



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA
MESTRADO ACADÊMICO EM FÍSICA

RAUL PEIXOTO DA SILVA

CAMINHOS MÍNIMOS EM CIDADES: PADRÕES E CARACTERÍSTICAS

FORTALEZA

2018

RAUL PEIXOTO DA SILVA

CAMINHOS MÍNIMOS EM CIDADES: PADRÕES E CARACTERÍSTICAS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Física do Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Física. Área de Concentração: Física da Matéria Condensada

Orientador: Dr. Ascânio Dias Araújo

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

D11c da Silva, Raul Peixoto.

Caminhos mínimos em cidades: padrões e características / Raul Peixoto da Silva. – 2018.
60 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física, Fortaleza, 2018.

Orientação: Prof. Dr. Ascânio Dias Araújo.

1. Transporte. 2. Centralidade de intermediação. 3. Caminhos mínimos. I. Título.

CDD 530

RAUL PEIXOTO DA SILVA

CAMINHOS MÍNIMOS EM CIDADES: PADRÕES E CARACTERÍSTICAS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Física do Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Física. Área de Concentração: Física da Matéria Condensada

Aprovada em: 20/02/2018

BANCA EXAMINADORA

Dr. Ascânio Dias Araújo (Orientador)
Universidade Federal do Ceará (UFC)

Dr. Rilder de Sousa Pires
Universidade Federal do Ceará (UFC)

Dr. Luiz Ozório de Oliveira Filho
Universidade Federal do Ceará (UFC)

Dedico este trabalho à minha maior apoiadora,
Maria Analibia Peixoto de Moraes.

AGRADECIMENTOS

Primeiramente, agradecer aos meus pais e a minha família, em especial à minha mãe *Maria Analibia Peixoto de Moraes* que não está mais ao meu lado. A minha cōnjuge *Alice Felix da Silva* por ter me dado a melhor companhia que poderia ter e juntamente a ela o meu querido filho.

Agradeço ao professor *Ascânio Dias Araújo* pela dedicação, confiança e paciência para com minha pessoa ao longo desses dois anos que tive oportunidade de ser seu orientando de mestrado.

Devo parte das minhas conquistas a todos os meus amigos, sem a ajuda e apoio deles com certeza eu não teria chegado até aqui, em especial: *Danilo Borges da Silva* e *Nathanael Souza Bandeira*, pelo apoio, amizade e ajuda nas dificuldades em programação, além deles: *Emanuel Pinheiro*, *Nícolas Carvalho*, *Sofia Magalhães*, *Diego Rabelo*, entre outros.

Ao CNPq pelo apoio financeiro.

Enfim, agradeço a todos os que me ajudaram, direta ou indiretamente, para a elaboração e conclusão desse trabalho.

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

(Ariano Suassuna)

RESUMO

Entender a dinâmica de deslocamento das pessoas em uma cidade, é de fundamental importância para que haja um planejamento adequado e eficiente no sentido de trazer melhorias para a população. É fato que a dinâmica de deslocamento da população é regida pela forma como se estabelece a distribuição da população em bairros e também pela forma estrutural da cidade que é desenhada ao longo dos anos durante o seu desenvolvimento. Neste trabalho realizamos um estudo sistemático a cerca dos deslocamentos realizados pelos usuários de aplicativos de *GPS* em diversas cidades em continentes distintos do nosso planeta. Por meio de uma base dados, extraído do aplicativo *Google Maps*, analisamos os padrões de deslocamentos considerando vários pontos de origem e destino distribuídos aleatoriamente em regiões estabelecidas previamente de cidades como Boston, Lisboa, Porto, São Francisco, Rio de Janeiro e São Paulo. Nestas cidades, investigamos como são os padrões de deslocamentos entre origens e destinos tendo como base duas primícias básicas: menor distância e menor tempo. Estudamos as distribuições de tempo de trânsito considerando diferentes distâncias entre origem e destino, $r=1\text{km}$, 2km e 3km . Constatamos com base neste estudo, que essas se estabelecem na forma de uma distribuição Log-normal, independente da cidade considerada. Observamos que essas distribuições apresentam mudanças quando a distância entre origem e destino aumentam. Adicionalmente, realizamos um estudo acerca de quais vias são as mais utilizadas, ou seja, calculamos a *betweenness centrality* para as diferentes cidades tendo como base as rotas definidas entre os pontos escolhidos. Calculamos a distribuição de *betweenness centrality* para cada cidade, segundo os dois critérios da distância mínima e tempo mínimo e verificamos que as distribuições seguem uma função do tipo lei de potência seguida por um decaimento exponencial. Este comportamento apresentou-se invariante quanto ao critério e quanto a cidade, embora tenha sido observado variações nos valores para os expoentes que controlam o comportamento em lei de potência. Finalizamos o estudo, apresentando uma análise quanto a eficiência das cidades e quanto a orientação de suas ruas que formam a rede da cidade, na tentativa de explicar os padrões no comportamentos observados para os deslocamentos nessas cidades.

Palavras-chave: Transporte. Centralidade de intermediação. Caminhos mínimos.

ABSTRACT

Understanding the dynamics of displacement of people in a city is of fundamental importance for adequate and efficient planning in order to bring improvements to the population. It is a fact that the dynamics of commuters is driven by the way in which the population is spread in the city and also by the structural form of the city that is drawn over the years during its development. In this work, we perform a systematic study about the displacements carried out by users in different cities in different continents of our planet. Using a database, extracted from the Google-Maps application, we analyse the patterns of displacements considering several points of origin and destination randomly distributed in regions of cities such as Boston, Lisbon, Porto, São Francisco, Rio de Janeiro and São Paulo. In those cities, we investigate the patterns of displacements between origins and destinations based on two basic rules: minimal path and shorter time. We studied the distributions of transit time considering different distances between origin and destination, $r = 1\text{km}, 2\text{km}$ and 3km . From this study, we found that those distributions follow a Log-normal distribution, independent of the city considered. We also found, that those distributions present a change when the distance between origins and destinations considered increases. In addition, we performed a study about which routes are most used, that is, we have calculated the betweenness centrality for the different cities based on the routes defined between the chosen points. We calculate the distribution of betweenness centrality for each city, according to the two criteria of minimum distance and minimum time and we verify that those distributions are power law followed by an exponential decay. This behaviour was invariant in terms of the established criteria and the city, although changes were observed in the power law exponents. We conclude the study, by presenting an analysis about the cities efficiency and the orientation of their streets network, in an attempt to explain the patterns in the behaviour observed for the displacements in those cities.

Keywords: Transport. Betweenness centrality. Minimal paths.

LISTA DE FIGURAS

Figura 1	– Exemplificação de um congestionamento em uma cidade, além do caos e dos visíveis problemas causados por ele (INTERNET, 2017a).	14
Figura 2	– Ilustração do aplicativo <i>Waze</i> da empresa Google que indica o melhor caminho baseado em notificações de usuários e dos tempos das vias (INTERNET, 2017b).	16
Figura 3	– Imagem descrevendo a configuração das sete pontes de Königsberg, o rio Prególia está na cor azul enquanto as pontes estão na cor vermelha, o mapa descreve a cidade de Königsberg no ano de 1736 (INTERNET, 2017c).	22
Figura 4	– Imagem descrevendo a configuração dos pontos e ligações da cidade de Königsberg, sendo assim provavelmente o primeiro grafo da história (INTERNET..., 2017a).	23
Figura 5	– Imagem ilustrando a cidade de Königsberg de uma forma direcionada, as direções foram distribuídas aleatoriamente apenas para ilustração (INTERNET..., 2017a).	23
Figura 6	– Resultado mostrando a variação da distribuição log-normal mantendo o parâmetro $\mu = 0$ fixo e alterando apenas o parâmetro σ	27
Figura 7	– Direção e orientação das vias das cidades. Nas figuras do lado esquerdo temos representado a distribuição para as direções e orientações da cidade de Boston, já no lado direito temos o diagrama de rosa para o mesmo caso.	29
Figura 8	– Grafo da cidade de Lisboa. Grafo representando toda a região de estudo da cidade de Lisboa obtida utilizando o método de extração do OSM© descrito. Neste grafo estão sendo descritos 11890 nós e 13465 links.	32
Figura 9	– Demarcação da área de estudo. Amostra da região selecionada em que se é realizado o estudo. Conforme pode ser visto, é selecionado um quadrado com 5 km de lado e a partir dele são consideradas todas as ruas que possuem a <i>tag = highway</i> . A escala está demonstrada de modo qualitativo, apenas para melhor compreensão (INTERNET..., 2017b).	36

Figura 10 – Imagem demonstrando a diferença entre o caminho mais curto geometricamente e o caminho mais rápido na cidade de São Paulo. O caminho se inicia no ponto em azul e termina no ponto amarelo. O caminho mais rápido é o caminho em verde e o mais curto em vermelho. Perceba que em certos pontos ambos os caminhos compartilham os mesmo trechos, mas não acontecem em grande parte de nenhum dos dois percursos.	37
Figura 11 – Distribuições lognormal das cidades estudadas. Nela estão descritos o histograma da distribuição e sua curva de ajuste representada pelas linhas. Analisamos três distâncias entre os pares OD, $r = 1$ km, $r = 2$ km e $r = 3$ km. . .	40
Figura 12 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Boston. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	43
Figura 13 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Lisboa. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	44
Figura 14 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Porto. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	45
Figura 15 – Ambas as figuras descrevem a centralidade de intermediação na cidade do Rio de Janeiro. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	46
Figura 16 – Ambas as figuras descrevem a centralidade de intermediação na cidade de São Francisco. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	47
Figura 17 – Ambas as figuras descrevem a centralidade de intermediação na cidade de São Paulo. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.	48
Figura 18 – Gráfico em escala logarítmica para as distribuições de centralidade de intermediação com relação as cidade estudadas.. . . .	50
Figura 19 – Orientação das vias das cidades. Nas figuras do lado esquerdo temos representado o histograma da orientação de cada cidade, já no lado direito temos o diagrama de rosa para a mesma.	54

Figura 20 – Orientação das vias das cidades. Nas figuras do lado esquerdo temos representado o histograma da orientação de cada cidade, já no lado direito temos o diagrama de rosa para a mesma. 55

LISTA DE TABELAS

Tabela 1 – Tabela apresentando o comparativo entre as taxas de semelhança dos trechos percorridos no grafo das distâncias e no grafo dos tempos.	39
Tabela 2 – Tabela apresentando todos os parâmetros que ajustam os pontos que descrevem a centralidade de intermediação relacionada a distância.	49
Tabela 3 – Tabela apresentando as eficiências e as entropias das seis cidades estudadas e alguns parâmetros que contribuem direta e indiretamente, são eles o percurso médio, P_m , e o número de ruas, E	53

SUMÁRIO

1	INTRODUÇÃO	14
2	BIG DATA E MODELOS ANALÍTICOS.	18
2.1	<i>Big data.</i>	18
2.1.1	<i>Volume.</i>	19
2.1.2	<i>Variedade.</i>	20
2.1.3	<i>Velocidade.</i>	20
2.1.4	<i>Veracidade.</i>	20
2.1.5	<i>Valor.</i>	21
2.2	Teoria de grafos.	21
2.2.1	<i>Redes direcionadas e não direcionadas.</i>	22
2.2.2	<i>Grau de conectividade.</i>	24
2.2.3	<i>Centralidade de intermediação.</i>	24
2.2.4	<i>Eficiência</i>	25
2.3	Distribuição log-normal e sua aplicação na natureza.	26
2.4	Entropia e a organização de vias urbanas.	28
3	DADOS E SUA OBTENÇÃO.	30
3.1	Openstreetmaps - OSM©	30
3.2	Google Maps.	33
4	RESULTADOS E DISCUSSÃO	36
4.1	Distribuição de tempos e distâncias.	37
4.2	Centralidade: Padrões e distinções.	42
4.3	Entropia e Eficiência.	52
5	CONCLUSÕES E TRABALHOS FUTUROS	56
	REFERÊNCIAS	57

1 INTRODUÇÃO

Com o crescimento irrefreável das cidades ao redor do mundo, acabamos por nos deparar com desafios e dificuldades, tudo isso graças ao crescimento demográfico e de interesse em relação a cidade, necessitando de uma intervenção do governo em relação a fornecer uma melhoria na qualidade de vida, na eficiência dos meios de transportes e na locomoção desses meios. Por causa desse crescimento acelerado e de certa forma agressivo, as cidades espalhadas pelo mundo tem cada uma sua própria forma e características, sendo únicas em relação a topologia e dinâmica. Essas características implicam em vias com ou sem congestionamentos (TREIBER *et al.*, 2000).



Figura 1 – Exemplificação de um congestionamento em uma cidade, além do caos e dos visíveis problemas causados por ele (INTERNET, 2017a).

Boa parte dos congestionamentos é devido ao excesso de veículos que circulam nas cidades e também devido a forma de determinadas vias que não contribuem de forma positiva (BANDO *et al.*, 1995; WU *et al.*, 2004), ou seja, existem avenidas e ruas que recebem um fluxo de carros e estas não possuem capacidade para tal demanda ou levam uma certa quantidade de veículos a pontos na cidade que não estão preparados para receberem tal demanda, como por exemplo a figura 1. Em alguns casos, essa demanda além da capacidade ocorre devido a topologia da cidade que acaba por gerar gargalos (JIANG, 2007), elevando a densidade de veículos em alguns pontos da cidade, enquanto que em outras regiões esta permanece baixa.

Para definir qual o tipo de intervenção deverá ser realizada e em qual região da cidade, é preciso conhecer o padrão de deslocamento que ocorre na cidade (ÇOLAK *et al.*, 2016).

Desde o advento dos meios de transporte, principalmente os automotores, o conceito de mobilidade urbana vem se tornando cada vez mais importante de ser estudado (SIMINI *et al.*, 2012; YANG *et al.*, 2014). Com o gigantesco aumento de veículos nas ruas nas últimas décadas, o estudo da mobilidade em grandes cidades se tornou não só algo importante mas de fato necessário. Solucionar problemas de engarrafamentos (ÇOLAK *et al.*, 2016), construção de novas vias e ruas, modelos de descrição do tráfego (HELBING, 2001), além de muitos outros estudos que utilizando as complexas redes de ruas de uma cidade vem sendo realizados devido o entendimento sobre mobilidade urbana e a respectiva análise de dados.

Devido a tecnologia dos dias de hoje, muitos se utilizam de ferramentas para facilitar a escolha de qual caminho tomar (HANDY *et al.*, 2005), o mais conhecido e utilizado é o GPS (*Global Positioning System*) de aplicativos que mostra o caminho para determinado destino, seja ele casa, trabalho ou qualquer lugar a passeio. Ilustramos um exemplo de caminho traçado por um GPS na figura 2 em que temos uma rota do aplicativo *Waze*. O GPS mostra ao usuário o caminho mais rápido até o seu destino atualizando esse mesmo caminho a todo instante, ou seja, mesmo que você pegue uma via que não esteja traçada na rota dada pelo GPS ele se atualiza e logo lhe mostra o caminho mais rápido a ser tomado.

Entender a complexidade da organização das ruas não é uma tarefa fácil e lidar com o fluxo de carros e pedestres é ainda pior (MURAMATSU *et al.*, 1999; SEYFRIED *et al.*, 2006; TREIBER, 2013), esses sistemas são completamente complexos, diversificando-se rapidamente do seu estado inicial. Sistemas como esses são amplamente encontrados em física, biologia e em muitas outras áreas, onde seus constituintes evoluem e se adaptam, ocorrendo após várias interações o surgimento de padrões (ALBERT; BARABÁSI, 2002). As formas tomadas por esses sistemas que crescem e evoluem são comumente estudadas a partir da teoria de grafos (NEWMAN, 2003; BARABÁSI, 2016).

