

RELAÇÃO ENTRE A OFERTA DO MODO ÔNIBUS E A ESCOLHA MODAL PARA ACESSO A AEROPORTOS: UMA ABORDAGEM ATRAVÉS DE *DATA MINING*

M. M. Thebit, A. L. B. N. Cunha, C. S. Pitombo

RESUMO

O presente artigo busca analisar a relação entre a oferta de ônibus e a escolha do modo de transporte no acesso aos principais aeroportos do estado de São Paulo (Brasil) por meio da técnica de *data mining*, conhecida como Árvore de Decisão (AD). O banco de dados utilizado compõe a Caracterização e Dimensionamento da Matriz Origem-Destino do Transporte Aéreo no Brasil. O resultado obtido com a utilização do algoritmo CHAID mostrou que a oferta de ônibus por parte da companhia aérea apresenta uma contribuição positiva no aumento da escolha desse modo, em viagens de menor custo e curta duração e de custo e duração intermediários. A oferta de ônibus urbano apresentou influência assertiva com maior relevância em viagens de custo intermediário e duração longa. Em contrapartida, observou-se que a oferta de ônibus interurbanos e interestaduais não apresentou grande influência na escolha pelo modo ônibus.

1. INTRODUÇÃO

O acesso terrestre a aeroportos é importante objeto de estudo no cenário atual, uma vez que esses são grandes polos geradores de viagens e impactam a circulação do entorno. No caso específico do Brasil, tais lugares tiveram seus impactos potencializados nos últimos anos, devido ao aumento significativo do número de usuários. Segundo o Anuário do Transporte Aéreo, publicado pela ANAC em 2013, no período entre 2003 e 2012, ocorreu um crescimento médio anual de 14,35% ao ano, em termos de passageiros-quilômetros pagos transportados (RPK), pela aviação doméstica.

Nesse contexto, faz-se necessário estudar a característica do acesso a tais terminais e possíveis providências para promover medidas políticas que incentivem a escolha, por parte dos usuários, de modos alternativos ao transporte individual particular motorizado ou semipúblico (carro próprio, carona e táxi), os quais têm sido os mais utilizados nas cidades brasileiras e no exterior.

Atualmente, são poucos os aeroportos nacionais que apresentam opções de modos de acesso, além dos rodoviários, como, por exemplo, metrô e trens urbanos, e, apenas alguns oferecem acessibilidade por meio de serviços de transporte público, como ônibus urbanos e interurbanos. Também são poucos explorados os ônibus de companhias aéreas que, em grande parte, são utilizados apenas para conexão de passageiros entre dois aeroportos internos a um mesmo centro urbano e não para o acesso ao aeroporto inicial.

O presente estudo busca avaliar a relação entre a oferta do modo de transporte ônibus, seja ele urbano, interurbano ou de companhias aéreas e a escolha desse modo no acesso dos principais aeroportos do estado de São Paulo.

Este artigo é formado por cinco seções, além desta introdutória. A seção 2 apresenta a contextualização do tema a partir de uma síntese da revisão da literatura. A seção 3 descreve os materiais e método. Os resultados são apresentados na seção 4. Finalmente, as principais conclusões e recomendações para trabalhos futuros são descritos na seção 5. A seção 6 traz as referências bibliográficas citadas no texto.

2. CONTEXTUALIZAÇÃO

A literatura engloba um extenso número de estudos voltados para as questões relacionadas ao acesso terrestre a aeroportos. Alguns desses buscam relacionar tempo de viagem, custo e características socioeconômicas dos usuários, com suas escolhas por determinado modo de transporte, por meio da análise de dados obtidos pela aplicação de questionários dentro dos aeroportos em estudo (Harvey, 1986; Alhusseim, 2011; Gulsah Akar, 2013)

Harvey (1986) apresenta um estudo da caracterização da escolha pelo modo de acesso a aeroportos para a área da baía de São Francisco. A técnica de análise utilizada foi o modelo *Logit multinomial*, o qual foi aplicado separadamente em dois grupos: viagens a trabalho e viagens por outros motivos. Observou-se que os usuários são altamente sensíveis ao tempo da viagem de acesso, o que tende a tornar os modos relacionados ao automóvel particular, como carro próprio, carona e táxi, mais atrativos e, por essa razão, dominantes. Foram apresentadas propostas com potencial para aumentar a competitividade do ônibus como: a utilização de tecnologias e infraestruturas que permitam maior velocidade operacional e consequente redução no tempo de viagem; aumento das tarifas nos estacionamentos dos aeroportos, o que torna o usuário que mora longe e pretende ficar um longo período fora da cidade, mais tendencioso a adotar o ônibus, uma vez que não terá facilidade em obter carona e o uso do carro próprio exigirá gastos elevados com estacionamento.

