

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**REQUISITOS PARA A COLETA AUTOMATIZADA DE
DADOS HÍBRIDOS APLICADA AOS TRANSPORTES
PÚBLICOS**

TIAGO HENRIQUE FRANÇA BARONI

ORIENTADOR: PASTOR WILLY GONZALES TACO

DISSERTAÇÃO DE MESTRADO EM TRANSPORTES

**PUBLICAÇÃO: T.DM-005/2017
BRASÍLIA / DF: MARÇO/2017**

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**REQUISITOS PARA A COLETA AUTOMATIZADA DE DADOS
HÍBRIDOS APLICADA AOS TRANSPORTES**

TIAGO HENRIQUE FRANÇA BARONI

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO PROGRAMA DE PÓS-GRADUAÇÃO EM TRANSPORTES DO DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE MESTRE EM TRANSPORTES.

APROVADA POR:

**PROF. PASTOR WILLY GONZALES TACO, Dr. (PPGT/UnB)
(ORIENTADOR)**

**ALAN RICARDO DA SILVA, Dr. (PPGT/UnB)
(EXAMINADOR I)**

**ARTUR CARLOS DE MORAIS, Dr. (SEMOB)
(EXAMINADOR II)**

BRASÍLIA/DF, 28 de março de 2017.

FICHA CATALOGRÁFICA

BARONI, TIAGO HENRIQUE FRANÇA	
Requisitos para a Coleta de Automatizada de Dados Híbridos Aplicada aos Transportes Públicos [Distrito Federal] 2017.	
xvi, 99p., 210 x 297mm (ENC/FT/UnB, Mestre, Transportes, 2017).	
Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.	
Departamento de Engenharia Civil e Ambiental.	
1. Sistemas Inteligentes de Transportes	2.Coleta Automatizada de Dados
3. Planejamento da Operação	4.Transportes Públicos
5. Transportes Públicos Rodoviários	
I. ENC/FT/UnB	II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

BARONI, Tiago Henrique França. (2017). Requisitos para a Coleta de Automatizada de Dados Híbridos Aplicada aos Transportes Públicos, Publicação T. DM-005/2017, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 99p.

CESSÃO DE DIREITOS

AUTOR: Tiago Henrique França Baroni

TÍTULO: Requisitos para a Coleta de Automatizada de Dados Híbridos Aplicada aos Transportes Públicos

GRAU: Mestre

ANO:2017

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Tiago Henrique França Baroni
Rua 25, Lote 05, Apto. 2007 – Edif. Vitali
Norte – 71.917-180 – Águas Claras - DF
tbaroni@gmail.com // tiago.baroni@hotmail.com

“Apega-te à instrução e não a largues; guarda-a, porque ela é tua vida”

Provérbios 4:13

DEDICATÓRIA

Não haveria injustiça maior do que não dedicar cada linha desta obra à pessoa que me apoiou em todos os momentos, bons ou nem tanto, durante a concepção desta.

Para você Priscila Hoehr Mostardeiro, minha esposa, por ser a mulher mais maravilhosa, forte, inteligente e paciente que eu conheço.

AGRADECIMENTOS

Agradeço, inicialmente, à minha família. Minha esposa desempenhou o papel mais complexo e com maior grau de renúncia de todos: ela me viu dedicar boa parte da minha vitalidade e dos meus esforços neste trabalho. Muito obrigado Priscila, você é sensacional! O companheirismo e conforto oferecidos por você, pela Madonna e o Ozzy certamente me tocam a cada dia.

Agradeço aos meus familiares que, do mesmo modo, me viram cada dia menos para que eu me dedicasse cada dia mais ao curso, em especial à minha irmã mais nova Paola, que renunciei em ver crescer para que este sonho se realizasse.

Ao meu irmão, que por um erro do destino nasceu de pais diferentes, mas que eu tive a fantástica oportunidade de chamar caminhar lado a lado. Angelo Simonato Sanches, obrigado pelo tempo despendido em discussões técnicas, divagações e fúria criativa. Seu suporte foi imprescindível.

Aos meus amigos recém conquistados em Brasília, Ronny Aliaga, Alexandre Silva, Priscila Roque, Wallace Vargas, a pequena Júlia, Priscila Garcia e Fernando Franco. Vocês tornaram essa caminhada muito mais fácil.

Ao professor Pastor Taco, quem me orientou e me motivou a executar este trabalho. Obrigado pelas oportunidades ofertadas e pela paciência em me transformar naquilo que sou hoje.

Ao Grupo de Pesquisas Comportamento em Transporte e Novas Tecnologias, que me acolheu tão calorosamente tanto como membro, quanto como coordenador. Agradeço também ao CEFTRU pelo fornecimento de espaço e de subsídios de maneira ímpar para a condução e a estruturação deste trabalho.

Aos professores do Programa de Pós-Graduação em Transportes, em especial ao professor José Augusto, por me fornecer a fagulha necessária para elaboração deste trabalho durante uma aula de Operação dos Transportes Públicos. Agradeço ao CNPq pelo apoio financeiro a parte desta pesquisa.

RESUMO

Através da última década métodos diferenciados de Planejamento da Operação dos Transportes Públicos Urbanos foram apresentados com o intuito de atender a intrincada relação entre a demanda de transporte moderna e a oferta de transportes públicos, como o Planejamento e Operação Dinâmica dos Transportes (*Dynamic Transportation Planning and Operations - DTPO*) e a Estimativa Dinâmica de Origens e Destinos (*Dynamic OD Estimation – DODE*). Embora sua eficácia seja repetidamente provada, esses métodos trazem consigo o inconveniente de demandar de um grande volume de dados, inclusive de alguns dados em tempo real. A demanda por dados acaba levando os operadores e planejadores dos transportes públicos a adotarem métodos automatizados de coleta de dados.

Até o presente momento, na seara dos sistemas de transportes inteligentes, os métodos existentes propõem-se a coletar dados de diferentes vertentes: coletores automatizados de dados capturam dados de telemetria dos veículos, tarifadores automáticos coletam dados sobre o ingresso de passageiros nos sistemas e localizadores automatizados de veículos coletam dados sobre o deslocamento dos veículos. Sempre que são necessárias coletas que interseccionem dois ou mais métodos, o formato de coleta é realizado ad hoc gerando um volume de dados com compatibilidade reduzida entre si.

Dessa maneira o objetivo deste trabalho é apresentar os requisitos para que a coleta automatizada seja considerada híbrida, ou seja, que tangencie todas as formas de coleta de dados solidificadas na literatura.

Para o atingimento dessa meta, a presente dissertação (1) realizou uma extensa revisão bibliográfica com o intuito de fundamentar o tema e determinar as tecnologias de coleta de dados automatizadas existentes e emergentes, (2) definiu os dados de maior relevância para o Planejamento da Operação dos Transportes Públicos Urbanos no modal rodoviário, (3) delimitou a interseção entre as tecnologias levantadas em (1) e as necessidades apresentadas em (2), formando um padrão de requisitos híbridos de dados, (4) validou a aplicabilidade comparativamente com estudos contratados pelo poder público, e ainda, emulando equipamentos que gerem o padrão de dados estipulado e verificando a aplicabilidade dos dados gerados. Os resultados obtidos por este trabalho apresentam um horizonte bastante promissor para o desenvolvimento prático de tecnologias baseadas em dados híbridos, além de uma vasta aplicabilidade no Planejamento da Operação dos Transportes Públicos Urbanos.

Palavras chave: Sistemas Inteligentes de Transportes, Coleta Automatizada de Dados, Planejamento da Operação, Transportes Públicos, Transportes Públicos Rodoviários

ABSTRACT

Throughout the last decade, different methods of public transportation operation planning have been presented in order to meet the intricate relationship between the demand for modern transportation and the supply of public transportation, such as Dynamic Transportation and Operations - DTPO and the Dynamic OD Estimation DODE.

Although their effectiveness is repeatedly proven, these methods carry with them the inconvenience of requiring a large amount of data, including some real-time data. Demand for data drives public transport operators and planners to adopt automated methods of data collection.

So far, in the field of intelligent transport systems, there are methods designed to collect data from different point of view: automated data collectors capture vehicle telemetry data, automated billers collect data on passenger entry into systems and locators Automated vehicles collect data on the movement of vehicles. Whenever collections that intersect two or more methods are required, the collection format is performed *ad hoc* generating a volume of data with reduced compatibility with each other.

In this way, the objective of this work is to present the requirements for the automated collection to be considered hybrid, that is, that tangencies all forms of solidified data collection in the literature.

In order to reach this goal, the present dissertation (1) carried out an extensive bibliographical review in order to base the theme and determine the existing and emerging automated data collection technologies, (2) defined the most relevant data for the planning of the operation (3) delimited the intersection between the technologies raised in (1) and the needs presented in (2), forming a pattern of hybrid data requirements, (4) validated the applicability comparatively with studies contracted by the State and by emulating equipment that generates the data standard and verifying the applicability of the generated data. The results obtained by this work show a very promising horizon for the practical development of technologies based on hybrid data, as well as a wide applicability in the planning of the public transportation operation.

Keyword: Intelligent Transport Systems, Automated Data Collection, Operation Planning, Public Transport, Public Road Transport

SUMÁRIO

1 -	INTRODUÇÃO	1
1.1 -	Apresentação.....	1
1.2 -	Delimitação do Problema.....	2
1.3 -	Hipótese	5
1.4 -	Objetivo Geral.....	5
1.4.1 -	Objetivos Específicos	5
1.5 -	Justificativa	6
1.6 -	Metodologia da Dissertação.....	8
2 -	SISTEMAS INTELIGENTES DE TRANSPORTES	12
2.1 -	Apresentação.....	12
2.2 -	Conceito de Sistemas Inteligentes de Transportes.....	12
2.3 -	Métodos de Coleta Automatizada de Dados	14
2.4 -	Tópicos Conclusivos.....	18
3 -	PLANEJAMENTO DA OPERAÇÃO DOS TRANSPORTES PÚBLICOS URBANOS 19	
3.1 -	Apresentação.....	19
3.2 -	Conceito do Planejamento da Operação dos Transportes.....	19
3.3 -	Técnicas para Estimativa de Matrizes de Origem e Destino	24
3.4 -	Aplicação de Técnicas para Obtenção de Matrizes OD Através de Coletas Automatizadas.....	26
3.5 -	Tópicos Conclusivos.....	29
4 -	O MÉTODO DA COLETA AUTOMATIZADA DE DADOS HÍBRIDOS PARA O PLANEJAMENTO E OPERAÇÃO DOS TRANSPORTES PÚBLICOS URBANOS.....	30
4.1 -	Apresentação.....	30
4.2 -	Coleta Automatizada de Dados Híbridos.....	30
4.3 -	A Coleta Automatizada e a Privacidade dos Usuários.....	33
4.4 -	Tópicos Conclusivos.....	34
5 -	DETERMINAÇÃO DE DADOS A SEREM COLETADOS PELO MÉTODO DE COLETA AUTOMATIZADA DE DADOS HÍBRIDOS.....	35
5.1 -	Apresentação.....	35
5.2 -	Critérios para a Seleção de trabalhos acadêmicos	35
5.3 -	Concentração de Tipos Semelhantes	37
5.3.1 -	Dados de telemetria	37
5.3.2 -	Estampa de tempo.....	37
5.3.3 -	Pontos geoespaciais	38
5.3.4 -	Transações do validador	38
5.4 -	Levantamento da Recorrência de Variáveis Coletadas.....	38
5.5 -	Variáveis de Qualidade dos Transportes Públicos.....	40

5.5.1 -	Pesquisa de Campo - Método de Coleta e Parâmetros Observados	41
5.5.2 -	Análise da Percepção dos Usuários do Transporte Público Rodoviário	42
5.6 -	Tópicos Conclusivos	45
6 -	DETERMINAÇÃO DA ARQUITETURA, SENSORES ASSOCIADOS E PADRÕES DE DADOS DO COLETOR AUTOMATIZADO DE DADOS HÍBRIDOS PARA O POTPU	47
6.1 -	Apresentação	47
6.2 -	Arquitetura Básica dos Equipamentos de AHDC	47
6.3 -	Equipamentos Mínimos do Sistema AHDC	51
6.3.1 -	Módulo Controlador Central do Sistema AHDC	51
6.3.2 -	Módulo de Gerenciamento de Energia	52
6.3.3 -	Módulo de Memória	53
6.3.4 -	Módulo Relógio de Tempo Real	53
6.3.5 -	Módulo GPS	54
6.3.6 -	Módulo de Transmissão de Dados	56
6.3.7 -	Módulo Registrador de Passageiros	56
6.3.8 -	Módulo Registrador de Eventos do Validador	58
6.3.9 -	Sensor Acelerômetro	59
6.3.10 -	Sensor de Abertura das Portas, Sensor de Acionamento do Botão de Parada e da Rampa para o Cadeirante	60
6.3.11 -	Sensores de Luminosidade, de Termohigrômetro, de Precipitação, de Ruído e de Qualidade do ar	60
6.4 -	Formatos dos dados	61
6.5 -	Tópicos Conclusivos	69
7 -	VALIDAÇÃO DO MÉTODO AHDC: COMPARAÇÃO, SIMULAÇÃO DE FUNCIONAMENTO E EMULAÇÃO DA ANÁLISE	70
7.1 -	Apresentação	70
7.2 -	Comparação com os Estudos de Mobilidade Recentes	70
7.2.1 -	Plano Diretor de Transporte Urbano do Distrito Federal	71
7.2.2 -	Plano de Mobilidade Urbana Sustentável da Grande Florianópolis	72
7.2.3 -	Plano de Desenvolvimento do Transporte Público Sobre Trilhos do Distrito Federal	72
7.2.4 -	Comparação entre os Dados Levantados nos estudos e os Requisitos Levantados	73
7.3 -	Simulação da Solução e Verificação de Conveniência	76
7.3.1 -	Características Técnicas da Plataforma de Teste	77
7.3.2 -	Teste de Volume de Dados	80
7.3.3 -	Teste de Simulação de fornecimento de dados	82
7.4 -	Tópicos Conclusivos	86

8 -	CONCLUSÕES E RECOMENDAÇÕES.....	88
8.1 -	Considerações Iniciais	88
8.2 -	Conclusões	89
8.3 -	Limitações da Pesquisa.....	90
8.4 -	Recomendações para Pesquisas Futuras	90
	REFERÊNCIAS BIBLIOGRÁFICAS	92
	ANEXO I – Transcrição do Formulário da Pesquisa Aplicada	98

LISTA DE FIGURAS

Figura 1.1 - Fluxograma Metodológico e Prático do trabalho	10
Figura 1.2 - Fluxo lógico para a construção desta dissertação.....	11
Figura 2.1 - Cadeia de concentrações dos equipamentos de ITS.....	14
Figura 3.1 - Cadeia de ITS considerando a presença dos métodos de coleta automatizada de dados híbridos.	28
Figura 4.1 - Interseção dos métodos e coleta atuais e o método híbrido.....	31
Figura 6.1 - Diagrama Multifuncional dos Equipamentos AHDC. (adaptado e modificado com base em Deibel e Zumwalt (1984))	49
Figura 6.2 - Diagrama esquemático da arquitetura dos AHDC.	50
Figura 6.3 - Representação esquemática dos "5v's" dos <i>big data</i>	62
Figura 6.4 - Formato do nome do arquivo	64
Figura 7.1 - Rede de transporte público coletivo urbano do Distrito Federal (Fonte: PDTU/DF, página 29 - SEMOB (2010))	78
Figura 7.2 - Carregamento da rede de transporte público coletivo do PDTU/DF, passageiros na hora de pico da manhã (Fonte: PDTU/DF, página 39 - SEMOB (2010)).....	79
Figura 7.3 - Retorno do comando de geração e busca dos dados.....	81
Figura 7.4 - Programa para simulação de Operação Rodoviária Online.....	82
Figura 7.5 - Resultado da simulação para o cenário padrão.....	84
Figura 7.6 - Resultado da simulação para o segundo cenário	85
Figura 7.7 - Resultado da simulação para o terceiro cenário	86

LISTA DE QUADROS

Quadro 6.1 - Retorno de 2 segundos de leitura de um GPS com padrão NMEA.	55
Quadro 6.2 - Pseudocódigo para o Sensor Acelerômetro	59
Quadro 6.3 - Estrutura usada para os registros síncronos	65
Quadro 6.4 - Versão reduzida dos dados coletados para ser enviada em tempo real pelo Módulo de Trasmissão de Dados	67
Quadro 6.5 - Estrutura usada para os registros assíncronos	68

LISTA DE TABELAS

Tabela 1.1 - Crescimento de publicações acadêmicas relacionadas com o tema <i>Automatic Data Collection</i> indexadas pelo serviço <i>Google Scholar</i>	4
Tabela 2.1 - Artigos mais relevantes no que tange a coleta de dados para os transportes nos últimos oito anos, segundo a inteligência artificial.	16
Tabela 2.2 - Tipos de dados coletados pelas diferentes metodologias clássicas de coleta de dados.....	17
Tabela 5.1 - Compilação dos dados coletados pelos artigos científicos e seu modo de coleta.	39
Tabela 5.2 - Análise da percepção do usuário de ônibus em relação ao deslocamento até o ponto de ônibus	42
Tabela 5.3 - Análise da percepção do usuário de ônibus em relação à viagem	43
Tabela 5.4 - Classificação de importância para itens relativos ao transporte público, na ótica dos usuários frequentes	44
Tabela 5.5 - Classificação dos três itens mais votados como importantes relativos à viagem, na ótica dos usuários frequentes.....	44
Tabela 5.6 - Compilação dos dados a serem coletados.	46
Tabela 7.1 - Comparativo entre dados coletados por estudos <i>versus</i> dados ofertados pelos AHDC.....	73
Tabela 7.2 - Comparação de dados coletados para aferição de qualidade com as coletas AHDC (adaptado de Martins (2015)).....	75

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES

ADC	- <i>Automatic Data Collection</i>
AFC	- <i>Automatic Fare Collection</i>
AHDC	- <i>Automatic Hybrid Data Collection</i>
AVL	- <i>Automatic Vehicle Location</i>
APCU	- <i>Automatic Passenger Counter Unit</i>
ATIS	- <i>Advanced Travel Information Systems</i>
ATMS	- <i>Advanced Travel Management Systems</i>
DFTrans	- Transportes Urbanos do Distrito Federal
DODE	- <i>Dynamic OD Estimation</i>
DTPO	- <i>Dynamic Transportation Planning and Operations</i>
GCTNT	- Grupo Comportamento em Transportes e Novas Tecnologias
GPS	- <i>Global Positioning System</i>
JSON	- <i>JavaScript Object Notation</i>
IFP	- <i>Iterative Proportional Fitting</i>
ITS	- <i>Intelligent Transportation Systems</i>
MCC	- Módulo de Controle Central
MGE	- Módulo de Gerenciamento de Energia
MGPS	- Módulo de GPS
MLE	- <i>Maximum Likelihood Estimator</i>
MM	- Módulo de Memória
MREV	- Módulo Registrador dos Eventos do Validador
MRP	- Módulo Registrador de Passageiros
MRTR	- Módulo Relógio de Tempo Real
MTD	- Módulo de Transmissão de Dados
NMEA	- <i>National Marine Electronics Association</i>
OD	- Origem-Destino
PLAMUS	- Plano de Mobilidade Urbana Sustentável da Grande Florianópolis
PDTT/DF	- Plano de Desenvolvimento do Transporte Público Sobre Trilhos do Distrito Federal
PDTU/DF	- Plano Diretor de Transporte Urbano do Distrito Federal
PMU/DF	- Pesquisa de Mobilidade Urbana do Distrito Federal
POTPU	- Planejamento da Operação dos Transportes Públicos Urbanos

PTP - Planejamento dos Transportes Públicos
SIG - Sistema de Informações Geográficas

1 - INTRODUÇÃO

1.1 - APRESENTAÇÃO

Os problemas relativos aos deslocamentos nas áreas de crescimento urbano tem aumentado proporcionalmente ao processo de antropização, reflexo do crescimento das cidades e do aumento da população urbana (Kneib, 2008).

Compreender, de maneira ampla e abrangente, o comportamento e os padrões de viagens é essencial para a previsão de tráfego, o planejamento da operação e a gestão dos centros urbanos (Wang *et al.*, 2015). Essa necessidade de compreensão se estende a todas as esferas dos transportes, públicos ou privados.

Abordagens diferenciadas vem sendo desenvolvidas e testadas nas últimas décadas, com o intuito de oferecer ferramentas eficazes de planejamento e operação dos transportes. Essas abordagens culminaram em modelos dinâmicos do planejamento da operação, entre os quais se destacam: (i) o Planejamento e Operação Dinâmica dos Transportes (*Dynamic Transportation Planning and Operations* - DTPO), no qual são descritas as características dinâmicas do transporte num sistema de dupla perspectiva, considerando a demanda e a oferta; e (ii) a Estimção Dinâmica de Origens e Destinos (*Dynamic OD Estimation* - DODE), que tenta inferir a matriz de origem e destino através de métodos indiretos dinâmicos, como a análise de fluxo das pessoas, coletada através dos deslocamentos de smartphones, por exemplo, permitindo um equilíbrio dinâmico entre a oferta e a demanda (Zou *et al.*, 2013).

Embora tenham se mostrado eficazes, para o correto funcionamento de tais métodos, é necessária a convergência de uma quantidade considerável de dados, sendo incompatível a coleta pelos meios tradicionais ou, ainda, através de demanda espontânea dos usuários. Assim, os métodos dinâmicos de planejamento das operações carecem de coletores automatizados de dados.

No que tange a coleta de dados para a realização de pesquisas de fundamento para o planejamento da operação, verifica-se uma iniciativa de modernização nos métodos e ferramentas aplicadas. Exemplo disso é a Pesquisa de Mobilidade Urbana do Distrito Federal (PMU/DF), conduzida pela Companhia do Metropolitano do Distrito Federal, como parte do Plano de Desenvolvimento do Transporte Público Sobre Trilhos do Distrito Federal (PDTT/DF).

No citado projeto, conduzido em 2016 e previsto para ser finalizado em dezembro de 2017, as entrevistas foram realizadas com o auxílio de *tablets*, que coletam automaticamente a geolocalização da unidade pesquisada, associando os domicílios às zonas de tráfego. Ainda, o dispositivo eletrônico georreferencia cada um dos pares origem/destino, posicionando-os automaticamente em zonas de tráfego, aumentando a acurácia e a confiabilidade do trabalho executado (Metrô/DF, 2016).

No entanto, parte dos dados a serem coletados independe de entrevistas e de coletas manuais. Esses dados, coletados dos operadores dos transportes públicos e dos órgãos responsáveis pela engenharia de tráfego, fornecem informações como: nível de oferta realizado, quantidade de viagens diárias, duração média da viagem, nível de saturação viário, distribuição de volume de veículos durante o dia, entre outras.

Em contrapartida, existe um volume considerável de ações necessárias para a modernização da forma de coleta dessas informações, sem qualquer iniciativa para planificação ou homogeneização dos dados coletados e seus formatos. Esse tipo de iniciativa se torna caótica e de difícil compatibilização, uma vez que cada interessado - seja autarquia governamental, instituições de ensino e/ou pesquisa ou entes privados - delimita um padrão de coleta de dados *ad-hoc*, que se altera a cada nova pesquisa e, nem sempre, permite a correta compatibilidade com as informações coletadas anteriormente.

Como resultado, tem-se inúmeras amostras, de diferentes cortes temporais, de dados extremamente similares, porém de difícil compatibilização, freando as tentativas de análise e de previsão das demandas de transportes como um cenário que evolui através do tempo.

Esse cenário com dados divergentes impacta também na forma com que os operadores dos transportes públicos tomam suas iniciativas táticas. Determinadas informações conflitantes, incompletas ou equivocadas podem transformar a tarefa de operar os transportes públicos em um ato de coordenação empírica.

1.2 - DELIMITAÇÃO DO PROBLEMA

O número de veículos em circulação nas cidades aumenta, sobretudo, a complexidade da análise das redes de transportes, em especial, o Planejamento da Operação dos Transportes Públicos Urbanos (Moreira-Matias *et al.*, 2015).

O problema se apresenta ainda mais crônico quando se observam as consequências oriundas dos congestionamentos.

No mesmo sentido, o cenário de perdas graças aos congestionamentos são igualmente impactantes no Brasil e estão estimados na casa dos 9,9 bilhões de reais para o custo pecuniário, com um custo de oportunidade que chega a 30 bilhões de reais, considerando apenas a cidade de São Paulo e sua região metropolitana (Cintra, 2014).

