

Enriquecendo Modelos de Mobilidade Urbana com Padrões Espaço-Temporais e Estratos Socioeconômicos

Priscila Santin¹, Fernanda R. Gubert¹, Mauro Fonseca^{1,2},
Anelise Munaretto¹, Thiago H. Silva^{2,3}

¹Pós Graduação em Engenharia Elétrica e Informática Industrial (CPGEI)
Universidade Tecnológica Federal do Paraná (UTFPR)
Av. Sete de Setembro, 3165 – Curitiba – Brasil

²Pós Graduação em Computação Aplicada (PPGCA)
Universidade Tecnológica Federal do Paraná (UTFPR)
Av. Sete de Setembro, 3165 – Curitiba – Brasil

³University of Toronto
55 St. George Street
Toronto, ON, Canada

santin.priscila@gmail.com, fernandagubert@alunos.utfpr.edu.br

***Resumo.** Este trabalho analisa a mobilidade urbana de Curitiba por meio do uso do transporte público. Com a população dividida em subconjuntos correspondentes a estratos socioeconômicos, foram estabelecidas redes onde os vértices correspondem aos bairros da cidade e as arestas representam a existência de um deslocamento entre dois bairros. Para estudar as principais propriedades estruturais dessas redes de mobilidade foram analisados os seus padrões espaço-temporais. Os resultados mostram que padrões espaciais e temporais variam entre os diferentes grupos socioeconômicos. Em particular, o conjunto de dados analisado mostra que, com o aumento da riqueza, a atividade matinal é adiada (em média por 2h) e a distribuição espacial das viagens torna-se mais localizada, com diferenças maiores que 10 km entre as classes.*

1. Introdução

A quantidade de dados gerados, coletados e compartilhados por administrações públicas, empresas, organizações sem fins lucrativos e até pela comunidade científica aumentou consideravelmente nos últimos anos [Agarwal and Dhar 2014]. O uso do conhecimento extraído dessas informações digitais fez uma diferença significativa nos negócios e nos meios de subsistência das pessoas [Labrinidis and Jagadish 2012, Mervis 2012, Nature Editorial 2008]. Nesse sentido, alguns estudos propõem maneiras pelas quais os governos podem usar dados estratégicos para melhor atender seus cidadãos e enfrentar outros desafios, como custos com saúde, criação de empregos, desastres naturais e mobilidade urbana [Manyika et al. 2011, Miranda and Silva 2012].

O debate sobre mobilidade urbana no Brasil tem sido um tema muito discutido entre os administradores públicos das grandes cidades. A mobilidade urbana refere-se à facilidade de movimentação da população e bens no espaço geográfico das cidades, constituindo um componente da qualidade de vida aspirada por seus habitantes [Araujo et al. 2011]. Segundo Miranda et al. [Miranda and Silva 2012], a principal

causa de problemas de mobilidade urbana no Brasil está relacionada ao aumento do uso de transporte individual sobre o uso de transporte público, embora este último também encontre dificuldades com a superlotação. Nesse contexto, problemas de qualidade de infraestrutura e transporte comprometem o deslocamento das pessoas para alcançar fisicamente os destinos desejados [Araujo et al. 2011].

Os estudos sobre padrões de mobilidade revelam princípios importantes em redes de mobilidade humana [Noulas et al. 2012]. Um possível modelo para representar a mobilidade humana é através de redes (ou grafos), que consiste em mapear elementos do sistema em vértices e arestas. A utilização de redes no estudo da mobilidade auxilia na codificação das informações sobre as muitas interações entre partes da rede, modelando e ajudando a explicar o mecanismo para o surgimento do comportamento coletivo [Newman 2010]. A identificação de elementos fortemente conectados, a distância entre eles e a relevância dos elementos formadores da rede são apenas alguns dos recursos que podem ser explorados nessa abordagem.

Nesse sentido, esse estudo apresenta como contribuições: (i) criação de uma rede que representa a origem e o destino dos cidadãos, para isso foi realizada a dedução dos dois bairros mais frequentados diariamente pelos usuários do cartão transporte; (ii) enriquecimento dos dados de mobilidade com a classe socioeconômica do usuário, utilizando dados do Censo Brasileiro; (iii) a caracterização dos padrões de mobilidade feitos no transporte público por moradores de Curitiba, por exemplo, os diferentes tempos entre viagens e distâncias percorridas. É possível observar um comportamento significativamente distinto nos padrões de mobilidade do transporte público para as diferentes classes socioeconômicas, o que habilita um melhor entendimento da cidade sob esse aspecto e a oportunidade de oferecimentos de serviços mais inteligentes aos cidadãos.

O restante do estudo está organizado da seguinte maneira. A Seção 2 apresenta os estudos relacionados. A Seção 3 mostra algumas informações relevantes sobre a cidade de estudo. Na Seção 4, apresentamos a metodologia aplicada neste estudo. A Seção 5 discute os resultados. Finalmente, a Seção 6 apresenta a conclusão deste estudo.

2. Trabalhos Relacionados

A mobilidade urbana é geralmente usada para se referir ao deslocamento de pessoas e veículos, seja por meio de transporte individual (carros, motos, bicicletas) ou por meio do uso de transporte público (ônibus, metrô). Quando se fala de mobilidade urbana, não se deve restringir apenas ao movimento realizados nos grandes centros urbanos. Pode-se estudar a mobilidade em diferentes escalas, seja global [Balcan et al. 2010, Van den Broeck et al. 2011], continental [Cardillo et al. 2013], nacional [Ganin et al. 2017] ou ainda a regional, que inclui cidades grandes ou pequenas [Loterio et al. 2016, Marques-Neto et al. 2018].

