60min.



A Figura 4.10 apresenta o coeficiente de aglomeração médio. Os maiores valores de aglomeração ocorrem às 7 h, 8 h, 17 h e 18 h e, no conceito de mobilidade urbana, isto significa que nesses instantes as zonas do grafo tendem a agrupar-se mais que nos outros horários, representando a conectividade entre zonas vizinhas.

- Diâmetro para cada hora do dia:

Figura 4.11 - Gráfico do diâmetro da rede com janela temporal de 60min.



O diâmetro de cada horário é apresentado na Figura 4.11 e refere-se ao conceito de caminho em grafos, ou seja, representa o maior menor caminho. Assim, durante a evolução temporal, 4 horários atingiram o valor máximo de 6 zonas indicando quais horários a rede são menos conectadas.

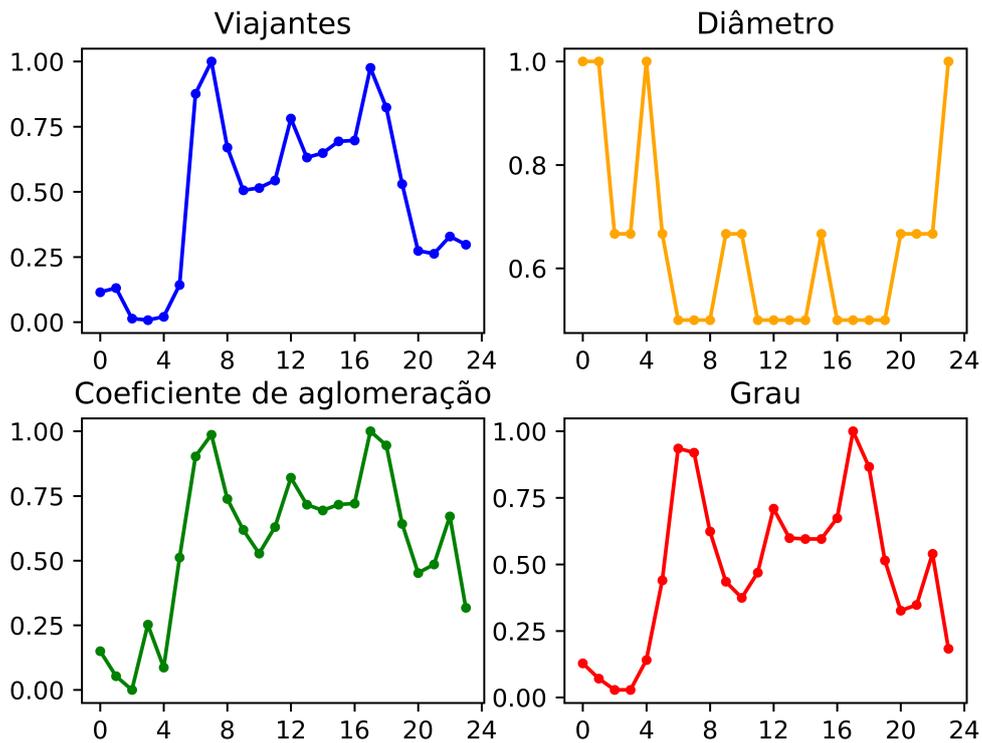
Para analisar e comparar as medidas topológicas calculadas os dados foram normalizados entre 0 e 1, a Tabela 4.1 apresenta o valor máximo de cada medida.

Tabela 4.1 - Fator de normalização (valor máximo) das medidas topológicas da rede

Medidas	Valor máximo
Grau médio	20.4
Coefficiente médio de aglomeração	0.594495
Fluxo Total	79664.8
Diâmetro	6

Com a normalização foi feito um agrupamento das medidas com os dados do fluxo total em um único gráfico, como na Figura 4.12.

Figura 4.12 - Gráficos gerados com os dados normalizados da medidas



A aglomeração representa o coeficiente de aglomeração médio da rede, o grau representa o grau médio da rede, viajantes o acumulado do fluxo, e o diâmetro re-

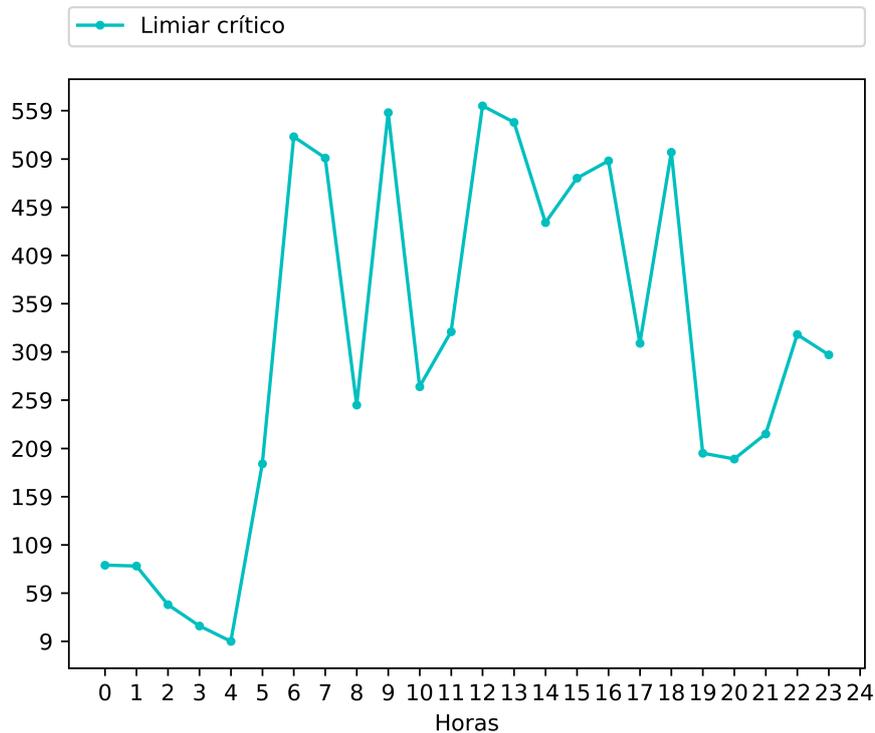
presenta o diâmetro da rede.

A aglomeração, grau e viajantes apresentam o mesmo comportamento na evolução temporal. Nos horários com alto fluxo de movimentação o potencial de conectividade e centralidade atinge altos valores, o que significa que as pessoas estão viajando para diversas zonas de tráfegos enquanto o diâmetro reflete o oposto do comportamento obtido no gráfico de grau, o que já é esperado, pois ao se ter altos valores de conexões entre zonas a probabilidade de existir um caminho menor entre elas é maior.

4.2.2 Limiar crítico

A partir do cálculo para encontrar o limiar crítico, cada hora do dia segue sua própria regra de criação das arestas. Na Figura 4.13 são apresentados os limiares para cada horário.

Figura 4.13 - Limiares crítico



A Figura 4.13 apresenta todos os limiares e 37,5% dos horários apresentam limiares acima de 400, enquanto o valor máximo observado é de 559.

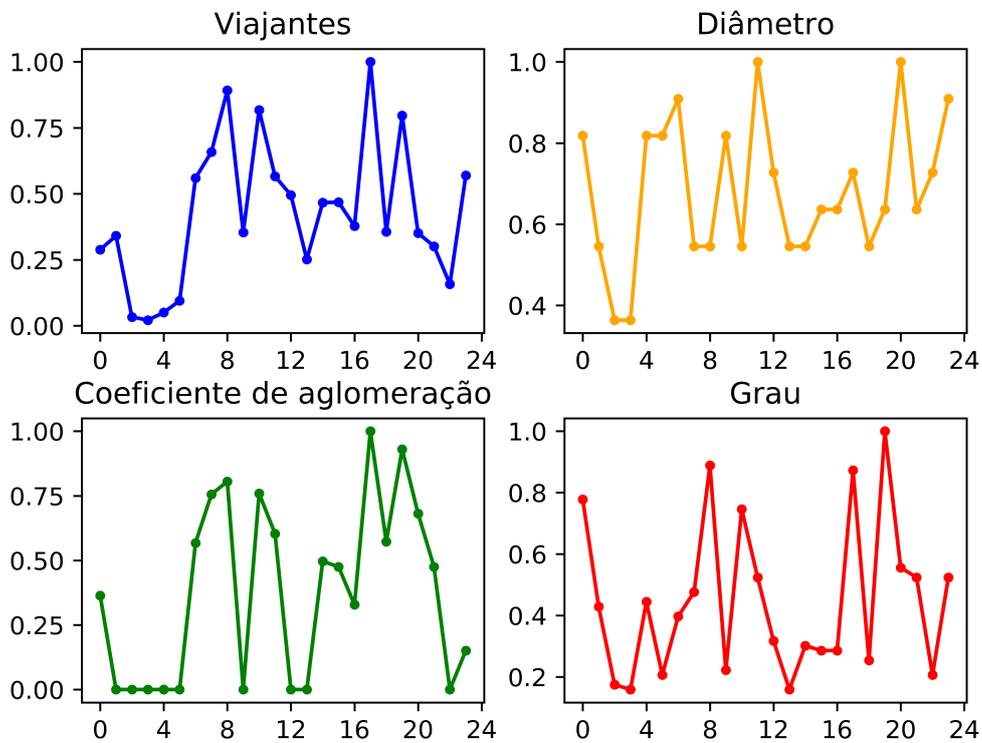
Para uma análise dos gráficos em conjunto, os dados foram normalizados entre 0 e 1. A Tabela 4.2 apresenta o valor máximo de cada medida.