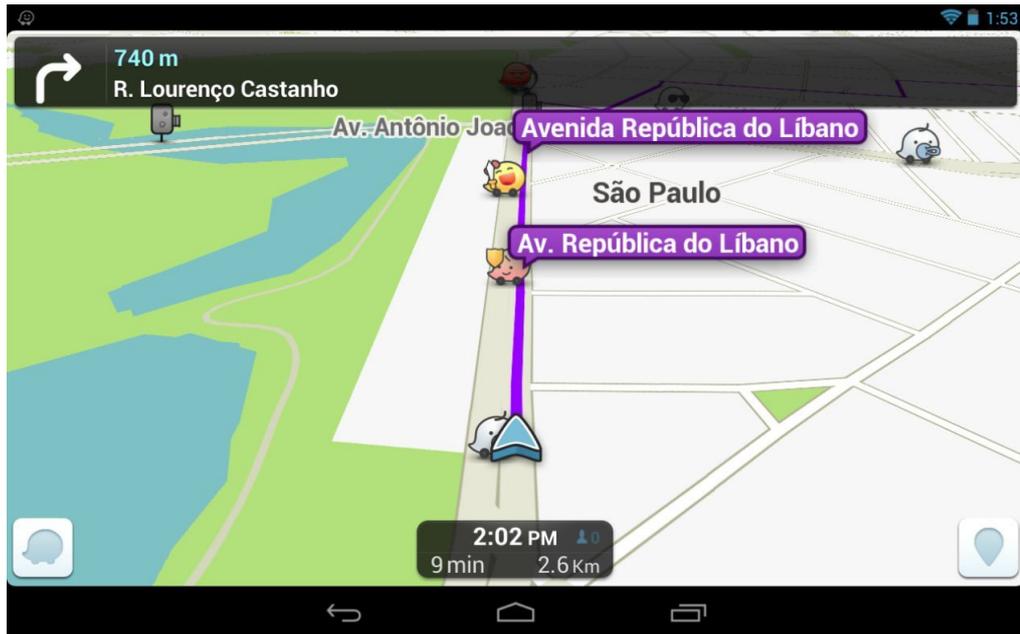


Figura 2 – Ilustração do aplicativo *Waze* da empresa Google que indica o melhor caminho baseado em notificações de usuários e dos tempos das vias (INTERNET, 2017b).

O estudo de redes é usado para representar diversos sistemas físicos, como empacotamentos, redes elétricas (MOREIRA *et al.*, 2000), percolação (NEWMAN; WATTS, 1999), além de diversas morfologias e, partindo para o nosso caso de interesse, descrição do formato e conexões das ruas de uma cidade (CARDILLO *et al.*, 2006). Cada cidade possui sua única e particular distribuição de ruas entre casas e prédios, e é nesse conjunto emaranhado de ruas que cada motorista procura a sua rota, onde, normalmente, este procura por uma rota que seja a mais rápida possível.

Neste trabalho realizaremos um estudo sistemático a cerca da forma como usuários utilizam a estrutura de ruas em cidades, durante os seus deslocamentos diários. Para isto, faremos uso da ferramenta do *Google Maps* (GM) para que possamos obter dados relativo ao deslocamentos realizados em diferentes cidades. Como base de dados para as redes de ruas das cidades investigadas, utilizamos a base de dados do *OpenStreetMaps* (OSM). A partir desses dados faremos uso de ferramentas usualmente aplicadas ao estudo de funções distribuições presentes nos diversos ramos da mecânica estatística.

Toda a abordagem sobre os conceitos de *Big Data* assim como teoria de grafos e análise dos dados extraídos são apresentados no capítulo 2. Nele também nos estendemos em distribuições, além de características relacionadas a organização das vias de uma cidade. No capítulo 3, é mostrado como é feita a retirada de dados do *openstreetmaps.com* e do *googlemaps.com* e a filtragem realizada para a montagem de grafos e do banco de dados dos tempos e das distâncias dos diversos segmentos de ruas e vias das cidades. Com as ferramentas adquiridas no capítulo 2, apresentamos no capítulo 4 os nossos resultados e discutimos os valores obtidos. Por fim, no capítulo 5, concluímos e encerramos nosso trabalho, além de apresentar as perspectivas para estudos futuros.

2 **BIG DATA E MODELOS ANALÍTICOS.**

Neste capítulo discutiremos todos os métodos e ferramentas matemáticas/computacionais utilizadas para a representação e análise dos dados obtidos neste trabalho. De forma geral, abordaremos neste capítulo o *Big Data* onde discutiremos os modelos que tratam com dados estatísticos disponíveis em aplicativos da internet e como estes podem ser utilizados na forma de distribuições.

2.1 **Big data.**

Com o avanço intenso da humanidade na direção de novas tecnologias surgiram novas maneiras das pessoas se agruparem e se interligarem. Estes agrupamentos podem ocorrer por motivos econômicos, sociais ou políticos, e devido a isto o crescimento e desenvolvimento da sociedade foi ocorrendo cada vez mais rápido. Demoramos milhares de anos para passarmos de nossas primeiras eras, era do fogo para a era do bronze, por exemplo, depois da era do bronze para uma era mais civilizada e agricultora. Posteriormente nos desenvolvemos para uma era industrializada e em menos de três séculos passamos para uma era totalmente informatizada e conectada.

Essa era extremamente conectada e informada, graças aos avanços da telecomunicação (internet, aplicativos), em poucos anos já não conseguia lidar com o volume, a rapidez e a variedade de informação criada e compartilhada em um curto período de tempo. Devido a esta demanda se tornou necessário o desenvolvimento de métodos eficazes para analisar, estudar, armazenar e organizar essa absurda quantidade de dados produzida pela humanidade. Assim, com o surgimento de ferramentas de *big data*, podemos manusear mais confortavelmente com uma quantidade de informação que antes não era nem possível imaginar se ter em mãos. Por exemplo, só em 2012 foram criados 2,8 zetabytes de dados e especialistas dizem que em 2020 teremos aproximadamente 40 zetabytes. A título de esclarecimento 1 zetabyte equivale à 10^9 kilobytes e, com base em estimativas, esse valor tende a dobrar a cada dois anos (MÜLLER MARCEL ROSENBACH, 2013).

Passando por um cuidadoso trabalho de armazenamento e higienização dos dados, pois devido a quantidade absurda de informação, muitas vezes esses dados apresentam equívocos que impossibilitam assim uma análise imediata. É possível encontrar com base nestes dados correlações e até realizar determinados tipos de predição do que pode ocorrer ou de como ocorre determinado comportamento para os mais variados sistemas, como por exemplo, bolsa de valores, prevenção de doenças, congestionamentos em cidades, relacionamentos e conexões em redes sociais, combate à criminalidade e etc.

Como exemplo de *big data* o sistema Google disponibiliza diversos de seus dados colhidos a partir de GPS ou dos telefones de usuários do sistema Android. Estes dados disponibilizados são totalmente anônimos e muitos já passaram por algum processamento e são tratados, organizados e oferecidos para serem utilizados seja em aplicativos ou realização de estudos científicos, como é o nosso caso. Os dados são retirados e organizados seguindo o que será explicado no capítulo 3. Porém mesmo com o pré-tratamento feito pelo Google e OSM os dados apresentam alguns conflitos e necessitam de um pós-processamento para termos certeza de quais dados podem ser úteis para o trabalho em questão, por isso a importância dos 5 V's (Volume, Variedade, Velocidade, Veracidade e Valor) no *big data*.

2.1.1 Volume.

Imagine todas as bilhões de mensagens, fotos e vídeos que circulam na rede por dia. Esse é um número gigantesco de informação e portanto precisamos lidar com esse grande volume de dados. Quando lidamos com *bigdata*, devido a esse volume de informação pertinente a esta técnica existe uma necessidade de desenvolvimento de ferramentas que possam lidar em tempo real com uma gama de dados. É preciso de ferramentas que sejam capazes de comprimir dados sem perda de informações e espaço.

2.1.2 Variedade.

Por volta de uma década, a maior parte dos dados eram estruturados, logo fáceis de lidar, além de poderem ser colocados em tabelas e relacionados por funções. Hoje, como mais de 80% dos dados do mundo não se comportam dessa forma, mensagens, fotos, vídeos e sons, que são dados não-estruturados, podem ser analisados e estudados por meio do uso de ferramentas específicas que permitem um tratamento juntamente com dados tradicionais.

2.1.3 Velocidade.

Pense na seguinte situação: Você está prestes a atravessar uma rua de olhos vendados e antes de atravessá-la você olhou uma foto da mesma rua a 5 minutos atrás, você acha que atravessará em segurança? Por que esse exemplo foi citado? Pois simplesmente a foto de 5 minutos atrás não diz nada de como estará a rua agora e precisamos de outra foto bem mais recente para poder dizermos algo. Isso serve para evidenciar a velocidade com que os dados são criados. Várias são as mensagens de redes sociais se viralizando em segundos, transações de cartão de crédito sendo verificadas a cada instante ou os milissegundos necessários para calcular o valor de compra e venda de ações. Tudo ocorre em uma tremenda velocidade que podemos caracterizar como em tempo real, sendo assim, precisamos lidar com um enorme número de informações instantaneamente e este tipo de tratamento exige ferramentas bastante eficientes.

2.1.4 Veracidade.

Para uma maior eficiência nos processos do *big data* é necessário obter dados verídicos, de acordo com uma dada realidade. O conceito de velocidade é bem alinhado ao conceito de veracidade uma vez que necessitamos de uma análise cada vez mais rápida, chegando a uma análise em tempo real. Logo, dados passados não podem ser considerados dados verídicos para o momento em que é analisado, igualmente com dados que passam por uma comparação e são considerados erros para determinadas situações, são exemplos de dados não verídicos.

2.1.5 Valor.

O último V é o que torna o *big data* tão importante e necessário. Não adianta ter uma grande quantidade de informação adquiridas e analisadas em tempo real que condizem com nossa realidade e não geram nenhum valor. Não seria interessante no mundo de hoje se não pudessemos adquirir algo com essa tecnologia que é o *big data*. Valor entende-se como informações que podem ajudar na previsão e no entendimento dos mais variados sistemas.

2.2 Teoria de grafos.

O problema das sete pontes de Königsberg é considerado como o primeiro problema relacionado com a teoria dos grafos. Este trabalho foi publicado em um artigo escrito pelo famoso matemático Leonhard Euler em 1736 (EULER, 1741; SHIELDS, 2012). Dado uma certa configuração espacial onde pontes estão disposta segundo a representação da figura 3, o problema discute a possibilidade de que um determinado caminhante consiga estabelecer um caminho que não passe pela mesma ponte duas vezes.

Para a resolução do problema Euler seguiu o seguintes passos: (i) transformou cada ilha em um nó e (ii) pontes em conexões, criando assim o primeiro grafo da história, conforme ilustrado na figura 4. Estudando esse desenho feito por ele, percebeu que só seria possível a existência de um caminho que atravessasse todas pontes uma única vez se, nesse caminho existissem dois ou nenhum ponto com conexões ímpares com os demais nós. Essa regra se deve ao fato de que os nós devem ter pelo menos duas conexões, devido ao fato da existência de uma saída e de uma chegada naquele mesmo ponto. Já os nós que possuem conexões ímpares terão uma ligação a mais que simboliza ele como ponto de partida e o de chegada.

A partir desse estudo a teoria de grafos foi difundida em diversos ramos, pois praticamente tudo pode ser representado usando a teoria dos grafos como por exemplo parentesco familiares ou similaridade genética (biologia), ligações em moléculas e redes cristalinas (química), tubulações hidráulicas e sistemas elétricos (engenharia), além de cidades e redes sociais (NEWMAN, 2003; BARABÁSI, 2016). A teoria de grafos nos permite retirar diversas informações do sistema/rede em estudo. Especificaremos algumas como o grau de conectividade e a centralidade de intermediação, além de diferenciar redes direcionadas e não direcionadas.

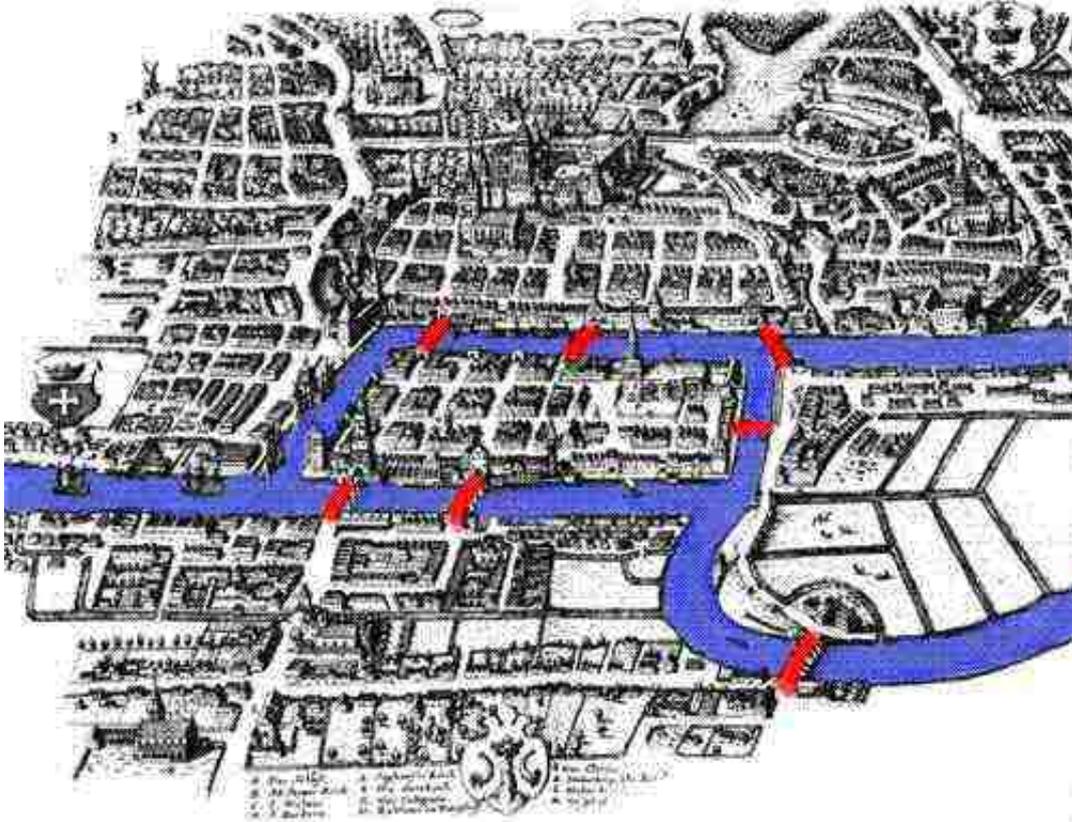


Figura 3 – Imagem descrevendo a configuração das sete pontes de Königsberg, o rio Prególia está na cor azul enquanto as pontes estão na cor vermelha, o mapa descreve a cidade de Königsberg no ano de 1736 (INTERNET, 2017c).

2.2.1 Redes direcionadas e não direcionadas.

Ao falarmos de redes uma informação importante sobre as redes diz respeito a sua direcionalidade, ou seja, se essas redes são direcionadas ou não-direcionadas. Dizemos que determinado grafo não é direcionado quando a ligação entre dois vértices os conectam nos dois sentidos, por exemplo, olhe a figura 4, perceba que todas as ligações não possuem nenhuma direção privilegiada, logo este grafo é dito não-direcionado (NEWMAN, 2003). Podemos representar o grafo da figura 4 da seguinte maneira:

$A - B, B - A, A - C, C - A, A - D, D - A, B - D, D - B, C - D, D - C.$

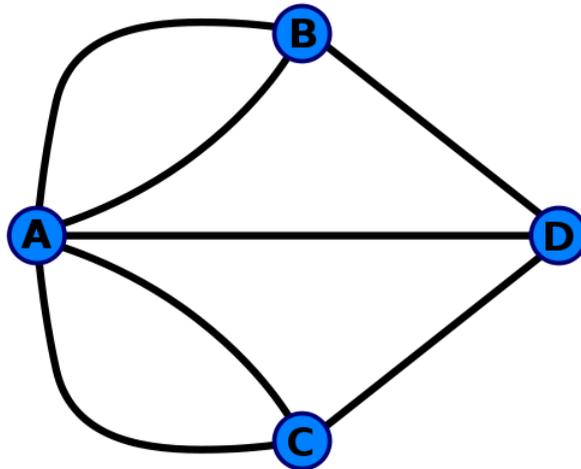


Figura 4 – Imagem descrevendo a configuração dos pontos e ligações da cidade de Königsberg, sendo assim provavelmente o primeiro grafo da história (INTERNET..., 2017a).