Ao estudar a mobilidade urbana regional, deve-se pensar no uso dos espaços da cidade e como garantir o acesso das pessoas ao que a cidade oferece (locais de trabalho, escolas, hospitais, áreas de lazer). Nesse sentido, Yang et al. [Yang et al. 2019] descrevem uma integração de dados de cartão inteligente (SCD) e dados de pontos de interesse (POI) em mídias sociais em um estudo de caso sobre a mobilidade na cidade de Shenzhen na China, com o intuito de deduzir o objetivo da viagem, quantificar os padrões de mobilidade espaço-temporal para diferentes grupos de viajantes e inferir como seus objetivos

de viagem mudam ao longo do dia. Já o estudo de Huang [Huang et al. 2018] analisa a relação entre POIs e comunidades em redes de mobilidade humana, aplicando diferentes métodos de detecção de comunidades e utilizando dados dos sistemas de táxi de duas cidades da China para construir as redes de transporte.

Focando no estudo da mobilidade durante eventos de grande escala, Marques-Neto et al. [Marques-Neto et al. 2018] exploram registros de celulares em três tipos diferentes de eventos: uma grande partida de futebol, um show de rock e uma celebração de Ano Novo. Ao analisar a dinâmica espaço-temporal dos padrões de movimento dos participantes, os autores foram capazes de melhorar a compreensão da mobilidade humana causada por eventos de grande escala e desenvolveram um aplicativo para ajudar as operadoras de telefonia celular a planejar sua infraestrutura para grandes eventos. Ainda usando informações de telefones celulares, Xu et al. [Xu et al. 2018] propõem uma estrutura analítica para entender melhor os padrões de mobilidade humana e sua relação com o status socioeconômico da população, acoplando bases de dados de telefones celulares (como o tipo de comunicação, informações temporais e espaciais de uso da rede) e informações de dados socioeconômicos urbanas.

Lotero et al. [Lotero et al. 2016] analisam a mobilidade urbana entre duas cidades colombianas usando uma pesquisa de origem-destino realizada com seus residentes e acoplando as informações de classificação socioeconômicas, mostrando que os padrões espaciais e temporais variam entre os diferentes grupos socioeconômicos. Os autores concluem que, com o aumento da riqueza, as atividades matinais acontecem mais tarde, o deslocamento do meio-dia se torna mais suave e as viagens se tornam menores.

Em um estudo para a cidade de Curitiba, Vila et al. [Vila et al. 2016] analisam informações de rotas de ônibus, dados de usuários e horários de ônibus para fazer uma primeira investigação sobre a identificação de cenários e implicações dos dados de transporte público de Curitiba para acomodar usuários e suas necessidades de transporte. Os autores caracterizaram vários padrões do sistema como um todo, como a reutilização de ônibus em diferentes linhas, capacidade dos ônibus e número de viagens diárias. Já Braz et al. [Braz et al. 2018] utilizam as informações das bases do transporte público da cidade de Curitiba para estimar a ineficiência nos tempos de viagem realizadas pelo sistema.

Este presente estudo também se concentra na mobilidade do transporte público. No entanto, diferente de todos os esforços anteriores, foram usados dados oficiais dos cartões transporte de Curitiba para criar uma matriz de origem e destino. Além disso, a rede estabelecida com a origem e destino foi aprimorada com os dados do Censo socioeconômico. Com isso, pôde-se estudar o deslocamento espaço-temporal de indivíduos de diferentes classes sociais em Curitiba.

3. Descrição dos Dados

Nesta seção são descritas brevemente informações relevantes sobre a cidade de Curitiba e como seu transporte público é estruturado. Além disso, informações sobre a estratificação social e econômica do Brasil são apresentadas.

3.1. Curitiba e seu Transporte Público

O município de Curitiba possui, de acordo com o IBGE¹ 1.751.907 habitantes em uma área total de 434.967 km^2 . A extensão Norte-Sul é de 35 km e a extensão Leste-Oeste é de 20 km. Administrativamente, Curitiba é dividida em 10 Administrações Regionais, cobrindo seus 75 bairros. Todos esses bairros são atendidos pelo sistema de transporte público de Curitiba. Tal sistema de transporte possui corredores de transporte exclusivos, onde circulam ônibus biarticulados. Esses ônibus biarticulados passam por vários terminais de integração que recebem os ônibus alimentadores da vizinhança, permitindo a integração do sistema. O sistema também possui linhas circulares entre bairros, que permitem a passagem de um bairro para outro, sem ter que percorrer a região central da cidade, além de linhas diretas, que oferecem viagens mais rápidas com menos paradas no itinerário. Essas linhas diretas têm um tipo de ponto de ônibus específico para elas, assim como os terminais de integração. Outra característica importante do transporte público de Curitiba é a tarifa integrada. Pagando apenas uma passagem, o cidadão pode compor sua própria rota, movimentando-se pela cidade.

Desde 2002, Curitiba utiliza o sistema de bilhetagem eletrônica para o transporte público. A implementação desse sistema foi necessária para reduzir a circulação de dinheiro em espécie no sistema de transporte, acelerar o embarque e a passagem de usuários pelas catracas, disciplinar e mensurar o uso do sistema de transporte por categorias que gozam de acesso gratuitos, além de reduzir os custos operação do sistema [Taniguchi and Duarte 2012].

