

CARACTERIZANDO A PROBLEMÁTICA DA BAIXA ACESSIBILIDADE DO TRANSPORTE PÚBLICO DE FORTALEZA COM O USO DE *BIG DATA*

Carlos Kauê Vieira Braga
Carlos Felipe Grangeiro Loureiro
Departamento de Engenharia de Transportes
Universidade Federal do Ceará

RESUMO

A compreensão da problemática da baixa acessibilidade se apresenta como uma etapa fundamental do planejamento urbano integrado, permitindo um melhor direcionamento na elaboração de objetivos e desenvolvimento de ações. Tradicionalmente, estudos feitos sobre essa problemática utilizam amostras pequenas para estimação de indicadores. O sistema de transporte público por ônibus de Fortaleza, por sua vez, coleta dados de validação e posicionamento dos veículos da frota, aqui chamados de *big data*, apresentando um grande volume de informações com custo baixo e alta confiabilidade. Essa pesquisa de mestrado se propõe a estabelecer e aplicar métodos para a caracterização dessa problemática com o uso do *big data*, sendo isso feito através da consolidação dos dados, da proposição e cálculo de indicadores e da descrição da situação atual. Com o uso da linguagem de programação R, os dados foram consolidados, e indicadores de acessibilidade, mobilidade e operação do transporte público foram calculados. Além disso, métodos de análise espacial em áreas e de *machine learning* estão sendo implementados para a análise dos indicadores. Espera-se estabelecer metodologias para a análise do *big data* e contribuir para o direcionamento de intervenções no sistema de transportes.

1. INTRODUÇÃO

A compreensão da problemática da baixa acessibilidade dos usuários cativos do transporte público se apresenta como uma etapa importante do planejamento urbano integrado. Soares *et al.* (2018) citam que os esforços de identificação, caracterização e diagnóstico da problemática dos transportes e uso do solo são fundamentais tanto para a tomada de decisões conscientes como para uma melhor negociação dos conflitos entre os atores que fazem parte do ambiente urbano. Dentre as etapas de compreensão da problemática, destacam-se as análises de caracterização embasadas na formulação e modelagem dos indicadores para a mensuração dos problemas estratégicos, táticos e operacionais identificados.

Tradicionalmente, a modelagem dos principais indicadores dessa problemática de acessibilidade e mobilidade urbana depende da realização de pesquisas domiciliares (Henrique *et al.*, 2005) que representam uma pequena amostra do padrão de deslocamento da população (geralmente 1%) e apresentam um custo elevado. Em Fortaleza, a última pesquisa domiciliar aconteceu em 1996 o que leva os modeladores e planejadores da cidade a calcularem indicadores atuais que carregam consigo grande incerteza. Em contrapartida, a crescente disponibilidade de dados advindos das mais recentes tecnologias de informação e comunicação (*Information and Communications Technology - ICT*) cria novas oportunidades e questões para o seu uso no planejamento urbano. No planejamento do transporte público, dados de *smartcard* e de GPS se apresentam como fontes promissoras de *big data*, fornecendo informações quase populacionais do comportamento dos usuários por um baixo custo e oferecendo um alto nível de desagregação.

Vários trabalhos foram feitos nos últimos anos utilizando dados provenientes de *smart card* de transporte público para a obtenção de indicadores (Barry *et al.*, 2002; Kusakabe e Asakura, 2014; Alsker *et al.*, 2016). Kurauchi e Schmocker (2016) compilaram uma grande gama de estudos referente a esse assunto. No Brasil, destacam-se os trabalhos de Farzin (2008) e Arbex e Cunha (2017), os quais realizaram esforços para estimar matrizes origem-destino a partir dos dados de *smart card* e GPS. Além disso, é possível a aplicação de ferramentas de *machine learning* (Gay, 2018), podendo representar avanços em outras etapas de caracterização da problemática que não somente a do cálculo dos indicadores.