Se tivéssemos um grafo como o mostrado na figura 5, teríamos um grafo direcionado e poderíamos representá-lo da seguinte maneira:

$A - B, A - C, A - D, B - A, C - A, D - B, D - C.$

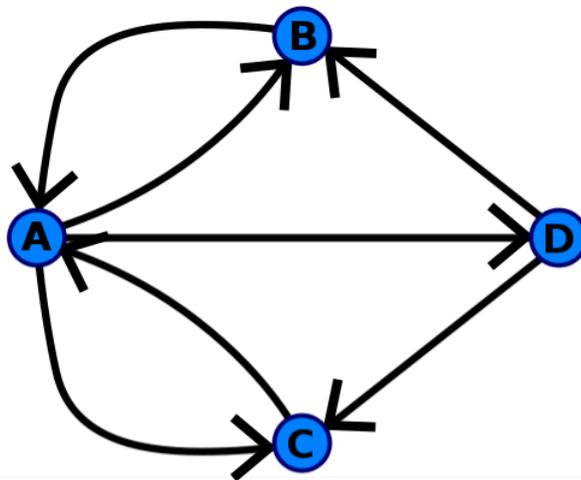


Figura 5 – Imagem ilustrando a cidade de Königsberg de uma forma direcionada, as direções foram distribuídas aleatoriamente apenas para ilustração (INTERNET..., 2017a).

Com essas diferenças entre os grafos, várias outras características também são modificadas, onde podemos citar por exemplo, o grau de conectividade de cada nó, a centralidade de intermediação dos nós e ligações de rede, dentre muitas outras. Estas características atribuídas aos grafos (redes) serão descritos em detalhes nas próximas seções.

2.2.2 *Grau de conectividade.*

Comumente conhecida na língua inglesa como *degree* de um grafo, o grau de conectividade de um nó mede o quão conectado este nó é no grafo, ou seja, o número de ligações conectados a ele (NEWMAN, 2003). Essa característica em um grafo é muito importante uma vez que a conectividade é basicamente uma de suas medidas mais importantes, podendo identificar nós como o mais importante em relação a sua conexão com o restante da rede. Quando lidamos com redes orientadas, ou seja, quando os nós não são apenas ligados entre si mas possuem uma determinada orientação, definimos duas maneiras distintas de centralidade de grau, são elas: *indegree* e *outdegree*. *Indegree* é a quantidade de links que se direcionam para o nó e *outdegree* é a quantidade de ligações que partem do nó em questão para os demais. Matematicamente a centralidade de grau de um grafo $G(V, E)$, onde o grafo G possui $|V|$ vértices e $|E|$ conexões é definida seguindo a equação (2.1).

$$C_D(G) = deg(G). \quad (2.1)$$

2.2.3 *Centralidade de intermediação.*

Do inglês *betweenness centrality*, a centralidade de intermediação é um conceito diferenciado de centralidade, ela relaciona a característica de que os nós pertençam a algum caminho entre outros nós, ou seja, define o número de vezes que determinado nó serve como ponto de acesso para todos os menores caminhos possíveis definidos entre um determinado par de origem e destino, ou par OD (NEWMAN, 2003). Nós com grande centralidade de intermediação podem vir a ter grande importância dentro de uma rede, visto que ele é responsável por uma grande passagem de informação por ele. Logo, as retiradas desses nós de uma rede implicam em uma quebra na comunicação entre os outros nós da rede. Essa intermediação pode ser descrita pela equação (2.2) que se estabelece por meio dos seguintes passos:

- Calcular todos os caminhos mais curtos entre determinados vértices (o, d) .
- Quantificar o número de caminhos entre os vértices (o, d) que utilizam um determinado nó v , $\sigma_{od}(v)$.
- Repetir as duas primeiras etapas para todos os pares de vértices (o, d) possíveis mantendo o mesmo vértice v como intermediador, σ_{od} .
- Realizar a soma sobre a razão entre todos os caminhos que contém o vértice v e todos os caminhos possíveis conectando os vértices (o, d) .

$$C_B(G) = \sum \frac{\sigma_{od}(v)}{\sigma_{od}}. \quad (2.2)$$

Tal medida de centralidade não é uma característica específica dos nós de um grafo, ela pode também ser estendida para as suas ligações, onde agora, medimos o quão importante tal ligação é para o grafo assim como a utilização de determinado *link* para os menores caminhos que conectam os pontos da rede. Podemos também ver a importância da ligação quando esta é removida da rede e observando se a falta desta ligação no grafo implica em alguma mudança significativa para o grafo. Como por exemplo, ocorrer uma grande mudança na conectividade entre todos os pontos assim como uma mudança efetiva na quantidade e/ou qualidade dos caminhos existentes na rede (grafo).

2.2.4 Eficiência

A eficiência de um grafo é dita como uma medida do quão eficiente determinado grafo troca informação entre seus pontos. O conceito de eficiência é aplicado em duas escalas em uma rede, são elas: locais e globais. Na escala global, a eficiência quantifica a troca de informações em toda a rede, onde a "passagem" de informações ocorre ao mesmo tempo. Na eficiência local temos quantificação da resistência de uma rede a falhas em pequena escala.

Por definição a eficiência média comumente utilizada é dada pela seguinte expressão:

$$E(G) = \frac{1}{n(n-1)} \sum_{i < j} \frac{1}{d(i,j)}, \quad (2.3)$$

onde n é o número de nós em uma rede e $d(i, j)$ representa o menor caminho entre os nós i e j .

Além disso, outra definição de eficiência é bastante utilizada, ela envolve distâncias reais percorridas em uma rede e a distância entre os dois nós em uma linha reta . Essa segunda definição é muito usada para redes planares e que assim podem representar alguma rede real segundo a expressão:

$$E(G) = \frac{1}{n(n-1)} \sum_{i < j} \frac{d^{Eucl}(ij)}{d(i,j)}, \quad (2.4)$$

2.3 Distribuição log-normal e sua aplicação na natureza.

Uma distribuição log-normal nada mais é do que uma distribuição contínua de uma variável aleatória cujo logaritmo é normalmente distribuída. A importância dessa distribuição ocorre devido a um resultado análogo ao teorema do limite central: assim como uma distribuição normal aparece quando são somadas várias variáveis independentes, a distribuição log-normal aparece naturalmente como o produto de diversas variáveis independentes (sempre positivas). Um simples exemplo pode ser visto se analisarmos o preço de uma ação no futuro, esse valor pode ser modelado como diversos ajustes independentes um dos outros de tal forma a gerar o preço final P_f ,

$$P_f = P_0 \times (1 \pm i_1) \times (1 \pm i_2) \times \dots \times (1 \pm i_n),$$

onde P_0 é o preço inicial da ação e i_n é a taxa. Portanto, aplicando o logaritmo é fácil ver que teremos a soma desses vários ajustes independentes, assim P_f pode ser aproximado por uma log-normal. A log-normal pode ser escrita pela equação (2.5) e ela aparece em diversos estudos, sejam eles na medicina, biologia e ciências sociais (MIZUSEKI; BUZSÁKI, 2013; MAKUCH *et*

al., 1979; BLACK; SCHOLLES, 1973), podemos ter uma noção da sua distribuição com a figura 6 que apresenta 5 distribuições log-normais variando seu parâmetro σ enquanto seu segundo parâmetro μ se mantém fixo em 0.

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left[-\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right]. \quad (2.5)$$

Para uma noção clara da forma de uma distribuição log-normal mostramos na figura 6 cinco exemplos de distribuições log-normal variando a sua variância σ mas mantendo a sua média μ fixa no ponto zero.

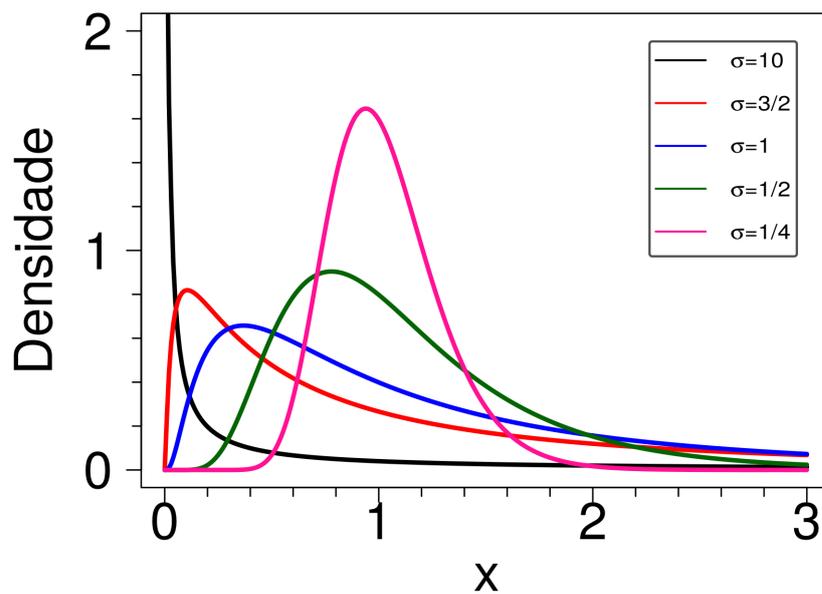


Figura 6 – Resultado mostrando a variação da distribuição log-normal mantendo o parâmetro $\mu = 0$ fixo e alterando apenas o parâmetro σ .

2.4 Entropia e a organização de vias urbanas.

Como já é conhecido da experiência diária de cada indivíduo em uma cidade, quando trafegamos em direção a destinos variáveis cada rua e avenida possui características próprias, sejam elas físicas ou não. Como exemplo de características das vias que abordaremos neste trabalho, podemos citar o seu comprimento e a sua orientação angular em relação a direção norte. Estas serão abordadas na tentativa de classificar a sua influência sobre os caminhos mais curtos e mais rápidos durante o deslocamento entre pontos distintos em uma determinada região. Para melhor quantificar essas diferenças quanto a orientação usaremos a equação de Gibbs/Shannon para a entropia (SHANNON, 2001; GUDMUNSON; MOHAJERI, 2013), definida da seguinte forma:

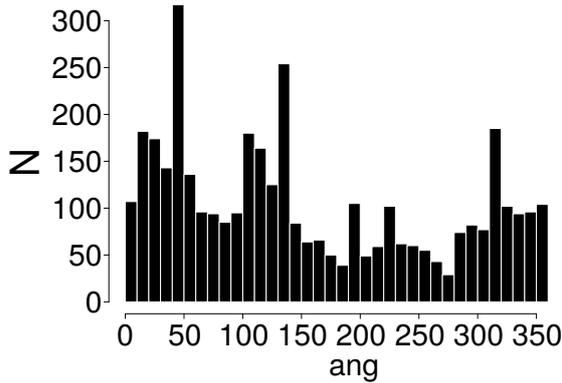
$$S = -k \sum_{i=1}^t p_i \ln(p_i), \quad (2.6)$$

onde S é a entropia, k é uma constante positiva, t é o número de bins com probabilidade diferente de zero e p_i é a probabilidade de que determinada rua com uma certa angulação pertença ao i -ésimo bin. Para a entropia de Gibbs, k é conhecida por constante de Boltzmann, $k_B = 1,38 \times 10^{-23}$ J/K, entretanto, para nossos cálculos, a entropia não possui dimensão, logo essa mesma constante possui valor unitário e sem unidades.

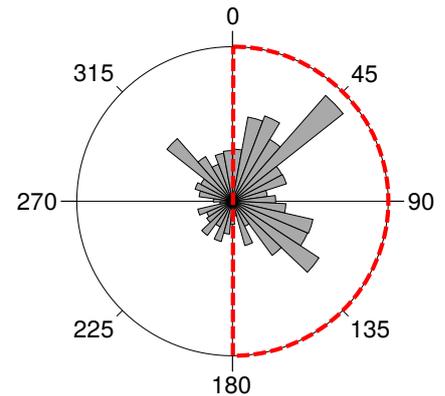
Ao tratarmos da orientação de uma rua, não citaremos em primeiro momento a direção da mesma. Portanto, ao invés de termos θ variando entre 0° e 360° ficaremos só com a metade desse intervalo não fazendo distinção entre Norte e Sul nem Leste e Oeste. Para caso de padronização na distribuição dos ângulos tomaremos a largura de todos os bins iguais a 5° .

A aparência da distribuição da probabilidade afeta totalmente a entropia, por exemplo, se todas as ruas ocuparem um único bin a entropia deve ser zero, caso contrário, se a distribuição for totalmente uniforme todos os bins terão o mesmo peso e assim a entropia atingirá um valor máximo que será:

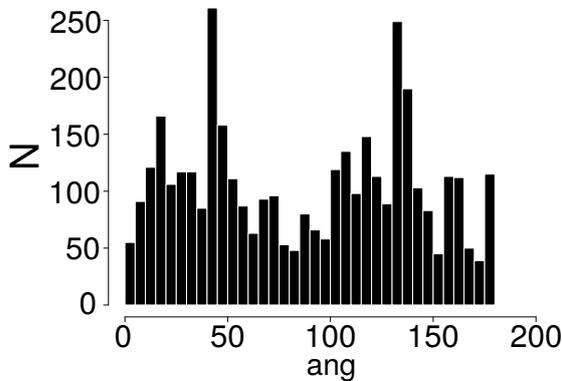
$$S = k \log(t). \quad (2.7)$$



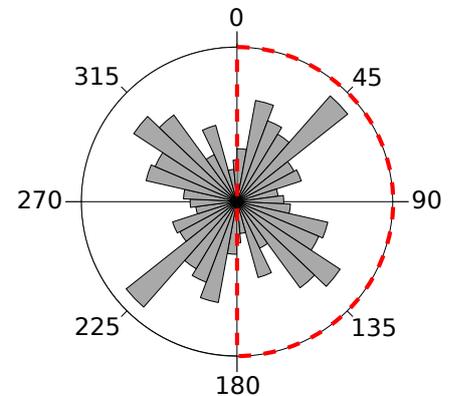
(a) Distribuição da direção das ruas de Boston.



(b) Distribuição em forma de diagrama de rosa das direções das ruas de Boston.



(c) Distribuição da orientação das ruas de Boston.



(d) Distribuição em forma de diagrama de rosa das orientações das ruas de Boston.

Figura 7 – Direção e orientação das vias das cidades. Nas figuras do lado esquerdo temos representado a distribuição para as direções e orientações da cidade de Boston, já no lado direito temos o diagrama de rosa para o mesmo caso.

Neste limite teremos a presença de ruas em todas as direções sendo portanto a configuração na orientação das ruas bem homogênea. No outro limite onde a entropia apresenta valor nulo, teremos todas as ruas apresentando a mesma orientação.

Na figura 7, vemos a distinção da distribuição das ruas para uma das cidades que estudaremos (Boston) quando consideramos a sua orientação e a sua direção. Perceba que há uma mudança clara na aparência dos diagramas de rosa para cada situação, no caso das direções temos preferencialmente um intervalo (o primeiro quadrante), enquanto que no caso das orientações temos uma igualdade entre dois intervalos (o primeiro e segundo quadrante).

3 DADOS E SUA OBTENÇÃO.

Neste capítulo abordaremos a forma e as ferramentas utilizadas para a obtenção dos dados usados em nosso trabalho. Explicaremos as ferramentas disponíveis na internet e como essas operam para a construção do nosso banco de dados como do grafo. Toda a retirada de dados foi realizada a partir de dois sites bastante difundidos na web nos dias de hoje, são eles: *openstreetmaps.com* e *googlemaps.com*.