Alhusseim (2011) destaca o menor enfoque dado ao acesso a aeroportos por modos públicos, pela literatura da área e foca seu estudo em um conjunto composto pelos modos carro e táxi. Após a realização de uma pesquisa por meio de questionário junto aos usuários do aeroporto, obteve-se, em média, 300 respostas, a partir das quais foi desenvolvido um modelo *logit binomial* para avaliar a escolha modal do aeroporto Riyadh, localizado na Arábia Saudita. Os resultados mostraram que, para o cenário estudado, renda, bagagem, tempo da viagem de acesso e nacionalidade afetam significativamente a escolha modal.

Gulsah Akar (2013) busca avaliar as questões relacionadas à escolha do modo de acesso ao aeroporto, bem como a propensão de diferentes perfis de passageiros em aderir a modos alternativos. O questionário desenvolvido, neste caso, englobou, além das questões relacionadas às características da viagem e socioeconômicas dos usuários, uma avaliação da

existência de potencial demanda por modos alternativos e foi aplicado a 642 indivíduos no Aeroporto Internacional de Port Columbus em Ohio, Estados Unidos. A análise foi realizada por meio da aplicação de um modelo de escolha discreta (*logit binomial*). O estudo mostrou que, embora o principal modo de acesso ao aeroporto seja o automóvel particular, existe uma demanda para modos alternativos, com destaque para o metrô e o *shuttle*. Observou-se que o modo de transporte deve oferecer maior confiabilidade, menor tempo de viagem, maior conforto e flexibilidade de horário de partida, a qual pode ser alcançada pelo aumento da frequência, para que os usuários considerem substituir o automóvel por esse modo alternativo.

Embora a literatura ressalte a importância do desenvolvimento de estudos exclusivos de escolha modal para cada aeroporto devido às especificidades de localização, cultura, dimensões, dentre outras questões relacionadas à inserção dos aeroportos na cidade, o presente estudo exploratório, inicialmente, buscou um padrão regional em relação ao comportamento do usuário diante da existência da oferta do modo ônibus de acesso ao aeroporto.

Sua contribuição está no maior enfoque dado aos modos públicos coletivos, aqui representados pelo ônibus, e pelo fato de utilizar um extenso conjunto de dados, com mais de 10 mil observações, o que se mostrou pouco frequente, em estudos anteriores, devido à dificuldade de obtenção. Destaca-se também a abordagem proposta, baseada em algoritmo de Árvore de Decisão, técnica não paramétrica de *data mining*.

3. MATERIAIS E MÉTODO

3.1 Banco de dados

O presente estudo utiliza o banco de dados obtido por meio de questionário aplicado aos passageiros dos 32 principais aeroportos brasileiros no período entre 30/07/2009 e 12/09/2009. A pesquisa em questão compõe a Caracterização e Dimensionamento da Matriz Origem-Destino do Transporte Aéreo no Brasil, elaborado pela Fundação Instituto de Pesquisas Econômicas (FIPE) e publicado no Estudo do Setor de Transporte Aéreo do Brasil: Relatório Consolidado (McKinsey & Company, 2010).

Para este estudo, foram selecionadas apenas as observações em que o aeroporto de origem do passageiro estava localizado no estado de São Paulo. Com essa seleção, obtiveram-se 11 aeroportos localizados em diferentes regiões do estado, conforme exposto na Figura 1

