

IDENTIFICAÇÃO E ANÁLISE DE PADRÕES DE VIAGENS DOS USUÁRIOS DO SISTEMA INTEGRADO DE TRANSPORTES DE FORTALEZA POR MEIO DO BIG DATA DE BILHETAGEM ELETRÔNICA

Renato Goersch Andrade Parente
Francisco Moraes de Oliveira Neto

Universidade Federal do Ceará (UFC)

Programa de Pós-Graduação em Engenharia de Transportes (PETRAN)

RESUMO

A compreensão da demanda por transporte público é essencial para orientar políticas de intervenção no sistema que priorizem a acessibilidade da população mais vulnerável no meio urbano. Nos últimos anos o sistema de transporte público de Fortaleza vem sofrendo várias mudanças operacionais, como implantação de um sistema eletrônico de bilhete único e intervenções viárias que permitiram a integração temporal do sistema e a priorização dos veículos de transporte público. Contudo, não se sabe como a demanda por transporte público vem mudando ao longo do tempo e no espaço e quais fatores contribuem para sua variação. Acredita-se que a vasta disponibilidade de dados coletados continuamente pelo sistema de bilhetagem eletrônica permite uma melhor compreensão sobre o perfil dos usuários que utilizam o sistema, o padrão de deslocamentos no meio urbano destes usuários, a variação da demanda no tempo e no espaço, e os fatores que contribuem para variação da demanda. Por meio da análise de um grande volume de dados, *Big Data* da bilhetagem eletrônica de Fortaleza, esta pesquisa tem como principal objetivo analisar a variabilidade da demanda realizada por ônibus e identificar fatores relacionados ao uso do solo, da operação do sistema, e outros fatores indiretos que possam interferir na variação da demanda.

1. INTRODUÇÃO

A compreensão da demanda por transportes é essencial para o planejamento e dimensionamento da oferta por transportes no meio urbano (Ferraz e Torres, 2004). A demanda em sistemas de transportes é resultado de interações entre elementos de oferta e elementos de demanda dos subsistemas que pertencem ao contexto urbano, em destaques podemos citar: Subsistema de Transportes, de Uso do Solo e de Atividades. A dinâmica das interações entre estes três subsistemas impacta na demanda, fazendo com que esta varie espacialmente e temporalmente. Conforme destacado por Balcombe *et al.* (2004), diversos fatores podem influenciar na variação da demanda por transportes público, tais como: eficiência e cobertura do sistema de transporte, disponibilidade de modos alternativos, características dos usuários, a diversidade de uso do solo, fatores políticos e culturais, entre outros. Dada esta diversidade de fatores de influência, a análise da demanda de transporte, portanto, é uma atividade essencial no planejamento de transportes. Ela permite estimar as necessidades de transporte a partir da previsão do volume de tráfego e de passageiros que usarão as instalações de transporte. Essa previsão é essencial para projetos de intervenções de transporte e para a avaliação de sua viabilidade econômica (Kanafani, 1983). Cascetta (2009) aponta que o conhecimento da variação no tempo e no espaço geográfico é essencial para o planejamento adequado da oferta de transporte, de modo a proporcionar um atendimento eficiente (economia de recursos) e de qualidade (satisfação dos usuários). Ou seja, não só a compreensão espacial da demanda num dado, mas também sua variação temporal, entre períodos do dia, entre dias da semana, entre meses do ano, e sua evolução entre anos, são essenciais para o planejamento estratégico, tático e operacional do sistema de transportes.

Outro aspecto importante a ser compreendido sobre a demanda por transporte público é o perfil dos usuários que utilizam o sistema, incluindo suas características sociais e padrões de deslocamento (diário, semanal e mensal). O advento dos sistemas automáticos de bilhetagem (*Smart Card*) e de localização geográfica veículos (GPS), vem permitindo uma melhor

compreensão sobre como os diversos usuários utilizam o sistema de transportes público e de como a demanda vem se alterando no tempo e no espaço. No planejamento de transportes, os dados de *Smart Card* e GPS vem se apresentando como novas fontes promissoras de dados, fornecendo informações indiretas quase populacionais do comportamento dos usuários por um baixo custo e um alto nível de desagregação. Pelletier *et al* (2011) faz uma revisão de literatura com várias aplicações de dados de bilhetagem eletrônica em transportes, inclusive para análises de demanda e sua variabilidade. Muito se vem avançando na utilização de *Big Data* no planejamento de transportes. No Brasil, destacam-se os trabalhos de Farzin (2008) e Arbex e Da Cunha (2017), que fizeram esforços para estimar matrizes origem-destino. Ainda com dados de bilhetagem, Pieroni (2018) avança na identificação de grupos e padrões de viagens e análise de variabilidade. Freitas (2014), através desse mesmo tipo de dado, tenta caracterizar a problemática do sistema de transporte público de Fortaleza olhando para um recorte no tempo desses dados. Assim como Freitas (2015), Braga (2019 – Em Andamento) busca construir um sistema de indicadores para a caracterização da problemática da acessibilidade e mobilidade no transporte público de Fortaleza.