Tabela 4.2 - Fator de normalização (valor máximo) das medidas topológicas da rede

Medidas	Valor máximo
Grau médio	2.30
Coefficiente médio de aglomeração	0.27
Fluxo Total	35050.25
Diâmetro	11

Com a normalização foi possível agrupar as medidas com os dados do fluxo limiarizado em um único gráfico, como na Figura 4.14.

Figura 4.14 - Gráficos gerados com os dados normalizados da medidas



Ao aplicar o limiar para cada horário os valores das medidas topológicas foram alterados, o valor para grau médio diminuiu para 2.30, o coeficiente de aglomeração médio para 0.27 e o diâmetro para 11, o que demonstra que uma alta redução de vértices conectados. Porém, mesmo com a alta redução de conexões as zonas de alto fluxo foram preservadas. O decaimento no fluxo, coeficiente de aglomeração e grau ao meio-dia, horário de almoço, significa que neste horário não há muitas conexões entre zonas com alto fluxo, muitas pessoas se locomovem dentro ou próximo da própria zona.

Alguns horários apresentaram características peculiares:

- Às 0h: Nesse instante apesar do baixo fluxo de viajantes e de coeficiente de aglomeração o índice de conectividade, o grau, apresenta altos valores. O diâmetro também apresenta um alto valor.
- Às 7h: Nesse horário o grau e o fluxo apresentam valores baixos, enquanto o diâmetro e o coeficiente de aglomeração estão altos.
- Às 12h: Ao meio-dia o fluxo de viajantes não está mais entre os maiores e o coeficiente de aglomeração reduziu para zero.
- Às 17h: O horário se manteve como o de maior fluxo, porém o grau diminuiu enquanto o coeficiente de aglomeração apresenta o maior valor do dia.
- Às 22h: As medidas de fluxo, coeficiente de aglomeração e grau apresentaram valores baixos e o diâmetro apresentou um valor médio.

Nos outros horários o fluxo, aglomeração e grau apresentaram um comportamento semelhante. Para enriquecer o estudo, os 5 horários diferenciados serão analisados.

4.3 Comunidades

Para identificar e analisar estruturas de comunidades, gráficos da modularidade e (geo)grafos com propriedades topológicas foram criados. Todas as análises seguiram as regras de fluxo completo e limiar crítico para a rede estática e para rede dinâmica representada por 5 horários: 0h, 7h, 12h, 17h e 22h.

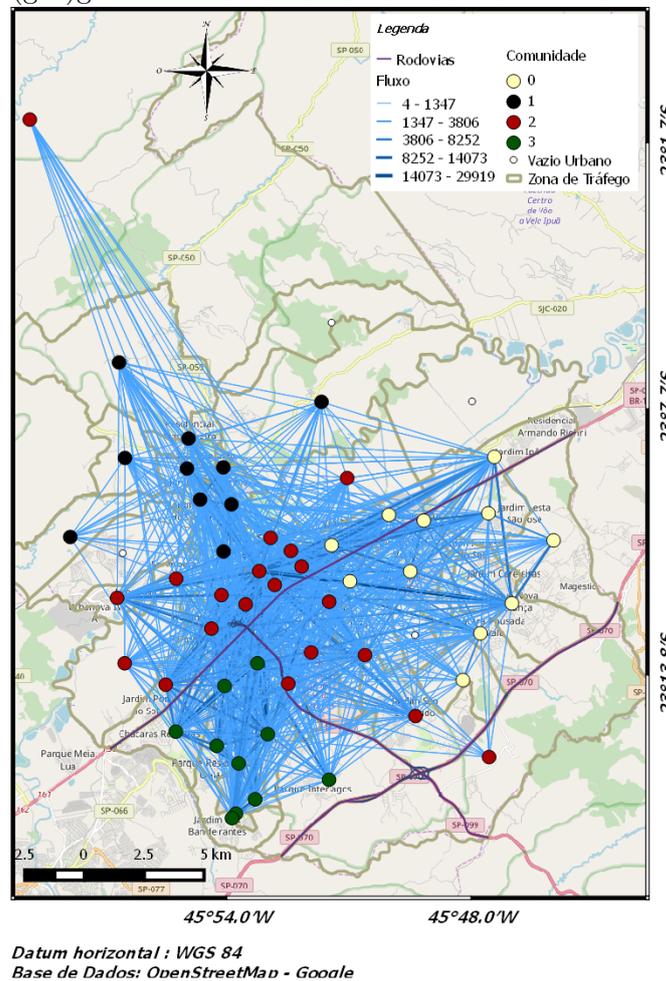
Os *shapefiles* foram criados pela GeoCNet. O vértice de cada zona está relacionado comunidade e cada cor está relacionada a sua macrozona. Se houver mais de uma

comunidade em uma macrozona elas serão diferenciadas por tons, como por exemplo, se na macrozona sul houver duas comunidades diferentes, uma será representada pelo tom de verde-escuro, enquanto a outra será verde-claro. Além disso, caso uma comunidade tenha membros em mais de uma macrozona, a cor estará relacionada a macrozona que tiver mais membros da comunidade.

4.3.1 Rede Estática

A Figura 4.15 representa as estruturas de comunidades encontradas para o acumulado do dia com o limiar 1 com $Q = 0.209$.

Figura 4.15 - (geo)grafo do acumulado do dia da cidade de São José dos Campos

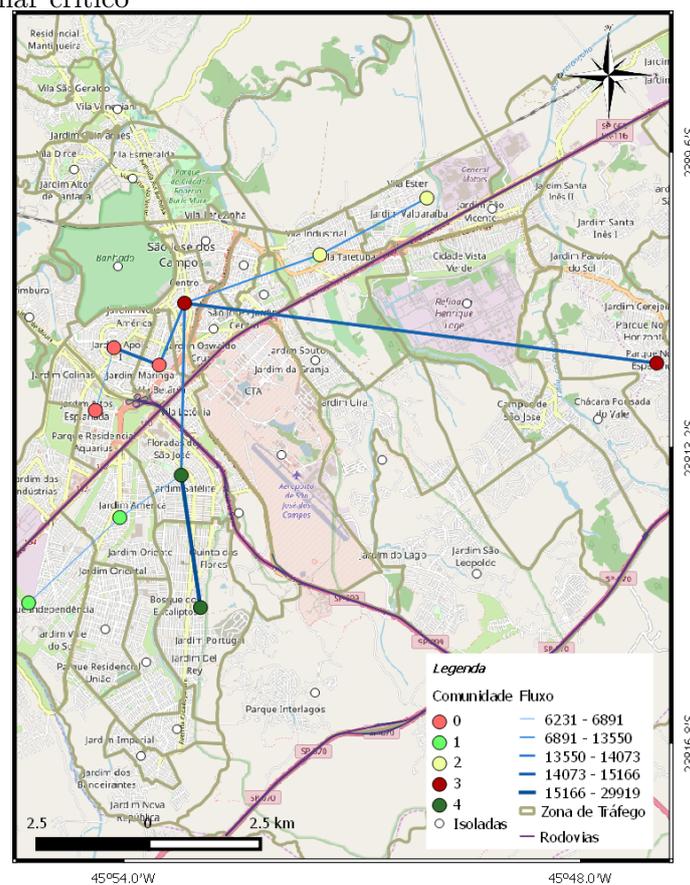


Apesar de não conseguir extrair medidas topológicas significativas, o algoritmo de detecção de comunidades para rede ponderada revelou comunidades coeren-

tes, mesmo com modularidade de 0.209. Quatro comunidades foram detectadas, mostrando a cidade bem dividida por suas regiões. Em cada comunidade há uma maior relação de deslocamento dentro da comunidade do que fora. A comunidade 0 (amarela) é composta por 11 vértices, todos da região leste da cidade. A comunidade 1 (preta) é composta por 10 vértices da região norte. A comunidade 2 (vermelha) é formada por vértices de 5 regiões distintas, extremo norte, oeste, centro, leste e sudeste. O vértice do extremo norte é separado geograficamente do restante da comunidade por uma outra comunidade. O vértice de maior grau é o da zona 1 (centro), com grau 49. A comunidade 3 (verde) é composta por 10 vértices da região sul da cidade.

A Figura 4.16 representa a comunidade do acumulado do dia para o limiar crítico com $Q = 0.410$.