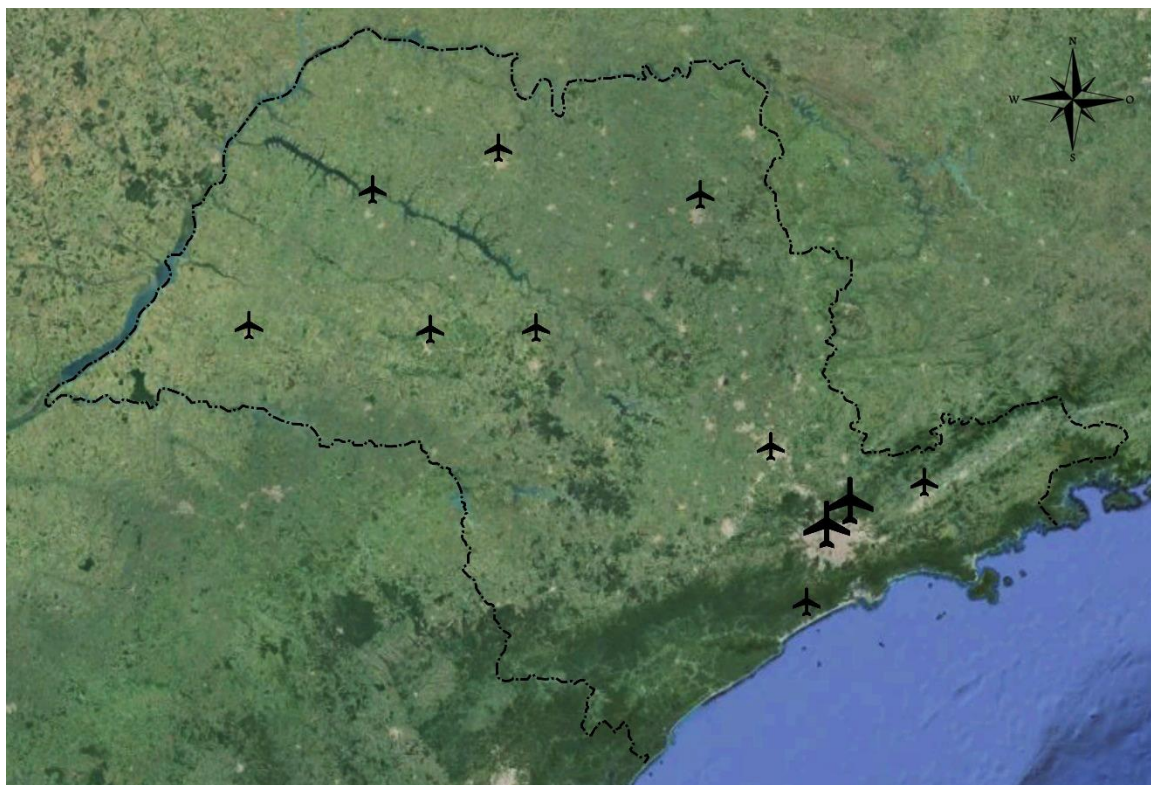


Fig. 1 Aeroportos estudados. Os ícones de maior tamanho correspondem aos aeroportos localizados na região metropolitana de São Paulo, no caso, os de Congonhas e Guarulhos.

A pesquisa original entrevistou um total de 44.558 pessoas, das quais 10.415 foram selecionadas para este estudo. Desse valor, 9.736 pessoas iniciaram sua rota em um dos dois aeroportos localizados na região metropolitana de São Paulo (Congonhas ou Guarulhos). A Tabela 1 apresenta o número de observações por aeroporto de origem.

Tabela 1 Observações por aeroporto

Município	Aeroporto	Número de observações
Araçatuba	Dário Guarita	7
Bauru	Moussa Nakahl Tobias	9
Campinas	Viracopos	476
Guarulhos	Gov. André Franco Montoro	5277
Itanhaém	Dr. Antônio Ribeiro Nogueira	1
Marília	Frank Miloye Milenkovich	8
Presidente Prudente	Adhemar de Barros	8
Ribeirão Preto	Leite Lopes	93
São José dos Campos	Prof. Urbano Ernesto Stumpf	2
São José do Rio Preto	Prof. Eriberto Manoel Reino	75
São Paulo	Congonhas	4459
Total		10415

Originalmente, o questionário aplicado pela FIPE permitia ao entrevistado apresentar um conjunto de modos de transporte utilizados para acessar o aeroporto de partida.

Considerando o objetivo do estudo, optou-se por avaliar apenas o último modo de transporte utilizado no acesso.

As 13 categorias apresentadas pela pesquisa inicial foram agrupadas da seguinte maneira:

0. Demais modos de transportes
1. Modos de transportes públicos individuais (semipúblicos)
 - Táxi
 - Carro alugado
2. Modos de transportes particulares individuais
 - Carona
 - Carro próprio
 - Carro da empresa ou evento
 - Motocicleta
3. Principais modos de transportes públicos coletivos
 - Ônibus urbano
 - Ônibus interurbano e interestadual
 - Ônibus da companhia aérea

Das 45 questões propostas aos entrevistados, foram selecionadas 17, das quais, 6 correspondiam a características socioeconômicas do usuário e 11 a características da viagem.

Variáveis relacionadas à caracterização do entrevistado:

- Nacionalidade brasileira
- Residência no Brasil
- Gênero
- Idade
- Número de dependentes
- Renda familiar

Variáveis relacionadas às características da viagem:

- Distância do local de origem ao aeroporto de início;
- Tempo de acesso ao aeroporto;
- Preço do deslocamento de acesso;
- Antecedência de compra da passagem;
- Valor da passagem aérea;
- Motivo da escolha do aeroporto;
- Tipo de local de partida;
- Horário de chegada ao aeroporto;
- Companhia aérea (cada companhia foi identificada como uma categoria);
- Motivo da viagem;
- Quem pagou pela passagem e
- Município de início do trecho.