Boa parte dos problemas inerentes às cidades de médio e grande porte estão diretamente relacionados com planejamento e operação deficientes ou embasados em estudos insuficientes, associados a redes de transportes públicos ineficazes. Infelizmente, parte das soluções de planejamento para os problemas de tráfego são adotadas sem a necessária fundamentação em estudos socioeconômicos e comportamentais. Embora erroneamente relegados ao status de supérfluos, a finalidade de tais estudos é, justamente, desvendar uma estrutura mais complexa do fenômeno de transporte e aprofundar uma análise da realidade cotidiana que as soluções triviais não conseguem captar (Pontes, 2005).

Em diversas áreas do conhecimento, que vão desde a física experimental a análise comportamental de humanos em aglomerações, são adotados métodos de coleta automática de dados, nos quais seria virtualmente impossível, ou excessivamente dispendioso, coletar dados da maneira tradicional (Mattos, 2004).

Na própria seara dos transportes a coleta automatizada de dados (*Automatic Data Collection – ADC*) é largamente utilizada, em especial para a coleta de dados de telemetria dos veículos, sendo a base de funcionamento de veículos autômatos e de instrumentos de controle.

Embora os trabalhos de Briskin *et al.* (1979) e Deibel e Zumwalt (1984) apresentem a base técnica para a maior parte dos métodos de ADC atuais, somente após o ano de 2005 houve um expressivo crescimento de estudos acadêmicos associando à coleta automatizada de dados e às pesquisas básicas ligadas ao planejamento dos transportes. A Tabela 1.1 apresenta os resultados das pesquisas pela sintaxe “*Transportation OR Transport OR Transports OR ADC "Automatic Data Collection"*” na ferramenta de buscas acadêmicas *Google Scholar*.

Entre os usos mais frequentes de ADC nesses materiais publicados está a construção de matrizes de origem e destino, bem como a atualização destas, seja para os transportes públicos ou para os transportes individuais. Cui (2006) apresenta como vantagens do uso de ADC para estimação

de matrizes de origem e destino: o custo reduzido, o uso de amostras maiores e menos enviesadas, as estimações mais frequentes e, ainda, a possibilidade de realização de pesquisas complementares mais focadas e diretas.

Tabela 1.1 - Crescimento de publicações acadêmicas relacionadas com o tema *Automatic Data Collection* indexadas pelo serviço *Google Scholar*.

Artigos publicados de	a	Quantidade	Crescimento em relação ao ano anterior
1950	1970	22	
1970	1975	50	127,27%
1975	1980	57	14,00%
1980	1985	92	61,40%
1985	1990	157	70,65%
1990	1995	258	64,33%
1995	2000	327	26,74%
2000	2005	439	34,25%
2005	2010	738	68,11%
2010	2015	1240	68,02%

Essas vantagens comuns, além das vantagens particulares a cada caso, têm criado um ambiente muito favorável e fértil para o crescimento de aplicações dos métodos de coleta automatizada em pesquisas de estimação para o Planejamento dos Transportes Públicos (PTP), assim como para o Planejamento da Operação dos Transportes Públicos Urbanos (POTPU).

Entretanto, por não haver um padrão de variáveis coletadas ou de formato dos dados coletados, os dados resultantes de cada uma dessas pesquisas seguem um padrão próprio, com resultados que não interagem naturalmente entre si.

Nesse sentido, Ben-Akiva e Morikawa (1989) apresentaram métodos para fusão de dados provindos de diferentes pesquisas. A estrutura de trabalho proposta por esse artigo se baseia na convergência de dados fixados sobre cinco pilares principais: (1) parâmetros primários desconhecidos; (2) medidas diretas; (3) medidas indiretas; (4) viés de aferição; e o (5) método de estimação. Embora o método de estimação adotado seja o estimador de máxima verossimilhança (*Maximum Likelihood Estimator* – MLE), sendo esse o mais popular para a estimativa de parâmetros, a aplicação indicada é de uma complexidade considerável, em especial quando se considera sua aplicação no dia a dia e por profissionais nem sempre treinados tecnicamente.

Mesmo se observando trabalhos mais recentes do que aquele indicado anteriormente, como os de Faouzi *et al.* (2011) e de Nachouki e Quafafou (2008), não há trivialidade na tarefa, sendo evidentes vários passos que podem induzir o usuário ao erro.

Sob essa perspectiva, este autor sugere diversas questões que podem ser aventadas: (a) como tornar os dados compatíveis entre si, descartando a necessidade de operações de fusão de dados e possibilitando a menor interferência humana possível? (b) como padronizar a coleta de dados de diferentes origens sem sobrecarregar os elementos de coleta? (c) quais dados devem ser considerados como prioritários, em uma estrutura básica, para que os bancos de dados formados pela coleta tenham um tamanho tangível pela maior parte das instituições? (d) como manter uma alta resiliência de dados, mesmo considerando as inevitáveis evoluções dos sensores e equipamentos de coleta?

Levando-se em consideração as preocupações indicadas por este autor e o cenário trazido à baila anteriormente, questiona-se: quais devem ser os requisitos mínimos para a coleta de dados de diversas fontes, que mantenham uma possibilidade de acoplamento simples em sua análise e uma boa resiliência através do tempo?

1.3 - HIPÓTESE

A análise documental de artigos científicos e a condução de uma pesquisa de campo com os usuários dos transportes públicos rodoviários fornecerão subsídios para a formulação de requisitos mínimos a serem observados nas coletas, com características comuns às mais diversas formas de coleta automatizada, formando um padrão híbrido de coleta de dados.

1.4 - OBJETIVO GERAL

Apresentar os requisitos mínimos, considerando a estrutura e os sensores associados, bem como as estruturas de dados geradas, visando a construção de um modelo unificado para a coleta automatizada de dados híbridos.

1.4.1 - OBJETIVOS ESPECÍFICOS

- (1) determinar as tecnologias de coleta de dados automatizadas existentes e emergentes, bem como suas aplicabilidades;
- (2) definir os dados de maior relevância para o PTU e para o POTPU no modal rodoviário, através da revisão sistemática de estudos acadêmicos - revisado por pares - já realizados;

(3) Formar um padrão de requisitos híbridos de dados a partir das tecnologias levantadas e as necessidades apresentadas de dados para o POTPU, e;

(4) validar o método comparativamente com base em pesquisas que se destinam ao planejamento da operação conduzida por órgãos públicos, além de simular da geração de dados para verificar o volume, variedade e usabilidade dos dados.

1.5 - JUSTIFICATIVA

A vascularidade dos transportes rodoviários confere a estes uma melhor distribuição no tecido urbano e, conseqüentemente, a capacidade de atender a um público maior. Basta observar que o sistema de transporte rodoviário do Distrito Federal transportou 1,2 milhões de pessoas por dia, em 2014, contra 160.000 pessoas transportadas pelo Metrô/DF. Mesmo em cidades com uma maior malha metroferroviária, a vascularização do transporte rodoviário possibilita uma maior captação de passageiros, como é o caso de São Paulo, onde são transportados 5 milhões de passageiros diariamente por ônibus contra 4 milhões de passageiros pelos transportes metroferroviários (Lobo, 2015).

Associada a essa predominância, a presença dos meios de transportes rodoviários em cidades de médio e pequeno porte conferem a estes um potencial latente para a geração de dados em todos os equipamentos e veículos envolvidos nos transportes públicos. Outrossim, seu aproveitamento está mais intimamente ligado à sua relevância e raridade do que à sua capacidade de fato.

Uma frota de ônibus que se desloca numa via pública com um equipamento que colete seu posicionamento a cada 30 segundos, enquanto está em movimento, e que colete e contabilize seu tempo e local imobilizado, pode gerar dados de extrema relevância para o POTPU, como: tempo real de ciclo, tempo médio de ciclo, velocidade média por trecho, tempo imobilizado, locais de parada e oferta real de transporte. Essas informações, em posse do gestor público, podem culminar num planejamento da operação mais eficaz em seu intuito primordial de servir adequadamente à população.

Tomando Brasília como exemplo, em consulta realizada à autarquia responsável pelos transportes urbanos rodoviários do Distrito Federal - Transportes Urbanos do Distrito Federal (DFTrans), em 01 de dezembro de 2016, de uma frota total de 3.156 veículos em circulação, 2.712 destes possuem equipamentos de GPS e mecanismo de ADC instalados.

Considerando que esses equipamentos coletam apenas as coordenadas instantâneas dos veículos, ou seja, 24 caracteres, acrescidas de um código identificador de 6 caracteres, com atualizações a cada 30 segundos, ter-se-iam 167,59Mb de dados gerados por dia.

Esses dados, combinados e transformados em uma série histórica, podem demonstrar claramente o funcionamento e a oferta dos transportes públicos rodoviários, especialmente, no que diz respeito à oferta de transportes aos cidadãos. Outros dados poderiam ser coletados com o objetivo de elucidar, por exemplo, a saturação real dos transportes públicos.

Embora o volume de dados gerados por esses potenciais coletores seja considerável, aproximando-se de um *big data* quando observados através do tempo, não há qualquer forma de consistência e padronização dessas informações, deixando-as “isoladas” em seus universos. Essa circunstância impossibilita a análise de dados reais quanto ao transporte público, induzindo ao uso de tabelas de dados sintéticos, esperados ou “médios”.

Adicionalmente, somente no ano de 2016, três capitais de estados brasileiros estão realizando pesquisas ligadas à mobilidade urbana, a saber: São Paulo (Rede Nossa São Paulo, 2016), RECIFE (2015) e Brasília (Metrô/DF - Companhia do Metropolitano do Distrito Federal, 2016), com um custo total superior a 15 milhões de Reais. Embora as pesquisas citadas apresentem um esforço de planejamento considerável, em nenhuma delas há uma preocupação explícita em manter uma constância no formato básico dos dados relativos às pesquisas anteriores, realizadas nas próprias cidades ou mesmo entre si, permitindo uma comparação objetiva entre os resultados de localidades diferentes.

Como parte dos projetos elencados acima diz respeito à coleta de informações relativas à rede de transportes públicos, em especial de dados que possibilitem reconstituir um nível de serviço médio dessas redes, a presença de coletores automatizados de dados, poderia permitir que os custos pecuniários e de tempo fossem reduzidos.

Ademais, a desagregação dessas diferentes fontes de dados e de suas características temporais praticamente estanques, podem exigir um esforço matemático/estatístico especializado para sua análise comparativa conjugada. Em contraponto, a integração dos formatos, dos dados e da sistemática pode trazer um determinado grau de trivialidade na congregação das informações de diferentes fontes, permitindo uma análise holística dos transportes, favorecendo, assim, o PTP e o POTPU. Da mesma forma, as análises e as comparações de cunho temporal poderão

ser realizadas com melhor facilidade e assertividade, permitindo o acompanhamento da evolução dos cenários e características.

Uma homogeneização nos tipos e nos formatos, pelo menos, dos dados recorrentes, entre os diferentes produtores, permitirá uma troca de experiências mais positiva e eficaz, vez que as adequações dos dados objetivos coletados serão mínimas.

Por fim, pode-se enxergar essa planificação das informações como um passo em direção a bases de dados abertas para pesquisas, de tal forma que novos caminhos, análises mais robustas e teorias complexas acerca do funcionamento e da evolução da mobilidade urbana possam ser realizadas de maneira cada vez mais transparente e objetiva.

1.6 - METODOLOGIA DA DISSERTAÇÃO

Uma vez aclarados os objetivos e os motivadores do presente trabalho, é necessário elucidar como os objetivos serão alcançados. Observando o postulado por Gil (2007), pode-se classificar o presente trabalho como uma pesquisa exploratória experimental, desenhado através de uma delimitação de cinco etapas, apresentadas conforme a Figura 1.1 e descritas a seguir.

Na Etapa 01 – Revisão Bibliográfica –, foram realizadas as pesquisas com intuito de construir as bases teóricas para a condução do presente trabalho. O levantamento bibliográfico incluiu conceitos e temas relacionados à coleta automatizada de dados, à estrutura básica dos sistemas inteligentes de transporte (*Intelligent Transportation Systems – ITS*), ao Planejamento da Operação dos Transportes Públicos Urbanos e ao planejamento dos transportes urbanos.

Na Etapa 02 – Definição do Método –, por meio das bases teóricas levantadas na Etapa 01, foi possível delimitar, com precisão, o método proposto por este trabalho. A definição concisa do método, sua cobertura e suas abstenções, com vistas à objetividade, foram descritas nesta etapa.

Na Etapa 03 – Pesquisa Exploratória –, foram observadas as pesquisas mais recentes que incluíram a coleta de dados dos transportes públicos rodoviários, com o intuito de caracterizar e de formar uma matriz de dados recorrentes e relevantes. A posteriori, foi realizada a compatibilização dos formatos dos dados coletados e o imediato paralelismo com possíveis famílias de sensores e/ou métodos consagrados, que poderão ser utilizados na coleta automatizada de dados.

Na Etapa 04 – Validação –, com intuito de verificar a possibilidade de aplicabilidade dos requisitos híbridos gerados na etapa anterior, foi desenvolvido um “programa tradutor” que

transcreve os dados coletados das empresas gestoras dos transportes públicos. Após a formatação destes, de forma a atender aos requisitos ora elencados, foram realizadas análises comparativas, fusões das tabelas, expansões e serialização dos dados, confirmando sua aplicabilidade.

Na Etapa 05 – Análise dos Resultados –, buscou-se a interpretação dos resultados obtidos no processo de simulação em cenários variados e ambientes conhecidos, a fim de identificar a assertividade na tarefa de fornecimento de dados relevantes para a Operação dos Transportes Públicos, validando o modelo proposto e os requisitos levantados.

Assim, a presente dissertação está estruturada em oito capítulos. O primeiro capítulo, conforme se observa, apresenta os elementos básicos que guiaram e motivaram o desenvolvimento da pesquisa.

O Capítulo 2, por sua vez, apresenta uma revisão bibliográfica acerca dos Sistemas Inteligentes de Transportes e das coletas automatizadas de dados aplicados aos transportes públicos. Em complementação, o Capítulo 3 traz o arcabouço teórico quanto ao planejamento dos transportes, segmentando o Planejamento da Operação dos Transportes Públicos Urbanos em seus meios tradicionais e os avanços nessa seara.

No Capítulo 4, é feita uma exaustiva revisão do método híbrido proposto, suas limitações e sua aplicabilidade. Já o Capítulo 5 apresenta uma revisão das pesquisas realizadas recentemente, traçando um paralelo entre as variáveis recorrentemente coletadas.

A partir das bases teóricas e da delimitação do método, o Capítulo 6 detalha a determinação de padrões dos formatos de dados a serem coletados, bem como as famílias de sensores que podem ser utilizados no intento de coletar tais dados. Esse capítulo apresenta ainda a estrutura central dos requisitos para uma coleta de dados híbrida.

No Capítulo 7 é feita a verificação da aplicabilidade dos requisitos através de dados sintéticos traduzidos a partir de dados reais fornecidos, bem como são descritos os simuladores criados, os cenários testados e os resultados obtidos para cada cenário. Por fim, no Capítulo 8 são apresentadas as conclusões e as recomendações finais quanto ao desenvolvimento da pesquisa.

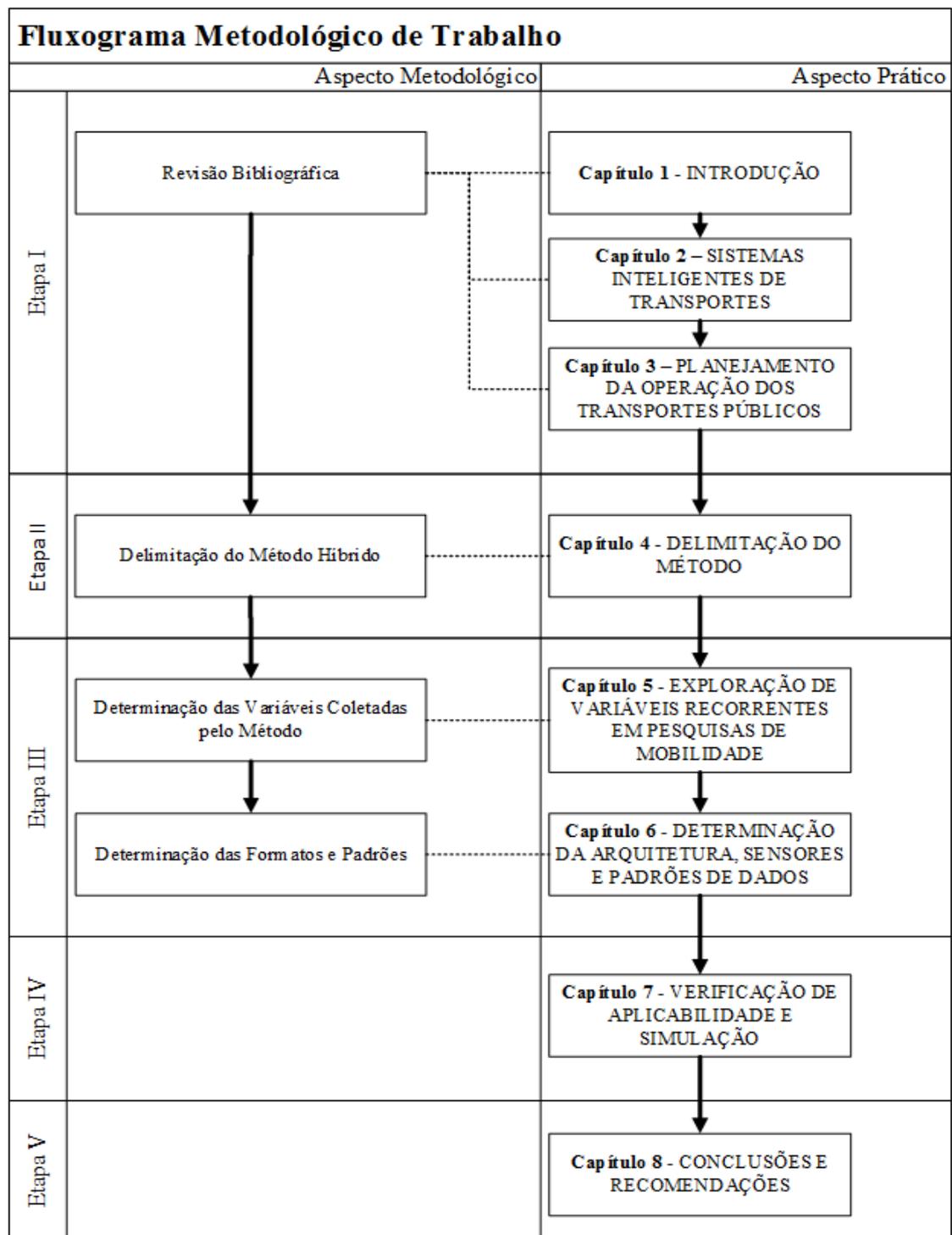


Figura 1.1 - Fluxograma Metodológico e Prático do trabalho

O presente trabalho pode ainda ser compreendido através do diagrama apresentado na Figura 1.2. Essa figura apresenta, na direção da seta de fluxo lógico, a sequência de passos adotados para a concretização do presente trabalho.

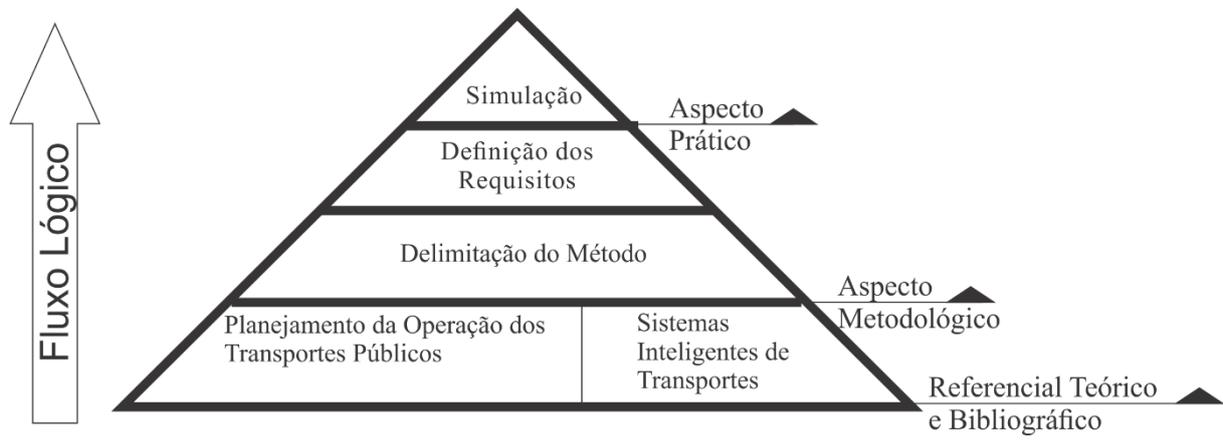


Figura 1.2 - Fluxo lógico para a construção desta dissertação

2 - SISTEMAS INTELIGENTES DE TRANSPORTES

2.1 - APRESENTAÇÃO

Os Sistemas Inteligentes de Transportes (*Intelligent Transportation Systems – ITS*) são um conceito amplo usado para indicar todo tipo de equipamento e/ou software que adicione facilidades e funcionalidades aos atores dos transportes.

Como a presente dissertação trata de apresentar os requisitos para um modelo de coleta automatizada de dados híbridos, é relevante caracterizar, de maneira clara, suficiente e objetiva os ITS e, assim, posicioná-lo de maneira assertiva em meio aos métodos de coleta automatizada.

Assim, este capítulo apresentará o conceito geral de Sistemas Inteligentes de Transportes, voltado para a ótica dos transportes. Posteriormente, serão caracterizados os métodos de coleta automatizada.

2.2 - CONCEITO DE SISTEMAS INTELIGENTES DE TRANSPORTES

Os Sistemas Inteligentes de Transportes se referem à aplicação de tecnologias de informação, tais como programas de computador, equipamentos, tecnologias de comunicação, dispositivos de navegação e eletrônica para melhorar a eficiência e a segurança dos sistemas de transportes (Hoel *et al.*, 2011).

Sistemas Inteligentes de Transporte é um nome usado para descrever o estado da arte de sistemas complexos para melhorar a experiência de mobilidade. Estes sistemas envolvem veículos, motoristas, passageiros, operadores rodoviários, gestores, e a interação entre eles e com o meio ambiente, integrada com toda a complexa infraestrutura de transporte.
(Williams, 2008)

Dada a pluralidade de atores (veículos, motoristas, passageiros, operadores e gestores) e do ambiente que os une, diversas formas de agregação podem ser realizadas com o intuito de compreender os ITS. Pelo viés de sua finalidade, os equipamentos de ITS podem ser divididos em duas grandes áreas: Sistemas Avançados de Gerenciamento da Viagem (*Advanced Travel Management Systems – ATMS*) e Sistemas Avançados de Informações da Viagem (*Advanced Travel Information Systems – ATIS*).

Por um lado, os sistemas e os equipamentos voltados para ATMS se preocupam em fornecer subsídios para o gerenciamento da infraestrutura, para a engenharia de tráfego e para o

planejamento, sejam dados ou informações processadas. Os equipamentos e os sistemas de ATMS, tratam tanto da coleta de dados quanto do processamento destes, com o intuito de decidir e/ou auxiliar a tomada de decisão por parte dos responsáveis pela Operação dos Transportes.

Por outro lado, os sistemas ATIS se voltam para a tradução das ações táticas geradas pelos ATMS e pelos operadores, sejam estas imperativas ou informativas, com o intuito de prover instruções úteis para os usuários dos transportes.

Uma maneira de elucidar o conceito sinérgico entre ATMS e ATIS é através dos sistemas semaforicos atuados. Os sistemas ATMS coletam dados sobre o fluxo de veículos num determinado conjunto de vias, a velocidade média, a dispersão, os pelotões formados e as possíveis anomalias, comparam com seu banco de dados histórico e calculam qual o tempo de ciclo ideal para um conjunto de semáforos. O cálculo, então, é validado pela equipe de engenharia de tráfego responsável pelo controle tático. O semáforo, por sua vez, que é um equipamento ATIS de informações imperativas, apresenta para o usuário os sinais luminosos com instruções de paradas ou de fluxo prioritário, com base nos tempos dimensionados pelos ATMS.

Inserido no conceito dos ATMS se observam dois tipos distintos de equipamentos e softwares: aqueles voltados para a análise e processamento e aqueles voltados para a coleta de dados. Cabe salientar que o segundo tipo alimenta o primeiro tipo com as informações reais dos atores.

Considerando-se a segmentação dos equipamentos de coleta para ATMS, três tipos distintos são perceptíveis, a saber: coleta automatizada de dados (*Automatic Data Collection – ADC*), localização automática de veículos (*Automatic Vehicle Location – AVL*) e equipamentos de tarifação automática (*Automatic Fare Collection – AFC*). Cada dos referidos equipamentos pode estar embarcado, ou seja, dentro de um veículo, ou em postos fixos de coleta.