Segundo a URBS² [Urbanização de Curitiba S/A - URBS 2018] em 2018 uma média de 1.365.615 passageiros foram transportados a cada dia útil. Além disso, de acordo com o mesmo relatório, cerca de 60,96% das tarifas foram pagas usando o cartão transporte. O número de cartões transporte ativos em 2018 foi de 1.928.184, dividido em três das quatro categorias de transporte existentes, a saber: Cartão de Usuário, Cartão de Isenção, Cartão de Estudante e Cartão Único (para esta categoria, não há informações nos dados fornecidos pela URBS).

3.2. Estratificação Social e Econômica no Brasil

No Brasil, existem vários critérios disponíveis para classificar a sociedade de acordo com a renda familiar. No entanto, neste estudo, foi utilizado o critério da Secretaria de Assuntos Estratégicos (SAE) do governo federal, estabelecido em 2012 [Ninis et al. 2012]. A Tabela 1 apresenta o resumo dessas classes.

Essas informações de renda foram adquiridas por meio do Censo Demográfico Brasileiro de 2010 realizado pelo IBGE [Instituto Brasileiro de Geografia e Estatística - IBGE 2010]. Esta pesquisa nacional realizada pelo IBGE teve como objetivo retratar a população brasileira com suas características socioeconômicas e também estabelecer uma base para todo o planejamento público e privado para a década entre 2010-2020.

No Brasil, a classificação social de acordo com a renda familiar é dividida em classe alta, classe média e classe baixa. Com base nos estudos realizados pela SAE, esses

¹IBGE: Instituto Brasileiro de Geografia e Estatística

²URBS (Urbanização de Curitiba S/A): empresa que gerencia o transporte público em Curitiba.

Tabela 1. Estratificação Socioeconômica de acordo com a SAE

| Grupo | Renda Familiar |
|-------------------------------|-----------------------|
| Classe 1 - Extremamente Pobre | até R\$ 324 |
| Classe 2 - Pobre | até R\$ 648 |
| Classe 3 - Vulnerável | até R\$ 1.164 |
| Classe 4 - Classe Média Baixa | até R\$ 1.764 |
| Classe 5 - Classe Média | até R\$ 2.564 |
| Classe 6 - Classe Média Alta | até R\$ 4.076 |
| Classe 7 - Classe Alta Baixa | até R\$ 9.920 |
| Classe 8 - Classe Alta | acima de R\$ 9.920 |

grupos foram subdivididos, ficando a classe baixa em três grupos, a classe média também em três grupos e a classe alta em dois grupos. De acordo com essa classificação e a renda estimada por bairro, Curitiba apresenta apenas quatro das oito classes apresentadas na Tabela 1, a saber: Classe Média Baixa, Classe Média, Classe Média Alta e Classe Alta Baixa. Vale ressaltar que existem divergências de rendas intra-bairro, porém a metodologia utilizada para a determinação da classe socioeconômica neste estudo não observa tal detalhe.

4. Metodologia

Nesta seção são descritos o conjunto de dados estudado e as etapas metodológicas mais importantes.

4.1. Bases de Dados de Estudo

Para a análise da mobilidade urbana desenvolvida neste estudo, foram utilizados os dados fornecidos pela URBS. A base de transporte público está disponível para uso pelo público em geral, sem restrições. Para ter acesso a esse conjunto de dados, é necessário fazer uma solicitação à URBS. As informações contidas nessa base de dados são dos registros diários de cartões transporte e cada um desses registros apresentam 7 campos: Código da Linha, Nome da linha, Código do Veículo, Número do Cartão, Data de Utilização, Data de Nascimento e Sexo.

Além disso, esse conjunto de dados possui vários arquivos de informações complementares, como as linhas de ônibus existentes e as paradas de ônibus que as linhas atendem, com localização GPS e tabela horária do veículo. Alguns dados são estáticos e atualizados uma vez por dia. Outros dados, como a localização dos veículos, são atualizados praticamente em tempo real, pois os veículos enviam sua localização em média a cada 5 segundos. Todas essas informações são disponibilizadas por meio de arquivos *.json*.

Neste estudo, foram utilizados dois arquivos complementares relacionados ao transporte público de Curitiba: Veículos e Pontos de Paradas. O arquivo de Veículos contém a localização de todos os veículos que circularam em Curitiba. As informações contidas nesse arquivo são distribuídas em 5 campos: Código do Veículo, Latitude, Longitude, Data de utilização e Código da Linha. Em seguida, o arquivo de Pontos de Paradas contém as informações de todos os pontos de ônibus existentes em Curitiba. Este arquivo possui nove campos de informações, a saber: Nome do Ponto, Número do Ponto, Latitude, Longitude, Sequência, Grupo, Direção, Tipo, Código da Linha.

4.2. Processamento das Bases de Estudo

Para a criação do banco de dados foram utilizados 58 arquivos referentes a entradas de cartão transporte e 67 arquivos de dados sobre os veículos, totalizando 69 dias analisados (53 dias úteis e 16 dias de finais de semana). As entradas de cartão transporte representam no total 17.511.710 registros e os arquivos de veículos 129.461.200 registros.