2. OBJETIVOS

Diante das lacunas observadas, o objetivo geral desta pesquisa de mestrado é caracterizar a problemática da baixa acessibilidade e mobilidade dos usuários do transporte público em Fortaleza através do uso de ferramentas de *big data*. Seus objetivos específicos são:

- Representar a problemática da baixa acessibilidade dos usuários de transporte público de Fortaleza às atividades de trabalho e educação;
- Consolidar a base de *big data* do transporte público de Fortaleza;
- Propor e calcular indicadores a partir de *big data* que sejam úteis no processo de compreensão da problemática;
- Propor método de descrição da situação atual da problemática utilizando ferramentas de *machine learning*.

3. REVISÃO DA LITERATURA E METODOLOGIA

A análise de revisão da literatura busca justificar a relevância das lacunas metodológicas e fenomenológicas, servindo também de base para o estabelecimento do método de caracterização aqui proposto.

3.1 Representação da problemática

A representação da problemática consiste na consolidação integrada dos problemas (Soares *et al.*, 2018), levantando hipóteses sobre suas relações de causa e efeito. Garcia *et al.* (2018) apresentam uma representação geral da problemática da acessibilidade e mobilidade urbana causada por dois fatores principais: oferta ineficiente de transportes e distribuição inadequada do uso do solo. No presente trabalho, busca-se analisar de forma mais específica como as ineficiências na oferta de transportes afetam a acessibilidade dos usuários de transporte público em Fortaleza. Lima (2017) fez uma representação da problemática do subsistema de transportes. Sua proposta serviu como base para o detalhamento das hipóteses de causalidade dos problemas inerentes à acessibilidade na rede de transporte público.

3.2 Métodos de caracterização da problemática

O método de caracterização proposto nesse trabalho de dissertação apresenta três etapas principais, diretamente ligadas aos objetivos específicos: a consolidação das bases de dados, a proposição e cálculo dos indicadores, e a descrição da situação atual. A consolidação dos dados consistiu na coleta, tratamento e integração do *big data* de transporte público. A base de *big data* consiste de cinco fontes principais: dados de bilhetagem (*smartcard*); dados de GPS (*Automated Vehicle Location*); dados *General Transit Feed Specification* (GTFS); dados dos endereços dos usuários do sistema; e dados do Google Maps. Os dados de bilhetagem incluem a validação na catraca de cada usuário do transporte público, registrando o momento exato da validação, a linha, o tipo de pagamento e o tipo da viagem (integração ou não). Os dados de GPS apresentam registros de localização de todos os veículos da frota a cada 30 segundos. Os dados de GTFS fornecem a localização de todas as paradas, itinerário das linhas e tabela de horários programados. Por conta do grande volume dos dados, a consolidação da base foi realizada com o uso da linguagem de programação R, levando em conta também seu grande potencial e caráter *open source*. Os dados foram processados das bases já disponíveis no grupo de pesquisa, da base de dados aberta da ETUFOR (Empresa de Transporte Urbano de Fortaleza) e de consultas ao API do Google Maps. A integração consistiu na criação de uma nova base de dados unindo o momento exato da bilhetagem com o instante exato do GPS para determinar a localização aproximada de cada validação, além da tradução das variáveis em formatos que permitisse uma integração facilitada entre as bases.

Com os dados consolidados, entrou-se na etapa da proposição e cálculo dos indicadores. Essa etapa foi realizada com a utilização da representação da problemática já proposta, com cada problema sendo mensurado por um indicador. A revisão da literatura permitiu avaliar quais indicadores podiam ser obtidos, sendo destacado o trabalho de Kurauchi e Schmocker (2016). A linguagem de programação R foi utilizada, sendo desenvolvida uma metodologia para o cálculo de cada indicador. Entre os indicadores propostos e calculados, destaca-se a matriz origem-destino (calculada a partir de dados de bilhetagem + GPS), que recebe uma grande atenção da literatura com vários métodos. Para o indicador de acessibilidade, van Wee (2016) indica questões e caminhos para a utilização de *big data*. Foi proposto então um indicador que utiliza a matriz OD estimada para a calibração de uma função de decaimento do método gravitacional. Uma função de custo generalizado também foi proposta, onde a incorporação de novos atributos foi possível com o uso do *big data* (a lotação do veículo, por exemplo). Indicadores operacionais, como frequência (a partir de GPS), cobertura (a partir de GTFS, Google Maps e endereços), e velocidade por link por linha (a partir de GPS) também foram calculados.