A cadeia de ITS, com foco em equipamentos de coleta para sistemas de ATMS, pode ser melhor compreendida através da Figura 2.1.

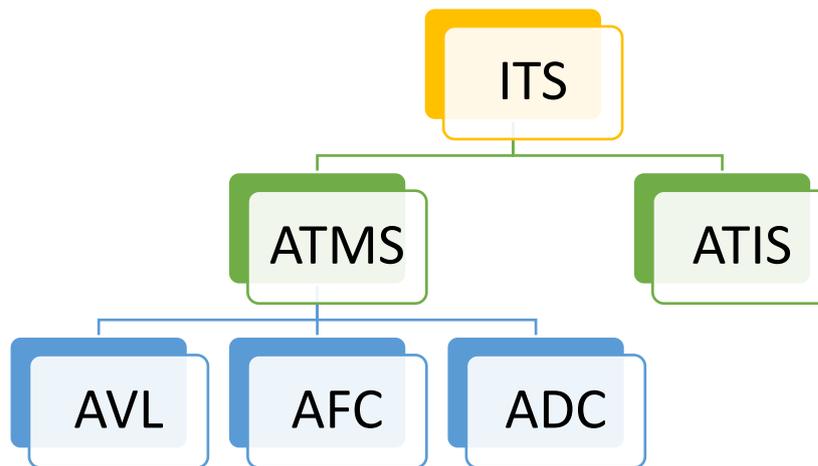


Figura 2.1 - Cadeia de concentrações dos equipamentos de ITS.

Nos instrumentos para a coleta de dados para ATMS embarcados se verifica que estes quase sempre possuem mais de uma especialização. Como exemplo, têm-se os instrumentos de ADC que coletam dados provenientes de AVL.

Entretanto, esse formato de arquitetura é *ad hoc*, ou seja, são casos isolados, sem padronização ou constância, ainda que dentro do mesmo órgão ou empresa. A arquitetura de equipamentos, os dados que devem ser coletados, o formato dos dados e as demais variáveis inerentes ao processo de coleta raramente se repetem, de modo a permitir uma análise através do tempo.

Evidencia-se, portanto, que os ITS podem se especializar em diversas ramificações dentro das suas finalidades *mater*, especialmente aqueles que se propõe a fornecer subsídios para o planejamento e a operação dos transportes. O item a seguir demonstrará as principais características dos sistemas AVL, AFC e ADC e a aplicabilidade destes para os transportes públicos rodoviários.

2.3 - MÉTODOS DE COLETA AUTOMATIZADA DE DADOS

Dadas as identidades encontradas na etimologia de ambas as palavras, não é um equívoco considerar os ADC como sinônimo de sistemas de telemetria. Conforme apresentado por Mattos (2004), a história da telemetria é deveras extensa e rica, sendo quase tão plural quanto suas aplicações.

Em linhas gerais e objetivas, pode-se considerar que a telemetria surgiu no início de 1910, com o intuito de monitorar a distribuição de energia elétrica em Chicago. Por volta de 1940, a tecnologia passou a ser aplicada ao setor aeroespacial, com o objetivo de captar os dados dos instrumentos aviônicos, para análise de desempenho posterior e, também, investigação de

acidentes. Nos anos de 1980, surgiram as primeiras soluções de telemetria para veículos terrestres de carga.

Com o avanço tecnológico, não demorou muito para que os primeiros equipamentos de ADC passassem a integrar os transportes públicos, com o mesmo intuito que os fez aparecer no setor aeroespacial. Atualmente, a quase totalidade de veículos comercializados possuem equipamentos embarcados para a coleta e análise de dados, sendo comumente chamados de “computadores de bordo”.

Embora haja vertentes que defendam a existência de unidades de contagem automática de passageiros (*Automatic Passenger Counter Unit – APCU*) como um método de coleta automatizada em si, como Perk e Kamp (2003), é muito mais comum que esse equipamento seja considerado como uma unidade pertencente ao ADC e, portanto, parte do mesmo. Dessa maneira, o presente trabalho tratará as unidades de APCU como módulo dos equipamentos de ADC.

Quanto aos sistemas baseados na tarifação, estes apareceram experimentalmente em 1964, nas estações de trem de Forest Hills e Kew Gardens, no bairro do Queens - Nova Iorque. A ideia inicial dos *faregates* (portais que faziam a cobrança automatizada das tarifas) era reduzir os custos operacionais do trem, eliminando os agentes ferroviários. Entretanto, as ideias iniciais foram superadas quando foi observado que, analisando-se a estrutura de dados coletados, poderia se obter informações importantes sobre o comportamento dos passageiros, auxiliando, inclusive, o planejamento das operações (Deibel e Zumwalt, 1984).

Sem muitas alterações no método, mas com modos cada vez mais complexos, os AFC são largamente usados nos transportes sobre trilhos, tendo ainda uma forte, porém não tão abrangente, vertente nos transportes públicos sobre pneus. Observam-se esforços, como os apresentados por Baroni *et al.* (2016) para que os métodos de AFC se tornem populares, dinâmicos e ofereçam uma grama de dados consideráveis, sendo aproveitáveis em maior escala para ônibus e metrô.

As tecnologias de posicionamento dos veículos baseadas no uso dos GPS embarcados, são as mais recentes das vertentes indicadas. O uso destas se iniciou após a liberação dos sistemas de GPS para fins civis. A miniaturização e a popularização dos equipamentos de GPS, o aparecimento dos GPS embutidos nos aparelhos de smartphone e a coleta de dados somente do posicionamento do veículo, sem dados adicionais de telemetria, deram início à AVL.

Os processos que manejam os dados fornecidos pelas AVL geradas por aplicativos de celular vem se tornando cada dia mais populares, especialmente para a busca de rotas alternativas para o trânsito.

Na seara dos métodos de AVL, as formas atualmente utilizadas para rastrear padrões de deslocamento podem ser divididas em três categorias: baseada em viagens, baseada em atividades e baseada em movimento.

Observa-se que o comportamento de viagens baseado em movimento, no qual as trajetórias são coletadas através do tempo, permite que vários recortes de interesse sejam realizados, como a reconstrução das viagens (Vlassenroot *et al.*, 2014), a identificação do final da viagem (Du e Aultman-Hall, 2007), a validação dos modos de viagem, bem como os propósitos (Bohte e Maat, 2009). Ainda, podem ser utilizados para a predição do destino de novas viagens através dos registros do GPS (Alvarez-Garcia *et al.*, 2010), conforme citado por Wang *et al.* (2015).

Transpondo a pontuação das características de cada um dos métodos de coleta, é importante verificar sua utilização no planejamento da operação de transportes públicos rodoviários e na publicação de material científico relevante acerca do tema.

Assim, utilizando-se do Sistema de Inteligência Artificial voltado para artigos científicos, desenvolvido pela *Allen Institute for Artificial Intelligence*, verificou-se quais os artigos mais relevantes nos últimos oito anos que tratam da prática da coleta de dados aplicada aos transportes. Os resultados podem ser visualizados na Tabela 2.1.

Tabela 2.1 - Artigos mais relevantes no que tange a coleta de dados para os transportes nos últimos oito anos, segundo a inteligência artificial.

Autores	Método de Coleta			Objetivo do Trabalho
	ADC	AFC	AVL	
Alvarez-Garcia <i>et al.</i> (2010)			X	Destino da viagem
Bohte e Maat (2009)			X	Propósito da viagem
Du e Aultman-Hall (2007)	X		X	Fim da viagem
Ma <i>et al.</i> (2012)		X		Origem dos passageiros
Munizaga e Palma (2012)		X	X	Inferência de matriz OD
Nitsche <i>et al.</i> (2012)			X	Reconstrução das viagens
Vlassenroot <i>et al.</i> (2015)			X	Comportamento de viagens
Wang <i>et al.</i> (2011)	X		X	Inferência de matriz OD

Observando-se os dados coletados nos artigos supracitados, foi possível identificar a aplicação de cada um dos métodos, trazidos na Tabela 2.2.

Tabela 2.2 - Tipos de dados coletados pelas diferentes metodologias clássicas de coleta de dados.

DADO DE INTERESSE	MÉTODO			
	ADC	APCU	AFC	AVL
Métricas veiculares (telemetria veicular)	X			
Dados de conforto relativos ao veículo	X			
Dados ambientais internos	X			
Dados ambientais externos	X			
Contagem de passageiros		X*	X	
Localização de embarque			X**	
Paradas do veículo (programadas ou não)				X
Tempo de paradas				X
Controle de tempo de circuito				X
Restrições comuns na rota				X
Restrições extraordinárias na rota				X

* - Considerado como uma subunidade do ADC e, portanto, dependente da existência do mesmo;

** - Dado coletado apenas quando a instalação está conjunta com um sistema de AVL e os *timestamping* estão completamente sincronizados.

De acordo com a Tabela 2.1 e com a Tabela 2.2, infere-se que a coleta da localização dos veículos aparece nos principais trabalhos. No entanto, cada um deles apresenta uma diferente finalidade, deixando clara a capacidade de um mesmo tipo de dado em ser utilizado para os mais diferentes objetivos. A principal diferença observada nas publicações apreciadas se trata da granularidade dos dados coletados via GPS, que variam sua frequência para atender aos objetivos declarados.

No trabalho conduzido por Ma *et al.* (2012) é coletada a identificação de um determinado cartão de ingresso no sistema de transportes públicos (passagem) no momento de seu acesso ao sistema, identificando, assim, sua origem para uma determinada viagem. Esse mesmo dado é coletado no trabalho de Munizaga e Palma (2012) que, associado à localização do veículo no momento do desembarque do passageiro, fornece a possibilidade da inferência da matriz de origem e destino de cada indivíduo.

Wang *et al.* (2011) e Du e Aultman-Hall (2007) coletaram dados provenientes da telemetria do veículo, como forma complementar de avaliar o início e o fim das viagens. Essa forma indireta é possibilitada pelos aspectos que denotam a movimentação dos veículos.

Pode-se observar que, conforme mencionado em itens anteriores, não há qualquer tipo de constância ou padronização nos dados coletados. Entretanto, há uma evidente propensão para esta ação, avultada pela recorrência de publicações que utilizam mais de um método para concretização dos seus objetivos, de acordo com o que se verifica na Tabela 2.1.

2.4 - TÓPICOS CONCLUSIVOS

O capítulo apresentou a teoria geral dos sistemas inteligentes em transportes, focando primordialmente em seu aspecto prático e comum no âmbito dos transportes. Foram levantados ainda, dentro dos ATMS, quais os modelos gerais de coleta de dados numa divisão pelo tipo de dado coletado e pela capacidade de coleta destes.

A capacidade ventilada neste capítulo através da revisão bibliográfica, demonstra que as formas de coleta existem de maneira estanque, sem que haja um modo híbrido declarado, sendo que a utilização conjunta de mais de um método, embora recorrente, ocorre de maneira a atingir os efeitos propostos no bojo do trabalho proposto. Esta predisposição, *per se*, é elemento suficiente para justificar o esforço em unificar os métodos distintos num padrão híbrido de maneira eficiente.

Para que esse esforço seja embasado do ponto de vista do Planejamento da Operação dos Transportes Públicos Urbanos rodoviários, o capítulo a seguir tratará da revisão bibliográfica sobre este tema, bem como do seu principal desafio: o fornecimento de dados confiáveis de maneira rápida e contínua.

3 - PLANEJAMENTO DA OPERAÇÃO DOS TRANSPORTES PÚBLICOS URBANOS

3.1 - APRESENTAÇÃO

A forma clássica de aplicação do planejamento e da operação dos transportes públicos, com pesquisas estanques e periódicas focadas na elaboração de planos de trabalho, embora resistam na maioria das situações, não cabem mais nas cidades modernas. Da mesma maneira, o modo arcaico da operação remediadora contrasta negativamente com o dinamismo apresentado pelas populações contemporâneas.

Mesmo as cidades com tecido urbano fortemente baseado em outras visões estão alterando sua forma de gerir os serviços prestados aos cidadãos, adotando o dinamismo na tomada de decisões e na transparência no fornecimento de informações. Essa realidade está presente, principalmente, na operação dos transportes públicos.

Da alteração de paradigmas emerge a ideia de *smart cities*, cujo conceito, de tão novo, não é unânime entre os pesquisadores, graças ao diversificado embasamento teórico apresentado por cada um deles (Prado e Santos, 2014). Embora não haja consenso sobre a definição, as premissas atreladas à imagem de *cidades inteligentes* passaram a permear as ações dos operadores dos transportes públicos em algumas cidades brasileiras.

Cabe, portanto, neste capítulo apresentar o conceito de planejamento da operação dos transportes e sua aplicabilidade na manutenção da fluidez do tecido urbano, bem como trazer à baila as benesses do uso da entidade de instrumentação para os mais diversos usos da operação dos transportes, mesmo quando aplicadas fora da ideia de *smart cities*.

3.2 - CONCEITO DO PLANEJAMENTO DA OPERAÇÃO DOS TRANSPORTES

A demanda por transporte é um item derivado e não um fim em si. Com a possível exceção de passeios específicos, como os safaris, as pessoas viajam a fim de satisfazer suas necessidades de deslocamento para locais e atividades específicas, como trabalho, lazer e saúde (Ortúzar e Willumsen, 2011).

A partir dessa noção primordial, infere-se que a necessidade de se transportar é derivada da vontade do transportado e, como subjetiva, podem se alterar de acordo com fatores externos e implícitos.

Assim, a maneira mais eficiente de prover a infraestrutura adequada para os deslocamentos é entender e quantificar as demandas por viagens de uma determinada população, mesmo nos

casos nos quais excepcionalidades mudam o quadro geral de comportamento. Essa forma de aprovisionar infraestruturas, ainda que em tempo real, se repete na esfera dos transportes públicos.

Conforme Pires *et al.* (1997) o planejamento de transporte é a atividade definidora das infraestruturas necessárias para garantir a mobilidade e organizadora dos sistemas que estão sujeitos à regulamentação pública de qualquer espécie, inclusive, quanto à tecnologia e ao nível de serviço ofertado.

Os autores apresentam, ainda, que o planejamento dos transportes procura responder a algumas questões básicas daqueles que administram os transportes:

- *Como é possível planejar o futuro do sistema de transporte da cidade? Quais informações são necessárias?*
- *Como se faz um plano de transporte?*
- *Como compatibilizar a oferta de transporte com o crescimento da cidade?*
- *Necessidade de organização do transporte urbano no município;*
- *Necessidade de organização de sistemas de transporte (coletivo, táxi, carga);*
- *Necessidade de compatibilizar ações urbanas, de transporte e de trânsito.*

Partindo do princípio de que o planejamento dos transportes pode variar quanto à abrangência, ao nível de serviço e aos prazos de execução, tem-se que o planejamento da operação, ou o planejamento tático, é aquele que compreende os procedimentos para que os objetivos sejam atingidos, através de medidas que consideram o sistema em operação, com baixo ou nenhum custo envolvido (Pires *et al.*, 1997).

Fica claro, portanto, que inclusive as ações de intervenção tática, como inversões no sentido de tráfego, atuação em semáforos, injeção ou recolhimento de veículos da frota de transportes públicos, são atos não aleatórios, ou seja, são planejados e fundamentados em métodos que visam a otimização do sistema de transporte.

Considerando-se o apresentado por Bruton (1979), o planejamento de transportes é baseado numa série de hipóteses e princípios, dentre os quais, os primordiais são: (a) os padrões de viagens são tangíveis, estáveis e previsíveis; e (b) as demandas por movimentos estão diretamente relacionadas com a distribuição e a intensidade dos usos do solo, que são possíveis de serem determinadas, com exatidão, em alguma data futura.

Embora o cerne dos princípios apresentados por Bruton permaneça válido, observa-se que o fenômeno de transporte tem se tornado mais dinâmico, assim como toda a sociedade. Nota-se, por exemplo, que os comportamentos esperados nas cadeias de viagens variam consideravelmente, dependendo de externalidades (culturais, sociais, territoriais, etc.) e dos fatores econômicos clássicos - tempo, distancia, custo, etc. (Medrano, 2016). Essa flexibilidade no comportamento esperado de viagens altera, em partes, os prolegômenos quanto aos padrões de viagens serem estáveis e previsíveis.

Dessa maneira, assevera-se o apresentado por Vuchic (2005) ao afirmar que um bom planejamento e continuidade na coleta de dados é necessária para o planejamento da operação de transportes e que a frequência e a abrangência das pesquisas de campo devem ser determinadas com base num compromisso entre a necessidade de informações precisas e o custo dos levantamentos.

Assim, intuindo a tomada de decisão tática para abarcar as possíveis mudanças de comportamento de um conjunto de usuários dos transportes, o dado que mais apresenta relevância é a matriz de origens e destinos. Essa matriz, em sua forma clássica, possui uma distribuição estimada do volume de viagens de uma determinada origem para uma lista de possíveis destinos predeterminados ou intuídos através do comportamento imediato das populações.

Esse molde, que satisfaz completamente a primeira hipótese de Bruton supracitada, mesmo considerando o dinamismo ora citado, permite que os operadores dos transportes públicos estimem qual a necessidade de deslocamentos num determinado sentido e num dado horário, permitindo a determinação da quantidade de veículos da frota, das linhas existentes, de intervalo de tempo entre veículos (*headways*) e da previsão de crescimento (ou de decréscimo) da demanda.

Em contrapartida do seu protagonismo, as matrizes de origem e destino são geradas através de pesquisas de origem e destino para os transportes públicos, que tendem a ser realizadas de forma

manual, através de entrevistas direcionadas. As coletas manuais, embora embasem a maioria dos estudos existentes até o momento, exigem grandes esforços para serem levadas a cabo, como o alto custo de realização e a morosidade na execução. Adicionalmente, essa forma de trabalho apresenta uma série de problemas como a possibilidade de erro na coleta e na transcrição/transmissão dos dados, que pode gerar uma série de inconsistências na expansão estatística, as limitações quanto à duração da pesquisa e uma gama de dados coletados bastante limitada.

Como forma de mitigar os problemas ocasionados pela coleta manual de dados para a formação das matrizes de origem e destino (OD), em especial, no que diz respeito à morosidade e à dispendiosidade, foram elaborados métodos para estimativa e atualização, dinâmica ou não, dessas matrizes.

Os meios de estimativa das matrizes OD, em geral, baseiam-se na expansão de uma matriz origem e destino menor e não necessariamente recente, chamada de matriz semente, utilizando dados provenientes de diferentes pesquisas complementares e menos dispendiosas, como pesquisas de sobe e desce (Ben-Akiva *et al.*, 1985). Uma vez que as pesquisas de sobe e desce são mais baratas e facilmente aplicadas, essa é uma saída amplamente utilizada pelos municípios com o intuito de atender suas necessidades de planejamento.

Quando os operadores e gestores dos transportes públicos trabalham em cenários sensíveis ou críticos, como os grandes centros urbanos, a estimação das matrizes de OD pode se tornar um trabalho dinâmico, através da coleta de dados automática de sobe e desce de passageiros nas paradas e nos terminais rodoviários.

Um exemplo de como essa atualização dinâmica pode ser usado eficientemente está relacionado com eventos inesperados e extraordinários, como greves que paralisem o sistema de transportes, mesmo que em partes. O comportamento dos usuários pode se alterar para evitar uma área que não esteja sendo servida por transportes públicos, mudando o perfil da matriz OD, forçando os operadores a aumentar o volume de viagens em outras direções, com o intuito de atender à demanda extraordinária. A partir da série histórica desse comportamento anormal, planos de emergência e contenção podem ser elaborados visando essas necessidades.

A geração e a atualização de matrizes de origem e destino perdurou por um período considerável trazendo resultados minimamente aceitáveis. Entretanto, a partir dos meados dos

anos 1990, a forma com que as pessoas interagem com o espaço urbano se modificou, tornando o dinamismo dos destinos das viagens algo fugaz e recorrente.

Esse movimento revel aos padrões estáticos passou então a contrastar com a morosidade e lentidão dos operadores em alterar sua atuação nos meios urbanos para a resolução de problemas. Do ponto de vista urbano, os conceitos de cidade foram revistos com o intuito de aplacar essa nova postura. Entre as novas formas de pensar o espaço urbano, ganhou espaço o conceito de cidades inteligentes (*Smart Cities*).

Komninos *et al.* (2011) conceituam que uma cidade inteligente é a congregação dos princípios de interconexão, instrumentação e inteligência, na qual a instrumentação é a entidade que permite a captura de dados em tempo real usando todo tipo de sensores, a interconexão é o elemento com o papel integrador, dando significância para os diferentes atores da gestão urbana e a inteligência, neste caso, refere-se à inclusão de técnicas analíticas complexas, à modelagem de dados, à otimização de processos e à visualização de resultados para que melhores decisões operacionais possam ser tomadas (Nam e Pardo, 2011).

Assim, formas dinâmicas de pensar na operacionalização dos transportes públicos nos centros urbanos passou a se modificar para acompanhar esse conceito. Neste ponto, os elementos de coleta de dados para ATMS passaram a colaborar para a formação de matrizes de origem e destino, com maior confiabilidade a um menor custo agregado, agindo como elementos de instrumentação.

Enquanto elementos de ADC passaram a fornecer dados ricos sobre o desempenho, saturação/ociosidade da frota e aspectos ambientais, equipamentos de AFC permitiram que fossem materializados vetores de embarque e equipamentos de AVL possibilitaram que a frota fosse monitorada. Esses avanços, embora isolados em suas técnicas, permitiram a aproximação da operação dos transportes públicos a meios dinâmicos e que melhor se aproximam da realidade.

Com finalidades didáticas e comparativas, os subitens a seguir tratarão de apresentar as generalidades tanto dos métodos de estimativa e expansão para as matrizes de origem e destino de maneira clássica quanto das características e das limitações dos métodos automatizados aplicados à obtenção dessas matrizes.

3.3 - TÉCNICAS PARA ESTIMATIVA DE MATRIZES DE ORIGEM E DESTINO

Dados para o planejamento, a análise, o desenvolvimento, o monitoramento e a operação dos serviços de transportes podem ser obtidos através de diversas fontes (Ben-Akiva, 1987). Essa pluralidade de fontes implica em matrizes de diferentes dimensões, com os mais diversos refinamentos dos dados colhidos. Ainda assim, existem inúmeras maneiras de combinar essas diferentes fontes, objetivando a viabilidade econômica e operacional de matrizes de OD (Cui, 2006).

Dada a considerável semelhança com os métodos para a estimativa das matrizes de origem e destino para os transportes individuais, durante certo tempo, fizeram-se aproximações dos métodos produzidos para atender às demandas de dados apresentadas pelos transportes públicos.

Um exemplo de estudo voltado à estimativa de matrizes OD de veículos particulares que foi adaptada para o transporte público, está presente no trabalho de Willumsen (1978). Nesse trabalho, o autor apresenta três abordagens para estimativa das matrizes de origem e destino. A primeira abordagem utilizou o modelo gravitacional, no qual o problema estava reduzido a calibrar a atração gravitacional dos polos, apresentando soluções por regressões lineares e não lineares, dependendo das premissas assumidas pelo modelo.

Posteriormente, Ben-Akiva *et al.* (1985), descreve uma diversa gama de tratamentos baseados em métodos estatísticos para o trabalho com matrizes de origem e destino de passageiros, dos quais se destacam as técnicas de Ajuste Proporcional Iterativo (*Iterative Proportional Fitting* - IFP) e de estimador de máxima verossimilhança (*Maximum Likelihood Estimator* - MLE). Tanto o método IFP quanto o método MLE são novamente descritos e devidamente ampliados por Ben-Akiva (1987) em outras publicações.

Por um lado, a principal dificuldade com a utilização da técnica de IFP é a probabilidade de que um número significativo de pares OD potenciais numa rede tenham pouca, ou nenhuma, viagem entre elas, o que pode gerar distorções nos resultados obtidos. O método é geralmente usado na estimativa de uma única matriz OD, através da combinação de uma matriz semente com pesquisas de embarque e desembarque e total de contagens. Este modelo requer, ainda, uma alta qualidade da contagem de embarque e desembarque, assim como o controle dos totais, pois a precisão da estimativa depende da confiabilidade dos totais de controle (Cui, 2006).

Por outro lado, a técnica de MLE necessita que cinco premissas básicas sejam assumidas para a validade do método:

1. Os passageiros chegam individualmente de acordo com uma distribuição de Poisson;
2. Os vieses encontrados na checagem de viagem via APCU são negligenciáveis;
3. Os comportamentos dos passageiros correspondem por um processo binomial;
4. Todos os passageiros podem embarcar no primeiro ônibus disponível; e
5. As contagens de embarque e desembarque são independentes umas das outras.