Para inferir em qual bairro cada entrada de cartão foi realizada, procurou-se uma correspondência nos arquivos dos veículos usando o código da linha, código do veículo, data e hora aproximadas, obtendo desta forma a latitude e longitude de cada entrada de cartão. Cada par de latitude e longitude foi associado ao ponto de ônibus mais próximo da linha específica da entrada do cartão transporte. Este procedimento foi realizado usando uma estrutura de dados hierárquica, conhecida como *Árvore-R*. A *Árvore-R* foi proposta por Antonin Guttman [Guttman 1984] e é amplamente usada como método de acesso ao espaço, permitindo a indexação de informações multidimensionais como coordenadas geográficas. No final desta etapa, foi possível inferir o bairro para 5.388.638 registros de cartão transporte.

Para cada cartão transporte foi necessário associar um bairro de origem, que denominamos a moradia do usuário, e um bairro de destino, o segundo bairro mais visitado com base em todas as atividades de um determinado usuário. Para isso, as entradas foram separadas por horários. Para o bairro de origem foram consideradas as entradas entre 5h e 10h, pois esse é um período no qual as pessoas geralmente usam o transporte público para ir trabalhar ou estudar. Para o bairro de destino, foram consideradas as entradas feitas a partir das 10h.

Entretanto, era necessário identificar quantas entradas mínimas o cartão precisava ter registrado para que pudesse ser considerado. Traçando uma CDF com as entradas realizadas entre 5h e 10h e agrupando os dados por cartão, esse limiar foi definido em 10 entradas. Esse valor representa um bom compromisso entre o número de usuários filtrados e dados disponíveis para cada usuário, pois é possível representar aproximadamente 22% dos cartões, permitindo análises subsequentes com maior confiança. O mesmo procedimento foi realizado para as entradas filtradas a partir das 10h, obtendo uma distribuição muito semelhante, apenas com um pouco mais de entradas nos cartões e, portanto, também foi possível considerar os cartões que obtiveram uma frequência de 10 ou mais entradas. Em seguida, recuperou-se o bairro mais frequente nos dois grupos de horário por cartão e, ao combinar cartões que possuíam bairro de origem com cartões que possuíam bairro de destino, foram obtidos 14.632 cartões transporte exclusivos.

4.3. Métricas de Análise

Para analisar a mobilidade do transporte público de Curitiba foram usadas as informações de origem e destino extraídas conforme apresentado na Seção 4.2. Com isso, uma matriz M da ordem $N \times N$ foi criada, onde $N = 74$, ou seja, o número de bairros de Curitiba³.

Os elementos M_{ij} da matriz indicam o número de viagens observadas do bairro de origem i para o bairro de destino j . Com essas informações, é possível calcular a força do bairro i , ou seja, o total de viagens que saem desse bairro.

³Embora Curitiba tenha 75 bairros, o bairro denominado Parolin não apareceu em nenhum registro usado para este estudo.

$$s_i = \sum_{j=1}^N M_{ij} \quad (1)$$

Outra análise possível de se fazer com as informações de origem e destino, é a conexão dos bairros da cidade. Para esta análise, utiliza-se a matriz adjacente A , onde $A_{ij} = 1$ se pelo menos uma viagem entre i e j foi observada, enquanto $A_{ij} = 0$, caso contrário. Com a matriz A foi possível calcular outras características dessa rede, como o grau dos vértices, representado pela equação 2 e o menor caminho entre os bairros i e j , que nesse caso é definido como a sequência mínima de n vértices $(i, l_1, l_2, \dots, l_{n-1}, j)$. Portanto, $A_{il_1} = A_{l_1l_2} = \dots = A_{l_{n-1}j} = 1$, que também define a distância entre i e j como $d_{ij} = n$.

$$k_i = \sum_{j=1}^N A_{ij} \quad (2)$$

5. Análises dos Dados

5.1. Informações Socioeconômicas sobre a População de Curitiba

A Tabela 2 apresenta de forma simplificada a distribuição da população de Curitiba de acordo com sua classificação socioeconômica e o número de usuários distintos de cartões transporte por classe, com o total de viagens analisadas neste estudo.

Tabela 2. Distribuição da população de acordo com a classe socioeconômica

| | Curitiba | Classe 4 | Classe 5 | Classe 6 | Classe 7 |
|--|-----------|----------|----------|----------|----------|
| população oficial | 1.751.907 | 97.910 | 530.683 | 602.653 | 520.661 |
| % | | 5,59% | 30,29% | 34,40% | 29,72% |
| usuários distintos de cartões analisados | 14.632 | 330 | 3.412 | 3.826 | 7.064 |
| % | | 2,26% | 23,32% | 26,15% | 48,28% |
| viagens analisadas | 721.910 | 14.966 | 163.776 | 180.425 | 362.743 |
| % | | 2,07% | 22,69% | 24,99% | 50,25% |

Ao analisar a Tabela 2 observa-se que o maior número de habitantes de Curitiba são pertencentes à Classe 6. Em termos de cartão transporte analisados é a Classe 7 que se destaca, com quase 50% dos usuários. Analisando o banco de dados para esse estudo, observa-se também que a maioria dos usuários do cartão transporte são da Classe 7, corroborando com as informações apresentadas na tabela.

5.2. Rede de Mobilidade

5.2.1. Rede Geral

A união dessas duas matrizes M e A , permite a visualização da rede de mobilidade estabelecida observada na Figura 1, onde a espessura das arestas representa sua força e o tamanho dos vértices representa o seu grau. Os vértices estão geo-localizados conforme o ponto central do bairro, de acordo com informação oficial disponibilizada pela prefeitura.