Figura 4.16 - (geo)grafo do acumulado do dia da cidade de São José dos Campos com limiar crítico



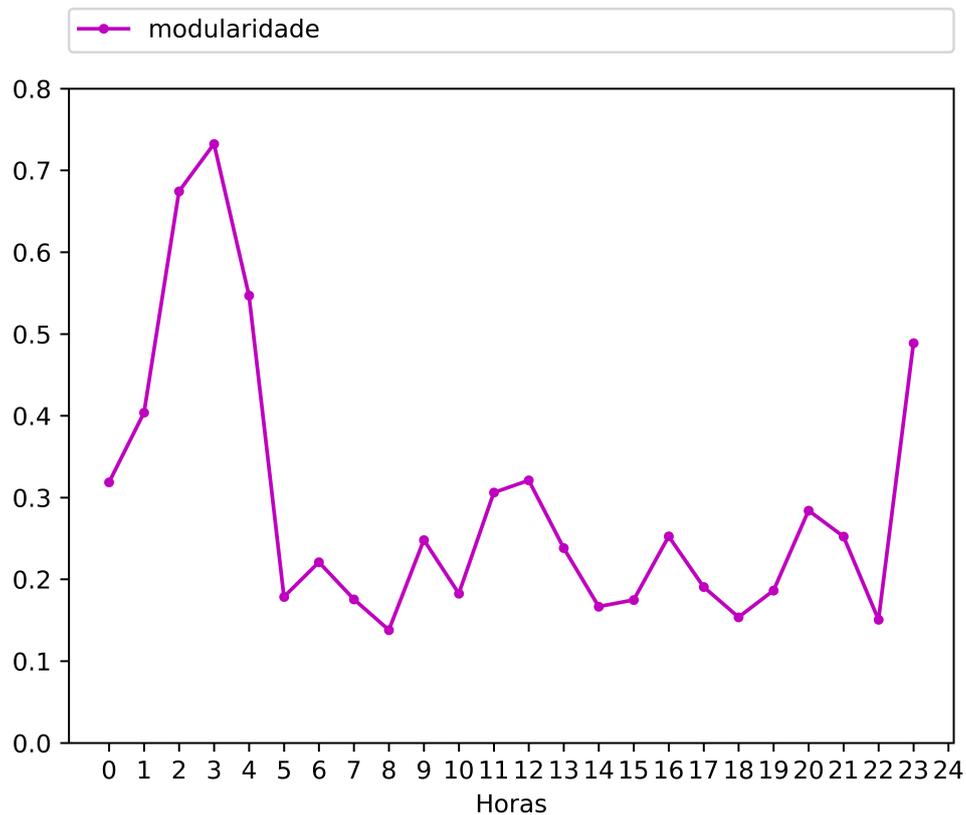
Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Google

Ao analisar as comunidades com limiar crítico de 6231, observa-se que há 11 vértices no (geo)grafo, divididos em 5 comunidades. Três comunidades apresentam pares conexos, dessas três, duas estão na região do sul, comunidade 1 e 4, e a outra na região leste, comunidade 2. Na região do centro há uma comunidade com um 3 vértices próximos, comunidade 0, dois vértices da região central e uma da região oeste. A comunidade 3 apresenta 2 vértices distantes, conexão entre os vértices 1 e 44, que representa a maior distância neste horário, com 11,6 km.

4.3.2 Rede Dinâmica

A modularidade de cada horário para o fluxo completo é apresentada na Figura 4.17.

Figura 4.17 - Modularidade para limiar 1



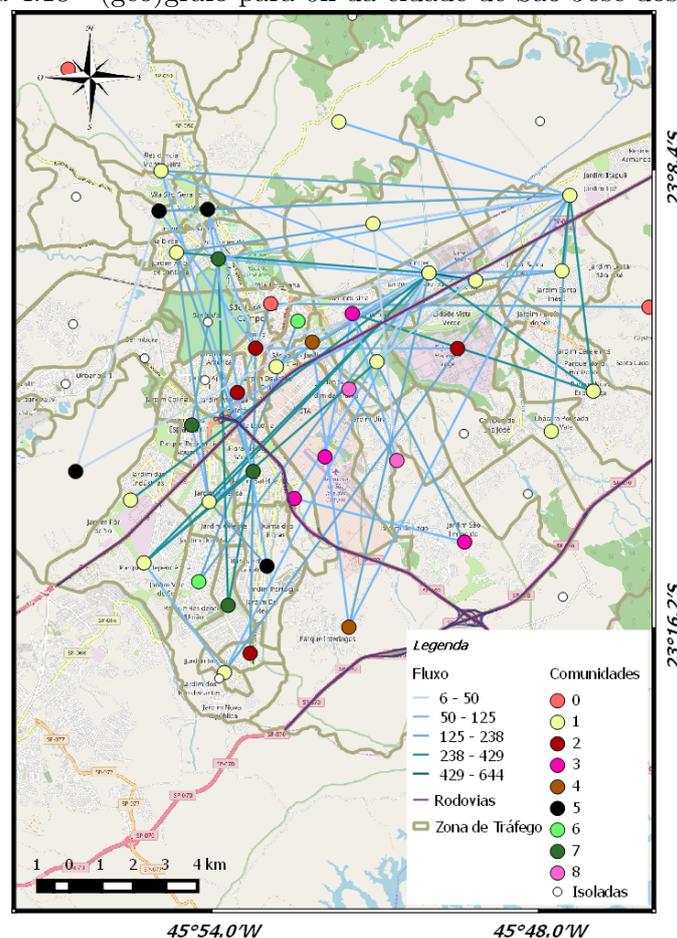
No gráfico da Figura 4.17 há uma clara presença de estrutura modular entre as 2h e 4h e às 23h. Nos demais instantes, as comunidades não estão claramente defini-

das pela modularidade. Apesar dos baixos valores da modularidade para as comunidades de cada horário, na literatura não foi encontrado trabalhos que avaliaram se a medida é uma boa medida de qualidade de comunidade para (geo)grafos.

- Fluxo completo:

A Figura 4.18 representa a comunidade para 0h.

Figura 4.18 - (geo)grafo para 0h da cidade de São José dos Campos



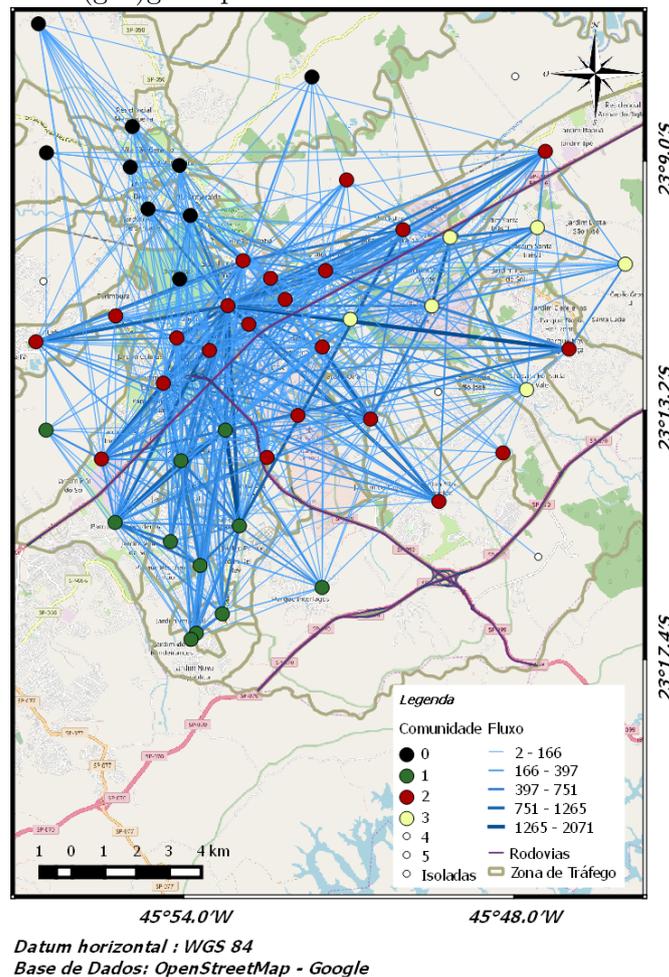
Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Google

Para a primeira hora do dia a comunidade com mais vértices é a comunidade 1, representada pela cor amarela, composta por 16 vértices, áreas industriais, centrais e residenciais fazem parte da mesma comunidade. Apenas a região sudeste não tem nenhum vértice nessa comunidade. O vértice da zona Vila Tesouro e Jardim

Universo, da comunidade 1, é o de maior grau: 18. A maior distância entre dois pares de vértices (31 e 42) é de 18,6 km, também membros da comunidade 1. A comunidade 3 tem 4 vértices, 3 no sudeste (11,12 e 23) e 1 no leste (9). As demais comunidades são conjuntos ou pares simplesmente conexos ou desconexos espalhados pelas regiões da cidade.

A Figura 4.19 representa a comunidade para 7h.