3.1 Openstreetmaps - OSM©

O OSM© é um site colaborativo onde diversas pessoas podem editar o mapa do mundo. Essas edições podem ser as mais variadas possíveis, partindo de onde se localiza a sua casa até delimitações de lojas, cidades e países. No OSM© tem a descrição de construções, especificando se elas são residenciais ou lojas e empresas; assim como as construções, também há a descrição de ruas e avenidas, especificando se elas são ciclofaixas, vias primárias (possuindo de três ou mais faixas), vias residenciais (que possuem um fluxo menor de veículos) ou até vias de pedestres.

Todas essas informações podem e são facilmente retiradas do próprio site ou de diversos extratos em outros sites que são regularmente atualizados. Os arquivos extraídos possuem formato *.osm* que nada mais é do que um tipo de arquivo análogo ao XML, portanto possuem uma estrutura simples e fácil de ser compreendida tanto por seres humanos quanto por máquinas. A estrutura desses arquivos é "encabeçada" por tags que ajudam muito a encontrar qualquer informação no arquivo. As divisões de um arquivo *.osm* são basicamente três: *node*, *way* e *relation*.

Em *node* temos todos os nós do mapa sejam eles conexões entre ruas, pontos marcando o local de uma loja ou até a localização da sua casa. Cada nó possui atributos que são sua latitude, longitude, a pessoa que modificou por último esse nó e, se for o caso, descreve o ponto representado seja ele uma loja ou um semáforo.

Em *way* temos a relação de todas as combinações dos nós que representam alguma linha. No mapa, essas linhas podem ser uma avenida, uma parede delimitando um condomínio ou até uma linha delimitando a entrada de um parque.

O agrupamento dessas linhas ocorre na seção *relation* analogamente ao agrupamento dos pontos na seção *way*, esse agrupamento é o responsável pelas delimitações e seleções mais abrangentes. Por exemplo, em *way* temos apenas uma avenida enquanto em *relation* temos toda uma BR, pois agrupamos diversas avenidas e vias que constroem e fazem parte dela. Outro exemplo é a delimitação de uma área, seja ela um parque, um conjunto residencial ou o país inteiro.

No arquivo exportado extraímos dele tudo relacionado a tag *highway* que aborda todas as vias e faixas de tráfego existentes em uma cidade. Para limitarmos essas vias exclusivamente ao tráfego de veículos escolhemos subtags específicas, são elas:

- *motorway* - Grandes vias com um grande fluxo de veículos, possuem divisão entre os dois sentidos, equivalente as *Autobahn*.
- *motorway_link* - Conexões entre as *highways motorways* e outro tipo de *highway*.
- *trunk* - As vias mais importantes de uma região que não são *motorways*, elas não são necessariamente divididas em duas vias.
- *trunk_link* - Conexões entre as *highways trunk* e outro tipo de *highway*.
- *primary* - Vias de grande importância em grandes cidades.
- *primary_link* - Conexões entre as *highways primary* e outro tipo de *highway*.
- *secondary* - Vias de grande importância em cidades.
- *secondary_link* - Conexões entre as *highways secondary* e outro tipo de *highway*.
- *tertiary* - Vias de grande importância em pequenas cidades ou vilarejos.
- *tertiary_link* - Conexões entre as *highways tertiary* e outro tipo de *highway*.
- *residential* - Ruas que servem como acesso para habitações/residências.
- *unclassified* - Vias que possuem a função de conectar e permitir acesso a propriedades.

Após essa restrição das vias existentes no arquivo *.osm*, selecionamos todos os nós onde essas vias se originam e terminam. Nessa seleção arquivamos em um banco de dados todas as latitudes e longitudes contidas no atributo da categoria *node*. Nas referências da categoria *way* são extraídos quais nós se conectam uns com os outros e juntamente com as latitudes e longitudes já arquivadas criamos o grafo da área a ser estudada.



Figura 8 – Grafo da cidade de Lisboa. Grafo representando toda a região de estudo da cidade de Lisboa obtida utilizando o método de extração do OSM© descrito. Neste grafo estão sendo descritos 11890 nós e 13465 links.

Com todos os pontos de latitude e longitude armazenados calculamos a distância entre esses pares para servir como distância percorrida por um veículo que se desloca de um nó i para um nó j que sejam diretamente ligados no mapa por meio de vias. Todas essas distâncias são calculadas utilizando-se da fórmula de Haversine (RÍOS, 1795)

$$\text{hav}\left(\frac{d}{r}\right) = \text{hav}(\varphi_2 - \varphi_1) + \cos(\varphi_1) \cos(\varphi_2) \text{hav}(\lambda_2 - \lambda_1) \quad (3.1)$$

onde,

- $\text{hav}(\theta) = \sin^2\left(\frac{\theta}{2}\right) = \frac{1 - \cos(\theta)}{2}$.
- d é a distância entre os dois pontos.
- r é o raio da Terra.
- ϕ_1, ϕ_2 são as latitudes do ponto 1 e 2, respectivamente.
- λ_1, λ_2 são as longitudes do ponto 1 e 2, respectivamente.

Para descobrirmos a distância d entre dois pontos podemos fazer a inversa e teremos:

$$d = r \text{hav}^{-1}(h) = 2r \arcsin\left(\sqrt{h}\right)$$

$$d = 2r \arcsin\left[\sqrt{\text{hav}(\phi_2 - \phi_1) + \cos(\phi_1) \cos(\phi_2) \text{hav}(\lambda_2 - \lambda_1)}\right]$$

$$d = 2r \arcsin\left[\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1) \cos(\phi_2) \sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right]. \quad (3.2)$$

É importante ressaltar que esta distância é uma distância euclidiana entre dois pontos, que na maioria dos casos difere da distância real percorrida quando se trafega em uma cidade durante o deslocamento entre dois pontos na rede.

3.2 Google Maps.

Com os conhecimentos e técnicas da seção anterior já obtivemos o grafo da cidade e o comprimento de todas as vias que a compõem. Mas que outro fator é tão importante em tráfego ou até mais que as distâncias envolvidas entre origem e destino? O tempo gasto para percorrer determinada distância. Existem diversas maneiras de descobrir o tempo gasto em uma viagem a partir de fontes como ligações de celulares ou através de dados de rastreamento de veículos. Existem também outras maneiras indiretas de encontrar esse tempo a partir de simulações e cálculos, como por exemplo o *traffic assignment*.

Em nosso trabalho utilizamos o tempo das viagens retornado pelo Google Maps. Neste contexto temos como princípio básico o fato de que motoristas possuem o intuito de chegar ao seu destino o mais rápido possível e, por não possuir conhecimento global sobre o tráfego na cidade, usufrui dos recursos oferecidos pelo GPS onde o mesmo retorna o caminho que leva menos tempo. Na maioria dos GPS esta é uma opção que pode ser facilmente alterado para o caminho mínimo. como o tempo de viagem é o mais importante estes surgem na maioria dos casos como o mais utilizado. A retirada do tempo das viagens em uma cidade é feita com uma API (*Application Programming Interface*), em português, Interface de Programação de Aplicativos) do GM chamada: *Distance Matrix*.

A *Distance Matrix* API é um serviço de web oferecido pelo Google Maps que retorna a distância e o tempo de um percurso dados origem e destino. A informação é passada por meio de uma URL (analogamente a parte do OSM© da seção 3.1) e os dados são retornados em um formato XML. Na URL enviada alguns parâmetros são obrigatórios, são eles: origem, destino e a chave. A origem e o destino podem ser dados em latitude e longitude assim como o nome do endereço. A chave é um código fundamental dado pelo Google Maps para poder realizar essas solicitações, ele é obtido com uma conta como desenvolvedor do Google e a cada projeto criado uma chave diferente é originada. Outros parâmetros opcionais são também selecionados para uma melhor extração de informação, são eles: o modo de viagem, o horário de saída do veículo da origem e o modelo de tráfego usado, ou seja, a estimativa do tráfego, nesse caso é utilizado o *best_guess* que é a melhor estimativa de trânsito possível considerando médias históricas e informações em tempo real. Um exemplo de solicitação pode ser visto na expressão a seguir:

[https://maps.googleapis.com/maps/api/distancematrix/xml?origins=37.79,-122.41&destinations=37.79,-122.41&mode=driving&units=metric&key=...\(GOOGLE, 2017\)](https://maps.googleapis.com/maps/api/distancematrix/xml?origins=37.79,-122.41&destinations=37.79,-122.41&mode=driving&units=metric&key=...(GOOGLE, 2017))

Com isso, conseguimos o tempo e a distância de todos os trechos guardados no banco de dados citado anteriormente. Infelizmente, como o OSM© é um mapa feito de forma colaborativa não é sempre que os pontos dele estão em perfeita sincronia com os pontos do mapa do Google Maps, portanto existe a necessidade de ajustes a serem feitos para o cálculo dos tempos. Inicialmente ignoramos essa diferença e extraímos do Google todos os tempos e as distâncias relacionadas as vias previamente armazenadas. Após essa retirada todos os dados

são comparados e nesse ponto levamos a diferença nas localizações dos pontos. Para pontos que possuem a diferença entre a distância obtida do OSM© e a distância obtida do Google Maps superior ou igual a 40 m o tempo de viagem é calculado a partir da média dos tempos das outras viagens que possuem uma distância entre origem e destino semelhante a da viagem em questão. Mais especificadamente é colocado uma margem de 5 m para mais e para menos, portanto, caso a distância entre a origem e destino seja de 20 m consideramos todas as viagens de 15 m até 25 m. Vale destacar aqui que a distância utilizada para calcular o tempo médio é a distância calculada com a fórmula de Haversine, equação (3.2), pois é ela que está presente como distância entre os pontos no grafo e não a distância retornada pelo Google Maps.

Um aspecto importante relacionado a obtenção dos dados diz respeito ao horário. Este tem uma forte influência nos dados obtidos pois dependendo do horário escolhido teremos uma taxa de ocupação das ruas bastante diferente. É de conhecimento geral das pessoas que existem horários de pico no movimento de carros nas cidades. Este aspecto é bastante relevante quanto ao tempo despendido durante um determinado deslocamento. Tendo em vista este aspecto inerente ao deslocamento realizado pelos usuários, optamos por fazer em um horário bem definido. O horário para a retirada dos tempos das viagens de todas as cidades estudadas ocorre às 18h, este horário foi escolhido pelo fato de nesse instante o tráfego na cidade está em seu "auge", ou seja, há nas ruas um grande número de veículos podendo implicar em um aumento de *delays* nas vias e assim intensificando o papel de vias e avenidas principais e também secundárias. O nosso estudo foi concentrado nesse horário, pois analisamos quase todas as vias existentes na região de estudo. Para uma análise da variação desse horário seria mais interessante considerar as vias principais uma vez que a grande quantidade de vias residenciais possui pouca influência nos resultados, quando os dados são obtidos fora do horário de "*rush*". Essas ruas são pouco utilizadas nos demais horários do dia. Claro que existem outros horários com grande intensidade de movimento de veículos. Porém, a nossa escolha contempla a maior concentração de usuários, haja visto que os horários de início das atividades normalmente exercidas pelas pessoas, ocorrem com uma certa distribuição de horários pela manhã, o que usualmente não ocorre no final do expediente que na grande maioria dos casos o expediente encerra às 18h e um volume grande de usuários inicia sua jornada de volta para casa.

4 RESULTADOS E DISCUSSÃO

Neste capítulo mostraremos nossos resultados obtidos a partir dos dados extraídos de acordo com que foi discutido no capítulo anterior. Nosso estudo é realizado em 6 cidades diferentes, são elas: Boston, Lisboa, Porto, Rio de Janeiro, São Francisco e São Paulo. Inicialmente descreveremos e analisaremos a distribuição de tempo e das distâncias em zonas quadradas com 5 km de lado nas cidades mencionadas. Como exemplo de uma região selecionada mostramos na figura 9 um mapa da região de São Paulo que foi estudada. Destacamos por meios de linhas tracejadas o quadrado de 5 km de lado que foi considerado em nosso estudo. Depois de definida as regiões a serem investigadas, analisaremos padrões e distinções entre as centralidades de cada região destacada e por fim, apresentaremos uma breve discussão sobre os resultados e sobre futuros aspectos a serem investigados. Devido ao fato da região selecionada possuir uma escala menor que uma cidade, procuramos escolher um local que compreendesse um tráfego mais intenso, como um centro ou algo do gênero.

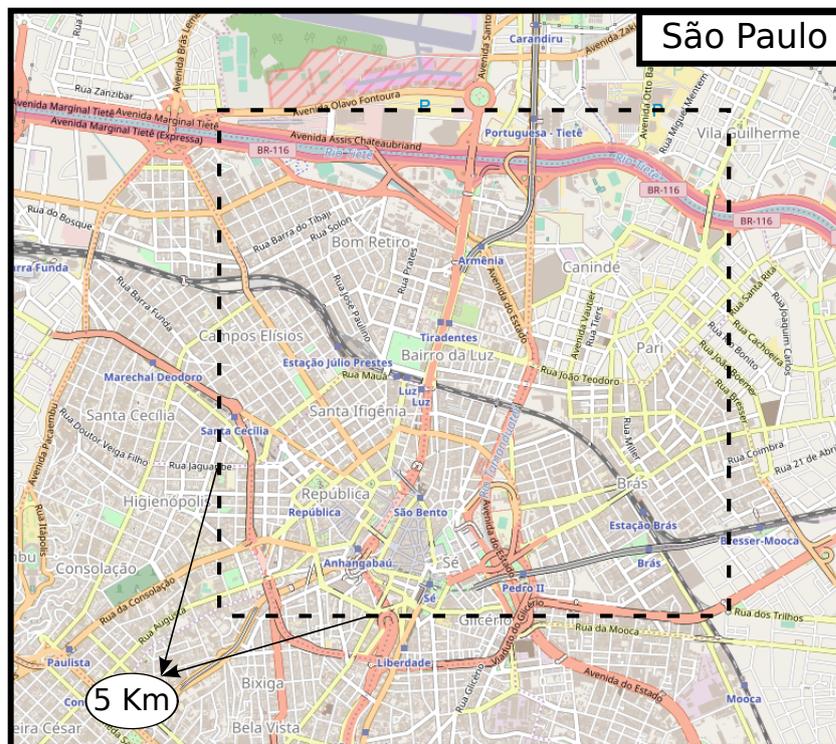


Figura 9 – Demarcação da área de estudo. Amostra da região selecionada em que se é realizado o estudo. Conforme pode ser visto, é selecionado um quadrado com 5 km de lado e a partir dele são consideradas todas as ruas que possuem a *tag = highway*. A escala está demonstrada de modo qualitativo, apenas para melhor compreensão (INTERNET. . . , 2017b).

4.1 Distribuição de tempos e distâncias.

Uma vez retirados o tempo e a distância de cada via do *googlemaps.com*, assim como as ruas e avenidas são extraídas do *openstreetmaps.com*, criamos um grafo com as latitudes e longitudes dos pontos que formam as ruas e avenidas. Após isso atribuímos a cada um desses links do grafo recém-formado seu comprimento e seu tempo de percurso. Para o caso do nosso estudo lidaremos esses dois grafos como distintos, denominando-os como grafo das distâncias e grafo dos tempos.