Ao analisar a Figura 1 é possível identificar um ponto (bairro) de extrema relevância na região central de Curitiba, conhecido como bairro 'Centro'. Esse é o bairro

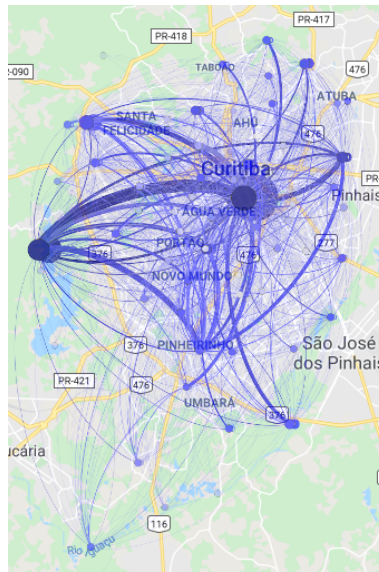


Figura 1. Visão Geral da Rede de Mobilidade de Curitiba

mais importante em quase todas as classes, porque é para essa região que ocorre o maior deslocamento dos usuários de cartões transporte estudados. Destaca-se ainda o bairro 'Cidade Industrial', região com o maior número de habitantes em Curitiba. Ainda pela rede estabelecida na Figura 1 é possível observar um grande deslocamento dos bairros mais periféricos.

5.2.2. Redes por Classes Socioeconômicas

A Figura 2 mostra as redes de mobilidade de transporte público modeladas neste trabalho, correspondentes às Classes 4, 5, 6 e 7 de Curitiba (as únicas observadas na cidade de acordo com o Censo de 2010).

Todas as redes contêm todos os vértices, ou seja, todos os bairros de Curitiba. No entanto não se observa atividades em todos os bairros para todas as classes. Há casos em que a classe α (sendo $\alpha = 4, \dots, 7$) não visita outros bairros da cidade. Nesse caso, o número efetivo de vértices não isolados na rede de mobilidade da classe α é $N_\alpha < N$. O tamanho da aresta reflete a força da ligação, ou seja, quanto mais grossa é a aresta maior é o número de viagens observadas entre os bairros.

Com a ajuda da Figura 2, é possível verificar padrões de mobilidade muito distintos entre todas as classes. Assim, cada rede de mobilidade de cada classe socioeconômica possui diferentes propriedades estruturais. Algumas das principais propriedades são apresentadas na Tabela 3, considerando a rede de forma agregada (observando toda a mobilidade, rede apresentada na Seção 5.2.1) e separadas por cada classe socioeconômica.

Na Tabela 3 são apresentados o número total de vértices não isolados (N) e o número total de arestas. Além disso, foram calculadas as médias (no conjunto de vértices não isolados) das medidas introduzidas na Seção 4.3. A primeira medida corresponde ao grau médio da rede de mobilidade, ou seja, os diferentes destinos alcançados pelos vértices das redes estabelecidas:

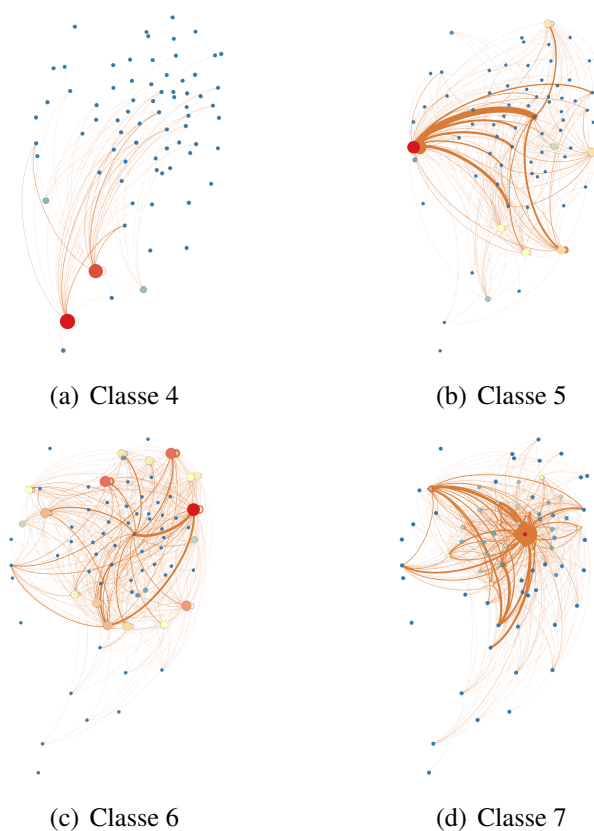


Figura 2. Redes de Mobilidade de Transporte Público por Classes Socioeconômicas

Tabela 3. Propriedades da rede de mobilidade de transporte público

| | Total | Classe 4 | Classe 5 | Classe 6 | Classe 7 |
|----------------------------|---------|----------|----------|----------|----------|
| Número de Vértices | 74 | 31 | 69 | 65 | 67 |
| Número de Arestas | 14,632 | 330 | 3,412 | 3,826 | 7,064 |
| Grau Médio - k_α | 21,227 | 1,912 | 4,028 | 8,014 | 10,269 |
| Força Média - s_α | 195,093 | 4,400 | 45,493 | 51,013 | 94,187 |
| Caminho Médio - d_α | 1,751 | 1,000 | 1,657 | 1,657 | 1,767 |

$$k_\alpha = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} k_{\alpha_i}, \quad (3)$$

onde α representa cada uma das classes ou a rede total. A segunda média está relacionada à força, ao número de viagens observadas entre os bairros:

$$s_\alpha = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} s_{\alpha_i}. \quad (4)$$