Figura 4.19 - (geo)grafo para às 7h da cidade de São José dos Campos

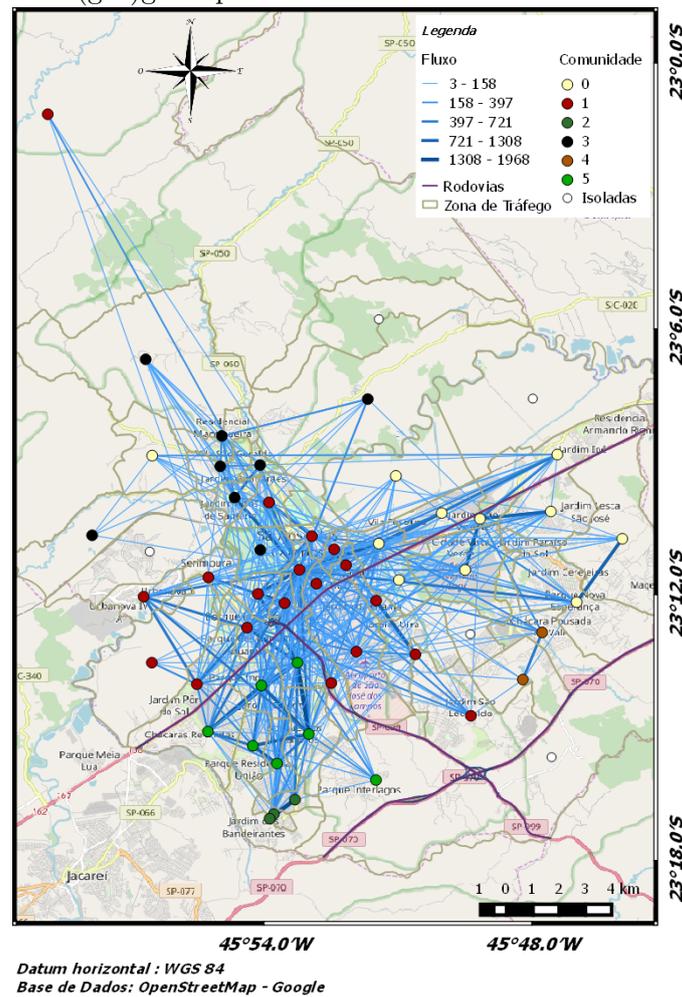


Nesse horário a cidade é dividida em 4 comunidades. A comunidade 0 é composta por 9 vértices, a 1 por 11, a 2 por 22 e a 3 por 6. A comunidade 0 é predominante na região norte da cidade. A comunidade 1 é na região sul, com apenas um vértice na região oeste, separada pela Rodovia Presidente Dutra (Dutra). A comunidade 2 vai de leste a oeste, crescendo a partir do lado norte da Dutra ao longo de toda

a cidade, possui também vértice no sudeste. Com grau 44, a zona 1, é a zona de maior grau e se conecta com todas as comunidades. Também fazem parte da comunidade as zonas 7,9,6 e 14 possuem grau 38,38,36 e 35 respectivamente, estes são os 5 maiores graus para este horário. A comunidade localiza-se na região leste da cidade e não apresenta membros separados pela Dutra.

A Figura 4.20 representa a comunidade para 12h.

Figura 4.20 - (geo)grafo para às 12h da cidade de São José dos Campos

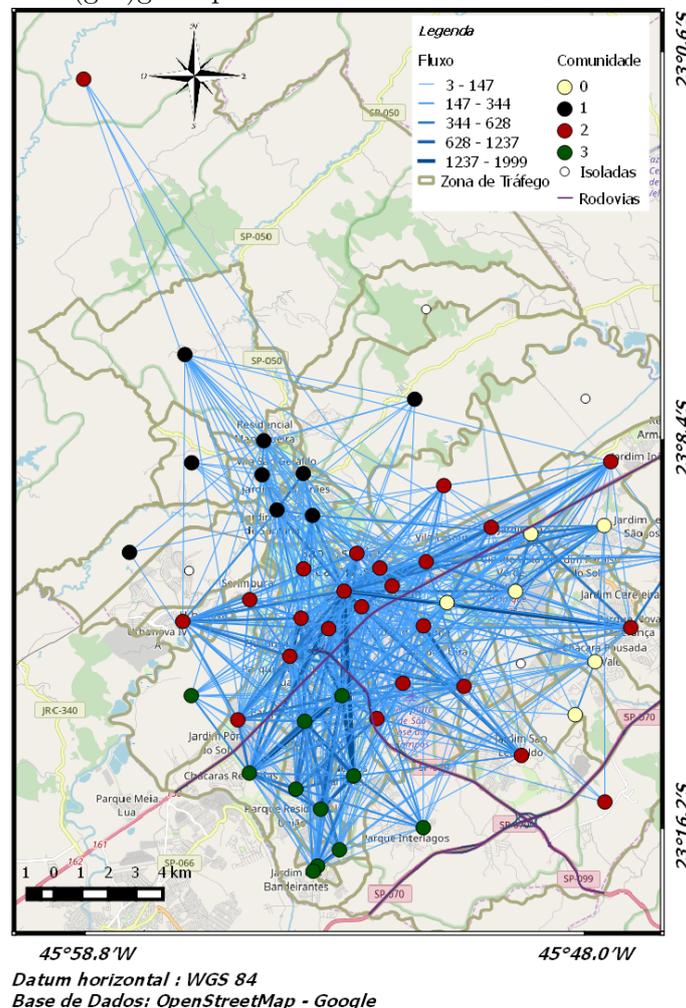


Ao meio-dia o cenário apresentado anteriormente muda: agora com 6 comunidades, conjuntos desconexos de regiões conexas são detectadas. A comunidade 0 representa a comunidade predominante na região leste, mas com vértices por ambos os lados da Dutra, 4 para o norte e 7 para o sul. Apenas um vértice está distante,

pertencendo à região do extremo norte. A comunidade 1 é predominante na região central, composta por 18 vértices, apenas 5 estão na região sudeste e 1 no extremo norte, que se conecta com a zona do centro, resultando na maior distancia entre dois pares em uma comunidade, 21,28 km. A região Sul possui duas comunidades, 2 e 5; a comunidade 2 é um pequeno conjunto simplesmente conexo, composta por 3 vértices, a 5 possui 7 vértices simplesmente conexos. A comunidade 3 é predominante na região norte, com apenas um vértice da região central. A comunidade 4 é um par simplesmente conexo da região leste.

A Figura 4.21 representa a comunidade para 17h.

Figura 4.21 - (geo)grafo para às 17h da cidade de São José dos Campos

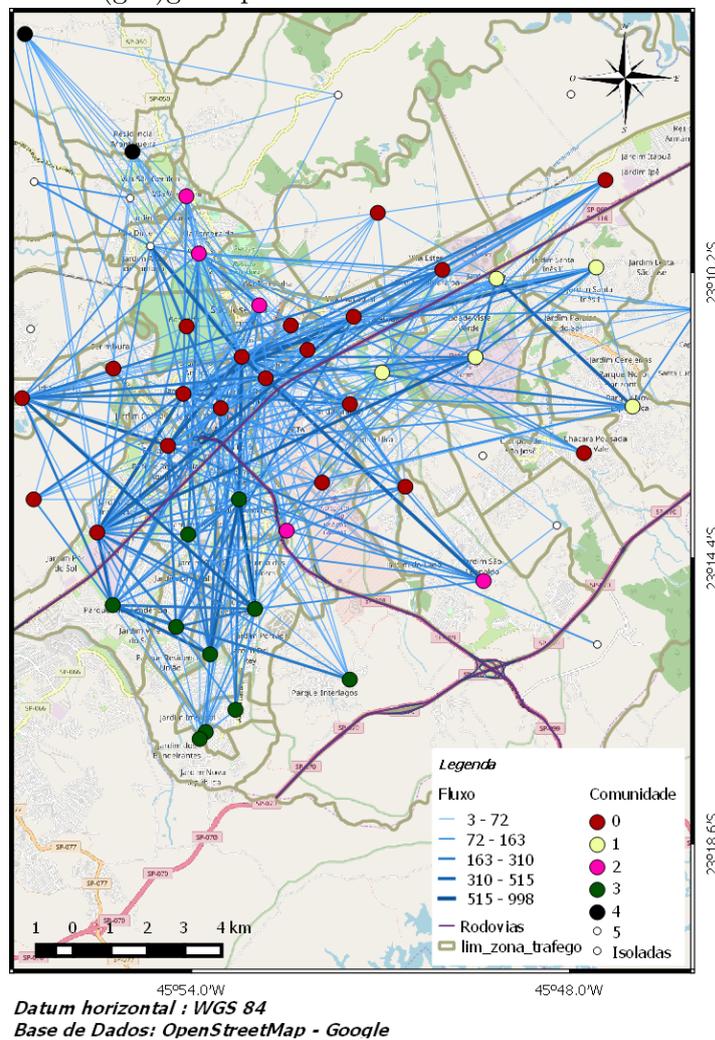


Neste horário a detecção de comunidade é semelhante ao que foi apresentado 07h.