Para complementar o banco de dados, foram elaboradas três novas variáveis referentes à oferta dos três tipos de ônibus identificados como possíveis modos de acesso ao aeroporto.

Três novas variáveis binárias elaboradas:

- Oferta de ônibus urbano;
- Oferta de ônibus interurbano e interestadual e
- Oferta de ônibus da companhia aérea.

Para a variável “oferta de ônibus interurbano”, foi atribuído como oferta existente apenas para os aeroportos de início do trecho, localizados na grande São Paulo: aeroportos de Congonhas e Guarulhos. Essa estratégia foi adotada por se tratar do único município, dentre os estudados, que, na data de realização da pesquisa em 2009, já apresentava um sistema de transporte público bem integrado com acesso facilitado entre os terminais rodoviários intermunicipais e interestaduais e os aeroportos.

Para a variável “oferta de ônibus urbano”, foi adotado o mesmo raciocínio da variável apresentada anteriormente, sendo atribuída a existência da oferta para o município de São Paulo. Além disso, foi considerada oferta existente também para os municípios em que, por meio do estudo do banco de dados, foi possível observar usuários que partiram da cidade em que estava localizado o aeroporto e acessaram o terminal utilizando apenas ônibus urbano.

Por fim, para a variável “oferta de ônibus da companhia aérea”, foi realizada uma pesquisa prévia, por meio de fontes secundárias, e identificadas as empresas que, na ocasião da pesquisa, já apresentavam esse tipo de serviço. Esse dado foi então contestado por meio da análise do banco de dados em que foram identificadas as empresas aéreas que ofereciam o serviço e qual o município de partida. Nesse caso, foram consideradas apenas as observações em que o usuário acessou o aeroporto, utilizando apenas esse modo de transporte. Por fim, foi atribuída como oferta existente de ônibus da companhia aérea a toda observação que apresentava o conjunto composto pela empresa e o município atendido, mesmo que esse não tenha sido o modo utilizado por aquele usuário.

3.2 Árvore de Decisão (AD)

O estudo utilizou a técnica de *data mining*, conhecida como Árvore de Decisão (AD). A escolha dessa técnica deve-se ao fato de ela ser não paramétrica, sem suposições matemáticas prévias ou restrições de tipos de variáveis explicativas ou dependentes.

A Árvore de Decisão permite classificar uma ampla base de dados em um número finito de categorias por meio de regras hierárquicas e da sua subdivisão em grupos, organizando os dados de maneira compacta, o que viabiliza uma visão geral da natureza do processo (Quinlan, 1983).

A hierarquia final obtida é denominada árvore e cada segmento é denominado nó. O segmento inicial corresponde ao nó raiz e apresenta o conjunto total dos dados analisados e que serão subdivididos em outros subnós (nós filhos) e em nós terminais ou folhas, quando não é mais possível subdividir aquele conjunto de dados com base em critérios pré-estabelecidos.

Os algoritmos utilizados para subdividir o conjunto de dados iniciais, nos modelos de árvore, selecionam as variáveis independentes que maximizam a segregação dos dados, segundo a variável dependente escolhida. Alguns algoritmos existentes são o CHAID (Kass, 1980), CART (Breiman *et al.*, 1984) e QUEST (Loh e Shih, 1997).

O algoritmo de AD realiza sucessivas subdivisões do banco de dados com o objetivo de tornar os subconjuntos resultantes cada vez mais homogêneos no que diz respeito à variável dependente. Tal subdivisão busca a redução da aleatoriedade na previsão da variável dependente por meio da minimização do desvio em relação a essa variável em cada divisão de nó da árvore. – (Breiman *et al.*, 1984). Neste artigo, é utilizado apenas o algoritmo CHAID (*CHi-squared Automatic Interaction Detection*) na avaliação da escolha do modo de transporte ônibus, no acesso aos aeroportos do estado de São Paulo, Brasil.

4. PROCEDIMENTO E RESULTADOS

O algoritmo CHAID foi executado a partir da parcela selecionada do banco de dados da FIPE, considerando 70% das respostas da pesquisa para treinamento e 30% para testes (validação) dos modelos. As variáveis admitidas na geração da AD foram somente as 20 apresentadas anteriormente, tendo como variável dependente o modo de transporte.

A Tabela 1 apresenta a porcentagem de acertos para a amostra de treinamento e teste do modelo elaborado, comparação entre observados e estimados. A porcentagem global de acertos, em torno de 77% em ambos os casos, pode ser considerada como satisfatória.

