



ESTIMAÇÃO DOS ATRIBUTOS RELEVANTES NA SATISFAÇÃO DE USUÁRIOS DE TRANSPORTE *INTERCAMPI* POR *RANDOM FOREST*

Tarssio Brito Barreto¹

Ícaro Bernardes dos Santos Coutinho²

Jorge Ubirajara Pedreira Júnior³

Karla Patricia Santos Oliveira Rodríguez Esquerre⁴

Resumo

A sustentabilidade de um serviço de transporte público depende da fidelização dos passageiros. Mesmo nas usuais condições vantajosas aos operadores, a adesão dos usuários não é garantida, uma vez que outros modais privados e públicos estão disponíveis aos usuários. Tal vínculo com os passageiros pode ser reforçado ao cumprir as expectativas de qualidade destes. Entretanto mensurar o status atual e importância de atributos de qualidade é tarefa difícil e subjetiva, usualmente realizada através de questionários ao público. O presente trabalho estimou as variáveis mais influentes na satisfação geral dos usuários com o serviço de transporte *intercampi* da Universidade Federal da Bahia, campus Salvador. Com o auxílio do R e bibliotecas relevantes foi possível extrair dados geográficos de um API, agregar novas variáveis ao banco de dados e efetuar a estimação indireta das variáveis mais relevantes na experiência geral no uso do serviço. Esta estimação foi realizada através do método da *random forest* e a precisão obtida foi cerca de 97%. As árvores tiveram seu parâmetro de complexidade para a poda estimado por meio de validação cruzada. Os resultados sugerem que a insegurança nos *campi* é fator significativo a favor da adesão dos usuários ao serviço, mas estes não estão satisfeitos com a segurança provida no mesmo. Ademais, a satisfação global dos usuários com o transporte é mais sensível à pontualidade, a frequência e a disponibilidade de informações sobre o serviço prestado. O curto intervalo entre aulas, ocasionalmente ministradas em diferentes *campi*, pode justificar tal preocupação com aspectos temporais na performance do serviço. Portanto, sugerem-se estudos que permitam compassar os horários de passagem dos veículos e as necessidades dos usuários levando em conta distúrbios externos como o tráfego local.

¹ Doutorando pelo Programa de Engenharia Industrial, Universidade Federal da Bahia (UFBA), tarssioesa@gmail.com

² Mestrando pelo Programa de Engenharia Industrial, Universidade Federal da Bahia (UFBA), asaicaro@gmail.com

³ Departamento de Engenharia de Transportes e Geodésia, Universidade Federal da Bahia (UFBA), jorge.ubirajara@ufba.br

⁴ Departamento de Engenharia Química, Universidade Federal da Bahia (UFBA), karlaesquerre@ufba.br



Palavras-chave: qualidade de serviço, transporte público, *random forest*, percepção do cliente

Abstract

The sustainability of a public transport service depends on the passengers' loyalty. Even on the usual advantageous conditions to operators, users' adhesion is not guaranteed, since other private and public means of transport are available to users. This connection with the passengers can be reinforced by fulfilling their quality expectations. However evaluate the present status and importance of quality attributes is a difficult and subjective task, usually made through public surveys. This work estimated the most influent variables on the general satisfaction of the users with the transport service within the Federal University of Bahia, in Salvador. With the aid of R and relevant libraries it was possible to extract geographic data from an API, aggregate new variables to the database and perform the indirect estimation of the most relevant variables on the global experience on the use of the service. This estimation was done using the random forest method and the attained accuracy was around 97%. The trees had their complexity pruning parameter evaluated through cross-validation. Results suggest that insecurity in the campi is a significant factor towards users' adhesion to the service; however, they are not satisfied with the security provided by it. Furthermore, global satisfaction of the users with the transport is more sensible to punctuality, frequency and information availability about the service. The short time between classes, sometimes given in different campi, may justify such concern with time aspects of the service performance. Thus, studies to synchronize the buses' times on the stops and the users' needs, taking into account external disturbances such as the local traffic, are suggested.

Keywords: service quality, public transportation, random forest, customer perception

Introdução

Serviços públicos de transporte de passageiros têm sua sobrevivência associada à adesão dos usuários ao modal (DE OÑA et al., 2013). Frequentemente tais serviços não são oferecidos diretamente pelas autoridades locais e sim, através de entes externos em condições de quase monopólio. Assim, a prioridade imediata destes operadores é a satisfação das necessidades da autoridade supervisora do serviço (MOUWEN, 2015). Os interesses dos passageiros e da autoridade podem se alinhar por pressão pública dos usuários ou por iniciativa orgânica da autoridade.

A fidelização dos passageiros alivia o tráfego local (GUIRAO; GARCÍA-PASTOR; LÓPEZ-LAMBAS, 2016), diminui a competição por espaços em estacionamentos e tem bom efeito positivo no desempenho econômico dos operadores a longo prazo (LAI; CHEN, 2011).