A comunidade 0, com 7 vértices, é predominante na região leste. A comunidade 1, com 9 vértices, é predominante na região norte. A comunidade 2, com 24 vértices, vai de leste a oeste, crescendo a partir da direção norte da Dutra ao longo de toda a cidade, possui também vértice no sudeste e possui apenas um vértice desconexo no extremo norte, os demais estão espalhados pela cidade entre as regiões oeste, centro, sudeste e leste de forma conexa entre todas as zonas. A comunidade 3, com 11 vértices, possui apenas um vértice da região oeste, todos os outros estão na região sul e do mesmo lado da Dutra.

A Figura 4.22 representa a comunidade para 17h.

Figura 4.22 - (geo)grafo para às 22h da cidade de São José dos Campos



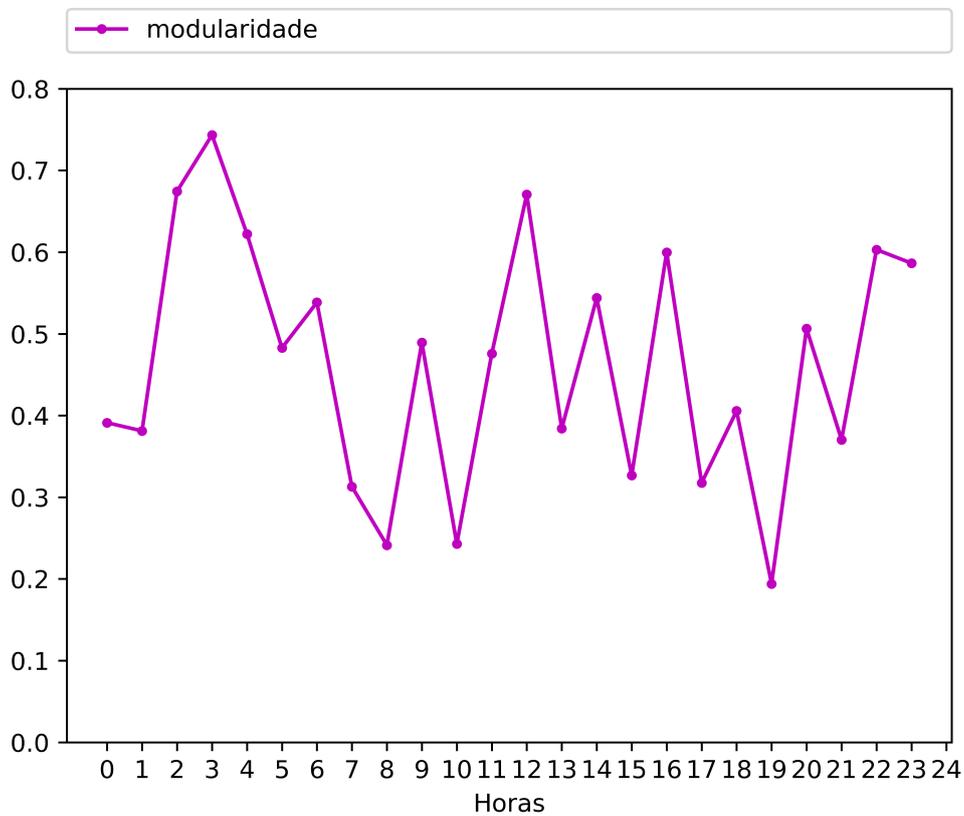
Neste horário são detectadas 5 comunidades, a comunidade 0, com 20 vértices, re-

pete o comportamento de ir de leste a oeste e crescer no centro a partir do lado norte da Dutra e seguir também para o sudeste. A comunidade 1 com 6 vértices se concentra na região leste abaixo da Dutra. A comunidade 2 sendo 2 conjuntos desconexos de regiões conexas. A comunidade 3 é composta por 10 vértices na região sul, todos conexos. A comunidade 4 apresenta um par conexo no extremo norte.

- Limiar crítico:

Modularidade de cada horário para o limiar crítico é representado na Figura 4.23.

Figura 4.23 - Modularidade para limiar crítico

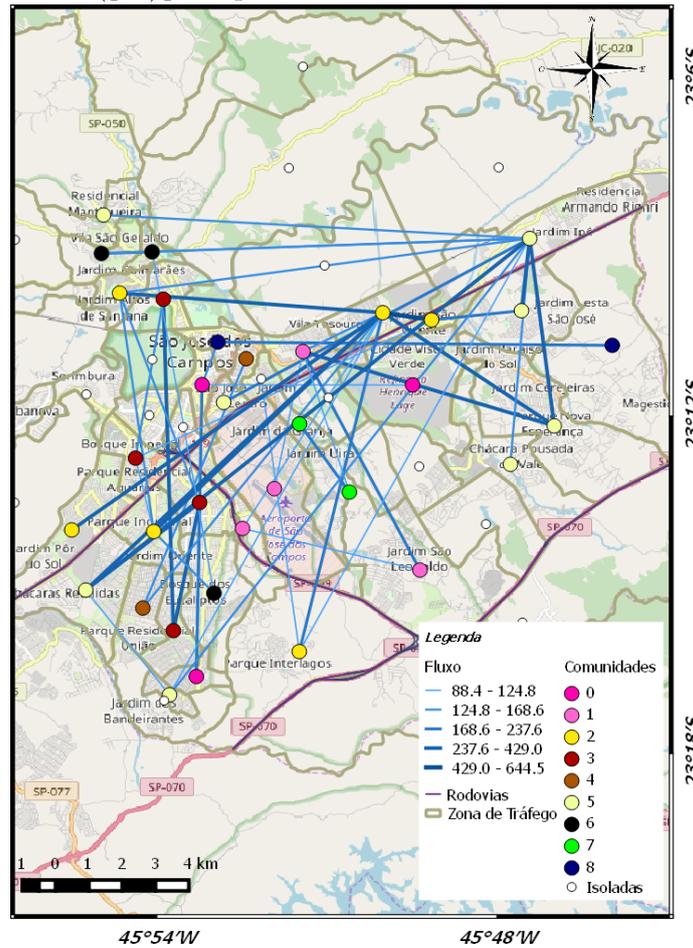


A Figura 4.23 ilustra a evolução temporal da modularidade para redes com o limiar crítico. Nesse cenário, a modularidade apresenta um valor superior a 0,3 para praticamente toda a evolução. Segundo Newman e Girvan, (NEWMAN; GIRVAN, 2004), valores de modularidade superiores a 0,3 podem indicar a existência de uma estrutura de comunidades na rede. Assim, pode-se dizer que, com exceção

de três momentos (8h, 10h e 19h), a rede apresenta comunidades. Em especial, destaca-se o alto valor de modularidade obtido em 3h, o que indica uma clara divisão da rede em comunidades.

A Figura 4.24 representa a comunidade para 0h.

Figura 4.24 - (geo)grafo para 0h da cidade de São José dos Campos.



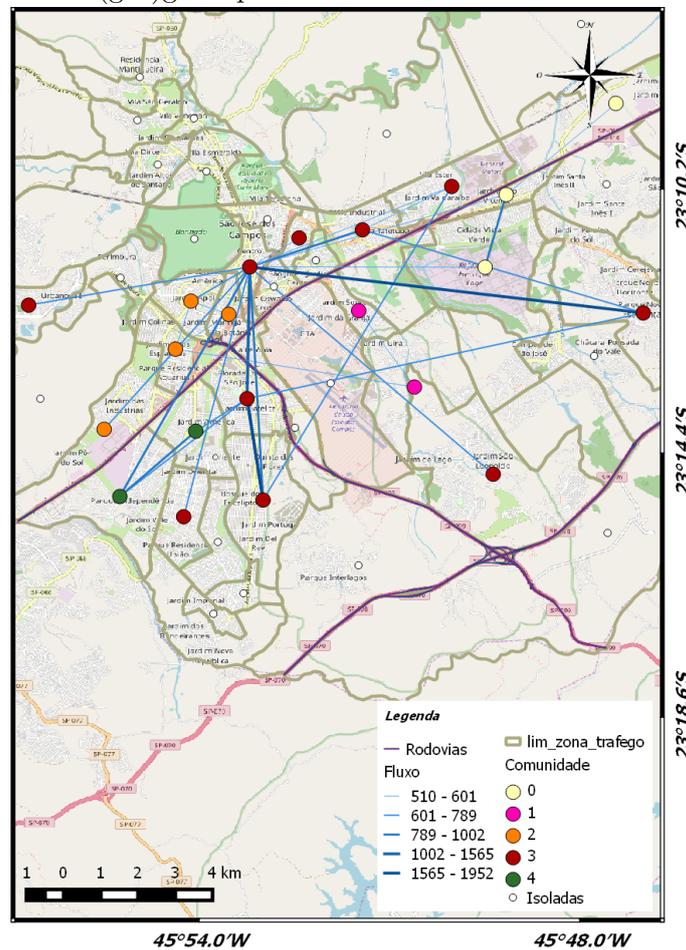
Datum horizontal : WGS 84
Base de Dados: OpenStreetMaps - Google

Na primeira hora do dia 9 comunidades são apresentadas, e duas zonas possuem grau 11: a zona de ID 18 (Vila Tesouro e Jardim Universo) e ID 42 (Eugênio de Melo e Galo Branco). A zona 18 apresenta alto fluxo entre regiões industriais como Jardim Motorama, Jardim Nova Detroit, Jardim Alvorada e Jardim das Indústrias; Eugênio de Melo e Galo Branco também é uma área industrial. Ambas possuem 5 conexões em comum. A zona 18 é a zona com mais conexões da comu-

nidade 2, a zona 42 é a zona com mais conexões da comunidade 5. A comunidade 2 é composta por 6 vértices e a maior distancia entre dois vértices é de 11,76 km. Três vértices estão do lado norte da Dutra e 3 estão do lado Sul. A comunidade 5 é composta por 8 vértices: 4 vértices próximos e 4 vértices distantes, a maior distancia entre os vértices nesta comunidade é de 18,49 km. Três vértices estão do lado norte da Dutra e 5 estão do lado Sul. Demais comunidades possuem conjuntos e pares conexos, mas sem grandes observações.