A última medida é o caminho médio:

$$d_\alpha = \frac{1}{N_\alpha(N_\alpha - 1)} \sum_{i=1}^{N_\alpha} \sum_{j>1}^{N_\alpha} d_{\alpha_{ij}}. \quad (5)$$

Analisando a Tabela 3, pode-se quantificar as diferenças aparentes observadas visualmente entre as redes de mobilidade existentes para as classes socioeconômicas de Curitiba. Embora a Classe 7 tenha quase o dobro de usuários únicos no conjunto de dados em comparação à Classe 6, ambos têm um alto valor no grau médio k_α , provando ser as redes mais conectadas. Ou seja, os usuários desses estratos socioeconômicos tendem a visitar uma maior diversidade de bairros. Ao comparar as Classes 5 e 6, que têm praticamente o mesmo número de usuários únicos no conjunto de dados e a mesma força média s_α , a Classe 5 tem uma rede muito menos conectada ao analisar seu grau médio. Observando o comprimento médio do caminho d_α pelas classes, descobre-se que as viagens da Classe 4 são mais concentradas em comparação com as outras classes, especialmente na Classe 7, corroborando com o que é observado na Figura 2.

Em seguida calculamos a assortatividade da rede geral levando em consideração as classes socioeconômicas. A assortatividade mede a semelhança das conexões na rede com relação a um determinado atributo e varia de -1 a $+1$ [Newman 2002]. Em um rede com assortatividade positiva, vértices com valores semelhantes do atributo considerado tendem a se conectar um ao outro. A classe socioeconômica foi o atributo considerado e o valor obtido foi: $r = 0.20011$. Isso indica que há uma tendência das pessoas se deslocarem entre bairros da mesma classe.

5.3. Padrões Temporais de Mobilidade

Nesta seção são apresentados os resultados sobre os padrões temporais de mobilidade do transporte público em Curitiba por diferentes classes socioeconômicas. Para essas análises foram consideradas todas as viagens feitas para os 14.632 usuários únicos do conjunto de dados.

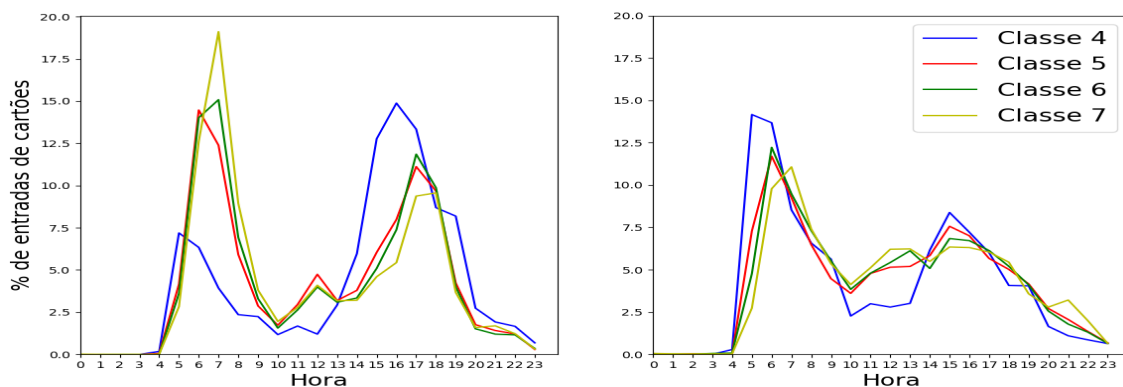


Figura 3. Esquerda: Dias da Semana – Direita: Finais de Semana

Analisando a Figura 3, que mostra a distribuição do número de cartões transporte ao longo do dia, é possível observar que existem dois picos durante os dias da semana, um entre as 5h e as 7h e outro entre as 16h e 18h. Ao meio-dia, há um terceiro pico, mas irrelevante em comparação com os outros dois. Ao analisar os dias de finais de semana observa-se um comportamento ligeiramente diferente, com um pico entre as 5h e as 7h, e depois uma distribuição um tanto uniforme no restante do período, exceto para a Classe 4, que tem uma curva mais acentuada no período da tarde também.

Observando os resultados das classes individualmente durante a semana com mais detalhes, é possível identificar que o pico da manhã atrasa à medida que a classe aumenta.

A Classe 4 atinge o pico às 5h, as Classes 5 e 6 às 6h e a Classe 7 às 7h. Ao meio-dia, há um pequeno pico, mas a Classe 4 não aparece nele. Um comportamento semelhante ao da manhã é observado à tarde. A classe mais baixa tem seu pico de deslocamento anteriormente em relação às demais classes. Enquanto na Classe 4 o pico ocorre às 16h, para a Classe 7 esse pico ocorre às 20h. Este comportamento evidencia que classes mais altas possam ter uma maior flexibilização na jornada de trabalho, o que não acontece numa linha de produção fabril, por exemplo. É importante observar que as janelas de picos da Classe 4 são mais amplas no período da tarde, mostrando uma maior dispersão na utilização do transporte para essa classe ao longo do dia.

Nos finais de semana é possível observar um comportamento muito diferente para a Classe 4 em comparação com as outras classes. A Classe 4 se comporta de maneira semelhante nos dias úteis e finais de semana, mas antecipa seu pico da tarde para às 15h. Já as Classes 5, 6 e 7 têm pela manhã o mesmo comportamento observado durante a semana. A partir das 10h, essas três classes fazem um deslocamento homogêneo, tendo também um pequeno pico às 15h.