Tabela 1 - Resultado obtido através do algoritmo CHAID

MÉTODO CHAID						
Amostra	Observado	Previsto				Porcentagem correta
		0 Outros	1 Público individual	2 Particular individual	3 Público coletivo	
Treinamento	0 – Outros	63	39	225	41	17,1 %
	1 – Públi. individual	15	2940	372	90	86,0 %
	2 – Part. individual	22	303	2335	160	82,8 %
	3 – Públi. coletivo	5	86	327	355	45,9 %
	Porcentagem global	1,4 %	45,6 %	44,2 %	8,8 %	77,2 %
Teste	0 – Outros	25	22	76	15	18,1 %
	1 – Públi. individual	9	1299	136	43	87,4 %
	2 – Part. individual	11	147	901	89	78,5 %
	3 – Públi. coletivo	0	32	123	119	43,0 %
	Porcentagem global	1,5 %	49,2 %	40,6 %	8,7 %	76,9 %

Como proposto, foram obtidos os nós terminais ou folhas, que evidenciam as classes de modos de transporte, segundo características dos entrevistados e das viagens e suas respectivas probabilidades de escolha. Embora tenham sido inseridas variáveis de caracterização do usuário, a AD obtida apresentou em suas folhas apenas aquelas referentes à viagem. A única variável referente ao usuário, que foi apresentada no resultado final, foi “renda familiar”, a qual deu origem a apenas dois nós finais.

A Figura 2 apresenta uma visão esquemática da AD obtida. Os nós destacados por cinza escuro são aqueles em que foi observada a interferência positiva da oferta do modo ônibus no aumento da escolha por esse ou na redução do uso dos individuais. Os nós indicados em cinza claro correspondem àqueles em que a oferta de ônibus não acarretou uma alteração positiva, com relação à escolha por modos sustentáveis de deslocamento. Os nós indicados em branco estão relacionados a outras variáveis.