A assiduidade dos usuários ao modal público é diretamente influenciada pela satisfação que estes têm ao utilizar o serviço (LAI; CHEN, 2011). Tal satisfação pode ser compreendida como a diferença entre a qualidade esperada pelo passageiro antes do uso do serviço e a qualidade percebida durante o uso (FREITAS, 2013). Assim, ainda que o monopólio seja



comum, deslocamentos a pé, em veículos privados ou outros modais públicos são usuais rivais da preferência do usuário. Portanto, mensurar a satisfação do usuário é uma atividade fundamental para a gestão de transporte de passageiros.

Entretanto, qualidade de serviços e satisfação dos clientes são conceitos subjetivos, portanto sua medição não é uma tarefa trivial. Três características dos serviços contribuem para a dificuldade do exercício: a) intangibilidade – a qualidade de um bem pode ser observada objetivamente em características como dureza, cor e flexibilidade. A qualidade de um serviço diz respeito à performance com que o mesmo é fornecido. b) heterogeneidade – a consistência de desempenho em serviços é rara. As práticas variam em função dos fornecedores, mão-de-obra ou por distúrbios aleatórios. c) inseparabilidade – enquanto bens podem ser produzidos em um local e entregues para consumo do cliente em outro sítio, serviços são gerados e consumidos simultaneamente no mesmo lugar (PARASURAMAN; ZEITHAML; BERRY, 1985).

A prática comum na literatura é a proposição de questionários ao público. Em tais pesquisas os clientes expressam sua satisfação global e sobre atributos do serviço. Há duas vias principais de estimação da ordem de importância destes atributos: importância declarada e importância derivada. Na primeira, os usuários são convidados a ordenar as características do serviço por importância. É um método mais direto e intuitivo, mas que amplia o tamanho do questionário e, por consequência, o fardo sobre o entrevistado podendo diminuir a precisão e a frequência de resposta da pesquisa. A importância derivada por sua vez estima o valor relativo das categorias mensurando estatisticamente a intensidade do elo entre cada atributo com a medida global de satisfação. Para esta tarefa, o indicador mais comumente utilizado é o *Customer Satisfaction Index* (CSI) devido à sua simplicidade. Usualmente as qualidades esperada e percebida destes atributos são medidas através de uma escala Likert. (GUIRAO; GARCÍA-PASTOR; LÓPEZ-LAMBAS, 2016).

Uma vez que a qualidade é uma construção social (PARASURAMAN; ZEITHAML; BERRY, 1985), a satisfação dos usuários e a importância dada a aspectos do serviço é função de características socioeconômicas como gênero, idade, ocupação, renda, nível educacional e outros (GUIRAO; GARCÍA-PASTOR; LÓPEZ-LAMBAS, 2016). Estes aspectos da experiência do serviço também variam com o tempo, tendo em vista que os usuários se adaptam às novas condições e consolidam seu conhecimento (ABENOZA; CATS; SUSILO, 2017). A percepção dos diferentes segmentos de mercado dentro da base de clientes permite uma gestão racional das demandas de diferentes usuários sem escalonamento desnecessário da complexidade do problema (ABENOZA; CATS; SUSILO, 2017). Técnicas e indicadores



complexos utilizados por pesquisadores podem se impor como barreiras para os gestores (GUIRAO; GARCÍA-PASTOR; LÓPEZ-LAMBAS, 2016).

O ambiente acadêmico é um interessante laboratório de experimentação dessas análises. A academia é um espaço privilegiado para a promoção da educação e sensibilização acerca da mobilidade sustentável em seu senso mais amplo (aspectos econômicos, ambientais e sociais). Além disso, o crescimento da oferta de cursos nas universidades federais do Brasil gerou uma forte demanda de viagens nesses ambientes. Neste cenário, Pires (2013) relata um crescente interesse na análise dos sistemas de transporte *intercampi*.

Esta constatação é marcante no cotidiano da comunidade acadêmica da Universidade Federal da Bahia (UFBA), no município de Salvador. Uma das estratégias para lidar com este problema foi a implementação de um sistema de transporte *intercampi* conhecido como Buzufba, licitado pela administração universitária em 2012. O Buzufba consiste em um serviço de transporte por ônibus que realiza percursos entre os campi espacialmente dispersos da UFBA. Inicialmente, a operação era realizada por 4 micro-ônibus que percorriam 3 linhas com roteiros distintos (denominadas B1, B2 e B3). Com o passar do tempo, o sistema se expandiu e atualmente opera com 6 veículos que realizam as rotas B1, B2, B3, B4, B5 e Expresso. Apesar da melhoria na mobilidade promovida pelo Buzufba, expressões de descontentamento com a qualidade do serviço ofertado são recorrentes, o que sugere que a satisfação do usuário com o sistema de transporte deva ser melhor investigada.

Objetivo

Identificar as variáveis mais influentes na satisfação geral dos usuários com o serviço do Buzufba.