Os pressupostos de 1 a 3 são consideravelmente razoáveis na maioria dos casos, enquanto que os demais, geralmente, não são totalmente satisfeitos sem grande prejuízo dos resultados obtidos. Conclui-se que a técnica de MLE é preferível em comparação a outras técnicas, como a MLE de binomial negativa, que requer menos pressuposições, porém é mais complexa para simulação computacional (Cui, 2006).

Essas duas técnicas são as mais usadas na estimativa de matrizes de origem e destino até hoje, sendo comum notar que métodos mais modernos se tratam de calibrações desses métodos específicos.

Com o advento da popularização de ferramentas computacionais para análise e equipamentos eletrônicos para a coleta de dados, o volume de dados coletados para as matrizes de apoio se tornou incompatível com a aplicação direta das metodologias estatísticas ora elencadas.

Ben-Akiva e Morikawa (1989), apresentam um método de fusão de dados, que visa a combinação de duas ou mais fontes de informações num banco de dados único, tangível e consistente, com matrizes com multicolumnas e bidimensionais. Embora o método apresentado seja multifuncional em conceito, sua aplicabilidade essencial está em reduzir os erros amostrais presentes nas coletas de dados individuais.

Os estudos citados apresentam um excelente embasamento estatístico, com resultados consideravelmente fiéis à realidade. No entanto, o trabalho para estimativa das matrizes requer mão de obra com um profundo conhecimento em estatística, uma vez que os cálculos envolvidos são complexos e podem levar a equívocos com facilidade. Essa necessidade profissional limita o acesso a tais métodos para pequenas e médias localidades,

O maior problema desses modelos é o fato estatístico em si. Embora os erros possam ser consideravelmente reduzidos em vista do volume de dados coletado, bem como das correções que os métodos apresentam sem a possibilidade do erro latente, essa possibilidade de erro cresce substancialmente quando considerada a não-linearidade da demanda por transportes. Isso significa que, em diferentes períodos, o comportamento dos usuários nos transportes muda, o que pode limitar, ou mesmo enviesar, o resultado das expansões das matrizes semente.

Ademais, os resultados dos métodos existentes estão contidos em uma única matriz, impedindo recortes e análises menos ortodoxas e/ou coordenadas por fatores externos, como períodos de chuva, períodos de greve de um determinado setor ou mesmo por períodos de um determinado dia.

Essas limitações poderiam ser vencidas se, ao invés de uma matriz estimada construída sobre uma matriz semente, fosse realizada uma coleta diária e permanente de dados. Isso faria com que as análises estimativas se transformassem em análises populacionais e, portanto, livres de erros e vieses das expansões, de fácil interpretação e manipulação, através de mineração de dados e outras ferramentas aplicáveis ao *big data* formado por essas coletas, além de possibilitar um número de recortes virtualmente ilimitados.

Haja visto as limitações apresentadas acerca das pesquisas de origem e destino, essa coleta perene somente se viabiliza através de métodos automatizados de coleta de dados. Os métodos existentes e suas limitações serão exploradas no item a seguir.

3.4 - APLICAÇÃO DE TÉCNICAS PARA OBTENÇÃO DE MATRIZES OD ATRAVÉS DE COLETAS AUTOMATIZADAS

Com o aumento da disponibilidade de dados de trânsito coletados automaticamente por meio de ADC, expandiu-se o uso das ferramentas estimativas apresentadas anteriormente. Em contraponto a esse aumento de disponibilidade de dados, diferentes agentes envolvidos na coleta de dados dos transportes têm disponibilidade de dados diferentes e com características e acurácias diferentes (Cui, 2006).

A pluralidade de fontes pode induzir os responsáveis pela determinação das matrizes OD ao erro, especialmente quando não se consideram os vieses e os erros gerados pela sazonalidade, pelas diferentes datas de coleta e pelos demais intervenientes.

Os potenciais benefícios da utilização de ADC para o planejamento de transportes públicos foram alvo da pesquisa de Wang *et al.* (2011), obtendo resultados consideravelmente positivos, especificamente usando dados ADC arquivados para inferir matrizes OD, a fim de avaliar o desempenho do serviço para o planejamento de operações.

Entretanto, a expansão de matrizes de origem e destino a partir das matrizes semente, usando dados provenientes somente de ADC, apresenta dois problemas essenciais. O primeiro deles é o erro estatístico, que aumenta cumulativamente a cada expansão realizada, uma vez que a coleta proposta no trabalho citado anteriormente foi estanque e limitada. O segundo problema se trata das informações relativas à localização e ao deslocamento dos veículos, que não são coletados, impossibilitando análises de tempo de parada, tempo de ciclo e demais assuntos associados a eles.

As tecnologias de AFC, por sua vez, quando aplicadas aos transportes públicos por ônibus urbano, apresentam o inconveniente de contarem, de maneira generalizada, apenas com os dados dos embarques dos passageiros (Ma *et al.*, 2012). Isso leva, impreterivelmente, à estimativa quanto ao local do desembarque dos passageiros, além da perda de dados precisos acerca do posicionamento e dos tempos de viagem dos veículos.

Pensando-se, então, em sistemas de AVL, a informação gerada somente passa a ter valor para o planejamento da operação quando associada aos outros métodos, uma vez que a localização de um veículo por si só é uma informação que diz pouco para o Planejamento da Operação dos Transportes Públicos Urbanos urbanos.

A integração dos dados do sistema AFC, que inclui características de cada transação com cartão de tarifa, com os dados do sistema AVL, que inclui locais de veículos, oferece uma solução através da correspondência entre a informação de localização de veículos e as informações de viagem do passageiro, ajudando os planejadores de trânsito a inferir locais de embarque de passageiros individuais (Wang *et al.*, 2011).

Entretanto, a convergência dos dados de AFC e AVL são extremamente frágeis e complexas. Pequenas discrepâncias na estampa de tempo dos dados coletados podem fazer com que o sistema reconheça que um determinado passageiro embarcou em um ponto de ônibus diferente do real, culminando em equívocos na geração das matrizes de origem e destino.

Moreira-Matias *et al.*, (2015), inferem que diferentes granularidades dos dados de AVL podem fornecer diferentes informações sobre o deslocamento dos transportes. No entanto, esses dados isoladamente não permitem nenhum tipo de inferência quanto às matrizes OD.

Evidencia-se, neste ponto, que os métodos de coleta de dados existentes apresentam lacunas bem delimitadas em seus métodos e dados coletados. Embora existam tentativas de vencer as lacunas deixadas, colocando esses sistemas para funcionar em paralelo, conforme Munizaga e Palma, (2012), Vlassenroot *et al.*, (2015) e inúmeros outros trabalhos publicados, não há um esforço para indicar um método de coleta automatizada de dados híbridos, que congregue em um único equipamento as características relevantes de ADC, AVL e AFC, conforme visualizado na Tabela 2.2.

Na referida tabela, nota-se que a lacuna num dos métodos é justamente a potencialidade de outro método. Isto deixa transparecer de maneira cabal a sinergia e a complementariedade existente entre os métodos, explorada apenas de maneira *ad hoc* até o momento.

Desta forma, evidencia-se a oportunidade para a construção de um padrão de coleta automatizada de dados híbridos (*Automatic Hybrid Data Collection – AHDC*), que congregue as potencialidades dos demais métodos já consagrados, tendo como finalidade básica o fornecimento de informações completas para o planejamento da operação dos transportes. Essa forma de coleta pode ser considerada como equivalente às demais, transformando a cadeia dos ATMS para o formato apresentado na Figura 3.1.

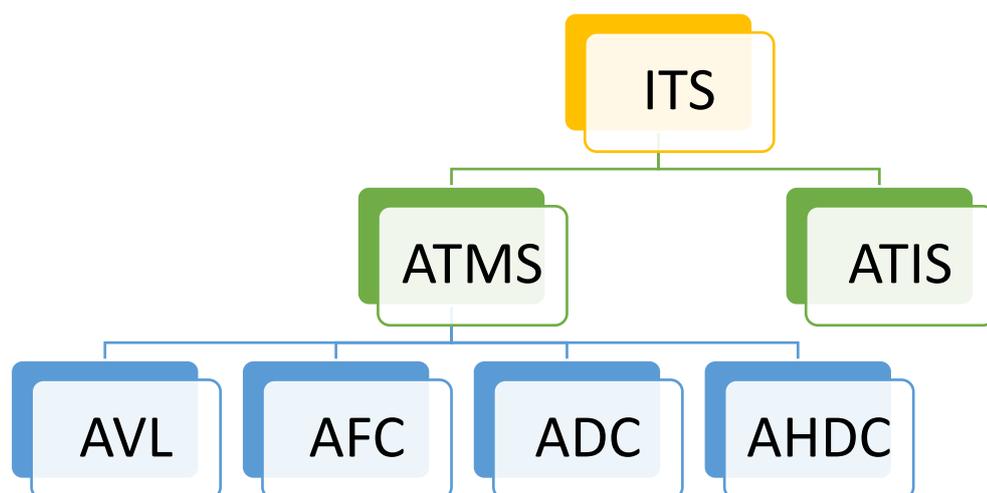


Figura 3.1 - Cadeia de ITS considerando a presença dos métodos de coleta automatizada de dados híbridos.

3.5 - TÓPICOS CONCLUSIVOS

O presente capítulo apresentou o conceito de Planejamento da Operação dos Transportes Públicos Urbanos, que pode ser sintetizado como a “operação tática e imediata” nos elementos passíveis de mediação dentro do tecido urbano. Entretanto, as intervenções devem ocorrer de maneira coordenada e baseada em dados confiáveis para que a tentativa de solução não agrave os problemas existentes.

Até o presente momento, mesmo para as cidades que não aderiram à corrente de evolução das cidades inteligentes, fica evidenciada a dependência de dados coletados pelos atores do sistema de transportes para fomento das tomadas de decisões, como fornecedores de dados. Especialmente, quando se considera que o dinamismo do comportamento das viagens observados impede a inferência de um “comportamento médio esperado” como uma forma plena de trabalho.

Cumpre, portanto, o que esta dissertação se propôs como seu primeiro objetivo: a determinação das tecnologias de coleta de dados emergentes e existentes, apresentadas no capítulo anterior, além de evidenciar essa sinergia no aspecto tecnológico e no aspecto de oportunidade para o Planejamento da Operação dos Transportes Públicos Urbanos na forma de suas aplicabilidades.

Com a evidenciação da coesão entre os assuntos, sem que haja uma forma de coleta híbrida possibilitando acesso aos diversos aspectos do comportamento dos usuários de transportes públicos, emerge a necessidade da definição de uma técnica declarada para tal. Desta maneira o capítulo seguinte apresentará e definirá o método proposto por este trabalho, trazendo à luz sua aplicação, limites e abstenções.

4 - O MÉTODO DA COLETA AUTOMATIZADA DE DADOS HÍBRIDOS PARA O PLANEJAMENTO E OPERAÇÃO DOS TRANSPORTES PÚBLICOS URBANOS

4.1 - APRESENTAÇÃO

A delimitação dos requisitos mínimos para a coleta automatizada de dados híbridos é precedida, necessariamente, pela definição da abrangência do método, de suas limitações e de suas aplicabilidades. Essa necessidade está pautada na ausência de literatura acadêmica que estabeleça um método de coleta automatizada conceitualmente híbrido.

Aclarar esses parâmetros norteadores afasta a interpretação equivocada de uma panaceia universal na temática das coletas automatizadas. Adicionalmente, o desenho dos limites e das potencialidades do método ora apresentado, permite a validação deste e, ainda, a verificação da sua aplicabilidade em escala real.

Dessa forma o presente capítulo apresenta, inicialmente, os conceitos relativos à coleta híbrida de dados e sua delimitação. Posteriormente, trata da sua aplicabilidade prática e das limitações e, finalmente, dos aspectos inerentes à privacidade dos usuários e à legalidade das coletas, no âmbito do planejamento e da operação dos transportes públicos.

4.2 - COLETA AUTOMATIZADA DE DADOS HÍBRIDOS

Pode-se conceituar, de forma geral, que a coleta de automatizada de dados híbridos é uma forma de coleta de dados referente aos atributos da viagem, à localização dos veículos e ao fluxo de passageiros, de modo síncrono e recorrente, sem qualquer tipo de interferência humana no processo de amostragem.

Conforme visto nos capítulos iniciais deste trabalho, é evidente uma predisposição para que os métodos de coleta explícitos pela literatura coexistam num único método com características comuns a cada um deles.

A pertinência da determinação de um método híbrido é demonstrada, ainda, pelas necessidades dos modelos dinâmicos de operação dos transportes públicos que, para seu bom funcionamento e desempenho, requerem dados plurais já cobertos por aqueles coletados nos métodos relatados anteriormente (Wang *et al.*, 2015; Zou *et al.*, 2013).

Tendo os métodos de coleta automatizada para o atendimento dos ATMS como universo de observação, as Coletas Automatizadas de Dados Híbridos (AHDC) podem ser delimitadas

como a interseção entre os métodos preexistentes, agregando as características de coletas de cada uma dessas formas distintas, conforme representado na Figura 4.1.

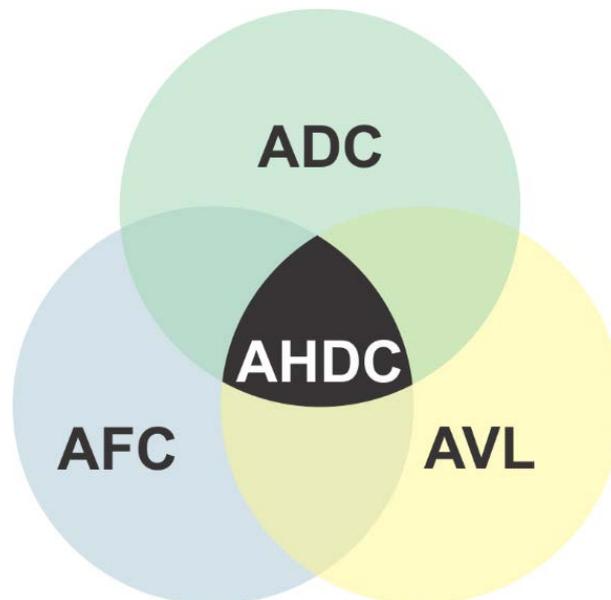


Figura 4.1 - Interseção dos métodos e coleta atuais e o método híbrido

Sob a luz do trabalho de Komninos *et al.* (2011) é perceptível que os equipamentos de AHDC são elementos de instrumentação para o desenvolvimento das cidades inteligentes. Da mesma forma, sua aplicabilidade se estende às cidades desenhadas e com tráfego operado com filosofias diversas.

Embora seja tangível que equipamentos de AHDC possam ser instalados em alguns veículos particulares para que sirvam de amostradores do comportamento desta categoria no trânsito, simplesmente removendo os módulos de AFC e de conforto interno, essa dissertação se abstém de tratar dessa possibilidade. Essa renúncia se justifica pela necessidade de manter o trabalho focado na operação dos transportes públicos e, assim, essa possibilidade deve ser tratada num trabalho a parte.

No contexto dos equipamentos para a coleta híbrida de dados, estas podem ser realizadas *online* ou *off-line*. A primeira forma transmite as informações obtidas (todas ou em partes) para um banco de dados, através de redes de dados móveis, canais de comunicação exclusiva ou internet. A segunda, por sua vez, reserva esses dados num banco de dados interno até que seja realizada uma coleta manual ou programada.

Considerando a finalidade como parâmetro de análise, equipamentos que trabalham conectados são mais propensos a desempenhar papéis correlatos à operação, enquanto a coleta *off-line* se apresenta como mais propensa para trabalhos de planejamento e que não necessitem de intervenção imediata em determinados cenários.

No conceito desenhado anteriormente, a forma síncrona com que os dados são coletados pelos equipamentos de AHDC dizem respeito, especialmente, aos dados provindos de sensores que visam acompanhar e caracterizar as viagens realizadas pelos veículos. Eventos excepcionais, como a entrada ou a saída de passageiros e a validação dos bilhetes nos equipamentos, geram registros considerados extraordinários. Cabe, assim, aos operadores, aos desenvolvedores dos equipamentos e aos demais participantes do processo de elaboração do sistema determinar se os eventos assíncronos geram registros adicionais de todos os sensores ou somente daqueles envolvidos nos eventos inesperados.

Observando a aplicabilidade do método aos transportes públicos, tem-se que a operação dos transportes metroferroviários de passageiros já monitora sua operação de maneira ativa com o intuito do controle e da segurança da circulação de trens. Com o mesmo intuito, os dados relativos aos aspectos de circulação trens, de telemetria e de suas condições são registrados por equipamentos denominados “registradores de eventos” e coletados em rotinas características de cada operador.

Dessa forma, a aplicação de um método similar ao aqui proposto para os transportes metroferroviários de passageiros trata muito mais da determinação de rotinas de trabalho (aspecto humano) e softwares integradores (aspecto de congregação) dos dados já existentes, do que da proposição de um padrão de coleta, como é proposto nesta dissertação.

Assim, concernente à aplicação do método nos transportes públicos, ressalta-se que, dadas as peculiaridades da operação dos transportes metroferroviários, a coleta automatizada de dados híbridos aqui apresentada se aplica somente aos transportes rodoviários. Essa aplicabilidade aos transportes públicos rodoviários se estende através dos veículos do tipo ônibus, atingindo também micro-ônibus, vans e demais categorias similares.

Em relação à operacionalização do método, urge observar que esta dissertação trata dos temas voltados aos equipamentos necessários para a coleta. Embora a aplicação prática desse método necessite dos outros dois pilares apresentados no trabalho de Komninos *et al.* (2011), nos quais a inteligência se correlaciona com o treinamento das equipes envolvidas na operação e a

interconexão seja dada por uma plataforma apropriada de trabalho, os requisitos aqui abordados se referem tão somente àqueles necessários para a instrumentação, em sua mínima forma.

4.3 - A COLETA AUTOMATIZADA E A PRIVACIDADE DOS USUÁRIOS

O Brasil não possui legislação específica para a proteção de dados pessoais, sendo que a maior proteção para esse tipo de dado se dá na esfera da privacidade e da transparência, não na proteção dos dados em si (Baroni e Souza, 2015).

Partindo dos moldes europeus de leis para a proteção de dados, em especial atenção às experiências portuguesas, o Brasil colocou para consulta pública um anteprojeto de lei para tratar sobre a proteção de dados pessoais em 2015 (Passos, 2015).

Entre os princípios apontados pelo anteprojeto é necessário enfatizar os seguintes: o princípio da necessidade, no qual só deve ser coletado o estritamente necessário para a análise; o princípio da transparência, pelo qual devem ser garantidas informações claras e adequadas aos titulares sobre o tratamento dos dados; e o princípio da segurança, que visa proteger os dados pessoais de acessos e usos não autorizados.

Comparando as premissas do método anteriormente aventadas neste Capítulo, assim como nas necessidades e oportunidades apresentadas nos capítulos anteriores, com a obra produzida por Baroni e Souza (2015), é nítido que os tipos de dados colhidos pelo método ora proposto se encaixam perfeitamente naqueles cuja coleta é permitida pelo projeto de lei.

Assim, não resta fato que possa, porventura, impedir ou trazer insegurança jurídica para a coleta no cenário de aprovação desse projeto de lei. Mesmo quando se compara com as leis em vigor, não há óbice formal na coleta, uma vez que essa execução se deriva do dever do Estado em prever e prover.

É oportuno, neste ponto, transcrever o quanto segue:

Abordando brevemente o teor filosófico da letra proposta e do mecanismo de funcionamento da sociedade, o direito individual de “permanecer invisível”, ainda que esse indivíduo ocupe espaço físico nos transportes, deve ser limitado pela necessidade coletiva de planejar, dimensionar, prever e prover as políticas de transporte público. A necessidade social de manter a acessibilidade dos serviços básicos vai além. (Baroni e Souza, 2015)

Por fim, a notoriedade presente no fato de que a letra legal não se atualiza com a mesma flexibilidade e abrangência que as novas tecnologias, fazendo com que, geralmente, haja um vale entre a situação corrente e o previsto pela lei, nesse aspecto não deve perfazer obstáculo para a gestão pública eficiente, em especial no que tange a gestão de serviços vitais, como os transportes e a mobilidade urbana.

4.4 - TÓPICOS CONCLUSIVOS

Neste capítulo dois assuntos relevantes para o entendimento formal do método híbrido de coleta de dados foram tratados. No primeiro, que diz respeito à formalização do método, foi possível observar sua interseção com os métodos existentes, bem como sua aplicabilidade. Ainda no bojo desse item foi definida sua aplicabilidade e os limites considerados para essa dissertação.

Aclarou-se que, embora a aplicação completa dos métodos dinâmicos de operação dos transportes públicos demandem de discussões sobre o treinamento dos envolvidos na tarefa e, ainda, da definição de uma plataforma agregadora de dados, o presente trabalho trata apenas dos requisitos mínimos para a coleta dos dados.

Já o segundo assunto abordado trouxe uma breve discussão sobre os aspectos legais que envolvem a coleta automatizada de dados e sua utilização para o planejamento e a operação dos transportes públicos. Conforme observado, não existe barreira legal que possa trazer insegurança jurídica ao ato de coletar dados dos usuários de transportes públicos por parte do ente gestor, com o intuito de planejar e operar.

Essa segurança se dá tanto pela inexistência de letra legal que proíba tal ato associada ao papel do Estado de prever e prover políticas públicas eficientes, quanto pela previsão de que os próximos preceitos a serem apreciados pelos legisladores não apresentam menção em contrário.

Tendo o método suficientemente definido neste ponto, o capítulo a seguir demonstrará como foram definidas as variáveis a serem coletadas por um equipamento que atenda aos preceitos de coleta automatizada de dados híbridos.

5 - DETERMINAÇÃO DE DADOS A SEREM COLETADOS PELO MÉTODO DE COLETA AUTOMATIZADA DE DADOS HÍBRIDOS

5.1 - APRESENTAÇÃO

Tendo definido o conceito de Coleta Automatizada de Dados Híbridos no capítulo anterior, faz-se necessária a definição de quais variáveis mínimas compõem a interseção entre os métodos existentes.

Por se tratar de uma nova abordagem sobre a coleta de dados para pesquisas relacionadas ao PTP e ao POTPU, optou-se pela sistematização do levantamento de variáveis recorrentes nos trabalhos acadêmicos publicados sobre esse tema, a fim de verificar quais deles são elegíveis para o AHDC. Para essa finalidade foram buscadas, através do uso de ferramentas de inteligência artificial, as publicações com maior impacto no meio acadêmico para compor tais variáveis.

Aproveitando a janela de oportunidade para a construção de um modelo que colete dados híbridos o presente trabalho buscou também compor uma série de variáveis que alterem a percepção do usuário quanto a qualidade dos transportes públicos e, de algum modo, possam influenciar no uso desse modal ou seu abandono.

Para essa tarefa, foi conduzida uma pesquisa de campo, com o intuito de verificar, na percepção dos usuários, quais fatores impactam na sua sensação de conforto nos veículos, além de quesitos ambientais. Essa medida liberou a literatura acadêmica como marco de comparação com os resultados obtidos, para confirmação ou refutação das variáveis escolhidas.

Nesse sentido, o presente capítulo demonstrará, sequencialmente, o procedimento adotado para a determinação de “variáveis recorrentes” na literatura para compor o rol de “variáveis mínimas” para a composição de um equipamento de coleta automatizada de dados híbridos.

5.2 - CRITÉRIOS PARA A SELEÇÃO DE TRABALHOS ACADÊMICOS

O primeiro critério de determinação foi adotado no sentido do uso estrito de materiais acadêmicos apreciados por bancas (dissertações, teses e artigos científicos). Ignorou-se, nesse momento, os estudos conduzidos por particulares e/ou autarquias governamentais, o que viabilizou o uso um parâmetro de comparação isento no momento de testar a funcionalidade do AHDC.

Desta forma, após a realização da modelagem dos dados e da caracterização dos seus tipos, foi possível verificar se o modelo é aplicável na prática sem que sua estruturação fosse “contaminada” ou conduzida pelos estudos a serem comparados.

A partir da adoção deste preceito, deu-se a seleção ampla de materiais que tivessem como objetivo declarado no desenvolvimento textual a tentativa de reconstituição de dados para fomento do Planejamento da Operação dos Transportes Públicos Urbanos.

Como limitador temporal, determinou-se que só seriam aproveitados materiais acadêmicos publicados a partir do ano de 2007. Esse limite temporal permitiu a realização de um corte quantitativo no universo de possibilidades, justificado pelo dever de atualidade dos tipos de dados coletados para fomento dos métodos de Planejamento da Operação dos Transportes Públicos Urbanos mais modernos.