Ainda existem diversas lacunas de pesquisa sobre o conhecimento da demanda, especialmente como esta varia no tempo e no espaço, e qual o perfil dos usuários que utilizam o sistema de transportes. Tal conhecimento pode ajudar a compreender a razão da queda no volume de passageiros observada nos últimos anos em Fortaleza, conforme apontado por gestores do sistema. Acredita-se, então, que o *Big Data* de bilhetagem eletrônica pode ajudar nesse processo de compreensão e análise da demanda para o caso da cidade Fortaleza, que conta com sistema de bilhete único desde 2013. Portanto, este estudo propõe um aprofundamento na identificação e análise da demanda do sistema de transporte público por ônibus em Fortaleza realizada por Freitas (2014), que buscou caracterizar essa demanda através da análise do retrato de um ano específico. Esse aprofundamento se dará principalmente em uma análise mais desagregada desta demanda, que terá como base a identificação de grupos de usuários que possam representar melhor as características sociais, assim como o comportamento da demanda no espaço e no tempo. É pretensão deste estudo também aumentar o período analisado de um ano para cinco anos, no caso de 2014 até 2018. Esse aprofundamento busca responder especificamente as seguintes questões: Como compreender a demanda através do uso dos dados de *Smart Card* do Sistema de Transporte Público por ônibus em Fortaleza? Quais são os padrões de deslocamentos que podem ser identificados com os dados de bilhetagem do sistema de transporte público de ônibus em Fortaleza? Como identificar e interpretar esses padrões? Como os grupos identificados são distribuídos temporalmente e espacialmente? Como essa distribuição muda ao longo do tempo? Quais os fatores que impactaram nessas alterações?

Dada as questões levantadas, o objetivo geral deste estudo é: compreender a demanda realizada do Sistema Integrado de Fortaleza e sua variabilidade espacial e temporal, através de análise do *Big Data* de bilhetagem eletrônica. Ou seja, compreender como essa demanda varia no espaço e no tempo, e quais são os fatores que impactam essas variações. Esse objetivo almeja ser alcançado através dos seguintes objetivos específicos:

- Identificar grupos e padrões de deslocamento na demanda transportada do sistema de transporte público de ônibus que auxiliem na caracterização da demanda;
- Analisar a variabilidade da distribuição espacial e temporal dos grupos identificados;
- Contextualizar os sistemas de transportes; uso do solo e atividades, de modo a levantar hipóteses sobre os fatores que contribuem na variação da demanda;

- Identificar os fatores que influenciam a demanda por transporte público.

2. METODOLOGIA E DADOS

Neste estudo, serão utilizados dados da operação e da demanda do Sistema de Transporte Público de Fortaleza, para um período de 2015 a 2018. Os dados incluem: Bilhetagem Eletrônica, rastreamento dos veículos por GPS, *General Transit Feed Specification* (GTFS) e dados de cadastro dos usuários do bilhete único. Os dados são cedidos pela empresa operadora dos ônibus e pela Prefeitura de Fortaleza. Os dados de bilhetagem são gerados quando os usuários efetivam o embarque através do seu *Smart Card*. A localização dos embarques será estimada a partir da utilização dos dados de GPS dos veículos, por meio de um método de associação entre as duas fontes de dados (GPS e Bilhetagem Eletrônica) desenvolvido no software R por Braga (2019 – Em andamento). Os dados de GTFS contém informações da operação do sistema de ônibus, como por exemplo: dados de horários, itinerários das linhas, local de paradas e tarifas, esses dados permitirão estimar e validar melhor as localizações de embarques que serão buscadas nos dados de GPS.

2.1 Contextualização do sistema e Análise descritiva

Esta etapa visa contextualizar os Sistemas de Transportes, de Uso do Solo e de Atividades de Fortaleza, por meio do levantamento de histórico de intervenções e mudanças ocorridas de em cada um desses subsistemas. Esta análise de contextualização dos subsistemas urbanos será posteriormente usada na identificação de fatores que influenciam a demanda por transporte público, especificamente relacionados ao desemprego, à qualidade da oferta do sistema de transporte, e ao surgimento ou crescimento de outras opções de modos de transportes (i.e., *ber* e a *motocicleta*). Uma análise descritiva da demanda do sistema de transporte público será realizada com intuito de levantar hipóteses iniciais sobre a relação da demanda com características do Sistema de Uso do solo e Sistema de Atividades. Para esta análise e as análises de variabilidade da demanda, indicadores da demanda serão propostos, como volume de passageiros transportados por corredor, por área, em períodos específicos do dia (i.e., *pico* e *entre-pico* das viagens, a serem definidos com base nos dados).