Material e Método

Para melhor entendimento da satisfação da comunidade acadêmica da UFBA com o sistema de transporte Buzufba, uma pesquisa de satisfação foi proposta no ano de 2018. Por meio de um questionário online aplicado via plataforma *Google Forms*, cerca de 1600 membros da comunidade acadêmica responderam perguntas relativas ao seu perfil sociodemográfico, às características dos deslocamentos realizados e qual importância e grau de satisfação conferem a certos atributos da qualidade no serviço do Buzufba. A Tabela 1 agrega o conjunto de variáveis coletados na pesquisa.



Tabela 1 – Variáveis coletadas na pesquisa.

Tipo de variável	Nome da variável
Tipificação demográfica	Idade
	Gênero
	Renda familiar
	Restrição de mobilidade
	Vínculo com a instituição (funcionário ou estudante)
	Curso (somente estudantes)
	Semestre (somente estudantes)
	PROAE (somente graduandos, recebe assistência estudantil) Órgão/unidade (somente funcionário)
Características de Mobilidade	Frequência de uso do sistema
	Origem da viagem (opção de inserir três deslocamentos)
	Destino da viagem
	Frequência da viagem Motivo da viagem (MOTIVO)
Atributos de qualidade	Satisfação geral com o serviço (S_SATS)
	Pontualidade (S_PONT e I_PONT)
	Frequência (S_FREQ e I_FREQ)
	Acessibilidade a pé (S_ACES e I_ACES)
	Conforto na viagem (S_CONF e I_CONF)
	Conforto na espera (S_ESPR e I_ESPR)
	Acessibilidade à informação (S_INFO e I_INFO)
	Tratamento dos funcionários (S_TRAT e I_TRAT) Condução do veículo (S_TRAN e I_TRAN) Segurança pessoal e patrimonial (S_VIOL e I_VIOL)

Fonte: Autores, 2019

Aos dados da pesquisa foram adicionadas variáveis relativas ao deslocamento mais frequente dos usuários. A Figura 1 ilustra a hipótese de como se dá uma viagem típica de um usuário. A viagem se dá entre os locais listados no formulário (origem e destino). Os passageiros partem da origem declarada e se dirigem a pé ao ponto do Buzufba mais próximo. Eles adentram o micro-ônibus cuja rota seja a mais direta para seu destino (trajeto mais curto em rodovia). Os passageiros se deslocam dentro do micro-ônibus até a parada mais próxima do destino desejado e mais uma vez se deslocam a pé da parada até o local declarado. Nota-se também que alternativamente os usuários podem ignorar o transporte e deslocar-se somente a pé.

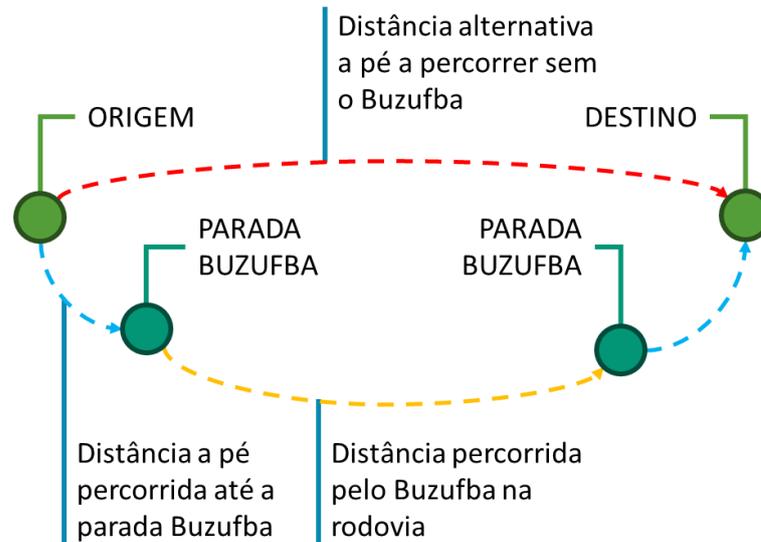


Figura 1 – Esquema hipotético dos deslocamentos dos usuários

Fonte: Autores, 2019

Estabelecido os mecanismos de deslocamento hipotéticos, os locais de trânsito dos usuários e os pontos de parada do Buzufba tiveram suas coordenadas coletadas em pesquisa manual no site do *Bing Maps*. Cada um dos 84 pontos de trânsito foi vinculado à mais próxima das 19 paradas. Em seguida, as distâncias percorridas pelo Buzufba entre cada parada na rodovia foram calculadas para as seis rotas, inclusive fazendo distinção entre as viagens de ida e volta. As distâncias foram estimadas utilizando o *Bing Maps API* para um veículo particular deslocando-se entre as coordenadas das paradas. A conexão ao API foi feita através de uma chave de usuário (versão gratuita) e mediada através do pacote **taRifx.geo** (FRIEDMAN, 2014). Logo após, para cada passageiro foi tomada origem e destino declarados e obtidas as paradas mais próximas. Dentre as doze rotas (seis pares de ida e volta) é selecionada a mais direta entre as paradas, empates entre rotas são decididos aleatoriamente.