Outro limitador adotado para a seleção dos trabalhos diz respeito à relevância calculada dos trabalhos no meio acadêmico. Para tal, utilizou-se um modelo de inteligência artificial como filtro qualitativo e quantitativo, por meio do qual foram determinadas as publicações consideradas influentes no período de tempo pré-determinado. Assim, foram avaliados fatores que vão da relevância do canal onde ao material acadêmico foi originalmente publicado à quantidade de vezes que este foi citado em outros trabalhos produzidos, além do contexto circundante para cada citação (Valenzuela *et al.*, 2015). Adicionalmente, para essa tarefa foi utilizado como plataforma de trabalho o site *Semantic Scholar* do *Allen Institute for Artificial Intelligence*.

Por fim, foram selecionados trabalhos que tratassem diretamente de tecnologias exclusivamente embarcadas. Essa medida foi adotada devido à evidente discrepância entre os métodos de coleta embarcados e não embarcados (como coletas de radares e em estações de metrô, por exemplo), que poderiam transformar a tentativa de apresentação de um método híbrido numa diligência improfícua. Dessa maneira, sempre que foram detectados trabalhos que abordavam coletas não embarcadas, estes foram descartados da revisão sistemática da literatura. A única exceção à essa regra são os sistemas de bilhetagem, incluindo validadores e bloqueios, presentes em estações e terminais de transportes públicos rodoviários.

Vencidos os critérios de seleção dos trabalhos elegíveis para a análise, fez-se necessária outra medida para levantamento dos dados de relevância: a concentração de tipos semelhantes em categorias únicas. Este fato foi evidenciado pelo volume de trabalhos selecionados para a

análise, que atingiu 87 produções científicas e pela semelhança dos dados coletados em cada um destes.

5.3 - CONCENTRAÇÃO DE TIPOS SEMELHANTES

O passo seguinte na determinação dos gêneros de dados relevantes foi a concentração de tipos semelhantes destes numa mesma nomenclatura. Essa medida foi extremamente relevante, por exemplo, para que a granulometria dos dados de GPS não os tornasse tipos únicos, apresentando uma falsa necessidade de dados, quando na realidade, trata-se de uma mesma informação subdividida de maneira diferente.

5.3.1 - DADOS DE TELEMETRIA

Assim como aventado nos capítulos iniciais deste trabalho, os dados de telemetria são tão plurais quanto abundantes em sistemas de transportes modernos. Esses dados abarcam um número bastante relevante de informações sobre o funcionamento dos veículos: peso nominal, quantidade de combustível utilizada por quilômetro rodado, temperatura dos elementos mecânicos, temperatura do motor, etc.

Essas informações são bastante relevantes para o Planejamento da Operação dos Transportes Públicos Urbanos, permitindo o controle do desempenho, da obsolescência da frota, das falhas e a da qualidade dos serviços prestados. Em veículos mais modernos, observam-se, ainda, parâmetros de importância ambiental como: emissão de gases, eficiência da queima de combustível e economicidade da forma de direção do condutor.

Assim, na seleção de variáveis, foi conveniente convergir toda a gama de dados coletados pela telemetria para um único tipo, observado em escala individual nos capítulos seguintes.

5.3.2 - ESTAMPA DE TEMPO

Estampa de tempo é o termo técnico utilizado para definir o ato de coletar os dados temporais do instante em que determinada ação ocorre. Esse dado, embora pareça trivial, é de extrema importância para as inferências e as análises realizadas, podendo ser utilizado, por exemplo, para determinar se um passageiro embarcou em uma determinada estação (Ma *et al.*, 2012) ou como balizador para a reconstrução dos propósitos das viagens (Bohte e Maat, 2009).

Desta maneira, a frequência com que esse dado é coletado e sua precisão são fatores relevantes para as possibilidades de uso dos dados coletados pelos diferentes sistemas de coleta.

Entretanto, nessa etapa, as frequências e os formatos serão aglomerados neste tipo único de dados.

5.3.3 - PONTOS GEOESPACIAIS

Os Sistemas de Informação Geográficas (SIG) lançam mão de recursos de representação computacional adequadas, com o intuito de capturar a semântica do conceito do mundo real (Câmara e Monteiro, 2001). Essas representações podem se dar através de pontos, linhas ou polígonos georreferenciados.

Considerando o elemento ponto como uma entidade unitária, sem área associada, formado meramente pelas coordenadas posicionais deste, podem-se considerar as informações coletadas pelos equipamentos de GPS como uma série de pontos geoespaciais.

A frequência com a qual esses pontos são coletados, denominada de granularidade dos pontos, permite que sejam inferidas informações tanto de posicionamento geral - para menores granularidades -, quanto de níveis de eficiência da rede de transportes - para granularidades maiores (Moreira-Matias *et al.*, 2015). Para o presente momento, as diferentes granularidades são dispensáveis, sendo todas as granularidades apresentadas contidas neste item.

5.3.4 - TRANSAÇÕES DO VALIDADOR

As transações realizadas pelos validadores podem fornecer dados de grande relevância para o planejamento da operação dos transportes, conforme apresentado por Baroni *et al.* (2016). Dados quantitativos de usuários acessando os sistemas de transportes públicos ou, até mesmo, a recorrência de um determinado grupo de usuários, indicam usos costumeiros de um determinado modal.

Os dados tipificados como “transações” independem da validação ou não dos tíquetes de entrada. Assim, mesmo uma tentativa de acesso infrutífera registrada pelo sistema de bilhetagem é considerada como “transação do validador”.

5.4 - LEVANTAMENTO DA RECORRÊNCIA DE VARIÁVEIS COLETADAS

Uma vez tipificados os dados coletados para grupos comuns, o presente trabalho se concentrou em verificar nos trabalhos científicos quais dados estavam presentes em cada uma das publicações.

Os trabalhos foram lidos cuidadosamente, a fim de identificar as variáveis utilizadas, os dados obtidos e, ainda, os resultados atingidos pela aplicação do método. Esses dados foram compilados na Tabela 5.1.

Os materiais acadêmicos que coletaram dados semelhantes dentro dos critérios apresentados anteriormente foram agrupados para que fosse possível sua contemplação. Dentro dos grupos formados por dados semelhantes, selecionaram-se os trabalhos com maior relevância, conforme o método descrito anteriormente. Desta forma, a Tabela 5.1 apresenta somente o artigo mais relevante para cada tipo de dado levantado.

Tabela 5.1 - Compilação dos dados coletados pelos artigos científicos e seu modo de coleta.

Autor	Objetivo do Trabalho	Grupo de dado coletado	Método de Coleta		
			ADC	AFC	AVL
Alsger <i>et al.</i> (2016)	Validação e melhoria das matrizes OD	Transações do validador; Estampa de tempo;		X	
Alvarez-Garcia <i>et al.</i> (2010)	Destino da viagem	Estampa de tempo; Pontos geoposicionais;			X
Bohte e Maat (2009)	Propósito da viagem	Estampa de tempo; Pontos geoposicionais;			X
Du e Aultman-Hall (2007)	Fim da viagem e divisor de viagens	Estampa de tempo; Pontos geoposicionais;			X
Fernández <i>et al.</i> (2008)*	Calibração das redes de transportes e equilíbrio das matrizes de origem e destino	Estampa de tempo; Pontos geoposicionais; Tempo de viagem; Transações do validador;	X	X	X
Huang e Levinson (2015)	Modelagem de destino	Estampa de tempo; Pontos geoespaciais;			X
Ma <i>et al.</i> (2012)	Origem dos passageiros	Transações do validador; Estampa de tempo; Estampa de tempo;		X	
Munizaga e Palma (2012)	Inferência de matriz OD	Transações do validador; Pontos geoposicionais; Dados de telemetria;	X	X	X
Nitsche <i>et al.</i> (2012)	Reconstrução das viagens	Estampa de tempo; Pontos geoposicionais;			X
Vlassenroot <i>et al.</i> (2015)	Comportamento de viagens	Estampa de tempo; Pontos geoposicionais;			X
Wang <i>et al.</i> (2011)	Inferência de matriz OD	Estampa de tempo; Pontos geoposicionais;	X		X
Zhao e Rahbee (2007)	Estimativa de matriz OD	Estampa de tempo; Transações do validador;	X	X	

* - Foram realizadas pesquisas tradicionais complementares para este trabalho.

5.5 - VARIÁVEIS DE QUALIDADE DOS TRANSPORTES PÚBLICOS

Além das variáveis que permitem a inferência do desempenho dos transportes públicos, as variáveis qualitativas do sistema podem explicar, em grande parte, o comportamento dos usuários deste modal.

Embora não sejam exploradas em conjunto com os dados de desempenho em nenhum dos trabalhos levantados acima, a prática no desenvolvimento de planejamentos para os transportes públicos demonstra que a percepção dos usuários acerca dos sistemas é um fator relevante a ser levado em conta no Planejamento da Operação dos Transportes Públicos Urbanos.

Ferraz e Torres, (2004), apresentam a preferência de escolha dos usuários pelo modo de transporte mais confortável: o carro. Segundo a visão dos autores, os transportes públicos são utilizados quando ocorre uma ou mais das seguintes situações: restrição econômica pessoal, impossibilidade de dirigir, existência de transporte público de boa qualidade, trânsito congestionado, dificuldade de estacionamento, custo elevado de acesso (pedágio) e estacionamento, restrições impostas pelo poder público ao uso do carro, conscientização da comunidade em relação aos problemas provocados pelo uso massivo do carro, dentre outras.

Observando o trabalho de Lima Júnior (1995), podem ser elencados como critérios de avaliação da qualidade para o transporte público na visão do usuário: confiabilidade (intervalo entre veículos, tempo de viagem, cumprimento do itinerário), responsabilidade (substituição do veículo em caso de quebra, atendimento ao usuário), empatia (disposição do motorista e cobrador em dar informações, atenção com pessoas idosas e deficientes físicos), segurança (condução do motorista, assaltos), tangibilidade (lotação, limpeza, conservação), ambiente (trânsito, condições climáticas), conforto (bancos, iluminação, ventilação), acessibilidade (localização dos pontos de parada), preço (tarifa), comunicação (informação sobre o sistema, relação entre os usuários), imagem (identificação da linha/serviço) e momentos de interação (contato com motorista/cobrador). Esses critérios de avaliação são repetidos quase que integralmente nos trabalhos de Dell'Olio *et al.* (2011) e Redman *et al.* (2013).

Ferraz e Torres (2004) apresentam as principais vantagens e desvantagens do transporte público urbano. Como vantagens, citam o custo acessível para a população de baixa renda, a democratização da mobilidade, o fato de ser uma alternativa para reduzir os impactos do transporte individual e de proporcionar uma ocupação mais eficiente do solo nas cidades.

Complementarmente, os autores supracitados apresentam 12 principais focos que influenciam a qualidade do transporte público por ônibus, a saber: acessibilidade, frequência de atendimento, tempo de viagem, lotação, confiabilidade, segurança, características dos veículos, características dos locais de parada, sistema de informação, conectividade, comportamento dos operadores e estado das vias.

De maneira similar, Eboli e Mazzulla (2008) apresentam como atributos de qualidade dos transportes públicos: a distância percorrida pelo passageiro, a frequência, a pontualidade, a estrutura das paradas de ônibus, a limpeza, a lotação, a tarifa, a disponibilidade de informações e a atitude do motorista.

Nesse sentido, foi conduzida uma pesquisa de campo com o intuito de levantar quais os quesitos relativos à qualidade dos transportes públicos são mais relevantes para os usuários. Com base nos resultados obtidos, foram determinados os dados a serem coletados, de forma a propiciar aos gestores dos sistemas de transportes públicos informações suficientes para a compreensão e a operação dos sistemas com foco nos usuários.

5.5.1 - PESQUISA DE CAMPO - MÉTODO DE COLETA E PARÂMETROS OBSERVADOS

Utilizando-se a ferramenta desenvolvida pelo Grupo Comportamento em Transportes e Novas Tecnologias - GCTNT, concebida para aplicação de questionários através da internet, chamada de *Plataforma Survey*, foi criado um questionário com 10 questões distribuídas em três conjuntos: identificação do participante, caminho até o ponto de ônibus e dentro do ônibus. Cada uma das seções, com exceção da primeira, foi elaborada com o intuito de observar a percepção sobre o conforto nos transportes públicos para usuários e não usuários.

Embora a ferramenta seja para uso através de internet, foram utilizados *tablets* com acesso à internet e pesquisadores para a coleta de dados. Esses dados foram coletados em pontos de ônibus, Rodoviária do Plano Piloto de Brasília, nas dependências da Universidade de Brasília e no Shopping Conjunto Nacional entre os dias 08 de fevereiro de 2016 e 18 de março de 2016.

A primeira seção coletou os dados sobre sexo, faixa etária e, ainda, identificou se o respondente é portador de necessidades especiais e se este se declara usuário frequente de transportes públicos. Em caso afirmativo para o último quesito, outras duas perguntas eram ativadas: (a)

com qual frequência o respondente toma o ônibus por semana, em média? e, (b) quanto tempo o respondente gasta dentro dos transportes públicos, em média, por viagem?

Observa-se que a pesquisa foi planejada no sentido de atender aos parâmetros indicados por Lima Júnior (1995) e Dell’Olio *et al.* (2011), mencionados anteriormente.

5.5.2 - ANÁLISE DA PERCEPÇÃO DOS USUÁRIOS DO TRANSPORTE PÚBLICO RODOVIÁRIO

Com a ferramenta de pesquisa foram coletados 182 formulários, todos de usuários do Distrito Federal, sendo 86 mulheres e 96 homens, dos quais 62,82% se declararam não usuários do transporte público e 37,18% se declararam como usuários frequentes. Desses usuários frequentes, 31,03% declararam que passam mais de uma hora por dia no trânsito.

A Tabela 5.2 e a Tabela 5.3 apresentam a compilação da percepção dos usuários frequentes em relação a afirmações predefinidas sobre o ônibus, o ponto de ônibus e o caminho até este. Nota-se que há o indicativo de desconforto dos respondentes sobre esperar mais que 10 minutos nas paradas de ônibus. Adicionalmente, nenhum dos respondentes concorda com a afirmação que o ponto de ônibus seja confortável.

Tabela 5.2 - Análise da percepção do usuário de ônibus em relação ao deslocamento até o ponto de ônibus

	<i>Concordo totalmente</i>	<i>Concordo Parcialmente</i>	<i>Indiferente</i>	<i>Discordo Parcialmente</i>	<i>Discordo Totalmente</i>
O caminho até o ponto de ônibus é seguro.	14,29%	39,29%	7,14%	17,86%	21,42%
O caminho até o ponto de ônibus tem calçadas em bom estado.	10,71%	21,43%	10,71%	21,43%	35,72%
Sinto-me tranquilo (tranquila) em acessar o ponto de ônibus a qualquer horário.	3,57%	10,71%	10,71%	35,71%	39,30%
O ponto de ônibus é seguro.	3,57%	21,43%	10,71%	28,58%	35,71%
O ponto de ônibus é confortável.	0,00%	10,71%	3,58%	25,00%	60,71%
O ponto de ônibus é coberto.	14,29%	39,28%	7,14%	14,29%	25,00%
No ponto de ônibus existem informações suficientes sobre horários e rotas que param ali.	7,14%	0,00%	3,57%	14,29%	75,00%
Fora do horário de pico há bancos suficientes.	25,00%	17,86%	10,71%	17,86%	28,57%
Não me sinto desconfortável em esperar mais que 10 minutos na parada.	25,00%	3,57%	10,71%	28,58%	32,14%

Tabela 5.3 - Análise da percepção do usuário de ônibus em relação à viagem

	<i>Concordo totalmente</i>	<i>Concordo Parcialmente</i>	<i>Indiferente</i>	<i>Discordo Parcialmente</i>	<i>Discordo Totalmente</i>
Os veículos cumprem seus horários.	0,00%	30,77%	3,85%	42,30%	23,08%
Os veículos estão em bom estado de conservação.	7,69%	57,69%	11,55%	15,38%	7,69%
Sinto-me seguro (segura) em relação a acidentes de trânsito enquanto estou no veículo.	11,54%	23,08%	3,85%	34,62%	26,91%
Sinto-me seguro (segura) em relação a assaltos/violência enquanto estou no veículo.	11,54%	26,92%	7,69%	19,23%	34,62%
O motorista dirige de maneira não agressiva e respeitando as leis de trânsito.	7,69%	42,31%	3,84%	23,08%	23,08%
Sinto-me confortável com a quantidade de passageiros nos ônibus.	3,85%	23,08%	7,68%	23,08%	42,31%
Sinto que posso decidir não embarcar em um ônibus muito cheio sem prejudicar meus horários.	3,85%	19,23%	0,00%	19,23%	57,69%
A qualidade do ar dentro do ônibus me parece boa.	3,85%	19,23%	11,54%	26,92%	38,46%
A temperatura dentro dos ônibus é sempre agradável.	3,85%	11,54%	15,38%	42,31%	26,92%
A luminosidade dentro dos ônibus me parece boa.	15,38%	26,93%	19,23%	23,08%	15,38%
A quantidade de ruído dentro dos ônibus me parece agradável.	3,85%	23,08%	15,37%	23,08%	34,62%
Não me incomodo em circular, por curtos trajetos, em pé.	42,31%	23,07%	11,54%	11,54%	11,54%
Os funcionários da empresa de transportes são sempre cordiais e prestativos.	7,69%	19,23%	38,47%	15,38%	19,23%

Observa-se, ainda, que é maciça a opinião de que não existem informações suficientes sobre linhas, rotas e horários. Fica nítido, sobretudo, que os usuários não se importam em circular

pequenos trechos em pé; entretanto, uma porcentagem considerável se declara desconfortável com a quantidade de passageiros viajando em pé.

Analisando-se a Tabela 5.4, fica evidente que os usuários dão grande importância para o tempo que ficam na parada de ônibus, sobrepondo-se, inclusive, a quesitos de segurança.

Tabela 5.4 - Classificação de importância para itens relativos ao transporte público, na ótica dos usuários frequentes

	1º	2º	3º	4º	5º	6º
Caminho seguro até a parada	25,00%	25,00%	14,28%	7,14%	21,43%	7,14%
Caminho confortável até a parada	3,57%	0,00%	7,14%	17,86%	28,57%	42,86%
Parada segura	14,29%	25,00%	17,86%	17,86%	7,14%	17,86%
Parada confortável	0,00%	7,14%	14,29%	39,29%	28,58%	10,71%
Esperar pouco tempo na parada	35,71%	28,57%	21,43%	3,57%	7,14%	3,57%
Ter informações atualizadas sobre horários e linhas	21,43%	14,29%	25,00%	14,28%	7,14%	17,86%

Por último, na Tabela 5.5, consegue-se observar que a maior preocupação dos usuários, em relação ao veículo, concentra-se nos itens de confiabilidade sobre horários e circulação e no comportamento do motorista enquanto dirige. Em segundo lugar, estão empatadas as preocupações sobre viajar sentado e o tempo de viagem até o destino.

Tabela 5.5 - Classificação dos três itens mais votados como importantes relativos à viagem, na ótica dos usuários frequentes

	1º	2º	3º
Viajar sentado	11,54%	19,23%	3,85%
Temperatura dentro do ônibus	7,69%	0,00%	7,69%
Quantidade de pessoas de pé	7,69%	19,23%	11,54%
Iluminação no ônibus	0,00%	3,85%	3,85%
Qualidade do ar (cheiro ou fumaça) dentro do veículo	3,85%	3,85%	3,85%
Comportamento do motorista (dirigindo)	15,38%	3,85%	7,69%
Polidez do motorista/cobrador	3,85%	3,85%	0,00%
Quantidade de barulho dentro do veículo	0,00%	0,00%	7,69%
Tempo de viagem	11,54%	15,38%	23,07%
Conservação do veículo	0,00%	3,85%	11,54%
Idade do veículo	7,70%	0,00%	0,00%
Confiança de chegar no horário	15,38%	15,38%	11,54%
Confiança que os ônibus irão transitar todos os dias	15,38%	11,53%	7,69%

As tabelas apresentadas demonstram que o ponto focal da percepção dos usuários está nos itens mais básicos, relacionados com a infraestrutura do sistema de transporte público por ônibus, seguida de aspectos operacionais, relativos ao tempo de viagem e à confiabilidade nos transportes. Os aspectos operacionais indicados pelos usuários podem fazer parte do conjunto de dados coletado pelo sistema híbrido sem prejuízo dos demais dados levantados.

Esses resultados são suficientes norteadores para a determinação de variáveis a serem coletadas com o intuito de demonstrar, mesmo que sinteticamente, a percepção dos usuários quanto aos transportes públicos.

5.6 - TÓPICOS CONCLUSIVOS

Nos capítulos anteriores, apresentou-se a revisão da literatura, caracterizando o planejamento da operação dos transportes, bem como os conceitos basais da coleta automatizada dos dados. Abordou-se, ainda, o levantamento das variáveis apresentadas em trabalhos científicos - relativas aos veículos, suas viagens e seus passageiros - e das variáveis inerentes à percepção de qualidade por parte dos passageiros.

Isto realizado, foi possível a construção do quadro de variáveis apresentado na Tabela 5.6. Na referida tabela, os dados desagregados, são apresentadas com sua relação direta com o método de coleta e sua justificativa, baseada nos estudos aventados neste capítulo. Essa organização num quadro geral é primordial para a condução dos demais capítulos dessa dissertação e, assim, encerra o objetivo de definir os dados de maior relevância para o PTU e para o POTPU.

A preponderância de dados provindos da telemetria do veículo era esperada, uma vez que os dados de posicionamento fornecido pelos sistemas de AVL são variados, quanto à granularidade. Entretanto, os dados coletados são basicamente os mesmos.

Já os dados de AFC, por sua vez, apresentaram-se como dependentes diretos da disponibilidade dos fabricantes e dos mantenedores dos sistemas de bilhetagem. Assim sendo, aparecem nesta tabela como um único bloco de informações.

No capítulo a seguir serão especificados os sensores envolvidos na coleta e os formatos que os dados de equipamentos com as características de AHDC devem ter para atendimento das necessidades do POTPU.

Tabela 5.6 - Compilação dos dados a serem coletados.

	<i>Dado a ser coletado</i>	<i>Quesito</i>	<i>Justificativa</i>
ADC	Aceleração gravitacional triaxial do veículo	Qualitativo	A aceleração gravitacional pode revelar a conduta agressiva ou imprudente do condutor;
	Concentração de CO2 interno ao veículo	Qualitativo	Indicar a qualidade do ar dentro do ônibus;
	Estampa de tempo	Geral	Posicionar temporalmente a sequência de eventos captado pelo sistema;
	Evento de abertura das portas	Qualitativo / Quantitativo	Indicar pontos onde e quais portas do veículo foram abertas, facilitando a reconstrução de sobe e desce em paradas;
	Evento de acionamento da rampa para cadeirante	Qualitativo / Quantitativo	Indicar pontos onde e quais portas do veículo foram abertas, facilitando a reconstrução de sobe e desce em paradas;
	Evento de acionamento do botão de parada	Qualitativo / Quantitativo	Indicar a solicitação dos passageiros para desembarque do veículo;
	Nível de luminosidade interna ao veículo	Qualitativo	Indicar o conforto luminoso dentro dos veículos;
	Nível de ruído interno	Qualitativo	Indicar o conforto sonoro dentro do veículo;
	Presença de chuva	Qualitativo	Indicar condição de trânsito externa ao veículo;
	Temperatura externa ao veículo	Qualitativo	Indicar condição externa ao veículo;
AFC	Temperatura interna do veículo	Qualitativo	Indicar o conforto térmico dentro do veículo;
	Umidade do ar	Qualitativo	Indicar as condições adicionais de conforto dentro do veículo;
AFC	Eventos do validador	Quantitativo	Indicar o volume real de usuários no serviço;
AVL	Pontos geospaciais	Geral	Posicionar geoespacialmente a sequência de eventos captado pelo sistema;

6 - DETERMINAÇÃO DA ARQUITETURA, SENSORES ASSOCIADOS E PADRÕES DE DADOS DO COLETOR AUTOMATIZADO DE DADOS HÍBRIDOS PARA O POTPU

6.1 - APRESENTAÇÃO

Conforme visto no Capítulo 1, o objeto da presente dissertação é a determinação de requisitos para que um equipamento eletrônico embarcado, instalado nos transportes públicos rodoviários, seja considerado um coletor automatizado de dados híbridos para o POTPU.

O método de coleta de dados híbridos é aquele que captura informações de diferentes fontes que, até então, eram utilizadas por métodos isolados, com o intuito de determinar o padrão de funcionamento e, ainda, fornecer os subsídios técnicos para a tomada de decisão nos transportes públicos rodoviários, conforme apresentado no Capítulo 4.

