

ANÁLISE COMPARATIVA DOS PADRÕES DE VIAGENS ENCADEADAS DE DIFERENTES GRUPOS DE INDIVÍDUOS PERTENCENTES A UMA MESMA ÁREA URBANA

Marcelo Figueiredo Massulo Aguiar
Eiji Kawamoto

Universidade de São Paulo
Escola de Engenharia de São Carlos – Departamento de Transportes

RESUMO

O principal objetivo da dissertação em andamento é comparar as viagens encadeadas realizadas por indivíduos que vivem em diferentes conjuntos de Zonas de Tráfego, pertencentes à mesma área urbana. As comparações serão feitas utilizando um minerador de dados (Árvore de Decisão e Classificação) aplicado ao conjunto de dados da Pesquisa Domiciliar coletada pelo Governo do Pará na Região Metropolitana de Belém (RMB). Espera-se que esta dissertação contribua para melhorar a compreensão dos aspectos subjacentes ao comportamento relacionado ao encadeamento de viagens.

ABSTRACT

The main goal of this master thesis under development is to compare the chained-trips made by individuals living in different sets of traffic zones, belonging to the same urban area. The comparisons will be made by using a data miner (Decision and Classification Tree) applied to the home-interview survey data set collected by Pará State Government in Belém Metropolitan Area (RMB). It is expected that this master thesis will improve the understanding of the aspects underlying the trip-chaining behavior.

1. INTRODUÇÃO

A análise da demanda por transporte baseada em atividades fundamenta-se no fato de que indivíduos tendem a organizar suas viagens em função, entre outras, das atividades a serem exercidas, dos locais onde estas atividades serão desenvolvidas, da prioridade das mesmas, do horário de início e término de cada atividade e das características dos modos disponíveis para ir do local de uma atividade ao local da atividade seguinte. Como resultado da decisão de cada indivíduo acerca da combinação mais adequada entre atividades e viagens, forma-se uma sequência de deslocamentos denominada encadeamento de viagens.

A dissertação em andamento se insere em uma linha de pesquisa que busca relacionar aspectos do encadeamento de viagens com características socioeconômicas dos habitantes de áreas urbanas. Ichikawa (2002), após revisão bibliográfica, testou a possibilidade de utilizar a Árvore de Decisão, um minerador de dados disponível no software S-Plus, para obter relações entre padrões simples de viagens encadeadas (somente motivo de viagem) e as características socioeconômicas dos viajantes. Uma vez que a experiência revelou-se satisfatória, foi seguida e ampliada por Pitombo (2003) que utilizou o minerador de dados para analisar o comportamento subjacente ao encadeamento de viagens, que desta vez apresentavam padrões mais complexos com a inclusão de atributos como modo, período do dia e duração da viagem. Sousa (2004) se valeu da mesma ferramenta para fazer comparações entre os padrões de viagens de diferentes grupos socioeconômicos de três áreas urbanas: Região Metropolitana de São Paulo (RMSP), Região Metropolitana de Belém (RMB) e cidade de Bauru.

Neste trabalho também pretende-se comparar os padrões de viagens encadeadas realizadas por diferentes grupos socioeconômicos, entretanto as comparações serão feitas dentro de uma mesma área urbana, neste caso a RMB. Serão comparados os padrões de viagens de indivíduos pertencentes a um determinado número de Zonas de Tráfego (ZT) escolhidas

aleatoriamente, com os padrões do grupo de indivíduos contidos nas ZTs restantes. As diferenças e/ou similaridades encontradas serão discutidas.

Analisando-se os resultados obtidos por Ichikawa (2002), Pitombo (2003) e Sousa (2004), verifica-se que as variáveis mais importantes para a definição de grupos homogêneos em relação ao padrão de viagem foram: Trabalha e Estuda. Na presente dissertação em andamento, aspira-se fazer uma análise mais detalhada destes parâmetros através da incorporação de novas variáveis que indiquem a distribuição espacial das vagas escolares e do emprego na RMB, permitindo assim capturar o efeito das localizações geográficas das oportunidades de trabalho e estudo nos padrões de viagens encadeadas.

2. OBJETIVO DO TRABALHO

O objetivo principal da dissertação em andamento é verificar e analisar as diferenças e/ou similaridades entre os padrões de viagens encadeadas de indivíduos pertencentes a um conjunto de zonas de tráfego escolhidas aleatoriamente e os padrões dos indivíduos pertencentes às zonas restantes da amostra.