As distâncias a pé e as distâncias usadas na alocação de locais de trânsito às paradas foram estimadas com auxílio da função **distGeo** do pacote **geosphere** disponível para o R (HIJMANS, 2017). Esta função calcula a distância mais curta entre dois pontos numa superfície elipsoidal. Os parâmetros de raio e achatamento da elipsoide utilizados foram o padrão, elipsoide WGS84, uma vez que este par é a melhor opção para condições genéricas.

Por fim, todas as distâncias foram estimadas em quilômetros, a saber: o trecho da viagem feito a pé (T_{pe}), o trecho da viagem por Buzufba (T_{buz}) e a distância puramente a pé entre os locais de viagem (d_{pe}). Foram também geradas duas variáveis adimensionais que avaliam a importância dos deslocamentos a pé. A fração da viagem percorrida a pé (f_{pe}) é dada pela Equação 1:



$$f_{pe} = \frac{T_{pe}}{T_{pe} + T_{buz}} \quad (1)$$

Também foi obtida a razão (r) entre a distância da viagem somente a pé e a viagem típica dos usuários conforme a Equação 2:

$$r = \frac{d_{pe}}{T_{pe} + T_{buz}} \quad (2)$$

Random Forest

Uma vez em posse de dados relativos ao deslocamento mais frequente dos usuários e das medidas de qualidade dos atributos foi utilizado o método da *Random Forest* para divisar quais variáveis têm maior poder explanatório sobre a satisfação geral dos passageiros com o serviço. A *Random Forest* é uma abordagem para modelagem preditiva que têm como base a construção de uma infinidade de árvores de decisões, completas, e a combinação destas em um único modelo misto, chamado de floresta.

O algoritmo da *Random Forest* considera, a cada divisão da árvore de decisão, um subconjunto das variáveis, ao invés da totalidade destas, utilizando "*bootstrap aggregation*". Desta forma, aumenta-se a variabilidade das árvores de decisão criadas, evitando a construção de árvores altamente correlacionadas.

Estas árvores são então criadas para seu tamanho máximo, sem poda, utilizando o algoritmo CART (BREIMAN et al., 1984), neste para cada divisão binária são escolhidas, aleatoriamente, n variáveis das M totais. Os melhores preditores são escolhidos utilizando o Índice Gini, calculado para cada nó da árvore.

Para realizar a poda das árvores criadas, com o objetivo de controlar possíveis efeitos causados por *overfitting*, é utilizado o critério de "*cost complexity pruning*". Abaixo está representado este processo de poda que se trata da minimização de uma função custo que têm como objetivo balancear o erro de predição e o número de "galhos" (Equação 3):

$$\text{mínimo}_T \sum_{m=1}^{|T|} \sum_{x_i \in R_m} (y_i - \bar{y})^2 + \alpha |T| \quad (3)$$

Onde α é determinado através da validação cruzada.

Por fim, a classificação é dada através da moda dos votos, resultado da predição, de cada uma das árvores de classificação.



Definidas as características do algoritmo *Random Forest*, é importante pontuar que se faz necessária uma função distinta para o índice Gini, uma vez que tratamos de dados ordinais. Classicamente, esta função assume a seguinte forma, dado um nó t (BRIEMAN et al., 1984):

$$i_{GG}(t) = \sum_{k=1}^J \sum_{l=1}^J C(\omega_k/\omega_l) p(\omega_k/t) C(\omega_l/t) \quad (4)$$

Supondo um conjunto s de scores que representam a variável Y podemos considerar o erro de classificação como uma transformação desejada para a diferença absoluta entre os pares de resultados, obtendo assim duas equações distintas para o $C(\omega_k/\omega_l)$ (GALIMBERTI; SOFFRITTI; MASO, 2015):

- a) Formulação Linear: $C(\omega_k/\omega_l) = |s_k - s_l|$
- b) Formulação Quadrática: $C(\omega_k/\omega_l) = (s_k - s_l)^2$

A modelagem foi realizada utilizando o pacote **caret** (*Classification And REgression Training*), biblioteca disponível em R que contém funções para: divisão dos dados, pré-processamento, seleção de modelos, validação cruzada, estimação de parâmetros e de nível de importância das variáveis (KUHN, 2019).

Segundo o autor supracitado, este pacote tem como finalidade prover uma interface uniforme que possua potencialidade de realizar diferentes modelos e processos comuns a esta modelagem como estimação e sintonia de parâmetros, validação cruzada e determinação da importância das variáveis.

Todas as análises realizadas no artigo foram realizadas no ambiente de programação R (R CORE TEAM, 2018).