A Figura 4.25 representa a comunidade para 7h.

Figura 4.25 - (geo)grafo para 7h da cidade de São José dos Campos



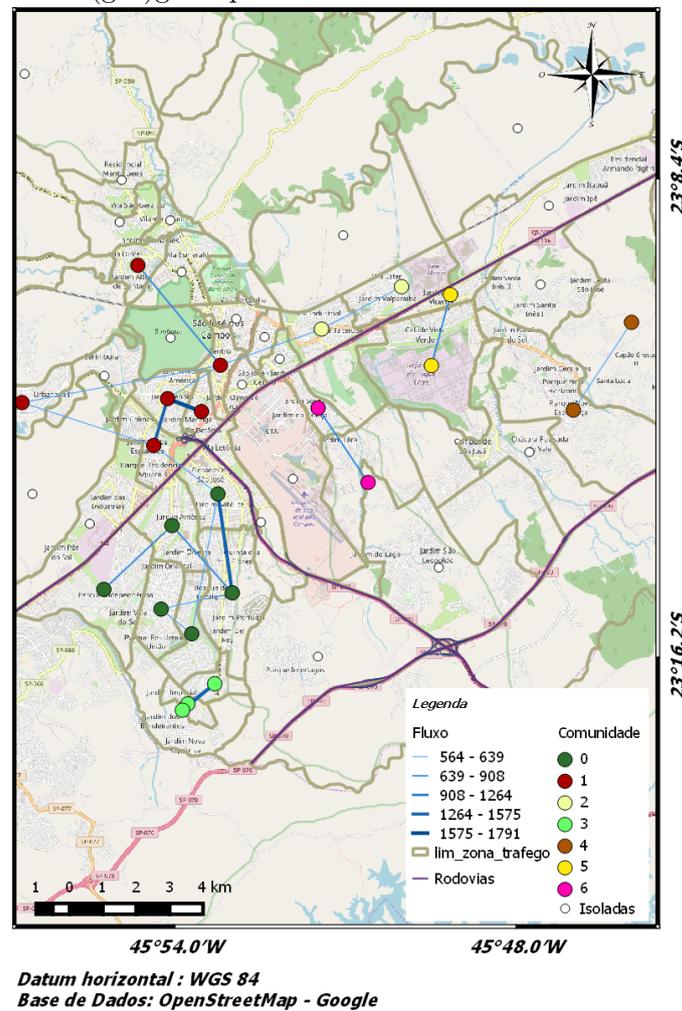
Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Google

Às 07h o número de comunidades é reduzido para 5 e com uma comunidade principal na área central da cidade, a zona de ID 1 (Centro), de grau 14, é a zona

mais conectada da comunidade verde, composta por 10 vértices. Possui um conjunto com 4 vértices próximos, incluindo o *hub*, um par na região sul e 4 vértices distantes da região central. A conexão com maior fluxo encontra-se entre a zona central e o parque novo horizonte e residencial Dom Bosco que também é a mais distante com 10,67 km entre seus centroides. A segunda maior comunidade, laranja, possui 4 vértices, todos conexos entre si. Dois vértices pertencem à região centro e dois a região oeste e todos à mesma margem da Dutra.

A Figura 4.26 representa a comunidade para 12h.

Figura 4.26 - (geo)grafo para 12h da cidade de São José dos Campos.

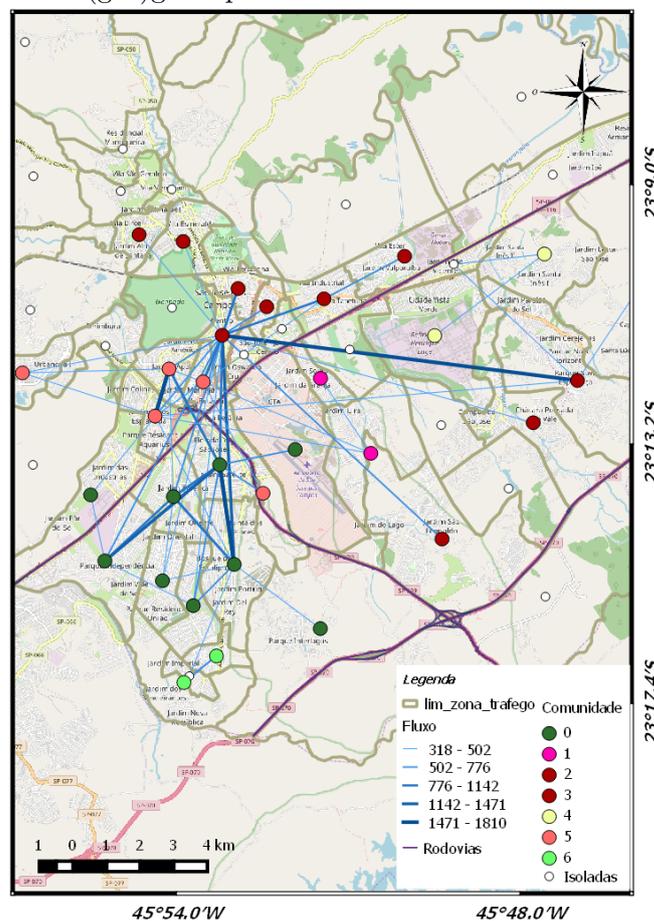


Ao meio-dia duas comunidades se destacam, ambas com o mesmo número de vértices: 6. A comunidade 0, verde escuro, apresenta um conjunto de vértices conexos

e todos na região sul da cidade. O vértice com maior concentração de fluxo é o vértice da Zona 25 (Bosque dos Eucaliptos e Quinta das Flores). A distancia máxima entre dois pares de vértices é de 4,3 km, (28 e 13). A comunidade 1, vermelha, é composta por vértices da região do centro (1, 6 e 7), oeste (34 e 14) e norte (16) . Os vértices 1, 6, 7, e 14 compõe o conjunto conexo da comunidade. Nessa comunidade a distancia de 6,6 km é a maior e ocorre entre o centro e o Urbanova.

A Figura 4.27 representa a comunidade para 12h.

Figura 4.27 - (geo)grafo para 17h da cidade de São José dos Campos.



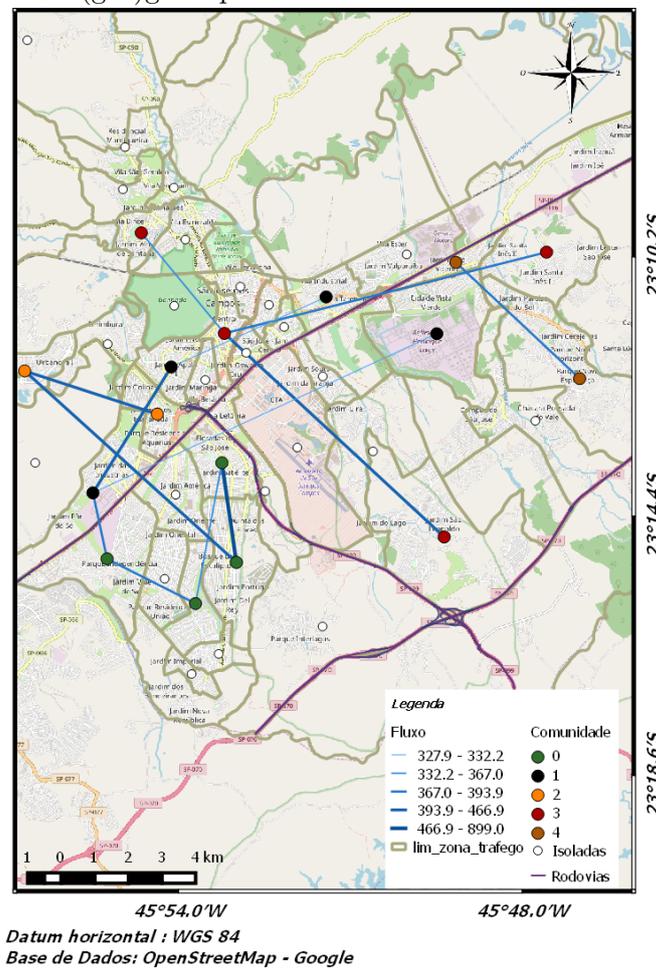
Datum horizontal : WGS 84
Base de Dados: OpenStreetMap - Google