A construção do conceito desse modelo de coleta de dados aponta para a necessidade de um conjunto mínimo de sensores para o funcionamento desses equipamentos. A estrutura mínima padronizada permite uma constância na captura dos dados, através do tempo e entre localidades diferentes. Oportunamente, favorece à realização de comparações até então dificultadas pela discrepância entre os dados capturados, seu formato de coleta e seu armazenamento. Esse conjunto de variáveis mínimas coletadas foi determinado por meio do confronto analítico da produção científica dos últimos 10 anos com maior relevância e citação, na forma apresentada pelo Capítulo 5.

Mediante o embasamento apresentado nos capítulos iniciais desta dissertação, que norteou e delineou o método desenvolvido, cabe a este capítulo apresentar a arquitetura básica dos equipamentos de AHDC e seu funcionamento básico. Ainda, os quesitos mínimos dos sensores para essa finalidade e demonstrar a forma de armazenamento temporário dos dados brutos coletados.

É importante ressaltar, destarte, que esse capítulo se absterá da abordagem das configurações e das características eletrônicas mais detalhadas, por se entender que estas não agregam profundidade à discussão proposta por esta dissertação, bem como por criarem uma espécie de engessamento tecnológico prejudicial à longevidade do método.

6.2 - ARQUITETURA BÁSICA DOS EQUIPAMENTOS DE AHDC

Embora a necessidade de definir a arquitetura básica possa remeter a uma tentativa de definição topológica ou de estruturação lógica, esse desenho detalhado dos componentes eletrônicos

poderia engessar a solução, transformando o AHDC num produto ao invés de ser um padrão para um método. Assim, os aspectos topológicos e práticos da construção dos equipamentos que atendam ao AHDC não serão tratados nessa dissertação, ficando a cargo de cada projetista que utilizará o modelo híbrido a definição da topologia que melhor se enquadre, além da escolha das tecnologias envolvidas mais adequadas. Essa liberdade de escolha é válida desde que sejam atendidos os pressupostos ora delimitados.

Entretanto, é importante abordar a forma hierárquica básica de como os sistemas que atendem ao padrão híbrido se comportam eletronicamente, evitando que apareçam discrepâncias irreversíveis na geração de dados.

Deibel e Zumwalt (1984) apresentam a estrutura de um sistema capsular para ADC, composto de treze módulos, divididos entre “módulos essenciais” e “módulos de aquisição/transmissão”. Em seu trabalho, esses módulos se focam em coletar dados internos aos veículos relacionados ao seu desempenho e à contagem categorizada de passageiros, através de leitura dos sistemas de arrecadação.

Por um lado, embora estes módulos sejam muito importantes para a indústria norte americana de sistemas embarcados, não refletem as necessidades do cenário atual, especialmente por determinar o uso de tecnologias que não cabem mais, como o uso de fitas cassete para gravação dos dados, e associar características topológicas ultrapassadas. Essa determinação de tecnologias fez com que o aparecimento de novas tecnologias deixassem o trabalho obsoleto. Ademais, esse trabalho não apresenta a caracterização das cadeias de dados para permitir sua transmissão em tempo real nem a coleta de dados de GPS, vez que não era a realidade à época. Dessa maneira, a aplicação desse modelo nos dias atuais não representa vantagem para o planejamento da operação.

Por outro lado, a estrutura básica dos módulos foi muito bem pensada, sendo referência para os equipamentos de ITS como um caso de sucesso, dentro do seu tempo. Dessa maneira, esta dissertação propõe, para a arquitetura e a hierarquia dos módulos, que seja utilizado uma adaptação dos módulos apresentados por Deibel e Zumwalt. A estrutura adaptada dos módulos pode ser vista na Figura 6.1.

Na estrutura proposta, os equipamentos são classificados em “essencial” e “não essencial” pois, em caso de problemas, como por exemplo, falta de fornecimento de energia externa, esses módulos precisam ser privilegiados e mantidos em funcionamento. Essa funcionalidade permite

que, no caso de uma parada inesperada por pane elétrica, o gestor tático dos transportes seja informado imediatamente, possibilitando a atuação ágil e com menor prejuízo para a população servida pelo sistema.

Os módulos são segregados, ainda, pela sua função na arquitetura do sistema. Os equipamentos de gerenciamento servem como regentes dos demais, tendo a função de gravar os dados, acionar o sistema de transmissão e sincronizar as coletas.

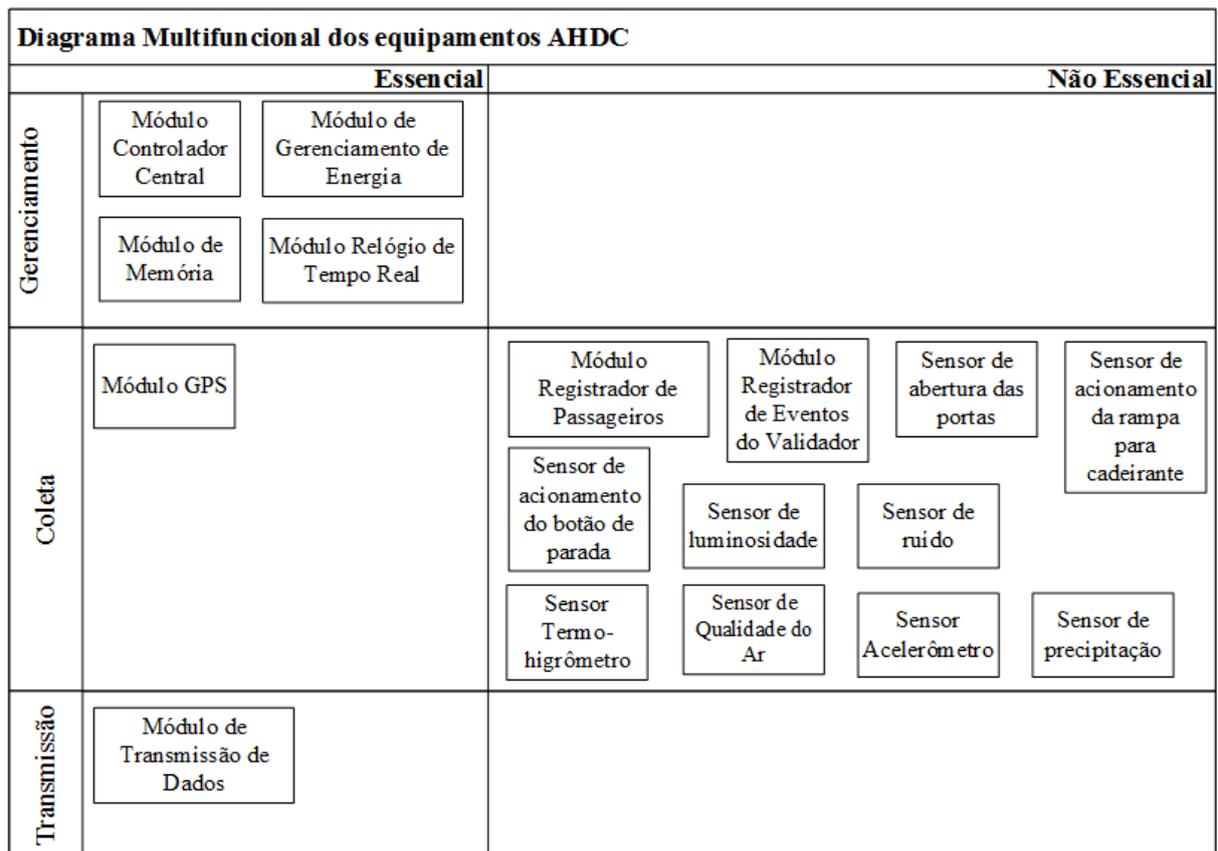


Figura 6.1 - Diagrama Multifuncional dos Equipamentos AHDC. (adaptado e modificado com base em Deibel e Zumwalt (1984))

Os dispositivos que compõem o equipamento foram divididos em módulos e sensores. Os módulos devem possuir operação independente, apenas convergindo os dados para o Módulo Controlador Central, podendo inclusive ser montado em placa de circuito impressa separada dos demais módulos. Os sensores, por sua vez, são equipamentos eletrônicos escravos, diretamente controlados pelo Módulo Controlador Central.

Um diagrama esquemático da arquitetura básica dos módulos AHDC pode ser visto na figura XXX à seguir.

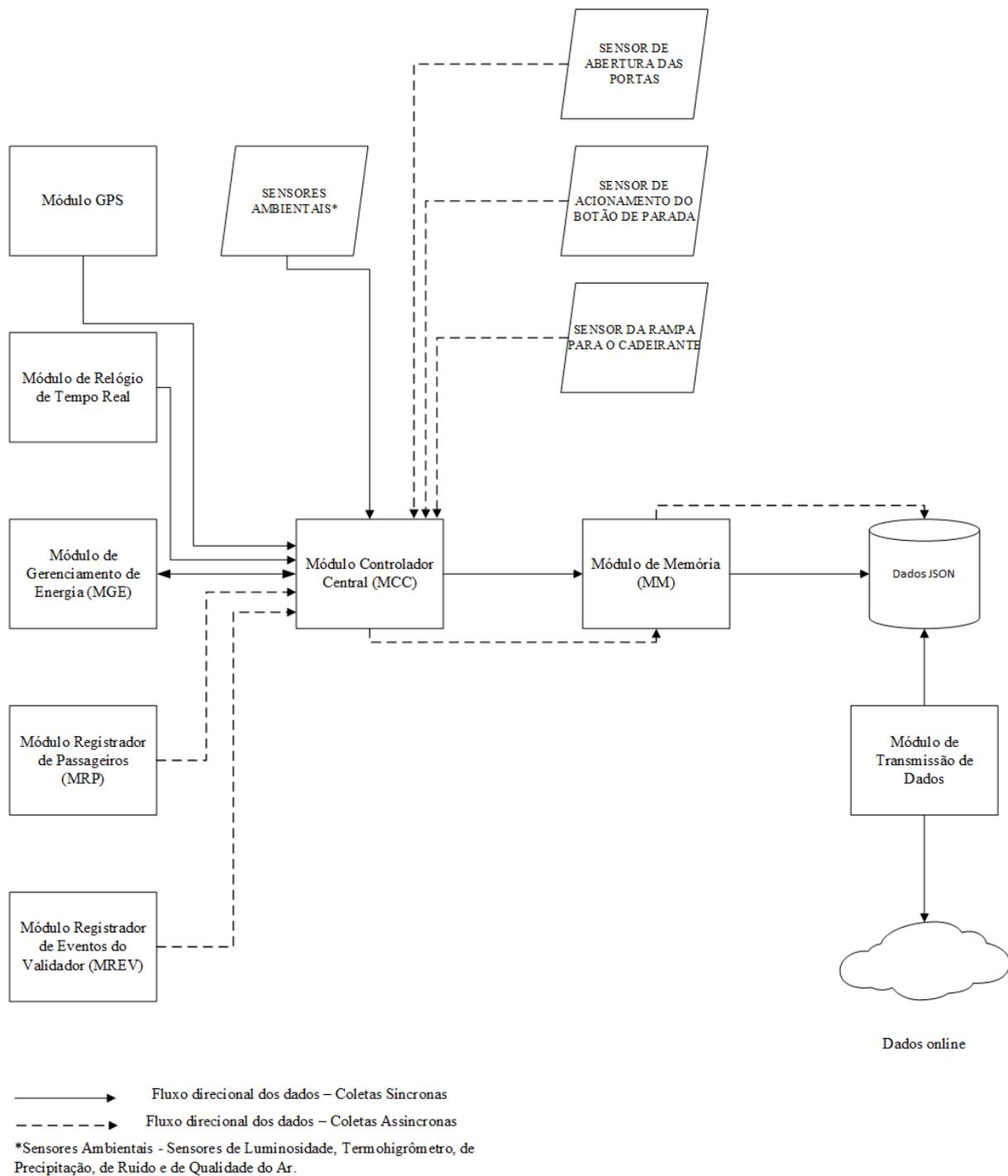


Figura 6.2 - Diagrama esquemático da arquitetura dos AHDC.

É importante frisar que essa quantidade de sensores é mínima, apenas para que o equipamento e os dados gerados por este se caracterizem como de coleta híbrida. Nada impede que, por necessidade ou vontade, os gestores da operação dos transportes públicos ou seus planejadores determinem sensores adicionais para compor o parque de máquinas. Outra alteração totalmente possível e indicada, em determinados casos, é a transformação de um módulo não essencial em essencial, e vice-versa.

6.3 - EQUIPAMENTOS MÍNIMOS DO SISTEMA AHDC

Com o intuito de atender à estrutura dos sistemas AHDC definida anteriormente, foram testadas diversas configurações de sensores e estruturas de ligação. Os testes foram executados com a montagem de matrizes em placas de prototipagem e de simulação computacional, utilizando-se o software *Eagle* da empresa AutoDesk.

As especificações a seguir compreendem as configurações mínimas que atenderam aos quesitos de estabilidade e de fiabilidade para a coleta de dados proposta, podendo ser alterada pelos fabricantes que se disponham a adotar o modelo, por equipamentos e configurações similares ou superiores.

Uma característica básica para todos os módulos apresentados, e que merece destaque, é a resistência a altas temperaturas. Dependendo do local de instalação do equipamento as temperaturas às quais os equipamentos estão sujeitos podem superar os 60°C, fazendo com que alguns sensores parem de trabalhar ou trabalhem de forma anormal.

6.3.1 - MÓDULO CONTROLADOR CENTRAL DO SISTEMA AHDC

O Módulo Controlador Central é o módulo concentrador das funções lógicas e de gerenciamento dos sistemas de AHDC. Entre suas funções estão: controlar as coletas síncronas e assíncronas, salvar os dados coletados no formato pré-estabelecido, disponibilizar para o módulo de transmissão online os dados necessários, gerenciando as respostas deste, entre outras.

Esse módulo pode ser constituído de um ou mais microcontroladores, *slots* de entrada e de saída de dados para comunicação com os demais. É imperativo que esse módulo trabalhe com uma frequência que permita uma coleta a cada síncrona de 10 segundos e permaneça em estado de prontidão para possíveis coletas assíncronas no restante do tempo. Sugere-se que esse módulo

trabalhe com, no mínimo, osciladores de frequência de 16MHz, com o intuito de evitar atrasos nas leituras.

Propõe-se, ainda, que o microcontrolador tenha, pelo menos, um banco de 32Kb de memória flash, 1Kb de memória EEPROM e 2Kb de memória RAM. Nos testes realizados, processadores com características inferiores a estas sofreram travamentos quando sua operação superou 6 horas ininterruptas.

Para o momento de produção, sugere-se a troca da memória EEPROM por memórias do tipo PROM. A primeira permite que o programa interno seja alterado sem a troca do microcontrolador. A segunda é dotada de um sistema de fusível/antifusível que impede qualquer tipo de alteração após a sua gravação, evitando que tentativas de fraudes ou de alterações indevidas tenham sucesso.

O módulo controlador central deve ser munido de um sistema de segurança, a fim de identificar falhas nos sistemas essenciais e não essenciais, encaminhando mensagens de alerta para o módulo de transmissão de dados.

6.3.2 - MÓDULO DE GERENCIAMENTO DE ENERGIA

O módulo de gerenciamento de energia deve prover fornecimento constante e estável de energia, tanto nas tensões quanto nas correntes necessárias para o correto funcionamento do sistema. Seu funcionamento deve incluir o necessário para o armazenamento de energia (bancos de bateria) suficientes para uma autonomia de, pelo menos, 20 horas e a recarga destes.

Não é indicado o funcionamento dos equipamentos ligados diretamente ao sistema de energia dos veículos, uma vez que as oscilações geradas pelo sistema de alternadores, aterramentos ineficientes e descargas eletroestáticas podem levar os sensores mais frágeis a uma operação com interferências. No entanto, o módulo de gerenciamento de energia pode se valer dessa fonte como fonte de recarga, desde que filtre eventuais ruídos. Os testes físicos realizados contaram com um módulo de gerenciamento de energia básico, recarregado com energia solar, obtendo um bom resultado, haja vista que manteve o sistema com bloqueio de carregamento reverso para os painéis fotovoltaicos.

É necessário observar que esse tipo de módulo tem uma grande dispersão de calor, sendo necessário o controle rigoroso para evitar acidentes, como queima de componentes eletrônicos ou mesmo combustão do equipamento.

6.3.3 - MÓDULO DE MEMÓRIA

O papel do Módulo de Memória é armazenar os dados gerados pelos sensores e processados pelo Módulo Controlador Central, deixando esses registros disponíveis para os Módulos de Transmissão Online e para o processo de coleta completa realizada no final da operação (ou em outro momento indicado pelo plano de trabalho).

Dessa maneira, o tamanho da memória é diretamente proporcional à quantidade de sensores instalados e em funcionamento. O sistema, em seu formato mínimo aqui proposto gera aproximadamente 110 bytes por coleta, ou seja, considerando uma frequência de coleta de 10 segundos, em 18 horas de operação, tem-se cerca de 1,3MB de dados coletados.

O tamanho dos módulos de memória deve ser suficiente para registrar a quantidade total de dados coletados por um período superior ao intervalo de coleta programado pela operação dos sistemas. Para os testes de conceito, utilizou-se um módulo de memória com capacidade de 1GB, sem maiores problemas.

Sugere-se que os Módulos de Memória sejam de estado sólido, ou seja, não armazenem dados através de sistemas eletromecânicos, uma vez que as vibrações e os choques mecânicos causados pelo trânsito dos veículos podem danificar equipamentos eletromecânicos.

Em contrapartida, equipamentos de estado sólido carecem de uma proteção contra descargas de eletricidade estática mais eficientes, reforçando assim o papel do Módulo de Gerenciamento de Energia em manter, também, um bom isolamento para os demais módulos. O invólucro usado para montagem dos módulos deve ser aterrado, evitando que descargas elétricas e eletroestáticas atinjam o equipamento.

Os Módulos de Memória devem ser do tipo removível, podendo ser usados pendrive ou cartão de memória para o equipamento, permitindo que este seja retirado no final da operação para a coleta e a descarga dos dados. Essa mobilidade traz, ainda, a vantagem de que esta seja facilmente substituída, caso ocorram problemas de corrupção ou danos na memória.

6.3.4 - MÓDULO RELÓGIO DE TEMPO REAL

Os equipamentos de coleta automatizada de dados devem conter um Módulo de Relógio de Tempo Real. A função desse módulo é fornecer a estampa de tempo a cada requisição realizada pelo Módulo Controlador Central.

O módulo de relógio deve possuir uma bateria interna que mantenha seu funcionamento normal mesmo quando o sistema é desligado. Essa bateria pode ser do tipo descartável, da qual se espera uma vida útil mínima de 18 meses, ou recarregável, com uma capacidade de sobrevivência de 8 dias sem novas cargas (Deibel e Zumwalt, 1984).

Este módulo deve incluir um sistema de comunicação de falha e/ou má operação da sua bateria interna, com o intuito de informar aos sistemas do Módulo Controlador Central quaisquer problemas.

O cristal oscilador acoplado ao módulo deve ser dimensionado de modo a prover a melhor acurácia possível, sendo inaceitáveis equipamentos com um atraso maior que 0,3024 segundo por dia, ou seja, 3,5ppm.

6.3.5 - MÓDULO GPS

O equipamento associado ao Módulo GPS deve ter uma precisão de 50 centímetros para as medidas de solo, sendo imprescindível, além da especificação do equipamento por parte do projetista, a indicação do correto local de instalação da antena para que este sofra o menor nível de interferências possível.

Sendo a NMEA¹ a autoproclamada força de unificação para a indústria eletrônica marinha e de localização, esta editou o padrão NMEA 0183, visando a conexão de dispositivos eletrônicos marinhos de localização. O padrão NMEA 0183 define, entre outras coisas, as características elétricas do sinal, o protocolo da transmissão de dados, o sincronismo e os formatos específicos de sentenças de transmissão.

Como se trata do padrão mais amplamente usado na coleta de mensagens dos satélites GPS, é imperativo que o equipamento de GPS usado para o módulo obedeça a tal padrão para a coleta dos dados.

As mensagens pelos equipamentos no padrão NMEA consistem em até 82 caracteres. No Quadro 6.1 é possível verificar o retorno gerado por um GPS Tripmate 850, coletado durante 2 segundos em Leixlip, na Irlanda.

¹ Associação Nacional de Eletrônicos Marinhos (*National Marine Electronics Association*), <http://www.nmea.org/>

Observa-se que uma quantidade de dados considerável é coletada num curto espaço de tempo, uma vez que o padrão NMEA determina que haja uma coleta de posição (\$GPGGA) e uma leitura de quantidade de satélites ativos (\$GPGSA) a cada segundo. Os demais dados coletados, que dizem respeito à constelação de satélites disponíveis e visíveis para o sensor GPS (\$GPGSV), são dados dispensáveis para os equipamentos de coleta automatizada de dados aqui propostos.

Como a quantidade de leituras realizadas por esse padrão supera à precisão de leituras mínimas indicada para o Módulo Controlador Central (10 segundos), o Controlador Central deve ignorar as leituras realizadas pelo equipamento e salvar as mensagens de posição estática a cada 10 segundos. Equipamentos com precisão superior às mínimas podem ser projetados, coletando cada leitura de posição realizada pelo GPS.

```

/*****
$GPGGA,092750.000,5321.6802,N,00630.3372,W,1,8,1.03,61.7,M,55.2,M,,*76
$GPGSA,A,3,10,07,05,02,29,04,08,13,,,,,1.72,1.03,1.38*0A
$GPGSV,3,1,11,10,63,137,17,07,61,098,15,05,59,290,20,08,54,157,30*70
$GPGSV,3,2,11,02,39,223,19,13,28,070,17,26,23,252,,04,14,186,14*79
$GPGSV,3,3,11,29,09,301,24,16,09,020,,36,,*76
$GPRMC,092750.000,A,5321.6802,N,00630.3372,W,0.02,31.66,280511,,,A*43
$GPGGA,092751.000,5321.6802,N,00630.3371,W,1,8,1.03,61.7,M,55.3,M,,*75
$GPGSA,A,3,10,07,05,02,29,04,08,13,,,,,1.72,1.03,1.38*0A
$GPGSV,3,1,11,10,63,137,17,07,61,098,15,05,59,290,20,08,54,157,30*70
$GPGSV,3,2,11,02,39,223,16,13,28,070,17,26,23,252,,04,14,186,15*77
$GPGSV,3,3,11,29,09,301,24,16,09,020,,36,,*76
$GPRMC,092751.000,A,5321.6802,N,00630.3371,W,0.06,31.66,280511,,,A*45

```

Quadro 6.1 - Retorno de 2 segundos de leitura de um GPS com padrão NMEA.

6.3.6 - MÓDULO DE TRANSMISSÃO DE DADOS

Os modos dinâmicos de operação do transporte público carecem de informações online, que têm por objetivo a apresentação de respostas e de medidas imediatas às possíveis perturbações causadas pelos mais diversos fatores.

Dessa maneira, é importante que parte dos dados seja transmitida para as centrais de operação, visando, assim, a alimentação dos sistemas de monitoramento e a conversão de parte desses dados em subsídios para os sistemas ATIS.

Assim, os equipamentos de AHDC devem possuir meios de envio de parte dos dados em tempo real. O Módulo de Transmissão de Dados deve funcionar de maneira totalmente independente do Módulo Controlador Central, buscando diretamente no Módulo de Memória os dados passíveis de transmissão online, bem como os dados de erros.

Sugere-se que o Módulo de Transmissão de Dados envie novos dados a cada 30 segundos, assegurando a economia no volume de dados transitados com a internet. A estrutura de enlace e transmissão de dados deve ser definida pelo gestor do sistema de operação dos transportes, podendo ser realizada através do envio de um conjunto de dados para um serviço na internet, pelo *upload* do arquivo de dados para um servidor de transferência de arquivos ou, ainda, por outro modo que o sistema de interconexão suporte.

6.3.7 - MÓDULO REGISTRADOR DE PASSAGEIROS

O Módulo Registrador de Passageiros tem como finalidade detectar a entrada e a saída de passageiros do veículo, através de métodos eletrônicos que não interajam diretamente com o passageiro, nem necessitem de intervenção do motorista ou do cobrador do veículo.

Uma vez que os dados de volume de passageiros podem servir para a auditoria de sistemas de bilhetagem, a contagem deste deve ser totalmente independente dos sistemas de arrecadação automática, por mais eficientes que eles sejam.

Cada novo evento registrado por esse módulo deve ser enviado para o Módulo de Memória, contendo dados de direção do passageiro (se este estava entrando ou saindo do veículo), geolocalização do evento e estampa de tempo, uma vez que estes registros terão características assíncronas. A necessidade de detectar a direção do movimento dos passageiros serve para evitar sobreposições no caso de passageiros que desçam pela porta da frente, por exemplo.