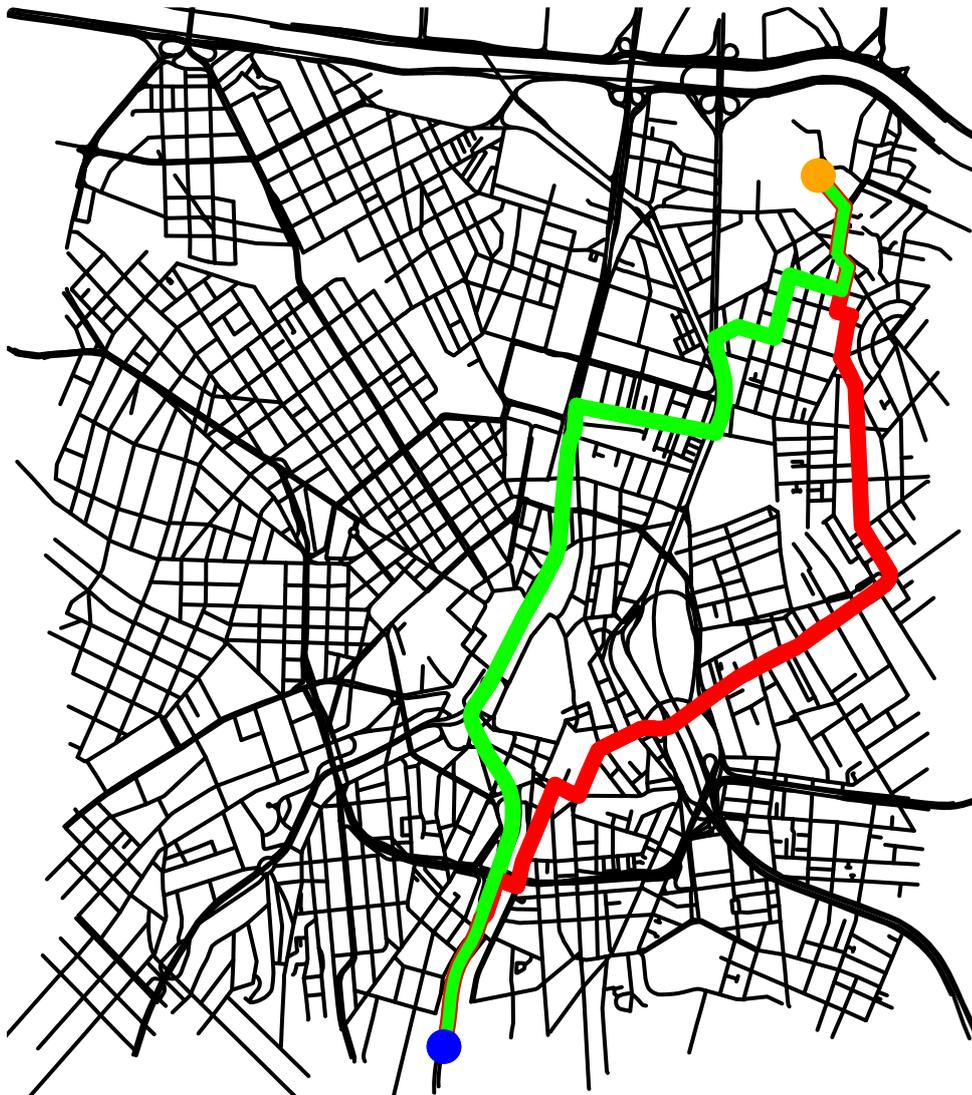


Figura 10 – Imagem demonstrando a diferença entre o caminho mais curto geometricamente e o caminho mais rápido na cidade de São Paulo. O caminho se inicia no ponto em azul e termina no ponto amarelo. O caminho mais rápido é o caminho em verde e o mais curto em vermelho. Perceba que em certos pontos ambos os caminhos compartilham os mesmos trechos, mas não acontecem em grande parte de nenhum dos dois percursos.

Após termos definido as regiões a serem estudadas em cada cidade, iniciamos a fase de obtenção dos dados sobre os tempos e distâncias entre diferentes pontos no interior da região de interesse. Para essa etapa optamos por gerar uma diversidade de pares OD com uma distância mínima de 2 km em linha reta entre os dois. Todos estes 2000 pares foram escolhidos aleatoriamente sempre dentro da zona de interesse, definida previamente.

A partir desta coleção de pares OD definido para cada cidade, iniciamos a coleta dos dados relacionados ao tempo de viagem e ao deslocamento mínimo necessário para se deslocar entre os dois pares OD. Ressaltamos que, em algumas ocasiões estas duas informações podem ser muito semelhantes, ou seja, o deslocamento que contempla o tempo mínimo pode ser coincidente com aquele que define a distância mínima. Fato que depende muito da quantidade de usuários utilizando as vias da cidade. Para podermos ter uma ideia desta coincidência ou não, mostramos na figura 10 duas trajetórias entre o mesmo par OD, com base no tempo e na distância mínima. É possível perceber que de fato existem trechos do caminho entre este par OD que são coincidentes e outros não. Este típico exemplo demonstra que a escolha de um tempo mínimo não necessariamente acarreta em uma distância mínima ou vice-versa.

Com base nesse procedimento foi possível se estabelecer parâmetros quantitativos, p_{dist} e p_{time} que definem quanto de tempo gasto existe em comum entre os dois tipos de deslocamentos entre dois pontos da rede e o quanto ele representa do tempo total gasto, ou seja, a razão entre o tempo em comum e o tempo gasto para percorrer o caminho mais curto define o p_{dist} , analogamente para o caminho mais rápido define p_{time} . Apresentamos na tabela 1 todos os percentuais médios considerados sobre 2000 viagens para as diferentes cidades e distâncias investigadas.

Percebemos que, por mais distintos que sejam as zonas de cada cidade, em média a semelhança entre os menores caminhos percorridos sejam eles em relação a distância ou ao tempo estão sempre na mesma faixa de porcentagem. Com base nesse estudo preliminar podemos verificar que a taxa de semelhança se reduz com o aumento da distância considerada, quase que em uma mesma proporção independente da cidade avaliada. Para distâncias maiores as opções de caminhos crescem rapidamente fazendo com que o fator trânsito possa ser contornado em muitas ocasiões. Sendo assim, a coincidência entre os caminhos de tempo mínimo e distância

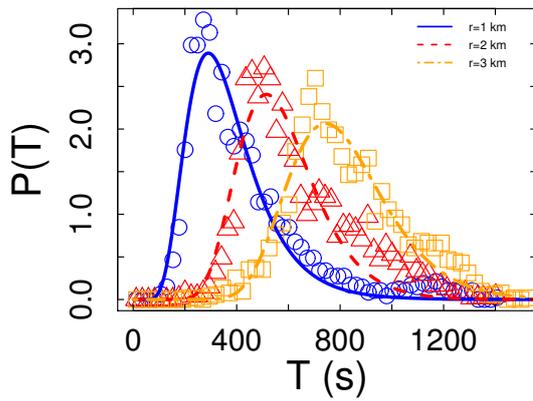
Cidade	1 km		2 km		3 km	
	P_{dist}	P_{time}	P_{dist}	P_{time}	P_{dist}	P_{time}
Boston	0,694	0,703	0,530	0,542	0,417	0,429
Lisboa	0,695	0,703	0,533	0,542	0,471	0,483
Porto	0,762	0,767	0,614	0,624	0,477	0,489
Rio	0,718	0,726	0,539	0,550	0,394	0,406
São Francisco	0,645	0,658	0,472	0,489	0,367	0,388
São Paulo	0,648	0,659	0,431	0,447	0,294	0,312

Tabela 1 – Tabela apresentando o comparativo entre as taxas de semelhança dos trechos percorridos no grafo das distâncias e no grafo dos tempos.

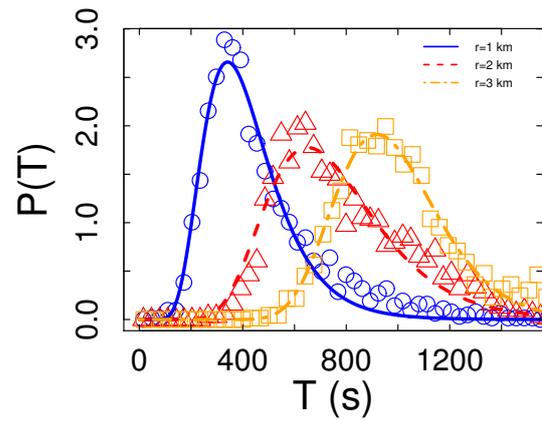
mínima passa a não ser tão frequente. Para o caso de distâncias pequenas entre os pares OD (1 km) as opções não são tão significativas e o trânsito passa a ser o fator determinante, pois não existe a possibilidade de mudança de rota ou não compensa fazer tamanha mudança.

A independência com as cidades para este comportamento pode ser explicado pelo fato de como as ruas e avenidas das cidades são organizadas. Embora sejam cidades com topologias distintas estas se apresentam organizadas em relação ao trânsito com características semelhantes típicas da engenharia de tráfego. Outro aspecto que pode contribuir para este comportamento diz respeito a escala do nosso estudo, que não abrange a cidade como um todo, fato que poderá levar a algumas pequenas alterações em nossa análise.

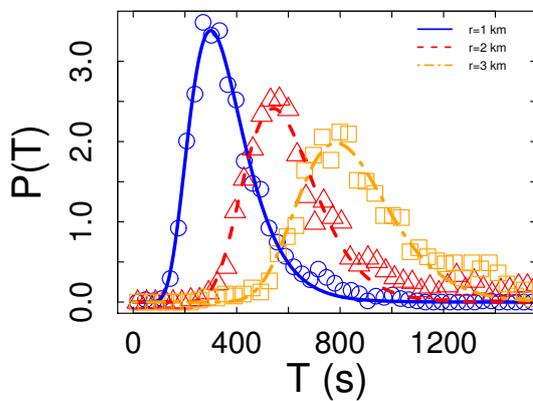
Um outro aspecto que foi analisado em nosso estudo diz respeito a distribuição de tempo demandado entre deslocamentos. Analisamos as distribuições de tempo considerando os deslocamentos de possíveis usuários nas cidades estudadas. Neste caso o critério para o deslocamento foi baseado no menor tempo de trânsito durante o deslocamento considerando o horário de 18h. Os resultados são mostrados nas figura 11. Os símbolos nas figuras correspondem aos resultados experimentais obtidos a partir das estimativas do site Google para três conjuntos de distâncias consideradas, $r = 1$ km, $r = 2$ km e $r = 3$ km. Considerando que cada trajetória estabelecida é feita independente das demais, ou seja, são fenômenos independentes que só dependem da quantidade de usuários utilizando o sistema no momento da requisição dos dados, optamos por tentar um ajuste para as curvas de distribuição utilizando uma função do tipo log-normal. Estes ajustes são representados por meio das curvas em linha contínua e tracejadas nas respectivas figuras.



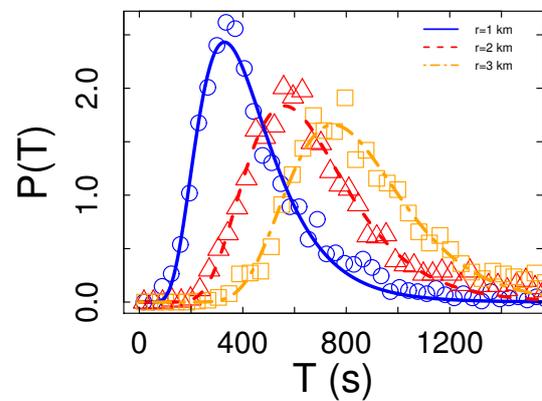
(a) Histograma da cidade de Boston.



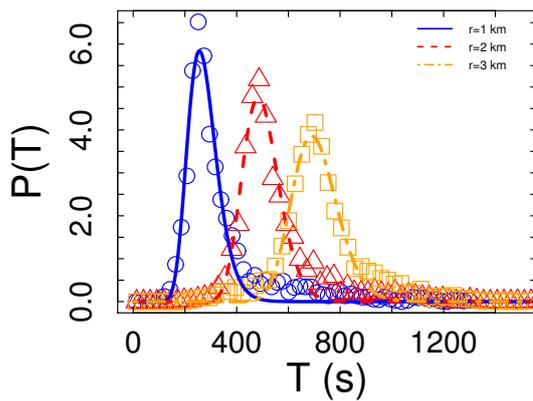
(b) Histograma da cidade de Lisboa.



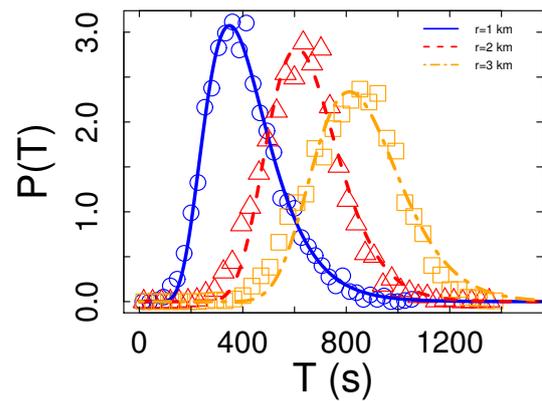
(c) Histograma da cidade de Porto.



(d) Histograma da cidade do Rio.



(e) Histograma da cidade de São Francisco.



(f) Histograma da cidade de São Paulo.

Figura 11 – Distribuições lognormal das cidades estudadas. Nela estão descritos o histograma da distribuição e sua curva de ajuste representada pelas linhas. Analisamos três distâncias entre os pares OD, $r = 1$ km, $r = 2$ km e $r = 3$ km.

Com base nos dados das figuras podemos constatar um excelente ajuste entre os dados experimentais e as curvas propostas para os ajustes. Embora as curvas apresentem um mesmo formato estas exibem algumas peculiaridades que são particulares de cada cidade. No caso do tempo mais provável para uma distância de $r = 1$ km fica em torno de 300s independente da cidade. Porém, quando olhamos para a calda da distribuição na direção de tempos mais longos, estas se apresentam muito semelhante, exceto no caso da cidade de São Francisco que exhibe um decaimento mais acentuado quando comparado com as demais cidades, talvez a configuração de suas vias, como veremos mais a frente, seja responsável por essa distinção.

Para o caso de deslocamentos entre distâncias maiores os aspectos inerentes das estruturas das cidades tornam-se mais evidentes. Este aspecto é reforçado quando olhamos para as distribuições de tempo para esta distância. Neste caso as distribuições ainda são do tipo log-normal, mas apresentam características distintas. Algumas cidades possuem seu tempo médio superior aos 800s, como é o caso de Lisboa e São Paulo, enquanto em outras a mudança de distribuição para caminhos maiores não se altera fortemente, como no caso da cidade do Rio de Janeiro. Já para o decaimento acentuado que ocorre na cidade de São Francisco temos a permanência desse compostamento mesmo para distâncias maiores. É fato que quando a distância em uma viagem é pequena, não existe muita variação no tempo de viagem. No entanto quando a distância entre os deslocamentos aumenta alguns aspectos da organização da cidade e de seu funcionamento com relação ao trânsito ficam mais evidentes, e estas características surgem na forma da distribuição de tempo para as viagens.

4.2 Centralidade: Padrões e distinções.

A centralidade de intermediação de uma ligação, b_c , é de suma importância em nosso trabalho pois nos informa o quão utilizada é determinada ligação (*edge*) no grafo, no nosso caso, como cada ligação simboliza uma via da cidade, a centralidade de intermediação nos mostra as ruas que são mais utilizadas quando usuários realizam deslocamentos pela cidade.

Uma vez que a centralidade de intermediação se estabelece a partir dos menores caminhos entre pontos do grafo, nós a calculamos com base na eq. (2.2) para cada elemento do grafo (ligação) considerando dois aspectos distintos: (i) centralidade de intermediação utilizando as distâncias e (ii) utilizando o tempo de deslocamento na via (ligação). Essa distinção é feita para entendermos individualmente a utilização de cada via e seu papel nos caminhos mais curtos e nos caminhos mais rápidos nas áreas estudadas, descobrindo assim o quão importante tal trecho é naquela região.

Visualizamos na figura 12 as intensidades de centralidade de intermediação para a cidade de Boston considerando os dois critérios previamente discutidos. A primeira vista já percebemos que a maioria das vias e avenidas principais estão em destaque azul (intensidade de utilização maior) e não mudam muito sejam elas representadas na figura 12a ou 12b, mostrando assim o fato de que vias principais são altamente utilizadas sejam pelo fato de serem mais rápidas ou por sua contribuição em menores caminhos. Além desse destaque temos que as vias que se afastam cada vez mais das vias principais se tornam cada vez menos utilizadas mostrando uma conexão com as vias principais, logo para uma grande utilização de determinada via significa proximidade de uma avenida principal. Como as avenidas normalmente conectam regiões estratégicas da cidade, estas inevitavelmente estão presentes seja nos caminhos mais rápidos ou mais curtos, fato que não ocorre com vias ditas secundárias.