Resultados e Discussão

Análise descritiva das variáveis de distância

A seguir, será apresentada uma breve discussão acerca das distâncias percorridas pelo Buzufba ou pelos alunos, buscando traçar um perfil do serviço quanto a estas variáveis. A Figura 2 condensa os resultados desta análise inicial.

Quanto ao trecho que seria percorrido caso o usuário tomasse o caminho a pé para um dos pontos de parada do veículo, têm-se como a moda cerca 200 metros. A maioria dos trechos percorrido pelos alunos para chegar a um dos pontos de parada para pegar um dos



Buzufba são curtos, não se destacando como um potencial problema para a maioria dos alunos. Quase 80% dos deslocamentos a pé feitos nos EUA com propósitos educacionais têm no mínimo 200m (YANG; DIEZ-ROUX, 2012).

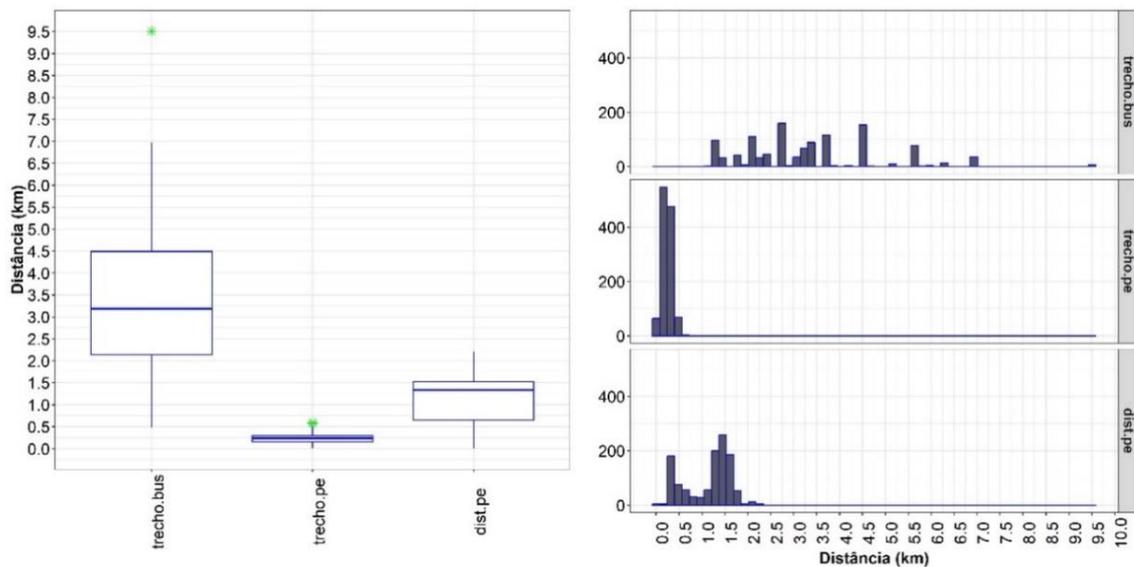


Figura 2 – Análise da distribuição das distâncias

Fonte: Autores, 2019

Quanto à distância a pé que seria percorrida, caso não fosse utilizado o Buzufba, observa-se dois padrões. O primeiro padrão versa sobre pequenas caminhadas, inferiores a 1 km de distância, provavelmente, ligadas a questão do desnível que existe entre alguns dos *campi*, outra questão importante também é a violência urbana que torna este sistema de transporte atrativo mesmo que a distância a ser percorrida seja curta. Quase 70% dos entrevistados em pesquisa de 2010 na UFPA dizem se sentir inseguros no campus, mesmo sob vigilância de câmeras (LEMOS et al., 2011).

Quanto à distância percorrida pelos veículos do Buzufba, têm-se em sua maioria pequenas distâncias, cerca de 3,5 km. Porém, deve-se pontuar que a maioria dos veículos trafega pela Avenida Garibaldi cujo o fluxo é intenso durante os horários de pico.

Por fim, é interessante observar, que se houvesse transportes internos capazes de vencer a questão da diferença de cota entre os *campi*, poderia ser pensado em sistemas mais inteligentes com rotas menores, tempos de percursos mais rápidos, causando maior sensação de segurança aos usuários e, provavelmente, melhor avaliação por parte dos usuários quanto a qualidade do serviço.



Análise descritiva das variáveis do questionário

O questionário agrupa as questões, quanto às características estudadas do serviço prestado, em relação a importância dada pelos usuários e pela sua satisfação. A Figura 3 apresenta os resultados quanto à importância dada aos atributos de qualidade.