Esse módulo deve ter, ainda, a capacidade de registrar, de maneira volátil, a quantidade total de passageiros dentro do ônibus, fornecendo para o Módulo Controlador Central essa quantidade a cada novo registro síncrono. A dupla gravação da quantidade de passageiros, tanto pelos registros diretos deste módulo quanto pelo cálculo matemático da diferença entre a quantidade de passageiros que entraram e saíram do veículo, permitem que os gestores da operação tenham um controle duplo do volume de passageiros registrado.

As informações sobre a quantidade de passageiros dentro dos veículos também podem ser enviadas de maneira online pelo Módulo de Transmissão de Dados, a fim de coordenar as ações de injeção e de recolhimento de veículos em operações dinâmicas.

Os testes realizados sugerem que a função de contagem da quantidade de passageiros somente deve estar ativa quando os sensores de abertura das portas e do acionamento da rampa para cadeirantes estiver ativo, evitando que a movimentação voluntária dos passageiros dentro dos veículos seja detectada e contada erroneamente.

Sendo estes os dados com maior protagonismo entre aqueles coletados pelos equipamentos de AHDC, os dados coletados devem conter uma precisão suficiente que lhes garanta característica censitária, permitindo sua utilização na expansão de matrizes de origem e destino.

Essa precisão segue conforme apontado por Deibel e Zumwalt (1984) e deve ser calculada para uma série de eventos de registro de dados, ou seja, logo após o ciclo de abertura e fechamento das portas, atendendo aos seguintes parâmetros:

1. Para 70% dos eventos de coleta, deve ser exatamente igual àquela determinada por verificadores manuais;
2. Para 85% dos eventos de coleta, nem a contagem de embarque, nem a contagem de desembarque deve exceder às contagens determinadas manualmente em ± 1 ; e,
3. Para 95% de todos os eventos, nem a contagem de embarque, nem a contagem de desembarque deve ser superior às contagens determinadas manualmente em ± 2 ;

Segundo os autores, ainda, para que o teste de acurácia seja considerado válido nas condições anteriormente aventadas, deve conter um mínimo de 500 testes consecutivos com as seguintes características mínimas:

1. O número médio de embarques por evento não deve ser inferior a 2; e,
2. O número médio total de passageiros para os eventos de embarque ou desembarque não deve ser inferior a 2.

6.3.8 - MÓDULO REGISTRADOR DE EVENTOS DO VALIDADOR

O Módulo Registrador de Eventos do Validador, por sua vez, tem como finalidade manter um registro, assíncrono, das respostas fornecidas pelos sistemas de bilhetagem, como a validação da entrada, o registro de gratuidade e também as recusas do sistema.

Conforme apresentado por Baroni *et al.* (2016), os sistemas de validação são bastante plurais, quadro que se repete para os transportes públicos rodoviários, sendo improfícua para a determinação de requisitos mínimos a imposição de qualquer tipo de coleta de dados ou a fixação do tipo de dado colhido.

Fica evidente, porém, que a coleta de dados dos registradores de eventos pode ser realizada com diferentes graus de profundidade, a ser definido com base na tecnologia adotada pelo sistema de transportes públicos rodoviários local:

- a) Coleta Completa: Cada nova ação do validador é transmitida para o Módulo Registrador de Eventos, através de comunicação serial, com as características completas da ação, indicando a entrada de gratuidade, o tipo de cartão (para localidades que diferenciam a tarifa pelo tipo de cartão) e se a operação obteve sucesso ou não;
- b) Coleta Parcial: Os dados detalhados da operação são perdidos pela ausência de acesso. Entretanto, é possibilitada a realização de uma conexão com o sistema de liberação eletromecânica, detectando a liberação do sistema de travas para que o passageiro avance;
- c) Coleta Indireta: Indicada nos casos em que todo o sistema de bilhetagem não permite a coleta de dados, sendo realizada pelo sistema de bloqueio infravermelho instalado no sistema de catraca (ou similar), permitindo a contagem unitária de passageiros que avançaram pelo bloqueio.

Por certo, a forma mais indicada é aquela que fornece uma maior variabilidade e volume de dados. No entanto, os métodos indiretos servem ao propósito de, no caso de impedimentos, ao menos apresentar dado de suposição de avanço através dos bloqueios.

6.3.9 - SENSOR ACELERÔMETRO

O acelerômetro pode ser definido como um dispositivo eletrônico usado para medir a aceleração própria num determinado eixo. A aplicação desse tipo de dispositivo é bastante ampla, sendo comum sua presença em *smartphones* e em equipamentos de robótica.

Um acelerômetro triaxial, ou seja, que coleta as informações de aceleração gravitacional nos eixos *x*, *y* e *z* pode ser usado para reconstruir o perfil de direção de um motorista, indicando comportamentos inseguros ou indevidos, como curvas e frenagens abruptas (Eren *et al.*, 2012; Li *et al.*, 2016).

Essa capacidade eletrônica viabiliza a análise do perfil de condução dos motoristas. O modo de captura desse sensor é melhor aproveitado quando se utiliza uma leitura a cada 300ms, com essas leituras convergindo para o Módulo Controlador Central, que deve ter a capacidade de coletar tais informações e processar usando o pseudocódigo apresentada no Quadro 6.2.

```
algoritmo calculo_forcaG
variável forcaG_x_medida: força gravitacional medida em x;
variável forcaG_y_medida: força gravitacional medida em y;
variável forcaG_z_medida: força gravitacional medida em z;
constante forcaG_x_padrao: força gravitacional padrão em x;
constante forcaG_y_padrao: força gravitacional padrão y;
constante forcaG_z_padrao: força gravitacional padrão z;

início
leia forcaG_x_medida, forcaG_y_medida, forcaG_z_medida
se forcaG_x_medida > forcaG_x_padrao ou forcaG_y_medida >
forcaG_z_padrao ou forcaG_z_medida > forcaG_z_padrao então
    registre forcag_x_medida, forcag_y_medida, forcag_z_medida no
    próximo registro ordinário
fim_se
fim
```

Quadro 6.2 - Pseudocódigo para o Sensor Acelerômetro

Vale ressaltar que a força gravitacional indicada como padrão para cada um dos eixos pode ser obtida através da literatura no quesito de conforto ou, ainda, de calibrações realizadas pelo operador dos transportes públicos.

6.3.10 - SENSOR DE ABERTURA DAS PORTAS, SENSOR DE ACIONAMENTO DO BOTÃO DE PARADA E DA RAMPA PARA O CADEIRANTE

Os sensores de abertura de porta, de acionamento do botão de parada e de acionamento da rampa para cadeirante são considerados “sensores de contagem simples”, coletando a alteração de estado destes elementos do veículo.

Todas as vezes que algum desses elementos citados altera seu estado, os sensores realizam uma coleta assíncrona, ou seja, o dado indicando o acionamento desses sensores deve ser salvo com sua estampa de tempo diretamente para o Módulo de Memória, em conjunto com os dados de localização.

Os sensores de acionamento das rampas para cadeirantes servem, ainda, para detectar população estatisticamente rara, mas que são de grande relevância para o planejamento da operação e para o diagnóstico da acessibilidade dos veículos.

6.3.11 - SENSORES DE LUMINOSIDADE, DE TERMOHIGRÔMETRO, DE PRECIPITAÇÃO, DE RUÍDO E DE QUALIDADE DO AR

Os dados ambientais ligados à viagem são coletados pelo conjunto de sensores que captam a luminosidade, a quantidade de ruídos, a qualidade do ar, a temperatura e a umidade interna, bem como a ocorrência de precipitação e a temperatura externa. Como os ônibus são veículos longos, é interessante a coleta em três pontos distintos registrando, assim, a característica média captada pelos sensores.

A qualidade do ar pode ser comprimida para a concentração de CO₂ no ar dentro dos veículos. Essa coleta pode ser realizada por sensores de baixo custo, com uma precisão de 400 a 10.000 ppm de CO₂.

Já o sensor de precipitação deve permitir a distinção entre chuvas leves e parada da chuva, preferencialmente, através de contato capacitivo. Essa característica é indicada pois os testes de conceito mostram que os sensores baseados em fechamento de circuito elétrico possuem uma

dificuldade considerável para a distinção do volume de chuvas, além de causar atrasos na detecção do término da precipitação.

Os sensores de precipitação fazem parte do circuito de sensores assíncronos, registrando a ocorrência do início e do final das precipitações.

6.4 - FORMATOS DOS DADOS

O volume de dados que convergem dos módulos e sensores para o Modulo Controlador Central e que, posteriormente, transita de todos os controladores centrais instalados numa frota para o banco de dados dos responsáveis pela operação dos transportes públicos, tem vultuosidade considerável, sendo imprescindível tratar esse assunto.

O primeiro ponto a ser observado diz respeito ao volume dos dados. Considerando apenas os registros síncronos, um único veículo gera 1,3MB de dados por dia, com 8.640 registros. Em uma frota como a de Brasília, com 3.156 veículos, diariamente seriam gerados 27.267.840 de registros síncronos (DFTrans, 2016). Por último, considerando 1,2 milhões de passageiros por dia, ter-se-ia, aproximadamente, 29.667.840 de registros diários (Lobo, 2015). Ao final de um ano, os mais de 10 bilhões de registros presentes no banco de dados poderiam consumir toda sua capacidade de armazenamento.

Mais do que gerar um banco de dados enorme, qualquer pequena alteração neste, proveniente de um sensor que parou de funcionar por alguns dias ou uma atualização do sistema, que insira um novo sensor e, portanto, que ocasione a criação de uma nova coluna, pode gerar um problema difícil de ser manejado pelos responsáveis.

O crescimento dos dados trafegados através da internet forçou a indústria da informática a enfrentar e resolver esse problema. O termo *big data*, muito usado quando se trata de grandes volumes de dados armazenados, tem permeado os maiores avanços tecnológicos e científicos da atualidade.

Os bancos de dados *big data* não se distanciam dos bancos de dados tradicionais apenas pelo seu gigantismo. Três características são suficientes para definir as características de um *big data*, a saber: o volume, a variedade e a velocidade (Mayer-schönberger e Cukier, 2013). Em publicações mais recentes, autores têm considerado, além destas, a veracidade e o valor, totalizando os chamados “5v's”, representados esquematicamente pela Figura 6.2.

Intuitivamente, a variável volume se refere, justamente, à grande quantidade de dados gerados e armazenados diariamente pelas diversas fontes de dados, associadas à variável variedade que indica a riqueza de fontes geradoras dos dados. A variável velocidade, por sua vez, refere-se à velocidade com que os dados são gerados e, principalmente, processados. Já os conceitos de veracidade e valor estão intimamente atrelados à confiança na fonte de tais dados e na usabilidade dos mesmos.

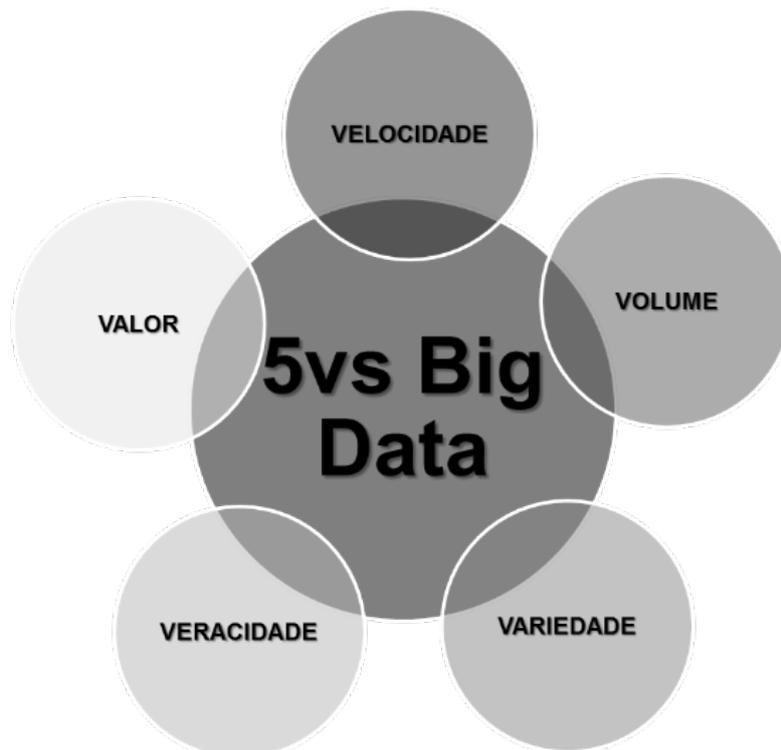


Figura 6.3 - Representação esquemática dos "5v's" dos *big data*.

A diferença entre um banco de dados tradicional e um *big data* transcende os quesitos filosóficos, atingindo questões práticas. Enquanto os bancos de dados tradicionais parecem muito com uma planilha de dados, na qual cada registro é inserido numa determinada coluna pré-definida para ele, no *big data* é completamente aceitável que uma ou mais colunas estejam faltando, ou sejam inseridas posteriormente (Mayer-schönberger e Cukier, 2013).

Esse fato leva de encontro com a resiliência dos dados pretendida pelos objetivos desta dissertação. Mesmo que os sistemas e os equipamentos mudem com o tempo, nenhuma alteração nos bancos de dados precisará ser realizada, uma vez que, para o *big data*, é completamente aceitável que determinados sensores e leituras não estejam disponíveis em parte dos arquivos.

Pensando-se na forma de armazenamento das informações coletadas, os bancos de dados apropriados para trabalhar com *big data* podem criar registros não relacionais de três maneiras distintas: baseado no esquema de chave/valor, baseado em documentos ou baseado em grafos.

Os bancos de dados que trabalham com chave/valor (*Key/Value*) são sistemas com um paradigma de armazenamento de dados projetado para armazenar, recuperar e gerenciar matrizes associativas, também conhecidas como dicionários de assinaturas.

Esses dicionários contêm uma coleção de objetos (ou registros) que, por sua vez, têm muitos campos diferentes dentro deles, como uma organização por matriz com n dimensões. A chave para encontrar cada uma dessas n dimensões está em um vetor a parte.

Embora possam remeter aos bancos de dados tradicionais, estes predefinem a estrutura das informações no banco de dados como uma série de tabelas contendo campos com tipos de dados bem definidos. Em contraste, os sistemas de valor-chave tratam os dados como uma coleção única opaca que pode ter campos diferentes para cada registro (Pokorny, 2013).

Bancos de dados orientados a documentos são aqueles nos quais as coleções de atributos (dados) são salvas num arquivo único. Usando-se a analogia de uma planilha eletrônica, é como se cada linha dessa planilha fosse salva num arquivo diferente.

Em geral, os bancos de dados orientados a documento não possuem esquema, ou seja, os documentos armazenados não precisam possuir estrutura em comum, embora a produção de esquemas aumente a potência e a velocidade das análises realizadas.

Diferentemente de outros tipos de bancos de dados ora apresentados, o modelo baseado em grafos está intimamente relacionado a um formato de dados previamente estabelecido. Esse modelo tem sua aplicabilidade potencializada quando se tratam de informações sobre relacionamento de dados, topologias e interconectividade.

Observando o caso dos dados gerados pelos sistemas de AHDC, os dados possuem uma propensão muito grande para o modelo baseado em arquivos. Pensando-se na potência utilizada pelos microcontroladores, os bancos de dados de chave/valor consumiriam uma capacidade de processamento desnecessária para serem gerados e mantidos, enquanto que os grafos não teriam aplicabilidade, uma vez que, neste momento, as relações entre os dados são dispensáveis. De maneira contrária, a maior parte dos microcontroladores comerciais possuem a capacidade nativa de escrever arquivos de texto puro, o que facilita a geração de arquivos para cada registro.


```

{ "structural_data" : [
  { "MCC": [1,1]      },
  { "MGE": [1,1]      },
  { "MM": [1,1]       },
  { "MRTR": [1,1]     },
  { "MGPS": [1,1]     },
  { "MTD": [1,0]      },
  { "MRP": [1,1]      },
  { "MREV": [0,0]     },
]
}
{ "collected_data" : [
  { "MRTR": "1433599200" },
  { "GPGGA":
"$GPGGA,092750.000,5321.6802,N,00630.3372,W,1,8,1.03,61.7,M,55.2,M,,*76"},
  { "GPGSA": "$GPGSA,A,3,10,07,05,02,29,04,08,13,,,,,1.72,1.03,1.38*0A"},
  { "MRP": "36" },
  { "SA": [0,0,0] },
  { "SL": "12000" },
  { "STH": [36,60] },
  { "STE": "34" },
  { "SR": "70" },
  { "SQA": "400" },
]
}

```

Quadro 6.3 - Estrutura usada para os registros síncronos

Os dados estão formados por dois blocos estruturais. Um deles, denominado “*structural_data*”, contém as informações de funcionamento de cada um dos módulos salvos. O outro, chamado de “*collected_data*”, trata dos dados coletados dos sensores síncronos. Para os dados estruturais, temos os seguintes valores e significâncias:

- MCC – Módulo de Controle Central, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento;
- MGE – Módulo de Gerenciamento de Energia, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento;

- MM – Módulo de Memória, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento;
- MRTR – Módulo Relógio de Tempo Real, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento;
- MGPS – Módulo de GPS, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento;
- MTD – Módulo de Transmissão de Dados, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele não está com seu funcionamento correto;
- MRP – Módulo Registrador de Passageiros, no qual o primeiro valor do vetor indica que o módulo está presente fisicamente e o segundo valor indica que ele está em completo funcionamento, e;
- MREV – Módulo Registrador dos Eventos do Validador, no qual o primeiro valor do vetor indica que o módulo não está presente fisicamente e o segundo valor, por consequência, indica que ele não está em completo funcionamento.

No caso dos dados coletados, tem-se os seguintes valores e significâncias:

- MRTR – É o valor coletado do Módulo Relógio de Tempo Real;
- GPGGA – São os dados provindos do Módulo GPS, contendo a coleta da posição do veículo;
- GPGSA – Apresenta os dados provindos do Módulo GPS, quanto à constelação de satélites ativos;
- MRP – Indica a quantidade de passageiros dentro do veículo, coletado do Módulo Registrador de Passageiros;
- SA – Registra se houve algum registro de aceleração acima dos padrões determinados para o veículo. Caso negativo, registra um vetor de zeros;

- SL – Registra a luminosidade dentro do veículo, em lux;
- STH – Registra a temperatura, no primeiro dado do vetor, e a umidade interna do veículo, no segundo dado do vetor;
- STE – Registra a temperatura externa do veículo;
- SR – Registra a quantidade de ruído no veículo em decibéis; e,
- SQA – Registra a quantidade de CO2 presente no ar, em partes por milhão.

Essa estrutura, proposta como requisito, pode ganhar registros de sensores e módulos adicionais. No entanto, esse formato básico deve ser mantido para que sua compatibilidade seja igualmente mantida.

Uma versão reduzida desses dados pode ser enviada pelo Módulo de Transmissão de Dados a cada 30 segundos, ou no prazo que melhor convier aos operadores dos transportes públicos, para os dados apresentados no Quadro 6.4 e com o mesmo significado e valores apresentados no Quadro 6.3.

```

{ "structural_data" : [
  { "MCC": [1,1] },
  { "MGE": [1,1] },
  { "MM": [1,1] },
  { "MRTR": [1,1] },
  { "MGPS": [1,1] },
  { "MTD": [1,0] },
  { "MRP": [1,1] },
  { "MREV": [0,0] },
]
}
{ "collected_data" : [
  { "MRTR": "1433599200" },
  { "GPGGA":
"$GPGGA,092750.000,5321.6802,N,00630.3372,W,1,8,1.03,61.7,M,55.2,M,,*76"},
  { "GPGSA": "$GPGSA,A,3,10,07,05,02,29,04,08,13,,,,,1.72,1.03,1.38*0A"},
  { "MRP": "36" }
]
}

```

Quadro 6.4 - Versão reduzida dos dados coletados para ser enviada em tempo real pelo Módulo de Transmissão de Dados

De forma similar, os registros extraordinários, gerados pelos eventos assíncronos, tiveram sua estrutura determinada da forma apresentada no Quadro 6.5.

```

{ "collected_data" : [
  { "*sensor ou módulo assíncrono*": "dado coletado"    },
  {"GPGGA":
"$GPGGA,092750.000,5321.6802,N,00630.3372,W,1,8,1.03,61.7,M,55.2,M,,*76"},
  {"GPGSA": "$GPGSA,A,3,10,07,05,02,29,04,08,13,,,,,1.72,1.03,1.38*0A"},

  ]
}

```

Quadro 6.5 - Estrutura usada para os registros assíncronos.

No exemplo apresentado no quadro, o campo delimitado por **sensor ou modo assíncrono** deve ser substituído pelo nome do sensor ou módulo que gerou o registro excepcional, sendo as possibilidades:

- MRP – Módulo Registrador de Passageiros – no qual o dado coletado deve ser guardado num vetor. O primeiro dado do vetor é a direção (sobe o desce) e, o segundo, é a quantidade de passageiros detectada no movimento;
- MREV – Módulo Registrador de Eventos do Validador – como não há consenso sobre os dados fornecidos pelos fabricantes dos sistemas de validação e a disponibilidade deles para os gestores dos transportes públicos, esse formato de dado deve ser definido por estes gestores, quando da implantação dos equipamentos AHDC;
- SAP – Sensor de Abertura das Portas – no qual o dado registrado é um vetor indicando o número da porta do veículo, contada sempre da frente para o fundo, e o movimento da mesma (se abertura ou fechamento);
- SBP – Sensor de Acionamento dos Botões de Parada – cujo dado enviado é o número 1;
- SARC – Sensor de Acionamento da Rampa para Cadeirante – cujo dado enviado é o número 1; e.
- SPr – Sensor de Precipitação – no qual o módulo salva como dado “1” o início da chuva e “0” o término desta.

6.5 - TÓPICOS CONCLUSIVOS

O presente Capítulo tratou de apresentar, de fato, os requisitos para a coleta automatizada de dados híbridos, incluindo a arquitetura básica dos equipamentos. Abordou, ainda, os equipamentos mínimos para a realização de tal tarefa, bem como o formato de dados esperado.

A arquitetura, embora desenhada de maneira a não engessar o desenvolvimento dos produtos derivados desse modelo, mostra que o equipamento de AHDC deve ser dividido em módulos e sensores, de forma muito similar àquela apresentada por Deibel e Zumwalt (1984).

Com o mesmo intuito de não engessar o progresso e a longevidade do método, a especificação de sensores contou com uma apresentação suficiente para nortear as necessidades básicas recebidas, sem o aprofundamento que pudesse imobilizar o método.

Por outro lado, pensando-se nas características esperadas para as coletas automatizadas de dados híbridos, como sua aplicação por um longo período de tempo e a resiliência esperada para os dados, o formato dos dados foi determinado de maneira a criar, efetivamente, um padrão.

O formato ora apresentado intuiu a resiliência dos dados, mantendo sua adaptabilidade para alterações futuras no modelo, como a inserção de novos sensores. Observa-se, ainda, que o formato adotado representa um dos mais modernos e promissores no âmbito dos grandes volumes de dados, possibilitando que os dados gerados pelos transportes públicos formem um *big data* e, assim, a previsão, o acompanhamento temporal e a comparação de tendências e de ações com o comportamento demonstrado nas redes.

Tal possibilidade se alinha com as previsões apresentadas pelos trabalhos de Cascetta *et al.* (1993), Jinhua (2004), Cui (2006), Zou *et al.* (2013), entre inúmeros outros. Ainda, alinha-se com a crescente modernização das cidades e com os novos conceitos de cidades inteligentes.

Finda a determinação dos requisitos, o capítulo seguinte apresentará uma comparação entre os dados coletados pelo modelo híbrido e as necessidades apresentadas pelas pesquisas mais recentes de mobilidade urbana. Posteriormente, no mesmo capítulo, será abordada uma simulação, em cenários distintos, do volume e da qualidade dos dados coletados.

7 - VALIDAÇÃO DO MÉTODO AHDC: COMPARAÇÃO, SIMULAÇÃO DE FUNCIONAMENTO E EMULAÇÃO DA ANÁLISE

7.1 - APRESENTAÇÃO

Sempre que uma inovação relativa a um método existente ou a um novo método é apresentada, a etapa de validação adquire um caráter coprotagonista, por demonstrar a aplicabilidade e o funcionamento daquilo proposto. A validação deve ser considerada como um teste preliminar para as tais novas abordagens, fundamentando sua aplicação em casos reais.

A coleta automatizada de dados híbridos não foge a essa regra, especialmente se for observado o esforço necessário para uma possível mudança de abordagem para aplicação desta forma de coleta numa localidade, ou mesmo, em parte de uma frota.