Já na figura 13 temos as distribuições de centralidade de intermediação para a cidade de Lisboa. Diferentemente da cidade de Boston as figuras 13a e 13b mostram que há algumas diferenças nas vias principais dessa zona. Enquanto na cidade de Porto, representada na figura 14, existe um local que há pouquíssima utilização. Essa localização sem utilização ocorre por conta de pontes que conectam duas regiões de Porto, mas na região superior de Porto destaca-se as vias principais com a cor azul. As figuras 15, 16 e 17 representam as cidades do rio de Janeiro,

São Francisco e São Paulo, respectivamente.



(a) Centralidade de intermediação em relação a distâncias.

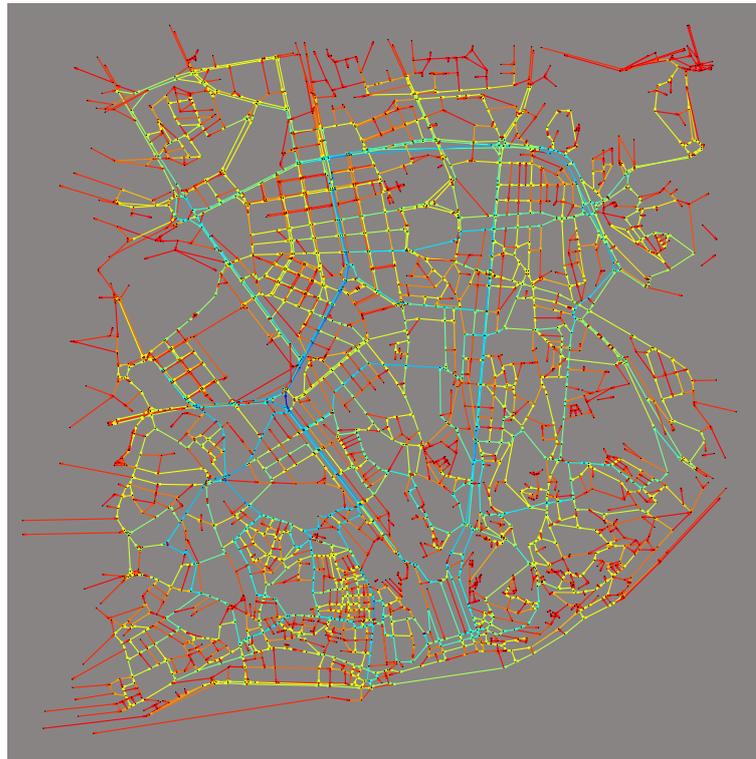


(b) Centralidade de intermediação em relação ao tempo.

Figura 12 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Boston. A escala de cores varia do vermelho para vias pouco utilizadas até azul para vias que são mais utilizadas.

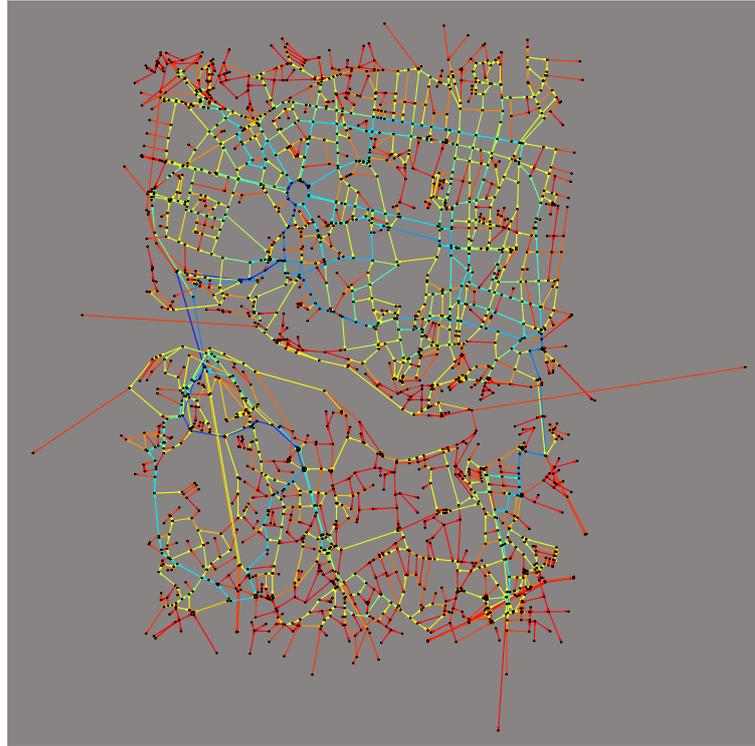


(a) Centralidade de intermediação em relação a distâncias.

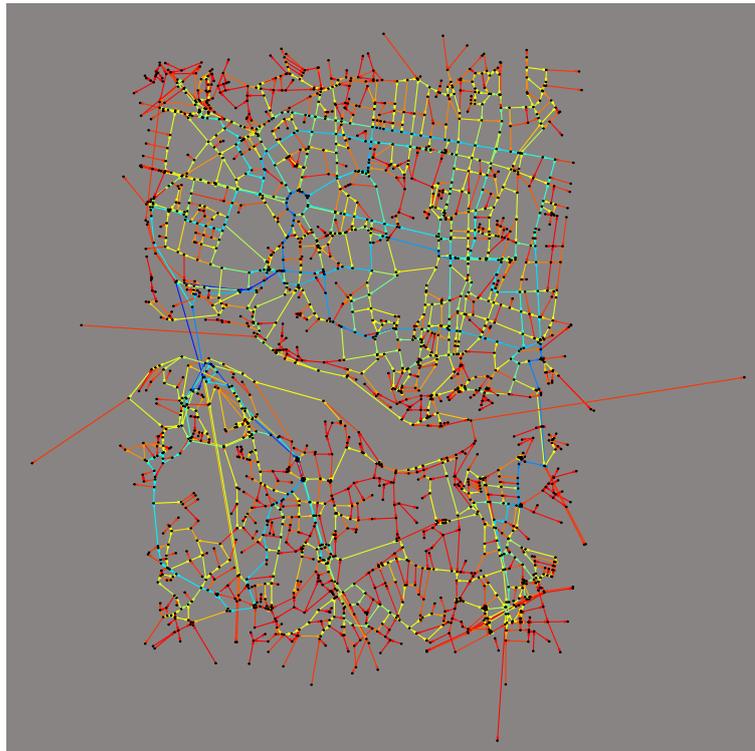


(b) Centralidade de intermediação em relação ao tempo.

Figura 13 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Lisboa. A escala de cores varia do vermelho para vias pouco utilizadas até azul para vias que são mais utilizadas.

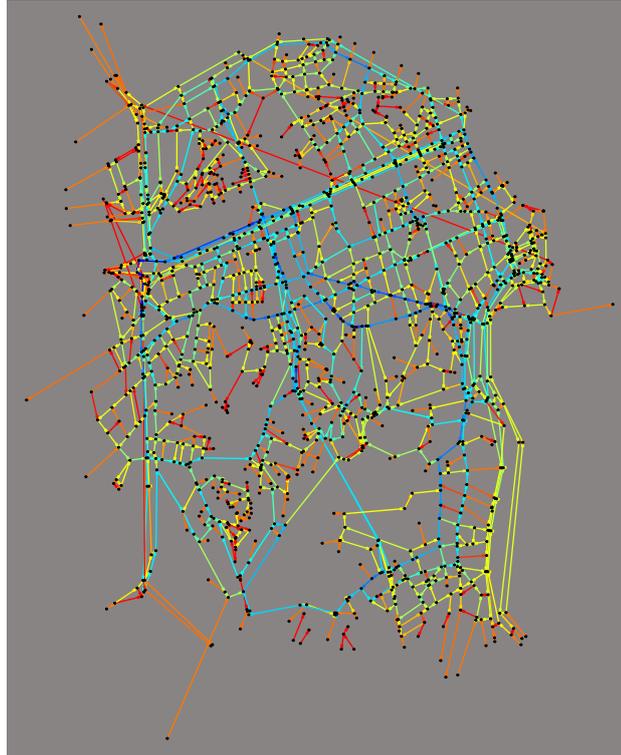


(a) Centralidade de intermediação em relação a distâncias.

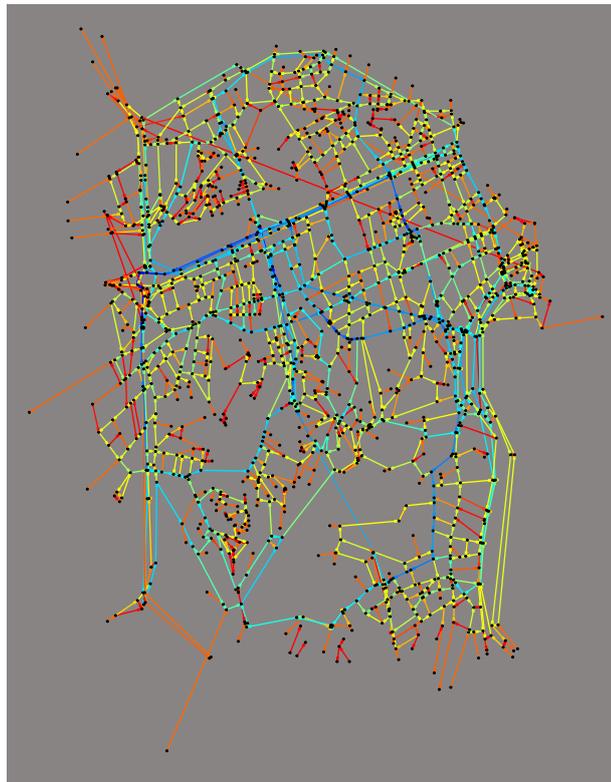


(b) Centralidade de intermediação em relação ao tempo.

Figura 14 – Ambas as figuras descrevem a centralidade de intermediação na cidade de Porto. A escala de cores varia do vermelho para vias pouco utilizadas até azul para vias que são mais utilizadas.

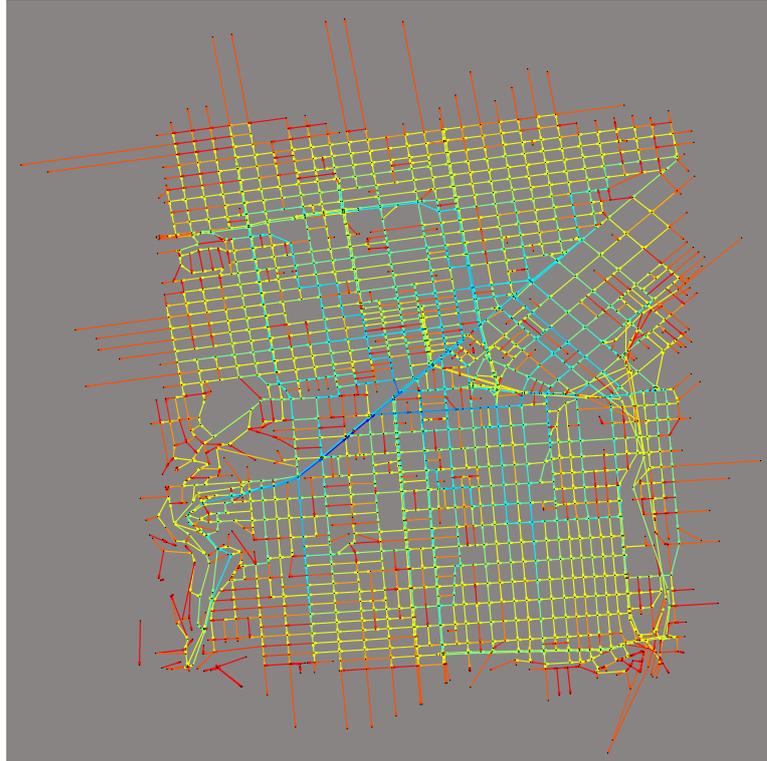


(a) Centralidade de intermediação em relação a distâncias.

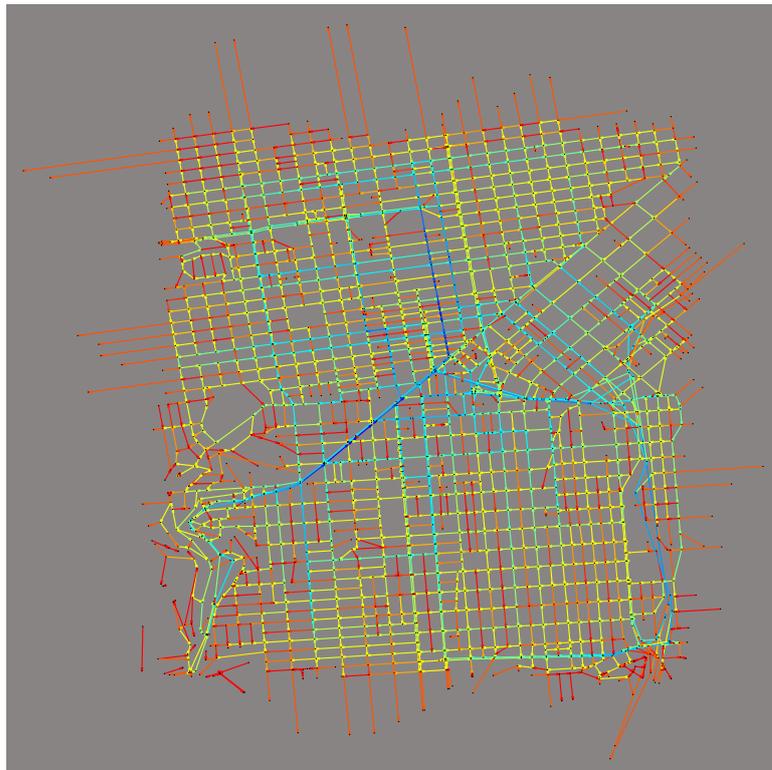


(b) Centralidade de intermediação em relação ao tempo.

Figura 15 – Ambas as figuras descrevem a centralidade de intermediação na cidade do Rio de Janeiro. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.

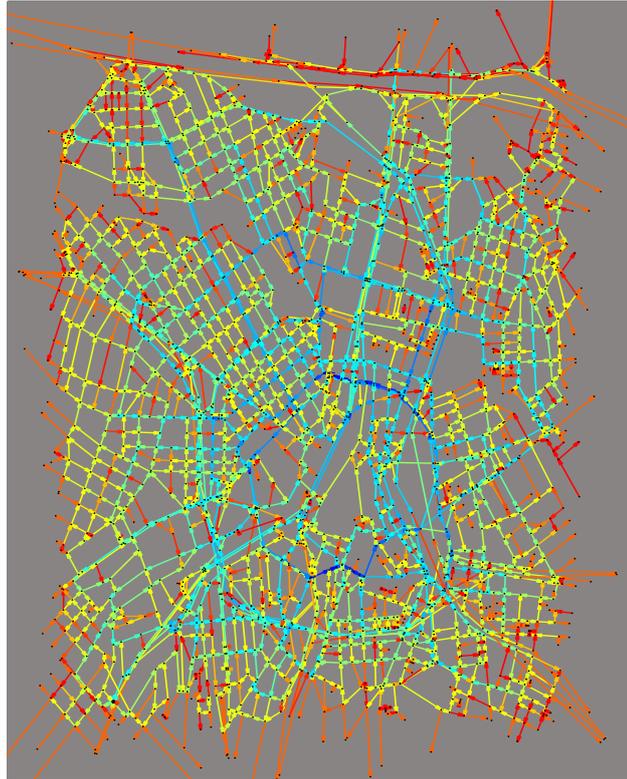


(a) Centralidade de intermediação em relação a distâncias.