3. MÉTODO

Os dados utilizados são os da Pesquisa Domiciliar, realizada pelo Governo do Estado do Pará em sete mil domicílios da RMB, no período de 24 de maio a 30 de junho de 2000, que registrou características socioeconômicas individuais, tais como: sexo, idade, grau de instrução, dentre outras; características do domicílio, como número de automóveis e renda familiar; características das zonas de tráfego, como número de empregos e número de vagas escolares; e atributos das viagens realizadas no dia anterior ao da entrevista domiciliar (motivo, modo, duração, período do dia, etc.). Na época a RMB contava com uma população de 1.782.395 habitantes distribuídos em cinco municípios (PDTU, 2001).

Para que se pudesse proceder às análises, foi necessário realizar a etapa de filtragem dos dados onde foram eliminados indivíduos: (i) cujas respostas estivessem incompletas; (ii) que realizaram uma ou mais de quatro viagens; (iii) cuja origem da primeira viagem ou destino da última viagem não fosse a residência. Após esta etapa partiu-se um universo de 32.467 pessoas para uma amostra de 28.245 indivíduos.

A codificação dos padrões de viagens encadeadas, passo importante para viabilizar o uso do minerador de dados, foi a mesma adotada por Pitombo (2003). Os atributos da viagem incluídos nos padrões são: Motivo (H-Residência, W-Trabalho, S-Escola e A-Demais Atividades), Modo (1 - Individual motorizado, 2 - Coletivo e 3 - Individual não-motorizado), Período (1 - 0:00h-8:00h, 2 - 8:00h-11:30h, 3 - 11:30h-14:30h, 4 - 14:30h-17:30h e 5 - 17:30h-24:00h), e Duração da viagem (1 - menos de 20 min., 2 - entre 20 e 40 min. e 3 - acima de 40 min.).

Os padrões de viagens constituem a variável dependente ou resposta, e o S-Plus permite a análise de 128 valores desta variável (padrões de viagem). Devido a esta restrição do software foi necessário excluir da amostra mais 5.507 indivíduos que realizavam padrões de viagens menos frequentes, obtendo-se uma amostra final de 22.738 indivíduos.

As características socioeconômicas constituem as variáveis independentes ou predictoras: *Trabalha* (sim – 1 ou Não – 0); *Estuda* (sim – 1 ou Não – 0); *Sexo* (Homem – 1 ou Mulher – 2); *Idade* (qualquer valor acima de seis anos); *Grau de Instrução* (Analfabeto / 4ª série incompleta – 1; Até 4ª série completa – 2; 1º grau completo – 3; 2º grau completo – 4; Superior completo – 5); *Renda familiar*, expressa em salários-mínimos da época; *Nº de pessoas na família e Nº de automóveis no domicílio* (nenhum – 0; um – 1, mais de um – 2).

Para algumas análises, também entraram como variáveis predictoras: *Grupo de Zonas* [Zonas escolhidas aleatoriamente (75% das ZTs) – 1 ou Demais ZTs – 2]; *porcentagem acumulada de empregos ao redor da ZT onde se localiza a residência do indivíduo e porcentagem acumulada de vagas escolares ao redor da ZT onde se localiza o domicílio do indivíduo*. Para se obter estas variáveis, calcularam-se as distâncias par a par entre os centróides de todas as ZTs. Estas distâncias foram categorizadas em nove classes (0-2 Km; 2-4 Km; 4-6 Km; 6-10 Km; 10-15 Km; 15-22 Km; 22-30 Km; 30-45 Km e acima de 45 Km), e em seguida foram calculadas as porcentagens acumuladas de empregos ou vagas escolares de cada ZT dentro destas classes. Por exemplo, há 9,05% das vagas escolares da RMB entre 0 e 2 Km do centróide da zona de tráfego 1, 21,99% entre 2 e 4 Km, 44,63% entre 4 e 6 Km e assim sucessivamente.

As relações entre a variável dependente (padrão de viagem) e as variáveis independentes (características socioeconômicas, distribuição espacial de emprego e vagas escolares e a variável Grupo de Zonas) serão obtidas através da Árvore de Decisão e Classificação, minerador de dados contido no software S-Plus. Este minerador classifica uma base de dados em um número finito de classes através de regras hierárquicas, fornecendo uma visão real da natureza da relação.