2.3 Identificação dos padrões de viagens

A identificação de grupos que possam melhor representar os diversos comportamentos diferentes da demanda se faz necessário. Assim, os dados serão agregados através de técnicas de clusterização (a serem definidas). De acordo com Jain *et al.* (1999), essas técnicas podem ser úteis em muitas situações de análises exploratórias de padrões, podendo ser aplicadas para identificar padrões de viagens, como destacado por Pieroni (2018), com base em atributos como distância de deslocamento, horário da primeira validação (hora em que o bilhete único é ativado), tempo entre validações no sistema, frequência de viagens semanais, entre outros. Estes padrões de viagens serão associados ao cadastro dos usuários do sistema, gerando perfis ou grupos de usuários conforme as suas características socioeconômicas e de viagens. Com isso será possível compreender como diferentes usuários do sistema se deslocam para realizar suas atividades no meio urbano.

2.4 Análise de variabilidade espacial e temporal

Nesta etapa, analisaremos como a demanda de cada grupo varia em diferentes níveis de agregação espacial (i.e., corredor de transportes, regiões geográficas da cidade e terminais de integração) e temporal (dentro de um mesmo dia, entre dias da semana, entre meses do ano, entre anos). Portanto, esta análise irá permitir caracterizar espacialmente e temporalmente a

demanda de cada grupo identificado, conforme descrito na Subseção 2.3, permitindo analisar as diferenças e similaridades de cada grupo de usuários. Além dos indicadores definidos na etapa de contextualização, outros indicadores da demanda em diferentes níveis de agregação temporal e espacial, serão definidos, tais como: viagens produzidas e atraídas por região geográfica, fluxos de passageiros nos terminais de integração, fluxos de passageiros em corredores de transportes, em períodos específicos do dia (i.e. pico e entre-pico das viagens, a serem definidos com base nos dados). Hipóteses sobre variação espacial e temporal da demanda serão verificadas a partir de análise de inferência estatística: análise de variância, ANOVA, para comparação de médias entre dias, meses do ano e entre diferentes entidades espaciais; análises espaciais (autocorrelação espacial) para analisar a variação espacial da demanda. Por fim, a variabilidade da demanda será confrontada com fatores identificados na contextualização, a partir de análises de correlação bivariada e correlação espacial bivariada.

3. RESULTADOS ESPERADOS

Ao fim desta pesquisa, espera-se compreender o padrão de deslocamentos dos usuários do transporte público de Fortaleza, especificamente, identificando quem são os diferentes usuários do sistema, onde eles se localizam no espaço, e como estes diferentes grupos utilizam o sistema de transporte público. Espera-se compreender como o padrão de deslocamento de cada grupo vem mudando ao longo do tempo e quais fatores possivelmente influenciam nessa mudança. Portanto, a pesquisa contribui para compreensão de como a demanda por transporte público varia no espaço e no tempo. Além disso, a pesquisa busca propor um método para extrair informação (identificação de padrões e análise destes padrões) de um *Big Data* de transporte público urbano. Por fim, a pesquisa contribui com indicadores da demanda e análises para a etapa de caracterização de problemas do planejamento nos níveis estratégico, tático e operacional.

Agradecimentos

Os autores agradecem o apoio financeiro do CNPQ para o desenvolvimento deste trabalho.

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*, v. 25, n. 3, p. 166, 2017.
- Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M. and White, P. (2004) *The demand for public transport: a practical guide*. TRL Reports. Report TRL 593, pp. 1-238.
- Cascetta, E. (2009) *Transportation System Analysis: Models and applications*. Ed. Springer (2ª ed.), New York, USA.
- Ferraz, A. C. P.; Torres, I. G.E. (2004) *Transporte público urbano*. São Paulo, Rima Editora.
- Farzin, J. (2008) Constructing an Automated Bus Origin-Destination Matrix Using Farecard and Global Positioning System Data in São Paulo, Brazil. *Transportation Research Record: Journal of the Transportation Research Board*, v. 2072, n. 2072, p. 30–37.
- Freitas, A. T. (2015) *Metodologia de caracterização da problemática do sistema de transporte público de passageiros a partir dos dados da bilhetagem eletrônica*. 98 f. Dissertação (Mestrado em Engenharia de Transportes) - Centro de Tecnologia, Universidade Federal do Ceará.
- Jain, A. K.; Murty, M. N.; Flynn, P. J. (1999) Data clustering: a review. *ACM Computing Surveys*, v. 31, n. 3, p. 264–323.
- Kanafani, Adib K. (1983) *Transportation demand analysis*. New York: McGraw-Hill College.
- Pelletier, M. P., Trépanier, M., and Morency, C. (2011) Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557–568.
- Pieron, C. B. V. (2018) *Analysis of Travel Patterns from Precarious Settlements Transit Users in São Paulo through Smart Card Data Mining*. Dissertação (Mestrado em Engenharia de Transportes) - Universidade de São Paulo.