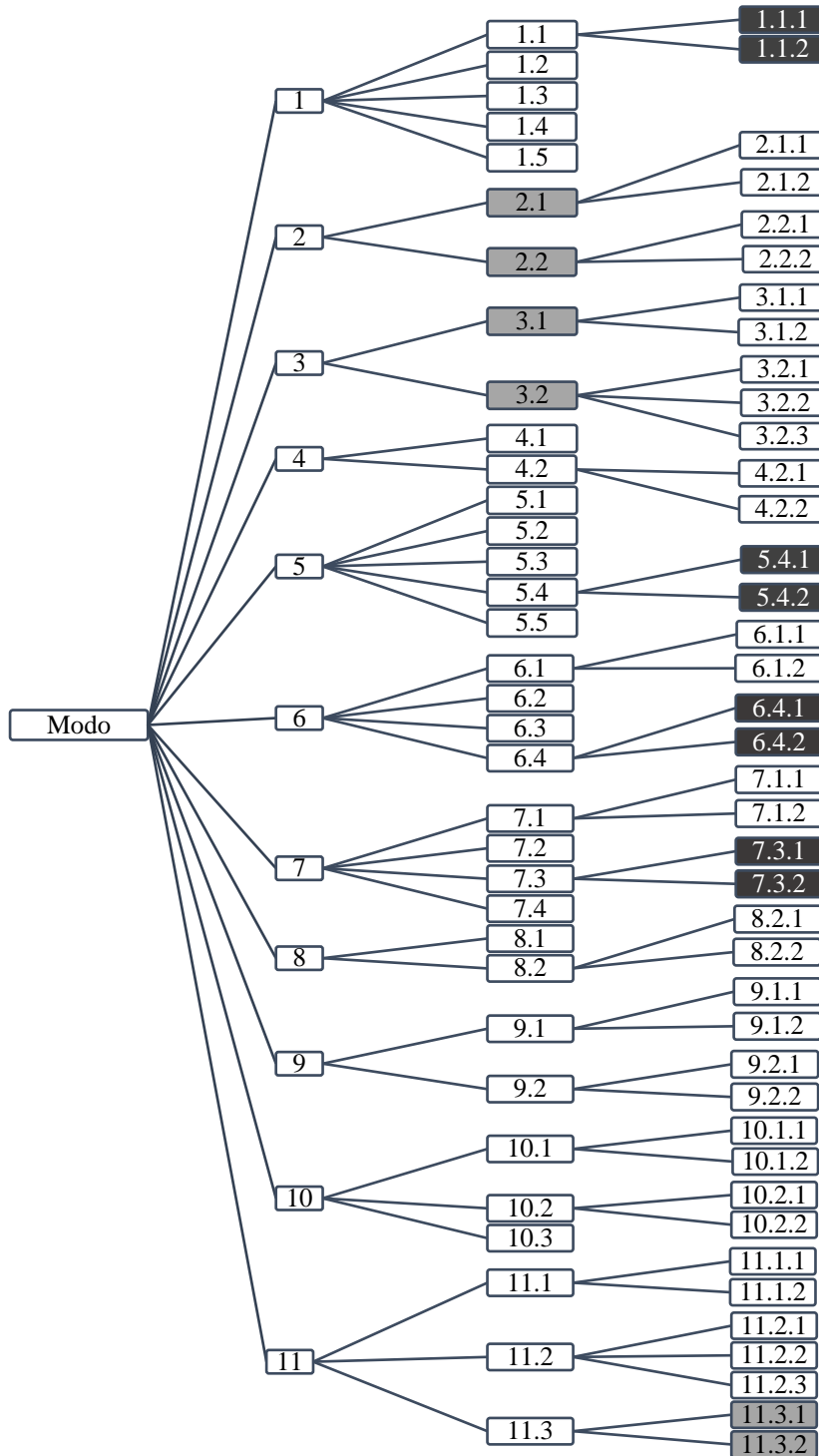


Figura 2 - Esquema da Árvore de Decisão obtida.

A seguir, serão expostos e detalhados três pares de nós finais em que a oferta do modo ônibus, aparentemente, influenciou, de forma positiva, na escolha do mesmo ou na redução dos modos individuais.

Os nós 1.1.1 e 1.1.2 são detalhados na Figura 3 e mostram que, para usuários, cuja viagem de acesso ao aeroporto tem o custo de até 8 reais e com tempo de viagem inferior a 15 minutos, a oferta do ônibus da companhia aérea não torna esse modo o principal a ser utilizado, mas contribui para o aumento de seu uso e para a consequente redução dos modos individuais públicos e particulares, sendo bastante significativa para o último caso.

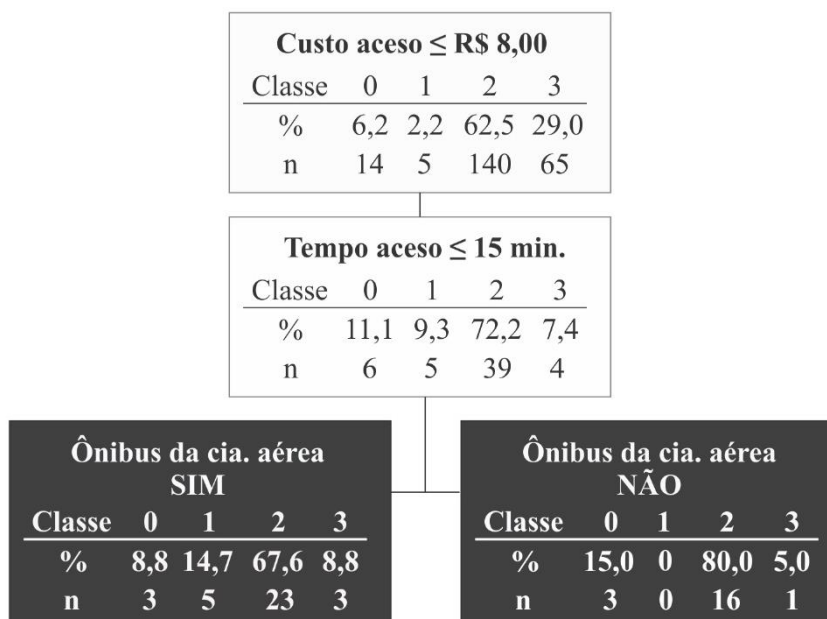


Figura 3 – Detalhe dos nós finais 1.1.1 e 1.1.2.

0 – Outros modos; 1 – Públicos individuais; 2 – particulares individuais;
3 – Públicos coletivos (ônibus).

Os nós 5.4.1 e 5.4.2 também apresentaram uma influência positiva da oferta de ônibus pela companhia aérea na escolha de modos de transporte mais sustentáveis, no caso, o ônibus. Para usuários que gastam entre 29 e 32 reais no acesso ao aeroporto e levam entre 40 e 85 minutos nesse deslocamento, a oferta de ônibus pela companhia aérea levou a uma redução no uso de veículos particulares individuais (classe 2 na Figura 4) e a um aumento no uso do modo ônibus (classe 3). Para essa parcela dos usuários, o ônibus chegou a ser um dos dois modos mais utilizados no acesso ao aeroporto.

Nos nós 6.4.1 e 6.4.2 (Figura 5) também foi observada uma redução no uso de modos individuais (classes 1 e 2) e aumento dos modos públicos coletivos (classe 3) devido à existência da oferta. Nesse caso, porém, a oferta não foi da companhia aérea como nos nós anteriores e sim dos ônibus urbanos.