Para analisar as diferenças e/ou similaridades entre os padrões de viagens encadeadas de indivíduos pertencentes a um conjunto de zonas de tráfego escolhidas aleatoriamente e os padrões dos indivíduos pertencentes às zonas restantes da amostra, foram efetuados dois procedimentos distintos:

1. (i) Primeiramente, gera-se um relatório / árvore no S-Plus com os dados das zonas escolhidas aleatoriamente (75% do total das ZTs). (ii) Em cada nó da árvore observa-se um grupo homogêneo de indivíduos e a probabilidade de um indivíduo deste grupo realizar cada um dos padrões de viagens. (iii) Em seguida, aplicando-se a mesma estrutura do primeiro relatório / árvore gerado nos dados das zonas de tráfego restantes (25% das ZTs), obtém-se os mesmos grupos homogêneos, entretanto as probabilidades dos padrões de viagens serão diferentes. (iv) Comparam-se as distribuições de probabilidade dos padrões de viagens através do teste estatístico Kolmogorov – Smirnov;
2. Geração de um único relatório / árvore no S-Plus com a inclusão da variável categórica *Grupo de Zonas*, que permite identificar a qual grupo de zonas o indivíduo pertence.

Para analisar a influência da distribuição espacial das vagas escolares e do emprego nos padrões de viagens encadeadas, será gerado um relatório / árvore no S-Plus com a inclusão das variáveis: porcentagem acumulada de empregos ao redor da ZT onde se localiza a residência do indivíduo e porcentagem acumulada de vagas escolares ao redor da ZT onde se localiza o domicílio do indivíduo.

4. RESULTADOS PRELIMINARES

Em caráter de teste preliminar da viabilidade da pesquisa proposta, foram analisados os resultados obtidos das árvores e relatórios gerados pelo S-Plus sem a inclusão das variáveis que dizem respeito à distribuição espacial das vagas escolares e do emprego na RMB.

Na comparação entre os padrões de viagens encadeadas de indivíduos pertencentes a um conjunto de ZTs escolhidas aleatoriamente e os padrões dos indivíduos pertencentes às ZTs restantes da amostra, foi verificado que os mesmos são bastante semelhantes.

No procedimento que adotou a comparação através do teste Kolgomorov – Smirnov, foi verificado que em 23 dos 25 nós terminais da árvore, a máxima diferença entre a distribuição das probabilidades acumuladas dos dois conjuntos de dados era inferior ao valor crítico para um nível de significância de 5%, o que indica que os padrões de viagens dos grupos homogêneos dos diferentes conjuntos de zonas de tráfego podem ser considerados semelhantes.

No procedimento que incorporava a variável Grupo de Zonas, verificou-se que a árvore não se ramificou em relação a esta variável, indicando que não há diferenças significativas entre os padrões de viagens de indivíduos pertencentes aos diferentes grupos de zonas considerados.

A manutenção da consistência destes resultados pode levar a conclusões encorajadoras no sentido de que se pode pensar em reduzir o custo de pesquisas domiciliares, uma vez que a coleta de dados pode vir a ser feita em apenas um conjunto de zonas escolhidas aleatoriamente, sem a necessidade de se coletar dados em todas as zonas de uma determinada área urbana.

Pretende-se ao longo da dissertação em andamento, efetuar mais verificações que corroborem com os resultados preliminares comentados acima, bem como incorporar a esta análise as variáveis das características da distribuição espacial das vagas escolares e do emprego na RMB, permitindo assim analisar sua influência nos padrões de viagens encadeadas.

AGRADECIMENTO

À CAPES, pelo apoio financeiro à pesquisa na forma de uma bolsa de mestrado.

BIBLIOGRAFIA PRELIMINAR

- Ichikawa, S. M. (2002) Aplicação de Minerador de Dados na Obtenção de Relações entre Padrões de Encadeamento de Viagens Codificados e Características Sócio-econômicas. Dissertação de Mestrado apresentada na Escola de Engenharia de São Carlos, USP. São Carlos – SP, Brasil.
- PDTU/2001 – Plano Diretor de Transporte Urbano da Região Metropolitana de Belém (2001). *Relatório Final*. Belém – PA, Brasil.
- Pitombo, C. S. (2003) Análise do Comportamento Subjacente ao Encadeamento de Viagens através do Uso de Minerador de Dados. Dissertação de Mestrado apresentada na Escola de Engenharia de São Carlos, USP. São Carlos – SP, Brasil.
- Sousa, P. B. (2004) Análise Comparativa do Encadeamento de Viagens de três áreas urbanas. Dissertação de Mestrado apresentada na Escola de Engenharia de São Carlos, USP. São Carlos – SP, Brasil.

EESC / USP, Avenida Trabalhador Sãocarlense, 400, Centro. CEP: 13.566-590. São Carlos – SP.
Eiji Kawamoto: eiji@usp.br Marcelo Figueiredo Massulo Aguiar: mmassulo@sc.usp.br