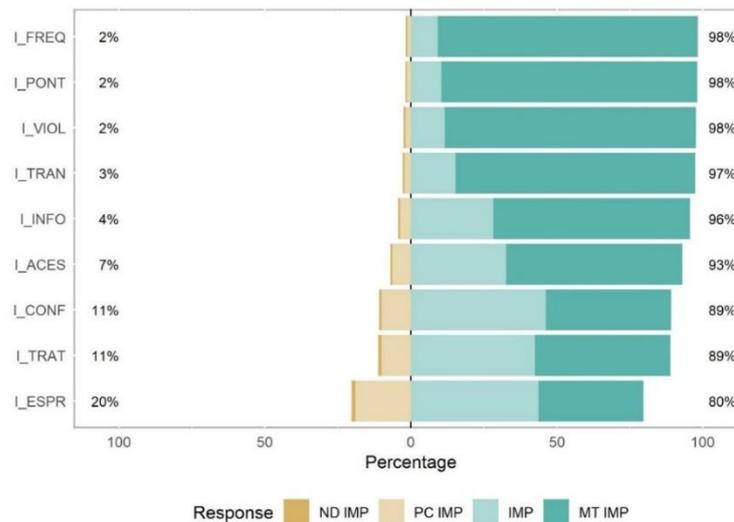


Figura 3 – Partição das respostas referentes à importância dos atributos de qualidade para os usuários do Buzufba

Fonte: Autores, 2019

Quanto a importância, exceto para a variável referente ao conforto existente nos pontos de espera, pelo menos 85% dos respondentes afirmaram que o atributo em questão era muito importante. Destaca-se, também, a importância dada a frequência dos veículos, a sua pontualidade e o sentimento de segurança pessoal que este promove, fato observado ao se perceber que 98% dos usuários apontam estas variáveis como muito importante. A análise da importância também aponta para a necessidade de se transportar de forma rápida e segura entre diferentes *campis* ainda que a distância total entre estes seja curta se realizada a pé.

Quanto à satisfação, têm-se um panorama de respostas mais diverso. A análise, a seguir, tem como foco as respostas negativas dadas acerca da satisfação com o serviço prestado. A Figura 4 apresenta os resultados quanto à qualidade percebida no serviço prestado.

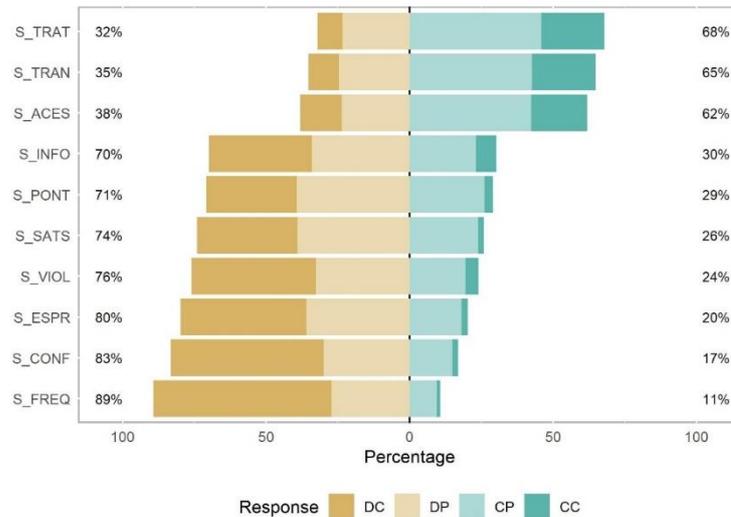


Figura 4 – Partição das respostas referente à qualidade percebida nos atributos de qualidade para os usuários do Buzufba

Fonte: Autores, 2019

Destaca-se, a avaliação dada quanto a falta de satisfação referente a frequência dos veículos, a confiabilidade do serviço e a qualidade dos pontos de espera, mesmo que o último seja avaliado como pouco importante por parte dos respondentes. É perceptível, desde já, que existe uma discrepância entre a importância dada a frequência e aquela que é oferecida pelo sistema, fato que deve influenciar de forma interessante a avaliação do sistema.

Enfim, foi estimada a correlação entre as variáveis de qualidade e a motivação para a viagem mais frequente usando a correlação de postos de Spearman que indica se há relação monotônica entre as variáveis. A Figura 5 apresenta a intensidade das correlações.

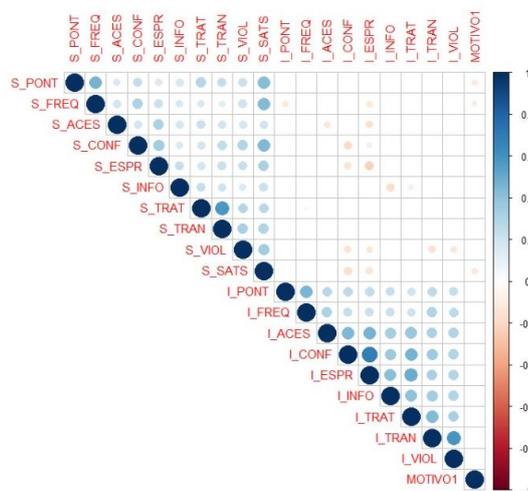


Figura 5 – Matriz de correlação entre as variáveis

Fonte: Autores, 2019



Notam-se dois grupos na Figura 5: O da importância e da satisfação com os atributos de qualidade. Dentre as importâncias, as mais correlatas são os pares conforto na espera e conforto na viagem (por sua natureza semelhante) e a condução do veículo e a sensação de segurança (uma vez que uma condução não suave incita temores nos usuários). Entre as variáveis de satisfação nota-se que a satisfação global se correlaciona bem com a satisfação com pontualidade, frequência e conforto na viagem.