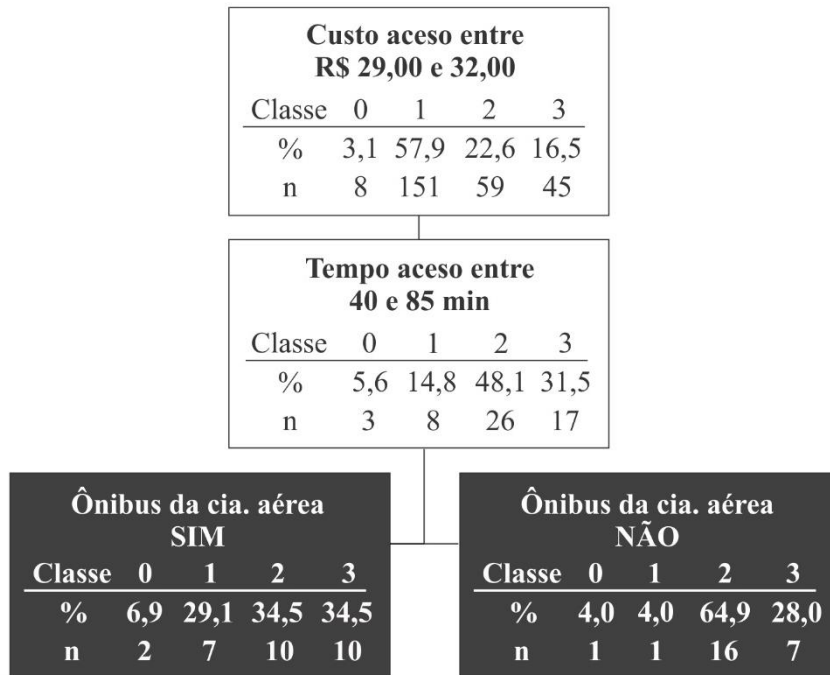


Figura 4 - Nós finais 5.4.1 e 5.4.2.

0 – Outros modos; 1 – Públicos individuais; 2 – particulares individuais; 3 – Públicos coletivos (ônibus).

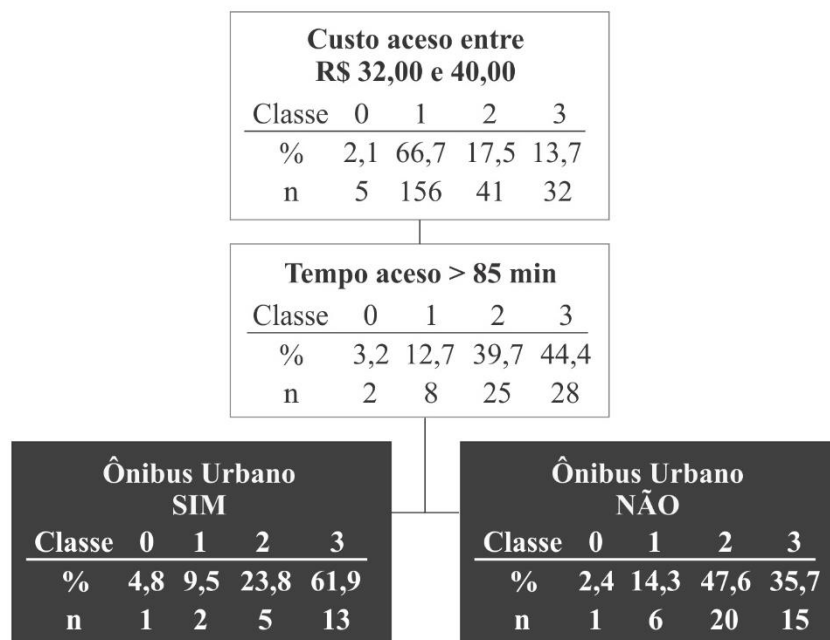


Figura 5 - Nós finais 6.4.1 e 6.4.2.

0 – Outros modos; 1 – Públicos individuais; 2 – particulares individuais; 3 – Públicos coletivos (ônibus).

A análise da Árvore de Decisão gerada mostra que, embora nem todos os tipos de usuários sejam sensíveis à oferta de transporte público, esse tipo de medida política aparenta ser capaz de influenciar a escolha do modo de transporte de algumas parcelas.

5. CONCLUSÕES E PESQUISAS FUTURAS

Este artigo descreveu uma proposta de análise da relação entre a oferta de transporte público coletivo, no caso o ônibus, e a escolha desse modo de transporte no acesso aos aeroportos do estado de São Paulo, utilizando técnicas de Árvore de Decisão. A escolha desse tipo de técnica deve-se ao fato de admitir qualquer tipo de variáveis (contínua, discreta e nominal), o que se mostrou essencial para a análise do banco de dados em questão, o qual apresentava um alto número de variáveis nominais. Além disso, ressalta-se a sua capacidade de representar relações não lineares e a sua não restrição relacionada a suposições matemáticas prévias.