Para todas as classes, a grande maioria dos usuários de cartões transporte usam o sistema apenas duas vezes por dia (mais de 80 %), com um intervalo maior do que 10 horas entre esses usos. Isso ressalta a hipótese de que os usuários de transporte público fazem uso do sistema para deslocamentos para fins específicos, como ir ao trabalho.

5.4. Padrões Espaciais de Mobilidade

Neste estudo foram analisadas também as características espaciais dos movimentos urbanos observados nos conjuntos de dados. Na Figura 4 são relatadas as distribuições das distâncias da viagem (em quilômetros), $P(d)$, de cada classe socioeconômica. Para esse cálculo foi utilizada a Fórmula de *Haversine* [Inman 1849], que retorna a menor distância entre dois pontos em uma esfera usando suas latitudes e longitudes medidas ao longo da superfície.

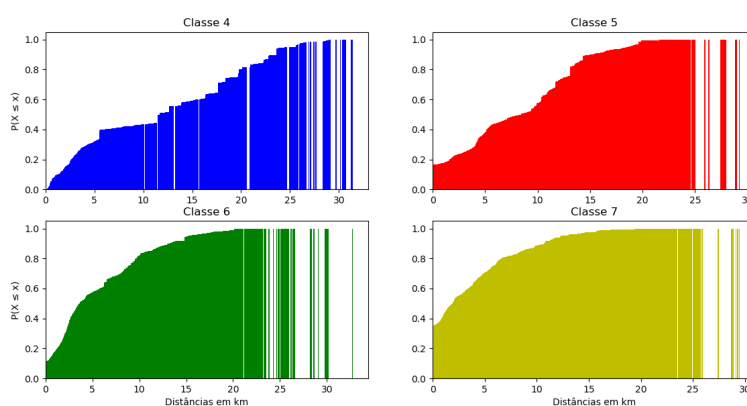


Figura 4. Distribuição do Comprimento d (em quilômetros) das Viagens, $P(d)$, realizada por cada Classe Socioeconômica

É possível observar que a distância percorrida diminui a medida que aumenta a classe socioeconômica. Tomando como base $P(d) = 80\%$, verifica-se que enquanto a Classe 4 percorre aproximadamente 20 km, a Classe 7, nesta mesma proporção, percorre

uma distância um pouco superior à 5 km. Olhando especificamente para a Classe 7, observa-se que quase 40% dos usuários do cartão transporte dessa classe realizam deslocamentos muito próximos a 0 km. Entende-se que esses usuários utilizam-se do sistema de transporte público para pequenos deslocamentos pela alta demanda de linhas de ônibus que permeiam esses usuários, visto que na Classe 7 estão alguns bairros da região central da cidade, que possui uma oferta de linhas de ônibus muito elevada se comparada às demais classes e onde a maioria das atividades disponíveis na cidade, por exemplo, econômicas e culturais, tendem a se conectar. Essa hipótese corrobora com o que pode ser observado na Figura 1, quando foi identificada que a região central, mais precisamente o bairro 'Centro', é o que mais possui ligações na rede estabelecida.

6. Conclusão

Neste artigo, foi apresentado um estudo sobre a mobilidade urbana em Curitiba com base em dados disponibilizados por órgãos públicos da cidade. Adicionando informações sobre o status socioeconômico dos usuários de cartões transporte, foi analisada a distribuição geográfica da origem e destino das viagens e o padrão temporal dessas viagens. As informações sobre classe socioeconômica fornecem uma nova dimensão a ser correlacionada com tempo e espaço. Explorando essas informações, o objetivo deste artigo foi descrever os padrões espaço-temporais mostrados por diferentes classes socioeconômicas.

As redes de mobilidade associadas a cada uma das classes socioeconômicas são bastante diferentes. As redes mais conectadas neste estudo são as representadas pelas Classes 6 e 7 (classes mais altas). Embora a Classe 5 tenha a mesma força que a Classe 6, ela se conecta a menos bairros da cidade ao considerar o seu grau. A Classe 4, por outro lado, tem deslocamentos mais concentrados ao se comparar o comprimento médio do caminho.

Observando o comportamento temporal da viagem, é possível observar que a grande maioria (80%) dos usuários estudados utiliza o sistema de transporte público duas vezes por dia, com intervalo médio de 10h. Existem basicamente dois picos de demanda ao longo do dia; um muito cedo e outro no final da tarde. Durante a semana, esses picos são adiados à medida que a classe socioeconômica aumenta. Nos finais de semana, enquanto a Classe 4 apresenta o mesmo comportamento em relação aos dias da semana, as Classes 5, 6 e 7 alteram o tempo de uso do transporte, sendo praticamente constante a partir das 10h.

Em uma análise espacial, observou-se que a distância percorrida diminuiu a medida que aumenta a classe social. Enquanto 80% dos usuários da Classe 4 percorrem aproximadamente 20 km, nessa mesma proporção os usuários da Classe 7 não ultrapassam 7 km de deslocamento. Sabendo disso e que os bairros de moradia dos usuários da classe 4 encontram-se geograficamente mais afastados do centro da cidade, pode-se formular alternativas de rotas diretas entre origem e destino mais frequentes com base em usuários desta classe, a fim de reduzir o tempo de viagens tão longas.

Identificar os bairros que participam dos deslocamentos mais abundantes e os horários de pico pode auxiliar as empresas de transporte público a ministrar a frota para atender a demanda da população adequadamente e oferecer um serviço com maior qualidade, o que conseqüentemente vai atrair mais usuários ao transporte público e diminuir congestionamentos, além de reduzir o consumo de combustíveis e a poluição ambiental.