Neste horário a comunidade da região sul aumenta para 9 vértices, acrescentando um vértice da própria região sul, um da região oeste e um da região sudeste. Os vértices 13 e 25 possuem o mesmo grau, 8, mas o vértice 13 apresenta a maior

concentração de fluxo passando a ser a zona de alta conexão da rede. A comunidade da região central também cresce para 8 vértices e se espalha para o norte (16 e 8), o centro mantém apenas o vértice 1 e acrescenta o 2 e o 3 da própria região, cresce para o sudeste (23) e para o leste (21 e 44), a maior distância entre os vértices é de 11,6 km entre zona 1 e a zona 44. Neste horário o centro apresenta grau 19, se conectando também com vértices de outras comunidades. Os vértices 6, 7, 34 e 14 não fazem mais parte da comunidade da zona do centro (1), mas continuam na mesma comunidade, acrescentando um vértice (12) da região sul.

A Figura 4.28 representa a comunidade para 22h.

Figura 4.28 - (geo)grafo para 22h da cidade de São José dos Campos.



Às 22h a comunidade da região sul, cor verde, apesar de diminuir para 4 vértices, continua conexa, com dois vértices se comunicando com comunidades externas

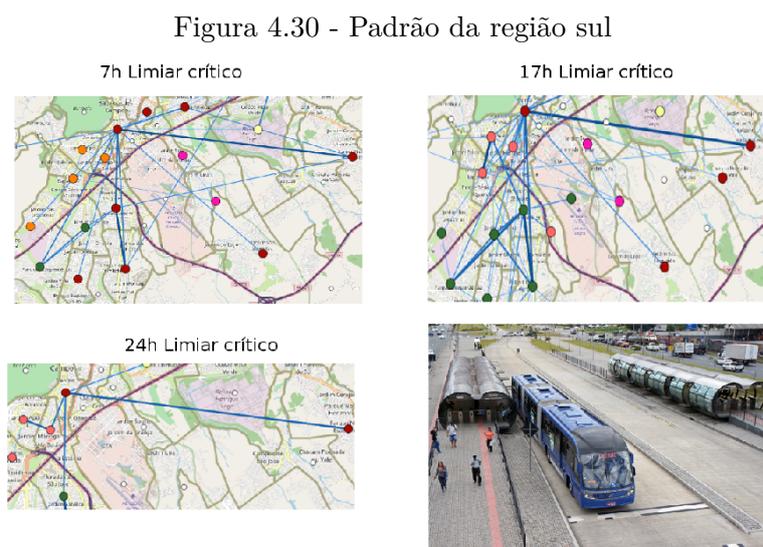
(Jardim industrial e Urbanova) . As zonas pertencentes a comunidade 1 estão na região norte (16), leste (43) e sudeste (23).

4.4 Planejamento urbano

Os resultados foram apresentados ao IPPLAN e duas observações foram feitas: uma em relação a região sul, apresentada na Figura 4.29 e outra para apoiar políticas de transporte, apresentada na Figura 4.30.



A Figura 4.29 demonstra que há um padrão na estrutura de comunidades ao longo do dia, o que pode ser justificado por ser a região que possui um expressivo subcentro comercial e de serviços, com a maior população e a maior densidade demográfica do município.



A Figura 4.30 refere-se a segunda observação relacionada à política de transporte. Existe um planejamento para a implantação de um sistema de transporte rápido por ônibus (BRT) entre zonas. O sistema BRT opera basicamente como o metrô, por meio de pistas exclusivas e o planejamento estuda implementar duas estações, uma da zona do centro com a região leste e outra da zona do centro com a região sul. Este estudo confirma o deslocamento entre as zonas.

5 CONCLUSÕES

Nesta dissertação foi realizado um estudo da mobilidade urbana da cidade de São José dos Campos, utilizando-se redes complexas

Duas modelagens dos dados foram implementadas: rede estática (acumulado do dia) e rede dinâmica (variação temporal). Na rede estática e dinâmica o processamento se dá em três fases: pré-processamento, processamento e pós-processamento.

Para estabelecer a topologia da rede, foram consideradas duas abordagens para regra de criação da conexão: a primeira permitindo qualquer viagem entre as zonas, representando o fluxo total. A segunda, estabelecendo um limiar crítico.

Para cada modelagem duas análises foram feitas. Em um primeiro instante, as medidas topológicas: grau, coeficiente de aglomeração e diâmetro, foram estudadas. Desses cálculos foi possível estabelecer o potencial de conectividade da rede e verificar o maior caminho mínimo entre os vértices conectados. Por fim, para determinar se existem e se há persistência de padrões espaço-temporais, o algoritmo de detecção de comunidade foi aplicado. A partir do estudo destas redes, o comportamento ao longo do dia pode ser observado.

A seguir, as principais contribuições desta dissertação, as respostas das perguntas científicas e as conclusões obtidas, os trabalhos futuros e as principais publicações são sumarizadas.

5.1 Principais contribuições

Com base no desenvolvimento realizado e descrito acima, podemos destacar como principais contribuições:

- exploração do uso combinado da componente espacial e a dinâmica temporal em redes complexas para estudo do comportamento da mobilidade. Este estudo focalizou-se em explorar as propriedades topológicas presentes ao longo do dia com uma estrutura dos objetos representados no espaço geográfico com relação espacial entre eles.
- desenvolvimento de uma nova ferramenta, a GeoCNet, que permite a criação de um *shapefile* com as propriedades topológicas do grafo para aplicar o conceito de (geo)grafos.

- verificação do comportamento ao longo do dia e da existência dos padrões espaço-temporais.
- discussão dos resultados com agentes especializados em mobilidade urbana do IPPLAN.

5.2 Perguntas científicas

- a) Como a rede se comporta ao longo do dia? Ou seja, qual a dinâmica espaço-temporal das propriedades topológicas da rede?

A abordagem unificando a componente espacial e a dinâmica temporal com uma ferramenta computacional que auxilia na visualização da rede (vértices, arestas e propriedades topológicas) se mostrou importante para entender o comportamento da mobilidade urbana. Nas medidas de centralidade e conectividade para o fluxo total e rede estática observou-se uma rede altamente conectada. Para o limiar crítico a conectividade da rede diminuiu e destacou a zona do centro como um *hub*. Para o fluxo total com a rede dinâmica os maiores valores ocorrem às 7hr, 12hr, 17hr (maior valor) e 22h. Os horários apresentados são horários chaves de movimentação. Para o limiar crítico a conectividade da rede diminuiu e destacou a zona do centro como um *hub*.

- b) Existem padrões espaço-temporais? Ou seja, há uma estrutura espacial dos deslocamentos?

A estrutura modular mostrou que no momento de alto fluxo a rede se torna muito conectada e a rede, como um todo, se torna um módulo. Nos momentos de baixo fluxo, a detecção de comunidade revelou padrões quando a cidade se divide em sub-regiões representadas pelas comunidades que persistem ao longo do dia (região sul).

- c) Se existirem: há persistência ao longo do dia, ou seja, a estrutura é a mesma para diferentes instantes da amostragem dos dados de mobilidade urbana?

A região sul da cidade teve poucas variações ao longo do dia, mostrando ser uma área que sofre pouca influência das outras regiões da cidade e um vértice do extremo norte se mostrou altamente dependente da região central da cidade. Da estrutura modular pode se destacar que os valores da modularidade para o limiar crítico são superiores aos do limiar 1.

Conclui-se que com o estudo por redes complexas é possível observar a mobilidade urbana com uma diferente perspectiva, o que permite verificar padrões espaço-temporais e apoiar estudos de planejamento urbano.

5.3 Trabalhos futuros

Embora as redes desenvolvidas apresentem diversas características interessantes, outras métricas e cenários podem agregar mais informações:

- Janela de tempo flutuante
- Cálculo de outras métricas como, por exemplo, força, *betweenness*, vulnerabilidade e árvore geradora mínima
- Aplicar os modelos desenvolvidos para outras cidades: São Paulo e Rio de Janeiro
- Incluir componente espacial na detecção de comunidades, para ser uma contribuição para um método de detecção de comunidade
- Análise de cenários de riscos de desastres com base nas análises de redes complexas.