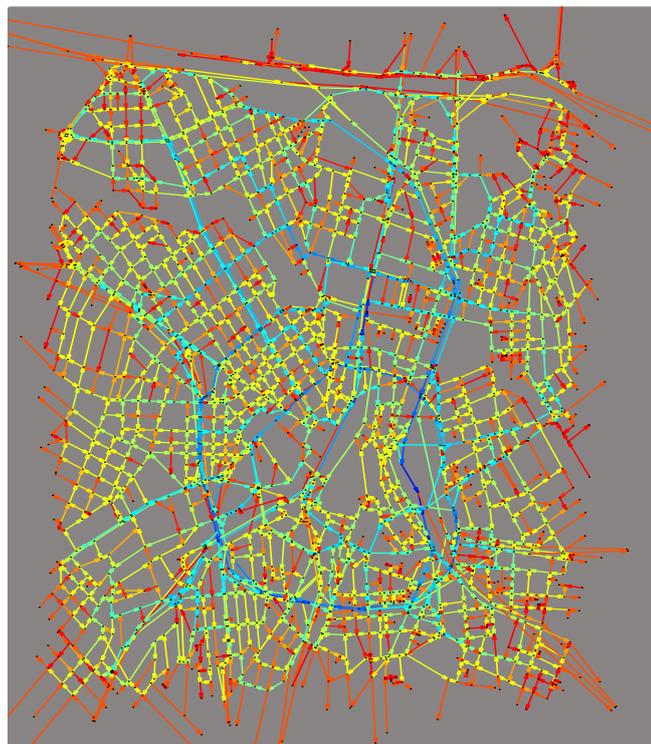


(b) Centralidade de intermediação em relação ao tempo.

Figura 16 – Ambas as figuras descrevem a centralidade de intermediação na cidade de São Francisco. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.



(a) Centralidade de intermediação em relação a distâncias.



(b) Centralidade de intermediação em relação ao tempo.

Figura 17 – Ambas as figuras descrevem a centralidade de intermediação na cidade de São Paulo. A escala de cores varia do vermelho para vias poucos utilizadas até azul para vias que são mais utilizadas.

Uma vez estabelecendo as mudanças na centralidade de intermediação considerando o critério para o cálculo da mesma, com base na distância e no tempo, iniciamos um estudo sobre as distribuições para as diversas cidades. Como uma tentativa para descrição da forma destas distribuições separamos em dois regimes, assim, propomos duas funções teóricas que contemplam dois comportamentos bastante evidentes nas distribuições calculadas nessas cidades. A presença de um comportamento em lei de potência e de uma lei de potência com um decaimento exponencial. Este último comportamento já foi observado na distribuição de massa de agregados de percolação entre dois pontos quando há presença de correlação de longo alcance (ARAÚJO *et al.*, 2003).

Nesse contexto, propomos como função que descreve o comportamento observado para altas centralidade de intermediação a seguinte expressão:

$$f(x) \propto x^{-\delta} \exp \left[- \left(\frac{x}{x_0} \right)^\gamma \right], \quad (4.1)$$

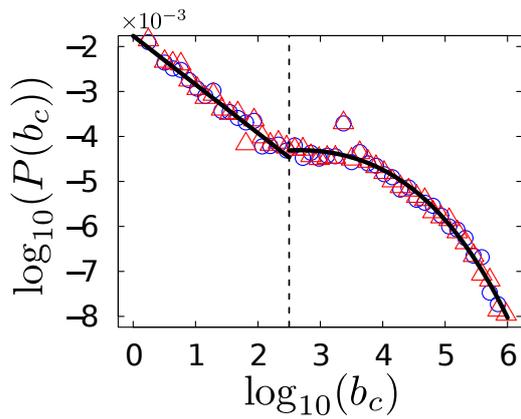
onde x_0 nos informa o *cut-off* e δ e γ os expoentes das respectivas funções. Para pequenas centralidades de intermediação, temos:

$$f(x) \propto x^{-\alpha}. \quad (4.2)$$

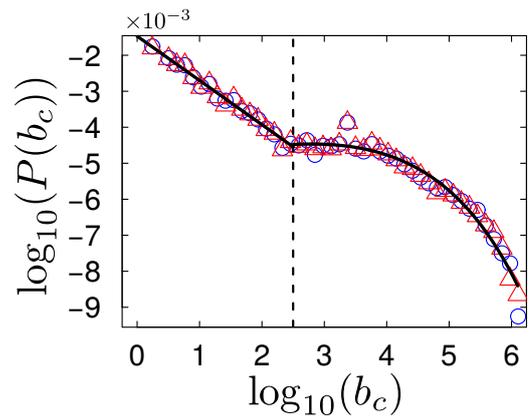
Observamos na figura 18 a *betweenness* para três das seis cidades em estudo, são elas: Boston, Lisboa e São Francisco.

Cidade	α	δ	x_0	γ
Boston	-1,08	-0,989	1,624	0,407
Lisboa	-1,23	-0,545	23,11	0,628
Porto	-1,01	-0,572	5,494	0,476
Rio de Janeiro	-1,12	-0,599	28,08	0,792
São Francisco	-1,05	-1,298	1,693	0,4453
São Paulo	-1,05	0,076	128,9	1,419

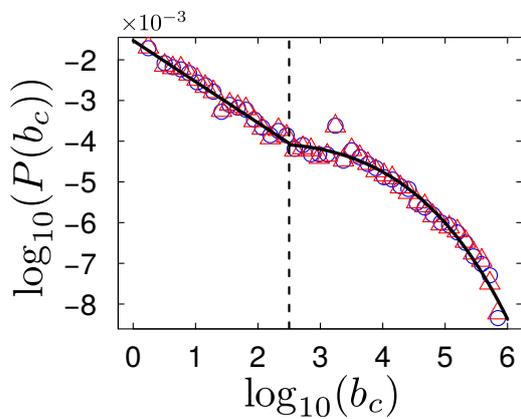
Tabela 2 – Tabela apresentando todos os parâmetros que ajustam os pontos que descrevem a centralidade de intermediação relacionada a distância.



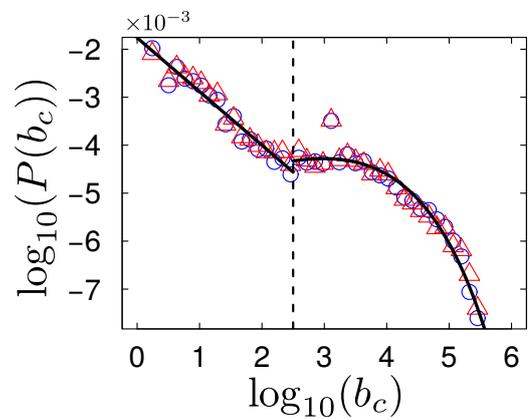
(a) Distribuição da centralidade de intermediação de Boston.



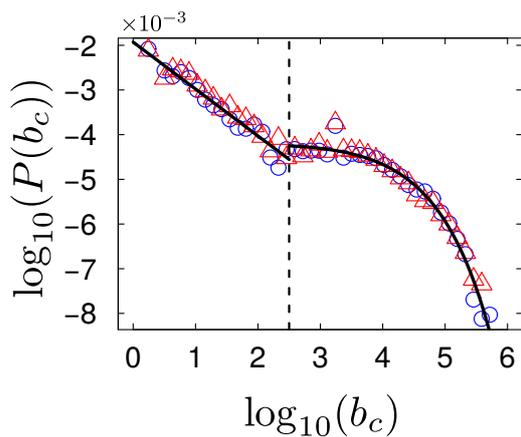
(b) Distribuição da centralidade de intermediação de Lisboa.



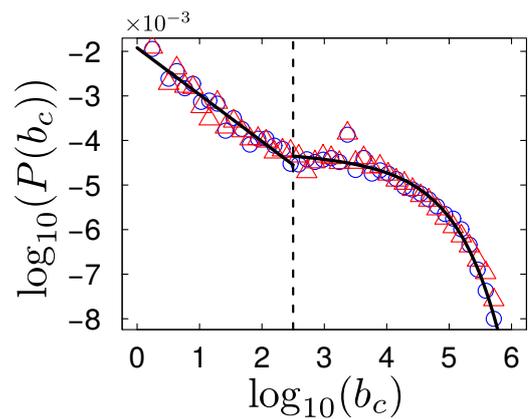
(c) Distribuição da centralidade de intermediação de Porto.



(d) Distribuição da centralidade de intermediação do Rio de Janeiro.



(e) Distribuição da centralidade de intermediação de São Francisco.



(f) Distribuição da centralidade de intermediação de São Paulo.

Figura 18 – Gráfico em escala logarítmica para as distribuições de centralidade de intermediação com relação as cidade estudadas..

Utilizando a função proposta fizemos uma análise minuciosa a cerca dos parâmetros que compõem a função, na tentativa de obtermos o melhor ajuste possível para as distribuições calculadas. Observando as distribuições e os respectivos ajustes, podemos constatar que a função teórica proposta apresenta uma boa concordância para os resultados obtidos por meio do *GoogleMaps*. A curva exibe um bom ajuste tanto para a centralidade de intermediação calculada a partir do critério do tempo quanto para o da distância mínima, visto que ambas são bastante semelhantes para as cidades investigadas. Ocorrem algumas pequenas mudanças nos valores dos expoentes que controlam o comportamento em lei de potência entre cidades, mas as mesmas são muito próximo quando olhamos para uma mesma cidade e critérios diferentes. Como sumário do estudo realizado utilizando as funções propostas, mostramos na tabela 2 os valores dos parâmetros para cada cidade, visto. Todas as curvas da centralidade de intermediação possuem um ponto em comum se disto de todas as curvas, ele é localizado logo após $b_c = 10^3$.

O comportamento em lei de potência observado em nosso estudo já foi reportado em trabalhos anteriores (LÄMMER *et al.*, 2006) onde o autor encontra que a distribuição de centralidade de intermediação segue um comportamento em lei de potência para várias cidade na Alemanha. O tamanho da cidade considerado pode ser um fator limitante quanto a uma análise mais contundente sobre como as ruas e avenidas de uma cidade são utilizadas durante os deslocamentos de seus usuários. Mudanças nos expoentes entre cidades, refletem aspectos locais que são características inerentes do processo de desenvolvimento de cada cidade, bem como do nível de organização das ruas e avenidas que compõem a rede de cada cidade.

4.3 Entropia e Eficiência.

Apresentamos as distribuições das orientações das vias de duas formas distintas: a primeira é como um histograma comum, onde o eixo-x equivale aos ângulos e o eixo-y equivale ao número de ruas que possuem aquela ângulação com o Norte; a segunda é uma diagrama de rosa, muito utilizado por geógrafos e geólogos em estudos. Ela consiste de um diagrama polar facilitando muito a representação e a visualização de certas distribuições. Como lidamos com ângulos que variam de 0° até 180° quando utilizamos o diagrama de rosa, espelhamos e invertemos a metade preenchida de modo que fique algo simétrico como visto na figura 19.

Para a utilização da equação (2.2) em ambiente computacional tivemos que obter previamente duas matrizes. A primeira matriz trata do termo que se encontra no denominador, os termos relacionados a distância entre cada vértice mas restrito a topologia da rede, portanto lidamos com todos os caminhos mínimos possíveis relacionando todos os nós. Obtemos essas distâncias a partir do algoritmo de *Dijkstra* implementado na linguagem R. A segunda matriz trata do numerador da equação (2.2), ou seja, são as distâncias entre os nós da rede em linha reta, logo, sem a limitação do espaço imposta pela topologia da rede. Essas distâncias são obtidas utilizando a equação (3.2) e, uma vez que sempre terá conexão com outro nó, não existe a possibilidade de haver uma distância infinita como ocorre com alguns resultados da primeira matriz já que existem nós que não se conectam de forma alguma. A eficiência é obtida a partir da divisão dessas duas matrizes uma pela outra e após isso ocorre a soma de todos os termos da matriz resultante que não são valores numéricos além de dividir pelo fator $N(N - 1)$, onde N é o número de nós do grafo.

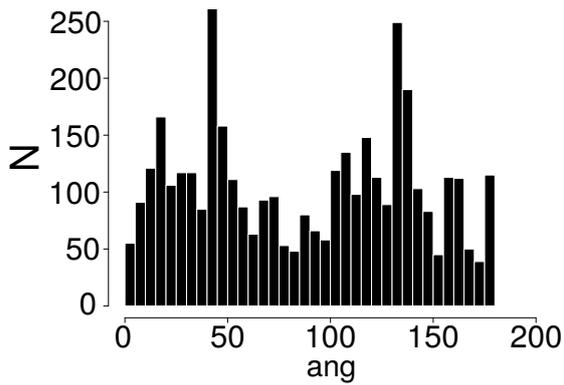
Tanto a eficiência das regiões extraídas das cidades assim como sua entropia organizacional é dada na tabela 3. Ao analisar o valor da entropia percebemos que ela tem pouquíssima variação, diferentemente da eficiência. Mas se olharmos para o percurso médio e relacionarmos com a eficiência, temos que para altos valores de P_m a eficiência tende a ser cada vez menor.

A eficiência apresenta quatro de seis valores abaixo de 0,5 e isso implica em redes bastante ineficientes, algumas razões para isso podem ser explanadas aqui. Primeiro, nem todas as ruas existentes no mundo real participam na construção do nosso grafo, existem vias de acesso e "indefinidas" que descartamos em nosso estudo. Além disso, o tamanho da região estudada, como dito anteriormente, possuem uma escala de 5 km e quanto maior mais discrepâncias há entre a distância que une dois pontos em uma linha reta e a distância que une os mesmos pontos no grafo.

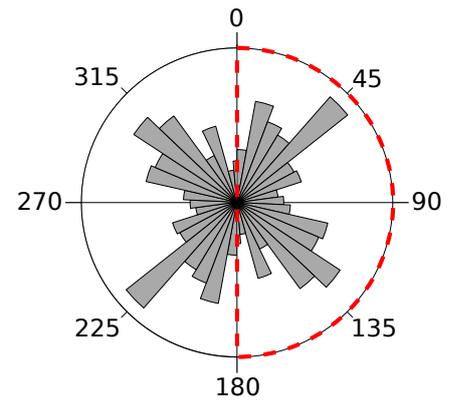
A pouca variação da entropia está relacionado ao fato de que na maioria das regiões estudadas a distribuição da orientação das ruas não é muito diferenciada uma da outra (possuem certas direções em maior quantidade mas as demais, em média, são iguais, implicando assim em entropias com valores aproximados), exceto nas cidades do Rio de Janeiro e principalmente na cidade de São Francisco, como pode ser visto na figura 19.

Cidade	Eficiência	Entropia	P_m	E
Boston	0.4189	1.018	3.858 Km	3.833
Lisboa	0.4192	1.025	3.552 Km	4.707
Porto	0.2484	1.026	4.940 Km	3.183
Rio de Janeiro	0.3967	1.015	4.352 Km	2.366
São Francisco	0.5325	0.8062	3.619 Km	3.462
São Paulo	0.5174	1.0235	3.259 Km	3.890

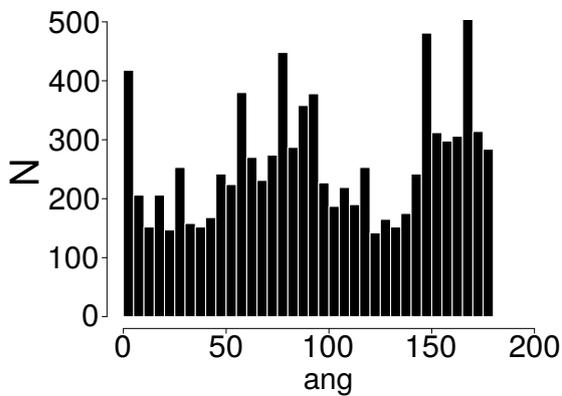
Tabela 3 – Tabela apresentando as eficiências e as entropias das seis cidades estudadas e alguns parâmetros que contribuem direta e indiretamente, são eles o percurso médio, P_m , e o número de ruas, E .



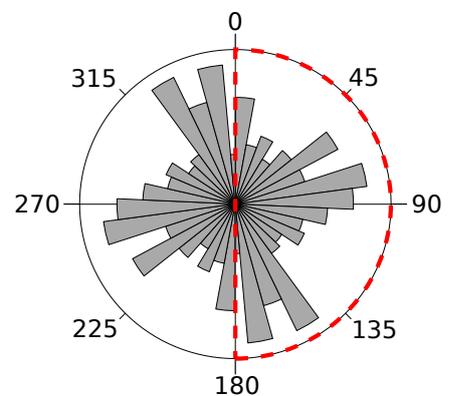
(a) Distribuição da orientação das ruas de Boston.