Esse tipo de técnica é de fácil aplicação e interpretação dos resultados, embora não descreva a significância das variáveis como ocorre nas técnicas usuais paramétricas. O poder preditivo do modelo gerado, pode ser mensurado pela quantidade global de acertos.

A partir da aplicação do algoritmo CHAID de Árvore de Decisão, foi possível observar uma provável influência positiva da oferta de ônibus urbano e da companhia aérea na escolha do modo ônibus no acesso aos aeroportos paulistas para algumas classes de usuários. No entanto, em outras parcelas de usuários, essa mesma oferta não apresentou a resposta esperada, ocorrendo casos em que foi observado um aumento dos modos individuais e consequente redução do coletivo.

Outra questão observada foi a possível não relevância da oferta de ônibus interurbano e interestadual para a escolha do modo de acesso ônibus. Essa questão pode estar relacionada à seleção da variável dependente referente apenas ao último modo de transporte, o que pode acarretar uma não representação adequada dos usuários dos ônibus interurbanos, visto que, normalmente, esse tipo de acesso é realizado com transbordos.

Como objetivo de pesquisas futuras, poderão ser realizados novos modelos utilizando outros algoritmos da AD e consequente comparação dos resultados. O mesmo processo pode ser realizado com o uso de técnicas usuais como, por exemplo, o método *Logit Multinomial*, e a subdivisão dos tipos de viagem em “a trabalho” e em “outros motivos”, como realizado em trabalhos progressos da área.

Outra análise interessante seria a seleção de apenas um aeroporto, em que se tem um conhecimento prévio de uma melhoria na oferta de ônibus, seja urbano, interurbano ou da companhia aérea, para uma nova aplicação do questionário e comparação da influência da variável por meio da relação entre o período de 2009 (ano do banco de dados utilizado) e o atual.

AGRADECIMENTOS

Os autores agradecem à FAPESP (Processo nº 2016/14056-6) o suporte financeiro para a realização desta pesquisa.

6. REFERÊNCIAS

- Akar, G. (2013) **Ground Access to Airports, Case Study: Port Columbus International Airport**. Journal of Air Transport Management 30, 25-31.
- Alhussein, S. N. (2011) **Analysis of Ground Access Modes Choice Kig Khaled International Airport, Riyadh, Saudi Arabia**. Journal of Transport Geography 19, 1361-1367.
- ANAC (2013) **Anuário do Transporte Aéreo. Dados estatísticos e Econômicos de 2012**. Brasília – DF, Brasil.
- Breiman, L; J. Friedman; R. A. Olshen e J. Stone (1984) **Classification and Regression Trees**. Wadsworth International Group, Belmont, CA, USA.
- Fundação Instituto de Pesquisas Econômicas (2009) **Caracterização e Dimensionamento da Matriz Origem Destino do Transporte Aéreo no Brasil**. São Paulo. 2009. Disponível em <http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/Apoio_a_estudos_e_pesquisas/BNDES_FEP/prospeccao/chamada_aereo_anexos.html>. Acesso em 10 de agosto de 2015.
- Harvey, G. (1986) **Study of Airport Access Mode Choice**. Journal of Transportation Engineering 112 (5), 525-545.
- Isler, C. A. e Pitombo, C. S. (2014) **Avaliação da Escolha Modal para o Transporte Ferroviário de Passageiros na Região Sudeste Através de Árvores de Decisão**. XXVIII Congresso de pesquisa e ensino em transportes. Curitiba, PR, Brasil.
- Kass, G. V. (1980) **An Exploratory Technique for Investigating Large Quantities of Categorical Data**. Applied Statistics, v. 29, p. 119–127.
- Loh, W. Y. e Y. S. Shih (1997) **An Exploratory Technique for Investigating Large Quantities of Categorical Data**. Applied Statistics, v. 29, p. 119-127.
- McKinsey & Company (2010) **Estudo do Setor de Transporte Aéreo do Brasil: Relatório Consolidado**. Rio de Janeiro RJ, Brasil. Disponível em < http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/Apoio_a_Estudos_e_pesquisas/BNDES_FEP/prospeccao/chamada_aereo.html >. Acesso em 14 de agosto de 2015.
- Psaraki, V. e Abacoumkin, C. (2002) **Access Mode Choice for Relocated Airports: The New Athens International Airport**. Journal of Air Transport Management 8, 89-98.
- Quinlan, R. (1983) **Learning efficient Classification Procedures and their Application to Chess end Games in Machine Learning: An Artificial Intelligence Approach** p. 463-482. Tioga, Palo Alto.