5.4 Publicações

A Seguir as principais publicações:

- Santos, L. B. L.; Jorge, A. A. S.; Rossato, M.; Santos, J. D.; Candido, O. A. ; Seron, W. ; Santana, C. N. de, (geo)graphs - Complex Networks as a shapefile of nodes and a shapefile of edges for different applications, 2017, arXiv:1711.05879
- Santos, J. D.; Santos, L. B. L.; Bacelar, R. B.; Desastres e mobilidade urbana - um recorte da literatura, XXII Símposio Brasileiro de Recursos Hídricos (2017)
- Santos, J. D.; Santos, L. B. L.; Quiles, M. G.; Macau, E. E. N; XXXVII Congresso Nacional de Matemática Aplicada Computacional. Série temporal de índices de Redes Complexas em mobilidade urbana. 2017

- Santos, J. D.; Santos, L. B. L.; Quiles, M. G.; Macau, E. E. N; I
Workshop Brasileiro para Avaliação de Ameaças, Vulnerabilidades, Exposição e Redução de Risco de Desastres. Redes Complexas de mobilidade Urbana. 2017.
- Santos, J. D.; Santos, L. B. L.; Quiles, M. G.; Macau, E. E. N; II IPTMU
- Encontro sobre Impactos Potenciais de Inundações em Infraestruturas de Transporte e Mobilidade Urbana. Primeiros passos em análise espaço temporal de redes complexas de base territorializada. 2016.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARABÁSI, A. L.; ALBERT, R. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, p. 47–97, 2002. 8
- BARTHELEMY, M. Spatial networks. **Physics Reports**, v. 499, n. 1, p. 1–101, 2010. 1
- BESSA, A.; SANTOS, L.; MARTINEZ, L.; COSTA, M.; CARDOSO, P. **Introdução às redes complexas. Salvador: UFBA, 2013.** 8
- BORODIN, O. V.; IVANOVA, A. O.; MONTASSIER, M.; OCHEM, P.; RASPAUD, A. Vertex decompositions of sparse graphs into an edgeless subgraph and a subgraph of maximum degree at most k . **Journal of Graph Theory**, v. 65, n. 2, p. 83–93, 2010. ISSN 1097-0118. 6
- COSTA, L. d. F.; RODRIGUES, F. A.; TRAVIESO, G.; BOAS, P. V. Characterization of complex networks: a survey of measurements. **Advances in Physics**, v. 56, n. 1, p. 167–242, 2007. 1, 7
- COSTA, L. F.; N, O. J. O.; TRAVIESO, G.; RODRIGUES, F. A.; BOAS, P. V.; ANTIQUEIRA, L.; VIANA, M.; ROCHA, L. E. C. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329–412, 2011. 1
- DENSAR, U.; SPATENKOVÁ, O.; VIRRANTAUŠ, K. Identifying critical locations in a spatial network with graph theory. **Transaction in Gis**, v. 12, n. 1, p. 61–82, 2008. 7, 10
- DONGEN, S. van. **Graph clustering by flow simulation.** 2000. 173 p. Tese (Doutorado) — University of Utrecht, Smalingerland, 2000. 10
- EMMERICH, T.; BUNDE, A.; HAVLIN, S.; LI, G.; LI, D. Complex networks embedded in space: Dimension and scaling relations between mass, topological distance, and euclidean distance. **Physical Review**, v. 87,e032802, 2013. 11
- FERNANDES, V. A.; ROTHFUSS, R.; HOCKSCHILD, V.; SILVA, W. R.; SANTOS, M. P. Resiliência da mobilidade urbana: uma proposta conceitual. In: CONGRESSO NACIONAL DE PESQUISA EM TRANSPORTE DA ANPET, 29, Ouro Preto. p. 2759–2770, 2015. 1

FORTUNATO, S. Community detection in graphs. **Physrep**, v. 486, p. 75–174, feb. 2010. 9

GAO, S.; LIU, Y.; WANG, Y.; XIUJUN, M. Discovering spatial interaction communities from mobile phone data. **Transaction in Gis**, v. 17, n. 3, p. 463–481, 2013. 11

GUO, D.; JIN, H.; ANDRIS, C.; ZHU, X.; GAO, P. Discovering spatial patterns in origin-destination mobility data. **Transaction in Gis**, v. 16, n. 3, p. 411–429, 2012. 11

HOLME, P.; SARAMÄKI, J. **Temporal networks: understanding complex systems**. Springer, Heidelberg, 2013. 1

Instituto de Pesquisa, Administração e Planejamento. Atlas da pesquisa origem e destino - panorama da mobilidade em são josé dos campos. Editora Cubo, p. 1–144, 2014. 2, 13, 14, 16

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Plano Diretor do INPE 2016-2019: São José dos Campos, 2016**. São José dos Campos: INPE, 2016. 2

KINDERSLEY, D. **History of the world**. [S.l.]: Plantagenet S Fry, 2000. 1

LOPES S. B.; PFAFFENBICHLER, P. . S. A. Ma ferramenta para planejamento da mobilidade sustentável com base em modelo integrado de uso do solo e transportes. **Paranoá (UnB)**, v. 7, p. 93–102, 2012. 1

LOUAIL, T.; LENORMAND, M.; PICORNELL, M. ; CANTU, O. G.; HERRANZ, R.; MARTINEZ, E. F.; RAMASCO, J. J.; BERTHELEMY, M. Uncovering the spatial structure of mobility networks. **Nature Communications**, p. 1–8, 2015. 11

MONTIS, A. de; CHESSA, A.; CAMPAGNA, M.; CASCHILI, S.; DEPLANO., G. Modeling commuting systems through a complex network analysis. **Journal of Transport and Land Use** 2, p. 39–55, 2010. 1, 10

NEWMAN, M. Modularity and community structure in networks. **Proceedings of the National Academy of Sciences**, v. 103, n. 23, p. 8577–8582, 2006. 9, 10

NEWMAN, M.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E Covering Statistical, Nonlinear, Biological, and Soft Matter Physics**, v. 69, n. 2, p. 1–16, 2004. 9, 46

PONS, P.; LATAPY, M. Computing communities in large networks using random walks. **Computer and Information Sciences-ISCIS**, p. 284–293, 2005. 10

ROVALL, M.; BERGSTROM, C. aps of random walks on complex network reveal community structure. **Proceedings of the National Academy of Sciences**, v. 105, n. 4, p. 1118–1123, 2008. 10

SANTOS, L. B. L.; JORGE, A. A. S.; ROSSATO, M.; SANTOS, J. D.; CANDIDO, O. A.; SERON, W.; SANTANA, C. N. de. **(geo)graphs** - complex networks as a shapefile of nodes and a shapefile of edges for different applications. **arXiv**, v. 1, n. Disponível em: <https://arxiv.org/abs/1711.05879>, p. 1–5, 2017. 1, 16, 21, 22

SILVA, B. M. M.; SANTOS, J. D.; SANTOS, L. B. L. Primeiros passos em análise espaço temporal de redes complexas de base territorializada. In: ENCONTRO SOBRE IMPACTOS POTENCIAIS DE DESASTRE NATURAIS EM INFRAESTRUTURAS DE TRANSPORTE E MOBILIDADE URBANA, 2., 2016, São José dos Campos, SP. [S.l.]: Unesp, Cemaden, 2016. 16

SILVA, B. M. M. da. **Representação de redes complexas de mobilidade urbana em sistemas de informações geográficas**. Trabalho de conclusão de curso (Graduação) — Fatec, São José dos Campos, 2016. 16

SOFFER, S. N.; VÁZQUEZ, A. Network clustering coefficient without degree-correlation biases. **Physical Review E**, v. 71, p. 057101, 2005. 8

SOH, H.; LIM, S.; ZHANG, T.; FU, X.; LEE, G. K. K.; HUNG, T. G. G.; DI, P.; PRAKASAM, S.; WONG, L. Weighted complex network analysis of travel routes on the singapore public transportation system. **Physica A**, v. 389, p. 5852–5863, 2010. 11

TOMINAGA, L. K.; SANTORO, J.; AMARAL, R. d. Desastres naturais: conhecer para prevenir. **São Paulo : Instituto Geológico**, 2009. 1

TUCCI, C. E. M.; BERTONI C, J. **Inundações urbanas na América do Sul**. Porto Alegre: Associação Brasileira de Recursos Hídricos, 2003. 1

WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. **Nature**, v. 393, n. 23, p. 440–442, 1998. 8

WEINAN, E.; LI, T.; VANDEN-EIJNDEN, E. Optimal partition and effective dynamics of complex networks. **Proceedings of the National Academy of Sciences**, v. 105, n. 23, p. 7907–7912, 2008. ISSN 0027-8424. 10

YANG, B.; CHEUNG, W.; LIU, J. Community mining from signed social networks. **IEEE Transaction on Knowledge and Data Engineering**, v. 19, n. 10, p. 1333–1348, oct. 2007. ISSN 1041-4347. [10](#)

YANG, Z.; ALGESHEIMER, R.; TESSONE, C. A comparative analysis of community detection algorithms on artificial networks. **Scientific Reports**, v. 6, p. 1–18, 2016. [10](#)

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o International Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.