Os indicadores calculados permitem a definição da situação atual da problemática. Os indicadores apresentam níveis de agregação diversos, o que permite que os problemas sejam caracterizados em diferentes níveis de agregação e atuação, seja no nível operacional, tático ou

estratégico. Metodologias de análise espacial de padrão em áreas estão sendo aplicadas para a detecção de *clusters* de acessibilidade. Além disso, métodos de *machine learning* de detecção de padrões desconhecidos nos dados através de algoritmos *k-means* também estão sendo implementados (Gay, 2018).

4. RESULTADOS ESPERADOS

O trabalho busca contribuir tanto no âmbito metodológico como no fenomenológico. Espera-se estabelecer métodos de consolidação do *big data* de transporte público, principalmente na parte da integração entre as bases de dados disponíveis, permitindo a utilização da base gerada em trabalhos futuros. Espera-se também propor métodos mais robustos para o cálculo de indicadores a partir de *big data*, com a contribuição de técnicas de *machine learning* na descrição da situação atual, assim como na calibração e validação de modelos de demanda e oferta de transportes.

Para o fenômeno da baixa acessibilidade por transporte público em Fortaleza, entende-se que uma matriz OD estimada pelo *big data* pode fornecer informações mais confiáveis sobre o padrão de deslocamento da população, além de um mapeamento e análise de indicadores de oferta que permitam identificar áreas e linhas problemáticas com maior precisão de medida e desagregação, contribuindo no direcionamento de intervenções táticas e operacionais.

REFERÊNCIAS BIBLIOGRÁFICAS

- Arbex, R. O.; Cunha, C. B. (2017) Estimação da matriz origem-destino e da distribuição espacial da lotação em um sistema de transporte sobre trilhos a partir de dados de bilhetagem eletrônica. *Revista Transportes*
- Alsger, A.; Assemi, B.; Mesbah, M.; Ferreira, L. (2016) Validating and improving public transport origin-destination estimation algorithm using smart card fare data. *Transportation Research part c: Emerging technologies*, 2016. v. 68, p. 490–506
- Barry, J. J.; Newhouser, R.; Rahbee, A.; Sayeda, S. (2002) Origin and destination estimation in New York City with automated fare system data. *Transportation Research Record*
- Farzin, J. (2008) Constructing an automated bus origin-destination matrix using farecard and global positioning system data in São Paulo, Brazil. *Transportation Research Record*
- Garcia, C. S. H. F.; Macário, M. R. M. R.; Menezes, E. D. A. G.; Loureiro, C. F. G. (2018) Strategic Assessment of Lisbon's Accessibility and Mobility Problems from an Equity Perspective. *Networks and Spatial Economics*
- Gay, J. S. (2018) Learning spatial inequalities: an approach to support transportation planning. Escola Politécnica da Universidade de São Paulo. Dissertação de Mestrado. São Paulo, 2018
- Henrique, C. S.; Loureiro, C. F. G.; Cavalcante, R. A. (2005) Caracterização Espacial da Mobilidade dos Usuários Cativos do Sistema Integrado de Transportes de Fortaleza. *Transporte em Transformação IX - Trabalhos Vencedores do Prêmio CNT Produção Acadêmica*
- Kurauchi, F.; Schmocker, J. D. (2016) Public transport planning with smartcard data. 2016.
- Kusakabe, T.; Asakura, Y. (2014) Behavioural data mining of transit smart card data: A data fusion approach. *Transportation research part c: Emerging technologies*, 2014. v. 46, p. 179–191.
- Lima, L. S. (2017) Espreadimento urbano por autosssegregação e seus impactos na acessibilidade urbana de Fortaleza. [S.l.]: Universidade Federal do Ceará, 2017. Dissertação de Mestrado.
- Soares, F. D. P.; Loureiro, C. F. G.; Miller, E. J. (2018) Diagnosis: a problem-oriented approach to urban transportation planning. *Transportation Research Board*
- van Wee, B. (2015) Accessible accessibility research challenges. *Journal of Transport Geography*, 2015