Análise do modelo gerado

Uma vez apresentadas as análises descritivas, utilizou-se o modelo de *Random Forest* a fim de inferir acerca das variáveis mais importantes para definição da qualidade do serviço do Buzufba por parte dos usuários que responderam ao questionário em Anexo.

A definição destas variáveis mais importantes é importante para sugerir novos caminhos que esta pesquisa pode seguir, a exemplo do monitoramento da pontualidade dos ônibus, análise do conforto dos veículos, modelagem da distribuição das rotas e dos pontos de acesso e aprimoramento dos funcionários deste serviço.

A *Random Forest* ótima apresentou os parâmetros exibidos na Tabela 2:

Tabela 2 – Parâmetros ótimos da *Random Forest*.

Parâmetro de complexidade	Método para definição da função custo	Método de poda
0,0559	Método Quadrático	Número total de erros de classificação

Fonte: Autores, 2019

Na Tabela 3, tem-se a precisão do modelo referida aos dados de teste:

Tabela 3 – Resultados de precisão do teste.

		Taxa (%)			
		1	2	3	4
Predito	1	122	0	0	0
	2	0	127	0	0
	3	0	0	85	11
	4	0	0	0	0

Fonte: Autores, 2019

Obteve-se precisão em torno de 97%, deve-se, porém, ressaltar que não foi possível identificar um padrão claro que implicasse numa resposta que indicasse alta qualidade do serviço (nível 4). Primeiramente, há um pequeno número de respondentes que classificaram



o serviço desta forma, além disto, ressalta-se que a qualidade exigida pelos usuários varia bastante frente a variabilidade das suas características.

A Figura 6 apresenta a relevância das variáveis na geração do modelo. Destacam-se aquelas que estão associadas à presente satisfação dos usuários, em detrimento de características geográficas como o percurso total ou a importância que estes dão ao sistema.

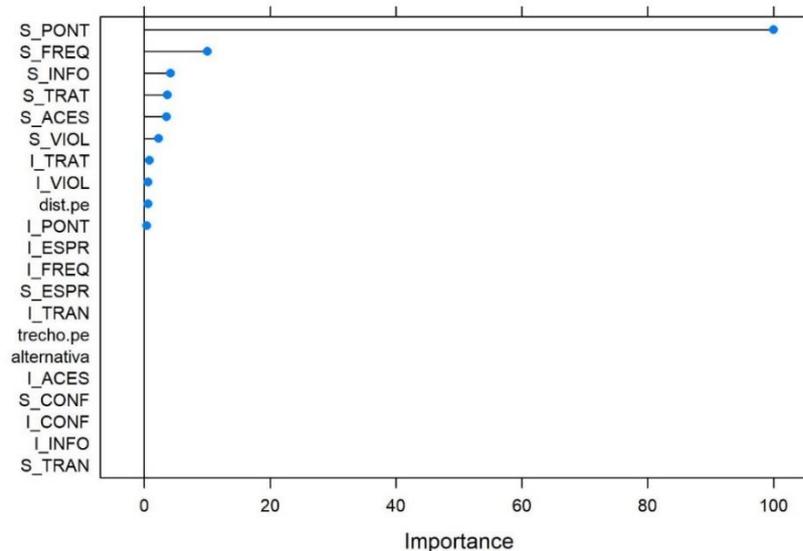


Figura 6 – Ordem de importância das variáveis na geração do modelo

Fonte: Autores, 2019

É interessante comparar estes resultados aos obtidos por Guirão et al (2016). Ao pesquisarem as variáveis mais importantes para trabalhadores que usam transporte público, obtiveram as mesmas encontradas para o modelo presente neste artigo.

Muitos alunos precisam se locomover entre os dispersos *campi* da UFBA em curto espaço de tempo para estar presente nas aulas, por isto a pontualidade e a frequência dos veículos se destacam de forma muito clara frente as outras variáveis. Desta forma, pode propor-se um estudo que catalogue os horários de chegada dos veículos a cada um dos pontos, correlacionando-os com a necessidade dos alunos de se deslocarem de um *campus* a outro, levando em consideração adequação das rotas e as características do trânsito local.

Em segunda instância, aparecem variáveis ligadas a gestão do serviço, nestas estão inclusas a disponibilidade de informação acerca do serviço, o tratamento que é dado pelos funcionários e a boa disponibilidade dos pontos de paradas dos veículos.