Uma visão da ocupação urbana considerando as diferentes classes socioeconômicas permite entender como as pessoas nessas classes se apropriam dos espaços urbanos. Embora a diferença de renda entre as classes seja muito sutil, de acordo com os dados do censo, é notável a diferença nos padrões de mobilidade entre essas mesmas classes. Nesse mesmo contexto, incorporar a presença de diferenças socioeconômicas ao estudar o desenvolvimento de processos de contágio ou a disseminação de informações nas áreas urbanas é um trabalho futuro.

Agradecimentos

Este trabalho foi parcialmente apoiado pelo projeto CNPq-UrbComp (processo 403260/2016-7), CAPES (código de financiamento 001) e projeto GoodWeb (processo #2018/23011-1 da agência de fomento FAPESP).

Referências

- Agarwal, R. and Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for is research. *Information Systems Research*, 25(3):443 – 448.
- Araujo, M., Melo de Oliveira, J., Jesus, M., Sa, N., Santos, P., and Lima, T. (2011). Collective public transportation: Discussing accessibility, mobility and quality of life. *Psicologia Sociedade*, 23:574–582.
- Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., and Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model. *Journal of Computational Science*, 1(3):132 – 145.
- Braz, T., Maciel, M., Mestre, D. G., Andrade, N., Pires, C. E., Queiroz, A. R., and Santos, V. B. (2018). Estimating inefficiency in bus trip choices from a user perspective with schedule, positioning, and ticketing data. *IEEE Transactions on Intelligent Transportation Systems*, 19(11):3630–3641.
- Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., del Amo, A. J. G., and Boccaletti, S. (2013). Modeling the multi-layer nature of the european air transport network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*.
- Ganin, A. A., Kitsak, M., Marchese, D., Keisler, J. M., Seager, T., and Linkov, I. (2017). Resilience and efficiency in transportation networks. *Science Advances*, 3(12).
- Guttman, A. (1984). R-trees: A dynamic index structure for spatial searching. *SIGMOD Rec.*, 14(2):47–57.
- Huang, L., Yang, Y., Gao, H., Zhao, X., and Du, Z. (2018). Comparing community detection algorithms in transport networks via points of interest. *IEEE Access*, 6:29729–29738.
- Inman, J. (1849). *Navigation and Nautical Astronomy, for the Use of British Seamen*. F. J. Rivington, 7 edition.
- Instituto Brasileiro de Geografia e Estatística - IBGE (2010). Censo 2010. <https://censo2010.ibge.gov.br>. [Online; acessado em 07-Novembro-2019].
- Labrinidis, A. and Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proc. VLDB Endow.*, 5(12):2032–2033.

- Lotero, L., Hurtado, R. G., Floría, L. M., and Gómez-Gardeñes, J. (2016). Rich do not rise early: spatio-temporal patterns in the mobility networks of different socio-economic classes. *Royal Society Open Science*, 3(10):150654.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. [Online; accessed 13-November-2019].
- Marques-Neto, H. T., Xavier, F. H. Z., Xavier, W. Z., Malab, C. H. S., Ziviani, A., Silveira, L. M., and Almeida, J. M. (2018). Understanding human mobility and workload dynamics due to different large-scale events using mobile phone data. *Journal of Network Systems Management*, 26(4):1079 – 1100.
- Mervis, J. (2012). Agencies rally to tackle big data. *Science*, 336(6077):22–22.
- Miranda, H. d. F. and Silva, A. N. R. d. (2012). Benchmarking sustainable urban mobility: The case of Curitiba, Brazil. *Transport Policy*, 21:141 – 151.
- Nature Editorial (2008). Community cleverness required. *Nature*, 455(7209):1–1.
- Newman, M. (2010). *Networks: An Introduction*. Oxford University Press.
- Newman, M. E. J. (2002). Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701.
- Ninis, A. B., Paes de Barros, R., Grosner, D., Franco, S., Rosalém, A., Bugarin, Karina Sayuri Sataka, G., and Jorge, J. (2012). *SAE - Relatório de definição da classe média*.
- Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., and Mascolo, C. (2012). A tale of many cities: Universal patterns in human urban mobility. *PLOS ONE*, 7(5):1–10.
- Taniguchi, G. and Duarte, F. (2012). Personal smart cards: From transportation to a city smart card—the database integration of public services in Curitiba. In *City Competitiveness and Improving Urban Subsystems: Technologies and Applications*, pages 217–232. IGI Global.
- Urbanização de Curitiba S/A - URBS (2018). URBS em números. <https://www.urbs.curitiba.pr.gov.br/institucional/urbs-em-numeros>. [Online; acessado em 07-Novembro-2019].
- Van den Broeck, W., Gioannini, C., Gonçalves, B., Quaggiotto, M., Colizza, V., and Vespignani, A. (2011). The gleamviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC infectious diseases*, 11:37.
- Vila, J., Koziévitch, N., Fonseca, K., Gadda, T., Rosa, M., and Gomez-Jr, L. (2016). Urban mobility challenges – an exploratory analysis of public transportation data in Curitiba. *Revista de Informática Aplicada*, 12:1.
- Xu, Y., Belyi, A., Bojic, I., and Ratti, C. (2018). Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Computers, Environment and Urban Systems*, 72:51 – 67.
- Yang, Y., Heppenstall, A., Turner, A., and Comber, A. (2019). Who, where, why and when? using smart card and social media data to understand urban mobility. *ISPRS International Journal of Geo-Information*, 8(6).