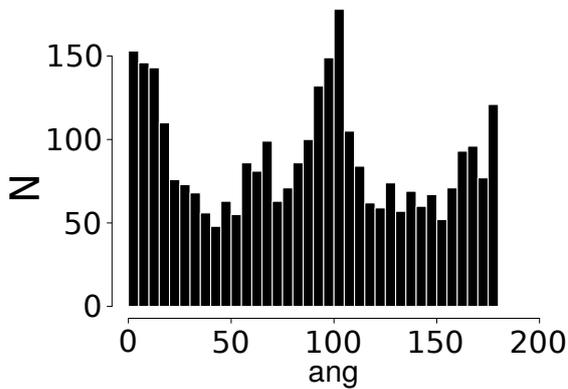
(b) Distribuição em forma de diagrama de rosa de Boston.



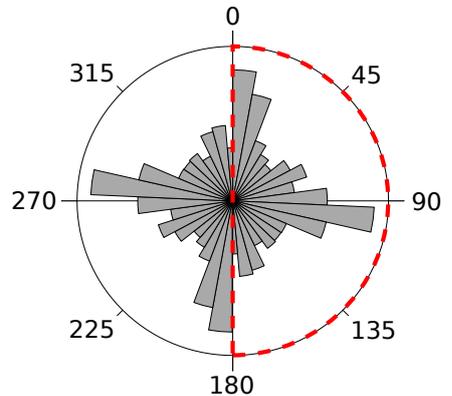
(c) Distribuição da orientação das ruas de Lisboa.



(d) Distribuição em forma de diagrama de rosa de Lisboa.

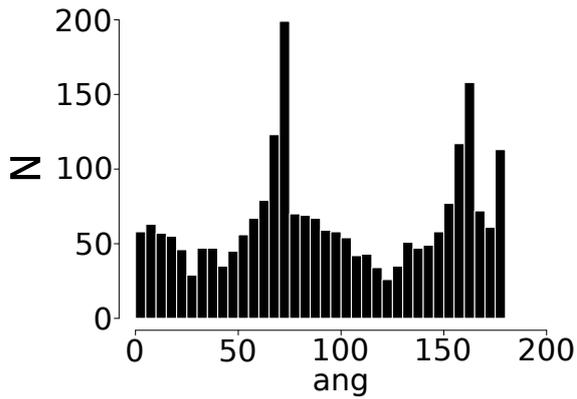


(e) Distribuição da orientação das ruas de Porto.

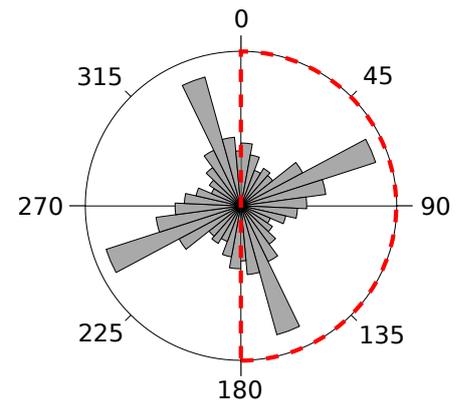


(f) Distribuição em forma de diagrama de rosa de São Francisco.

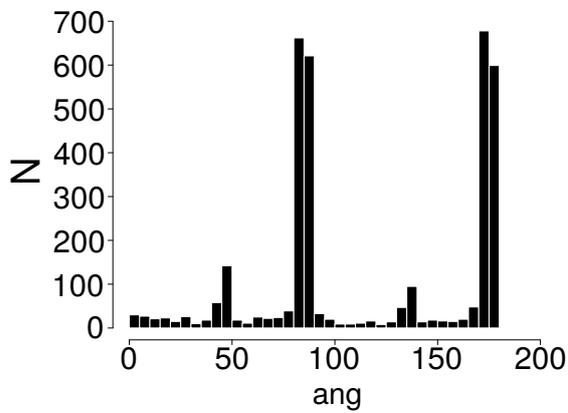
Figura 19 – Orientação das vias das cidades. Nas figuras do lado esquerdo temos representado o histograma da orientação de cada cidade, já no lado direito temos o diagrama de rosa para a mesma.



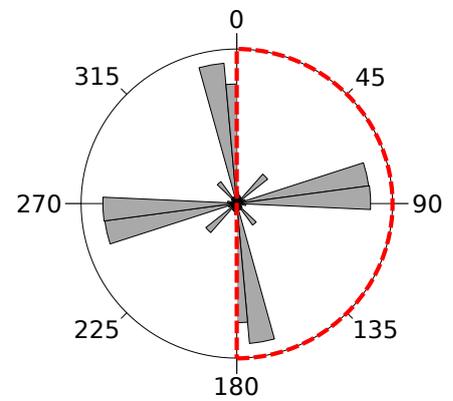
(a) Distribuição da orientação das ruas do Rio de Janeiro.



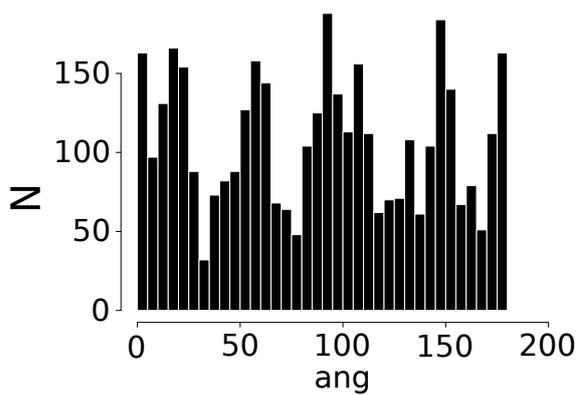
(b) Distribuição em forma de diagrama de rosa do Rio de Janeiro.



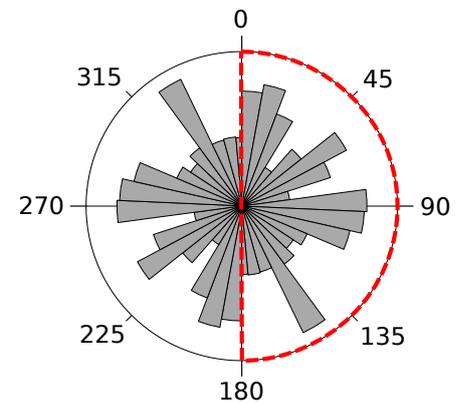
(c) Distribuição da orientação das ruas de São Francisco.



(d) Distribuição em forma de diagrama de rosa de São Francisco.



(e) Distribuição da orientação das ruas de São Paulo.



(f) Distribuição em forma de diagrama de rosa de São Paulo.

Figura 20 – Orientação das vias das cidades. Nas figuras do lado esquerdo temos representado o histograma da orientação de cada cidade, já no lado direito temos o diagrama de rosa para a mesma.

5 CONCLUSÕES E TRABALHOS FUTUROS

Com base neste estudo foi possível identificar que embora as cidades apresentem estruturas morfológicas distintas estas exibem padrões para o deslocamento dos usuários bastante semelhantes. Por exemplo, as distribuições de *betweenness centrality* calculada nas várias cidades seguiram, dentro de um certo limite, padrões similares com um decaimento em lei de potência para pequenos valores de centralidade de intermediação e para altos valores de intermediação temos uma lei de potência com um *cutoff* exponencial. Não foi observado mudança de comportamento significativo quando analisamos a centralidade de intermediação com base no tempo mínimo e distância mínima, elas apresentam o mesmo comportamento e distribuições na cidades semelhantes. Quando investigamos as eficiências e a orientação das ruas algumas diferenças significativas surgiram. Cidades que apresentavam uma boa eficiência e cidades com baixa eficiência, mas tal resultado ainda tem que ser melhor avaliado pois a zona da cidade é específica e dependendo da seleção as orientações das ruas alteram assim como a eficiência. Além disso, distribuições de viagens com o tempo mínimo ocorrem como distribuições Lognormais para três distâncias investigadas, algumas cidades possuem um decaimento após a média um pouco mais acentuado enquanto outras apresentam um achatamento na sua distribuição para altas distâncias. Por fim, temos a razão entre o tempo comum aos caminhos estudados com o tempo demandado em cada um deles. O resultado reforça o pensamento intuitivo de que quanto maior a distância entre dois pontos menor o tempo em comum, mas outro fator interessante é que o decrescimento dessa razão ocorre como uma progressão aritmética de razão entre .

REFERÊNCIAS

- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews of modern physics**, APS, v. 74, n. 1, p. 47, 2002.
- ARAÚJO, A. D.; MOREIRA, A. A.; FILHO, R. N. C.; ANDRADE, J. S. Statistics of the critical percolation backbone with spatial long-range correlations. **Phys. Rev. E**, American Physical Society, v. 67, p. 027102, Feb 2003. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.67.027102>>.
- BANDO, M.; HASEBE, K.; NAKAYAMA, A.; SHIBATA, A.; SUGIYAMA, Y. Dynamical model of traffic congestion and numerical simulation. **Phys. Rev. E**, American Physical Society, v. 51, p. 1035–1042, Feb 1995. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.51.1035>>.
- BARABÁSI, A.-L. **Network Science**. Cambridge University Press, 2016. ISBN 1107076269. Disponível em: <<http://barabasi.com/networksciencebook/>>.
- BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. **Journal of Political Economy**, v. 81, n. 3, p. 637–654, 1973. Disponível em: <<https://doi.org/10.1086/260062>>.
- CARDILLO, A.; SCELLATO, S.; LATORA, V.; PORTA, S. Structural properties of planar graphs of urban street patterns. **Phys. Rev. E**, American Physical Society, v. 73, p. 066107, Jun 2006. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.73.066107>>.
- EULER, L. Solutio problematis ad geometriam situs pertinentis. **Commentarii academiae scientiarum Petropolitanae**, v. 8, p. 128–140, 1741.
- GOOGLE. **Distance matrix API**. 2017. Disponível em: <<https://developers.google.com/maps/documentation/distance-matrix/intro?hl=pt-br>>.
- GUDMUNSON, A.; MOHAJERI, N. Entropy and order in urban street networks. **Scientific Reports**, American Physical Society, v. 3, 2013.
- HANDY, S.; WESTON, L.; MOKHTARIAN, P. L. Driving by choice or necessity? **Transportation Research Part A: Policy and Practice**, v. 39, n. 2, p. 183 – 203, 2005. ISSN 0965-8564. Positive Utility of Travel. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0965856404000941>>.
- HELBING, D. Traffic and related self-driven many-particle systems. **Rev. Mod. Phys.**, American Physical Society, v. 73, p. 1067–1141, Dec 2001. Disponível em: <<https://link.aps.org/doi/10.1103/RevModPhys.73.1067>>.
- INTERNET. 2017. Disponível em: <<https://www.pontorh.com.br/como-fugir-aproveitar-congestionamento/>>.
- INTERNET. 2017. Disponível em: <<https://waze-android.it.softonic.com/android>>.
- INTERNET. 2017. Disponível em: <<https://www.maa.org/press/periodicals/convergence/leonard-eulers-solution-to-the-konigsberg-bridge-problem>>.
- INTERNET ADAPTADA. 2017. Disponível em: <https://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg>.

INTERNET ADAPTADA. 2017. Disponible em: <<https://googlemaps.com>>.

JIANG, B. A topological pattern of urban street networks: Universality and peculiarity. **Physica A: Statistical Mechanics and its Applications**, v. 384, n. 2, p. 647 – 655, 2007. ISSN 0378-4371. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S0378437107006140>>.

LÄMMER, S.; GEHLSSEN, B.; HELBING, D. Scaling laws in the spatial structure of urban road networks. **Physica A: Statistical Mechanics and its Applications**, v. 363, n. 1, p. 89 – 95, 2006. ISSN 0378-4371. Information and Material Flows in Complex Networks. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S0378437106000938>>.

MAKUCH, R. W.; FREEMAN, D. H.; JOHNSON, M. F. Justification for the lognormal distribution as a model for blood pressure. **Journal of Chronic Diseases**, v. 32, n. 3, p. 245 – 250, 1979. ISSN 0021-9681. The Case-Control Study Proceedings of Bermuda Conference held April 1978. Disponible em: <<http://www.sciencedirect.com/science/article/pii/0021968179900705>>.

MIZUSEKI, K.; BUZSÁKI, G. Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. **Cell Reports**, v. 4, n. 5, p. 1010 – 1021, 2013. ISSN 2211-1247. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S2211124713004014>>.

MOREIRA, A.; ARAÚJO, A.; FARIAS, G.; MOREIRA, J.; JR, J. Charge transport in a dynamical system of interacting particles. **Physica A: Statistical Mechanics and its Applications**, v. 278, n. 3, p. 538 – 552, 2000. ISSN 0378-4371. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S0378437199005609>>.

MURAMATSU, M.; IRIE, T.; NAGATANI, T. Jamming transition in pedestrian counter flow. **Physica A: Statistical Mechanics and its Applications**, v. 267, n. 3, p. 487 – 498, 1999. ISSN 0378-4371. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S0378437199000187>>.

MÜLLER MARCEL ROSENBAACH, T. S. M. U. Big data enables companies and researchers to look into the future. **Spiegel Online**, SPIEGEL, 2013.

NEWMAN, M. E. J. The structure and function of complex networks. **SIAM Review**, Society for Industrial & Applied Mathematics (SIAM), v. 45, n. 2, p. 167–256, jan 2003.

NEWMAN, M. E. J.; WATTS, D. J. Scaling and percolation in the small-world network model. **Physical Review E**, American Physical Society (APS), v. 60, n. 6, p. 7332–7342, dec 1999.

RÍOS, J. de Mendoza y. **Memoria sobre algunos metodos nuevos de calcular la longitud por las distancias lunares y explicaciones prácticas de una teoría para la solución de otros problemas de navegación**. Imp. Real, 1795. Disponible em: <<https://books.google.cat/books?id=030t0OqIX2AC>>.

SEYFRIED, A.; STEFFEN, B.; LIPPERT, T. Basics of modelling the pedestrian flow. **Physica A: Statistical Mechanics and its Applications**, v. 368, n. 1, p. 232 – 238, 2006. ISSN 0378-4371. Disponible em: <<http://www.sciencedirect.com/science/article/pii/S037843710600118X>>.

SHANNON, C. E. A mathematical theory of communication. **SIGMOBILE Mob. Comput. Commun. Rev.**, ACM, New York, NY, USA, v. 5, n. 1, p. 3–55, jan. 2001. ISSN 1559-1662. Disponible em: <<http://doi.acm.org/10.1145/584091.584093>>.

SHIELDS, R. Cultural topology: The seven bridges of königsburg, 1736. **Theory, Culture & Society**, v. 29, n. 4-5, p. 43–57, 2012. Disponível em: <<https://doi.org/10.1177/0263276412451161>>.

SIMINI, F.; GONZÁLEZ, M. C.; MARITAN, A.; BARABÁSI, A.-L. A universal model for mobility and migration patterns. **Nature**, Nature Publishing Group, v. 484, n. 7392, p. 96–100, 2012.

TREIBER, A. K. M. **Traffic Flow Dynamics**. [S.l.]: Springer-Verlag Berlin Heidelberg, 2013. ISBN 978-3-642-32459-8.

TREIBER, M.; HENNECKE, A.; HELBING, D. Congested traffic states in empirical observations and microscopic simulations. **Phys. Rev. E**, American Physical Society, v. 62, p. 1805–1824, Aug 2000. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.62.1805>>.

WU, J.; GAO, Z.; SUN, H. Simulation of traffic congestion with sir model. **Modern Physics Letters B**, v. 18, n. 30, p. 1537–1542, 2004. Disponível em: <<http://www.worldscientific.com/doi/abs/10.1142/S0217984904008031>>.

YANG, Y.; HERRERA, C.; EAGLE, N.; GONZÁLEZ, M. C. Limits of predictability in commuting flows in the absence of data for calibration. **Scientific reports**, Nature Publishing Group, v. 4, 2014.

ÇOLAK, S.; LIMA, A.; GONZÁLEZ, M. C. Understanding congested travel in urban areas. **Nature communications**, Nature, v. 7, p. 10793, 2016.