Por fim, dentre as métricas de distância a único que obteve destaque foi a distância total a pé percorrida caso não houvesse o Buzufba. Tal resultado era esperado, porém, deve-se



colocar que para as maiores distâncias, provavelmente, a ausência do serviço levaria o usuário a optar pelo uso da frota de ônibus comum.

Conclusão

A partir da exploração dos dados do questionário sobre a satisfação quanto ao Buzufba, percebe-se, primeiramente, um bom distanciamento entre a satisfação dos usuários e a importância que eles dão a cada uma das variáveis referentes a qualidade deste serviço. Fato este que justifica a necessidade de uma pesquisa que busque compreender quais as variáveis são importantes para gerar esta insatisfação e propor estudos mais profundos para que se encontre soluções para os problemas latentes neste serviço.

As variáveis que mais influenciam na avaliação da qualidade do serviço são a pontualidade e a frequência. Isto pode indicar a necessidade de se repensar a estratégia de oferta dos veículos, realizando, inclusive, pesquisas *in loco* a fim de compreender tal questão.

Por fim, o uso de um ambiente de programação como o R foi fundamental para o acesso a uma API (*Bing Maps*) capaz de estimar o tamanho dos trajetos e as distâncias consideradas relevantes para este estudo, também, o seu uso foi essencial na geração gráficos que garantissem boa visualização e para criação modelos confiáveis a baixo custo computacional.

Referências

ABENOZA, R. F.; CATS, O.; SUSILO, Y. O. Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. **Transportation Research Part A: Policy and Practice**, v. 95, p. 64–84, 2017.

BREIMAN L, FRIEDMAN RA, OLSHEN RA, AND STONE CG. **Classification and Regression Trees**. Pacific Grove, CA: Wadsworth, 1984

DE OÑA, J. et al. Perceived service quality in bus transit service: A structural equation approach. **Transport Policy**, v. 29, p. 219–226, 2013.

FREITAS, A. L. P. Assessing the quality of intercity road transportation of passengers: An exploratory study in Brazil. **Transportation Research Part A: Policy and Practice**, v. 49, n. 2013, p. 379–392, 2013.

FRIEDMAN, A. B. **taRifx.geo: Collection of various spatial functions**. R package version 1.0.7, 2014. Disponível em: <<https://github.com/gsk3/taRifx.geo>>

GALIMBERTI, G.; SOFFRITTI, G.; MASO, M. DI. Classification Trees for Ordinal Responses in R: The rpartScore Package. **Journal of Statistical Software**, v. 47, n. 10, 2015.

GUIRAO, B.; GARCÍA-PASTOR, A.; LÓPEZ-LAMBAS, M. E. The importance of service



quality attributes in public transportation: Narrowing the gap between scientific research and practitioners' needs. **Transport Policy**, v. 49, p. 68–77, 2016.

HIJMANS, R. J. **geosphere: Spherical Trigonometry**. R package version 1.5-7, 2017. Disponível em: <<https://CRAN.R-project.org/package=geosphere>>

KUHN, M. **caret: Classification and Regression Training**. R package version 6.0-81, 2018. Disponível em: <<https://cran.r-project.org/package=caret>>

LAI, W. T.; CHEN, C. F. Behavioral intentions of public transit passengers-The roles of service quality, perceived value, satisfaction and involvement. **Transport Policy**, v. 18, n. 2, p. 318–325, 2011.

LEMOS, A. et al. Câmeras de vigilância e cultura da insegurança: percepções sobre as câmeras de vigilância da UFBA. **Revista Alceu - PUC-RJ**, p. 143–153, 2011.

MOUWEN, A. Drivers of customer satisfaction with public transport services. **Transportation Research Part A: Policy and Practice**, v. 78, p. 1–20, 2015.

PARASURAMAN, A.; ZEITHAML, V. A.; BERRY, L. L. A Conceptual Model Service Its Quality and Implications for Future Research. **Research Paper**, v. 49, n. 4, p. 41–50, 1985.

PIRES, L. S. **Mobilidade Sustentável em Campi Universitários: um estudo de caso na Universidade Federal Rural do Rio de Janeiro – Campus Seropédica**. [s.l.] UFRJ, 2013.

R CORE TEAM. **R: A Language and Environment for Statistical Computing** Vienna, Austria R Foundation for Statistical Computing, , 2018.

WICKHAM, H. **tidyverse: Easily Install and Load the 'Tidyverse'**. R package version 1.2.1, 2017. Disponível em: <<https://CRAN.R-project.org/package=tidyverse>>

YANG, Y.; DIEZ-ROUX, A. V. Walking distance by trip purpose and population subgroups. **American Journal of Preventive Medicine**, v. 43, n. 1, p. 11–19, 2012.

Anexo

O acesso ao script de comandos do R, bem como ao questionário aplicado podem ser encontrados no link a seguir: <<https://github.com/tarssioesa/Buzufba>>