Dessa maneira, o presente capítulo desta dissertação apresentará o trabalho de validação do método de coleta automática de dados híbridos, comparando os dados apontados como requisitos mínimos com os estudos recentes, contratados por órgãos públicos, cujo objeto tenha sido a descrição da mobilidade urbana e/ou que tenham servido de subsídio para o Planejamento da Operação dos Transportes Públicos Urbanos.

Posteriormente, este capítulo apresentará uma simulação computacional da geração de dados, realizada com o intuito de demonstrar a ordem de grandeza dos dados coletados e as características dos dados gerados, como volume e variedade. Essa simulação servirá para demonstrar as potencialidades dos dados gerados e transmitidos online, dentro da estrutura aqui proposta para os métodos dinâmicos de operação dos transportes públicos rodoviários.

7.2 - COMPARAÇÃO COM OS ESTUDOS DE MOBILIDADE RECENTES

Conforme apresentado no Capítulo 5, as variáveis que foram vertidas em requisitos para as coletas automatizadas de dados híbridos foram determinadas através de uma extensa revisão bibliográfica e uma pesquisa sobre a percepção dos usuários quanto à qualidade dos transportes públicos. Essa revisão realizada culminou numa estrutura de dados que, indubitavelmente, abrange as mais recorrentes necessidades de dados.

Todavia, nem sempre as aplicações em casos reais seguem aquilo indicado pela bibliografia. Comumente, observa-se a introdução de variáveis de interesse numa determinada pesquisa, em complementação àquelas indicadas pela literatura utilizada como embasamento técnico para o desenvolvimento desta, a fim de identificar um problema ou uma característica específica de uma localidade.

Isto posto, uma das análises que deve ser realizada nesta obra diz respeito à avaliação de sua utilização para o atendimento das pesquisas realizadas com intuito similar ou que poderiam, por suas características, se valer dos dados híbridos levantados.

Nesse sentido, foram selecionadas três pesquisas balizadoras, a fim de comparar as variáveis elencadas para o desenvolvimento destas com os requisitos evidenciados no presente trabalho. Desta forma, foi possível comparar, também, os dados demandados dos gestores públicos com aqueles apurados por dispositivos de coleta automatizada.

As três pesquisas selecionadas para o efeito comparativo são: (i) o Plano Diretor de Transporte Urbano do Distrito Federal (PDTU/DF), de 2010; (ii) o Plano de Mobilidade Urbana Sustentável da Grande Florianópolis (PLAMUS), de 2014; e, por último, o Plano de Desenvolvimento do Transporte Público Sobre Trilhos do Distrito Federal (PDTT/DF), ainda em elaboração.

7.2.1 - PLANO DIRETOR DE TRANSPORTE URBANO DO DISTRITO FEDERAL

O PDTU/DF, conduzido pela Secretaria de Estado de Mobilidade do Distrito Federal (SEMOB) teve como proposta a realização de um diagnóstico da situação do Distrito Federal, a época de realização do estudo, bem como a harmonização com o perfil urbanístico histórico da cidade (SEMOB, 2010). O documento final do referido estudo pode ser considerado como um instrumento norteador do planejamento da gestão dos transportes urbanos.

A pesquisa contou com uma coleta de 6.069 amostras domiciliares, que representava 3,26% da população considerada para o projeto. A amostragem relativamente pequena em relação à população do Distrito Federal à época, comparativamente aos demais estudos analisados, se justifica pela expansão realizada a partir dos dados da Pesquisa Domiciliar de Transportes, concluída pela Companhia do Desenvolvimento do Planalto Central (CODEPLAN) no ano 2000 (SEMOB, 2010).

De acordo com as premissas apresentadas nos capítulos introdutórios deste trabalho, em vários aspectos, o PDTU/DF demonstrou possuir características que mais o aproximam de uma expansão de matriz origem-destino do que de uma geração de nova matriz.

Para a expansão dos dados foram coletadas, também, informações quanto aos atributos dos transportes rodoviários, como a quantidade de linhas, a tabela horária, a quantidade média de passageiros, entre outras.

Os dados gerados por esse trabalho subsidiaram a redação da Lei Distrital nº 4.566, aprovada em 04 de maio de 2011, que dispõe sobre o Plano Diretor de Transporte Urbano e Mobilidade do Distrito Federal – PDTU/DF. O referido plano visa, ainda, o atendimento à exigência do Estatuto das Cidades, que determina que todas as cidades com mais de 500 mil habitantes desenvolvam seus Plano Diretor de Transportes.

7.2.2 - PLANO DE MOBILIDADE URBANA SUSTENTÁVEL DA GRANDE FLORIANÓPOLIS

O PLAMUS, por sua vez, foi desenvolvido com o intuito de apresentar as soluções pertinentes para os problemas de mobilidade urbana identificados nos 13 municípios da Grande Florianópolis, tendo como motivador o atendimento às diretrizes impostas pela Política Nacional de Mobilidade Urbana (PLAMUS, 2015).

Como resultado da elaboração do PLAMUS, foram elaboradas as matrizes de origem e destino, a repartição modal das viagens executadas, a determinação de linhas de desejo, entre outros. Diferentemente do PDTU/DF, as características dessa outra pesquisa são de coleta direta para geração dos dados e não de expansão da matriz origem-destino.

7.2.3 - PLANO DE DESENVOLVIMENTO DO TRANSPORTE PÚBLICO SOBRE TRILHOS DO DISTRITO FEDERAL

O Plano de Desenvolvimento do Transporte Público sobre Trilhos do Distrito Federal nada mais é do que o aprofundamento do PDTU/DF, no tocante dos transportes sobre trilhos. É um instrumento que possui como principal objetivo a definição de diretrizes e de políticas estratégicas para a gestão dos transportes urbanos.

A abrangência das pesquisas realizadas no bojo do PDTT/DF transcende os dados diretamente ligados aos transportes públicos sobre trilhos. Propõe-se ao levantamento dos dados sobre o perfil de mobilidade completo do Distrito Federal, além de características socioeconômicas.

Essas pesquisas, realizadas numa fase denominada Pesquisa de Mobilidade Urbana (PMU/DF), extraiu informações de, aproximadamente 18.000 residências. Ainda, coletou dados sobre o deslocamento de cerca de 50.000 brasilienses (Metrô/DF, 2016).

Adicionalmente, foram solicitados ao DFTrans e ao Metrô/DF dados sobre a constituição de suas linhas, o volume desagregado de passageiros, entre outros. Esses dados serviram de

subsídio para a modelagem e a simulação dos sistemas de transportes no Distrito Federal, bem como para a construção de cenários hipotéticos.

7.2.4 - COMPARAÇÃO ENTRE OS DADOS LEVANTADOS NOS ESTUDOS E OS REQUISITOS LEVANTADOS

Para a efetiva comparação dos dados levantados em estudos com aqueles oferecidos pela coleta automatizada de dados híbridos, foi realizada uma leitura detalhada dos estudos, em especial dos trabalhos intermediários, uma vez que estes trazem detalhes ricos sobre a forma de condução dos levantamentos.

No caso dos dados levantados do PDTT/DF, foi necessária a solicitação de informações aos gestores do contrato, uma vez que o mesmo continua em fase de execução, não tendo seus dados ou relatórios técnicos parciais publicados até o momento.

Os dados solicitados aos órgãos gestores dos transportes públicos rodoviários foram tabulados lado a lado com aqueles pretensamente disponibilizados pelos equipamentos de coleta automatizada aqui propostos. Essa estruturação tabular é apresentada na Tabela 7.1.

Tabela 7.1 - Comparativo entre dados coletados por estudos *versus* dados ofertados pelos AHDC

Dados Solicitados	PDTU (2010)	PLAMUS (2015)	PMU/DF (2016)	AHDC
Conjunto de linhas operadas e paradas oficiais	X	X	X	X
Tabela horária	X	X	X	X
Frequência de atendimento em dias típicos		X	X	X
Volume de passageiros transportado	X	X	X	X
Volume de ingresso de passageiros em recorte temporal	X		X	X
Horário de operação das linhas	X	X		X
Valor da Tarifa		X	X	

A primeira verificação que pode ser realizada é a de que os principais dados solicitados pelos estudos e fornecidos pelos órgãos gestores são cobertos pelos dados gerados através dos sistemas de AHDC.

O valor da tarifa, solicitado por todos os estudos, embora não coberto pela constelação de dados ofertada pelos sistemas de AHDC, não devem representar problemas no seu fornecimento,

tampouco queda de qualidade dos dados oferecidos, uma vez que são informações de domínio público, geralmente, apresentadas através de decretos municipais ou distritais.

Notoriamente, observa-se que os dados fornecidos por sistemas de AHDC apresentam qualidade e precisão superior àquelas fornecidas pelos gestores dos sistemas de transporte público rodoviário.

Em todos os casos, foram solicitados os dados sobre a tabela horária e a estimativa média de usuários transitando em cada uma das linhas oferecidas pelo sistema de transporte rodoviário. Por um lado, as respostas obtidas dos órgãos, invariavelmente, tratam da tabela horária planejada e não daquela executada para os dias típicos. De forma similar, as informações relativas ao volume de usuários são dados médios, em agregações de 30 minutos.

Por outro lado, os dados oferecidos pelos AHDC dão conta dos horários realmente executados, permitindo a verificação da aderência de horários, do volume real da oferta, das características de dias típicos e atípicos. Ainda, possibilita a análise do volume de embarque e desembarque de passageiros em sua granularidade mínima.

As vantagens dos dados ofertados pelos sistemas AHDC ficam ainda mais evidenciadas pela pluralidade dos dados oferecidos. Embora os estudos não solicitem diretamente para os operadores dos transportes públicos, todos eles fazem inferências para determinar o volume de usuários cadeirantes no ônibus, dados que podem ser ofertados pelos sistemas que atendem aos requisitos mínimos dos AHDC.

É possível destacar ainda, no escopo dos estudos apresentados, que o AHDC permite a realização de verificação de sobe e desce a cada parada, dado importante para a determinação das seguintes informações: taxa de renovação de passageiros nos veículos; aderência à tabela horária; tempo de parada em cada ponto de ônibus; tempo de imobilização não programado no percurso; paradas e abertura de portas não programadas, indicando o ingresso ou descida de passageiros em pontos não oficiais.

Nenhum dos planos estudados coleta dados relativos à qualidade percebida das viagens, apesar da referência direta em todos esses estudos à forma de prestação de serviços de transporte para a comunidade. Dessa maneira, não é possível comparar o fornecimento de dados de qualidade gerado pelos AHDC com os estudos, embora sua relevância seja perceptível para estes.

Uma comparação da pertinência dos itens de qualidade coletados é possível através do trabalho de Martins (2015). Nesse trabalho, focado na determinação de um índice para a avaliação da qualidade do serviço de transporte público por ônibus, o autor faz uma ampla revisão bibliográfica sobre os itens de qualidade percebidos pelos usuários.

Partindo-se da tabela elaborada pelo autor, na qual são apresentados os itens de qualidade relatados na revisão bibliográfica e, selecionando-se apenas os itens que são diretamente mensuráveis, passamos a ter a Tabela 7.2, contendo quesitos de qualidade mensuráveis e sua comparação com os dados colhidos pelos AHDC.

Tabela 7.2 - Comparação de dados coletados para aferição de qualidade com as coletas AHDC (adaptado de Martins (2015))

	Parasuraman <i>et al.</i> (1988)	Lima Júnior, (1995)	Lima Júnior e Gualda (1995)	Bertozzi e Lima Júnior (1998)	Gomide (2003)	Ferraz e Torres (2004)	Cardoso (2006)	Rodrigues (2006)	Antunes (2009)	Bubicz e Sellito (2009)	Couto (2011)	Ramos (2013)	Morais (2012)	AHDC
Cobertura da rede											X	X		X
Comportamento dos operadores						X		X	X	X				X
Conectividade		X	X	X		X		X						X
Confiabilidade	X		X	X			X	X					X	X
Conforto		X	X	X			X				X	X		X
Disponibilidade					X									X
Frequência						X		X	X					X
Lotação						X		X		X			X	X
Pontualidade										X	X	X	X	X
Respeito aos pontos de parada										X				X
Responsividade	X		X	X						X				X
Tempo de viagem						X	X	X	X					X

A cobertura da rede, a conectividade, a confiabilidade, a disponibilidade, a frequência, a pontualidade, o respeito aos pontos de parada e o tempo de viagem podem ser levantados através da análise do deslocamento dos veículos, através dos dados de GPS.

O comportamento dos operadores é passível de verificação através dos dados coletados pelo sensor acelerômetro, que mensura a agressividade de condução, assim como o respeito aos limites de velocidade podem ser aferidos pelo tempo de deslocamento.

Uma vez que alguns itens que alteram a percepção de conforto do usuário são coletados, como a temperatura, a umidade, a luminosidade, os ruídos e a quantidade de passageiros, é possível inferir, satisfatoriamente, o índice de conforto em veículo.

A lotação, por sua vez, é possível de ser verificada através dos dados processados pelo Módulo Registrador de Passageiros, trecho a trecho. Essa segmentação, por exemplo, permite a análise do volume de passageiros entre dois pontos determinados para melhor compreensão do funcionamento das redes de transportes.

Por fim, a responsividade pode ser observada através da análise do volume de veículos em operação efetiva num determinado momento *versus* a demanda esperada para o horário e trecho. Desta forma, é possível mensurar o quão responsiva a demanda é em relação à operação dos transportes públicos e, assim, ajustar possíveis distorções diagnosticadas.

Conforme verificado, os dados de mensuração direta levantados na obra de Martins são completamente satisfeitos, direta ou indiretamente, pelas coletas AHDC, com a vantagem da continuidade perene da coleta de dados, além da pluralidade dos mesmos.

Observa-se, deste modo, que os requisitos apresentados por este trabalho atendem às necessidades apresentadas pelo universo acadêmico e prático do Planejamento da Operação dos Transportes Públicos Urbanos rodoviários, comparativamente com os trabalhos apresentados até o momento.

Esse atendimento diz respeito tanto ao volume necessário quanto à pluralidade esperada dos dados, excedendo as expectativas quanto aos quesitos de continuidade e de granularidade da coleta.

7.3 - SIMULAÇÃO DA SOLUÇÃO E VERIFICAÇÃO DE CONVENIÊNCIA

A segunda análise apresentada para a validação do método proposto diz respeito à emulação dos dados coletados, à simulação da solução e à verificação da conveniência. Esta etapa tem o intuito de verificar o volume e a consistência dos dados gerados, além da sua pertinência para o uso na operação tática dos transportes públicos.

A fim de alcançar o objetivo mencionado para essa etapa, foram criados softwares com o intuito de simular a geração de dados, assim como as condições encontradas em casos reais. Essas condições foram moduladas para formar um cenário para o teste de volume de dados e, ainda, outros três para a pertinência da operação.

Os cenários construídos para teste levam em consideração as características apresentadas pela cidade de Brasília para a construção do PDTU/DF. Assim, foram usadas as redes de transporte e as informações sobre volume de tráfego, demanda por viagens, número de linhas e operação do transporte rodoviário utilizado para a construção do diagnóstico daquele trabalho.

A Figura 7.1 apresenta a rede de transportes públicos considerada para a elaboração do trabalho do PDTU/DF. Essa rede, absorvida para o programa através de uma camada de dados geográficos chamada *shapefile*, representa a rede para o ano de 2009.

Da mesma maneira, a Figura 7.2 é uma representação gráfica da demanda por transportes públicos no Distrito Federal no ano de 2009. Nessa figura é possível observar o carregamento da demanda por viagens nas redes de transportes públicos do Distrito Federal.

Cabe ressaltar que ambas as imagens citadas anteriormente, assim como os dados coletados de tal relatório, não distinguem a demanda por transportes rodoviários da demanda por transportes metroviários. Desta forma, para efeito deste trabalho, a rede de transportes metroviários foi desconsiderada e a demanda foi considerada como sendo integralmente focada em transportes rodoviários.

7.3.1 - CARACTERÍSTICAS TÉCNICAS DA PLATAFORMA DE TESTE

A simulação de uma solução desse porte demandou de uma capacidade de processamento e de armazenamento incomum para computadores pessoais. Com isso, fez-se necessário mobilizar dois servidores, ligados através de uma rede com fio, além de um equipamento de armazenamento de dados (*storage*) com capacidade de 10 terabytes.

Cada um dos servidores mobilizados para a tarefa possui 500 gigabytes de HD, 16 gigabytes de memória RAM, processadores de 3 GHz e cache de 8 megabytes. O sistema operacional instalado para suporte ao banco de dados baseado em documentos foi o *Linux Debian* - 64bits, versão 8.7.1.

O banco de dados baseado em documentos escolhido para o teste foi o *Apache MongoDB* e a linguagem dos programas desenvolvidos para teste foi *Java*. Nenhum outro programa foi instalado, além daqueles que fazem parte do pacote básico da Debian para servidores e os necessários para o correto funcionamento da referida linguagem de programação e banco de dados.

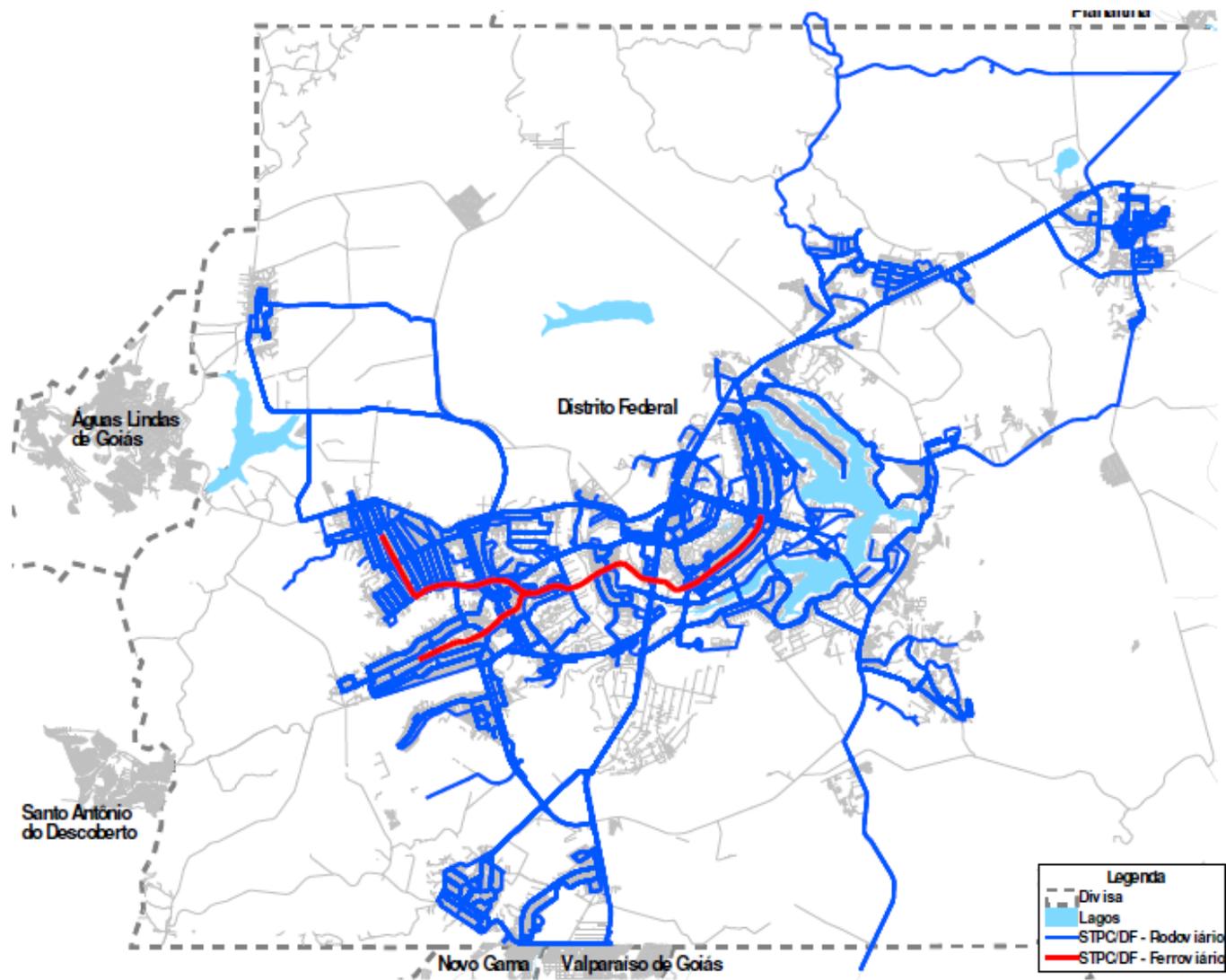


Figura 7.1 - Rede de transporte público coletivo urbano do Distrito Federal (Fonte: PDTU/DF, página 29 - SEMOB (2010))



Figura 7.2 - Carregamento da rede de transporte público coletivo do PDTU/DF, passageiros na hora de pico da manhã (Fonte: PDTU/DF, página 39 - SEMOB (2010))

Nenhuma consideração sobre a capacidade do servidor, do banco de dados escolhido ou da linguagem de programação utilizada ou, ainda, sobre melhorias de desempenho (*tuning*) será realizada no bojo deste trabalho. Esta dissertação manterá o protagonismo do método de coleta e seus requisitos, em detrimento de quaisquer outros aspectos, por mais interessantes que sejam.

Dessa maneira serão apresentados apenas detalhes essenciais para o entendimento e a validação do método, além daqueles já apresentados.

7.3.2 - TESTE DE VOLUME DE DADOS

O primeiro programa criado teve como finalidade a geração de um volume de dados correspondente a um ano de operação de um sistema similar àquele instalado em Brasília na ocasião da elaboração do PDTU/DF.

Em 2009, o Distrito Federal contava com 968 linhas de ônibus circulando, incluindo os desmembramentos operacionais, com uma produção quilométrica útil de 886.000km/dia e com a realização de 22.000 viagens/dia numa operação contínua de 18 horas (das 06h00m às 00h00m) (SEMOB, 2010). Ainda, segundo o referenciado documento, havia uma demanda de 1.571.917 viagens de ônibus por dia, desconsiderando o motivo da viagem.

A partir dos dados demonstrados, é possível inferir que o sistema de AHDC, com seu formato mínimo apresentado no Capítulo 6, respeitando a frequência mínima de coleta de 10 segundos, faria 6.480 leituras diárias por veículo em operação. Considerando uma frota de 2.359 veículos em circulação para atendimento das linhas, haveria uma geração de 15.286.320 arquivos de leitura ordinária por dia. Assim, ao final de um ano, seriam 5.579.506.800 de registros ordinários.

Com a finalidade justamente de construir o cenário mais “duro” e conservador possível, os sábados, domingos, feriados e dias de paralisação foram considerados como tendo volume e fluxo igual ao de um dia típico de circulação.

Em que pesem os registros extraordinários, haveria um total de 1.147.499.410 registros por ano, considerando que o sistema envia mensagens de embarque e desembarque e, ainda, mantidas as condições de estresse predeterminadas anteriormente. Dessa forma, ao final de um ano, ter-se-ia um banco de dados com um total de 6.727.006.210 registros, tanto de embarque quanto de desembarque.

Elaborou-se, então, um sistema que gerasse os 1.147.499.410 de registros extraordinários e os 5.579.506.800 de registros ordinários, no formato idêntico àquele apresentado no Capítulo 6. A geração de arquivos ordinários totalizou 2,6 terabytes de dados, contra 259 gigabytes de dados dos registros extraordinários, resultando numa massa de documentos não processados de 2,96 terabytes.

O referido volume, após a importação para o banco de dados, transformou-se em 3,01 terabytes de espaço ocupado em disco. Em se tratando de um volume de mais de 6,5 bilhões de registros, o tamanho do banco de dados é completamente aceitável, ocupando um espaço médio de 447 bytes por registro.

Nessa vertente de análise, uma segunda funcionalidade foi adicionada ao programa, que diz respeito à busca por determinados recortes de dados. Uma das principais finalidades de um sistema que congregue um completo registro do comportamento dos transportes públicos rodoviários é o fornecimento de dados dentro de recortes singulares destes.

Em que pese o conceito de *big data* o volume grande de dados armazenados pode representar um obstáculo considerável para a realização da seleção de dados e sua exportação, consumindo um alto custo computacional e de tempo.

Por conseguinte, foi adicionada a função de busca de um determinado conjunto de dados no banco de dados. Para dificultar ainda mais a tarefa de seleção, foi solicitado o retorno de dados de viagens realizadas às quarta-feira no horário das 17h00m às 20h30m.

Como não foi desenvolvida interface visual para o programa, o retorno da classe *Java* foi repassado, em resumo, para o console do Debian. A imagem do retorno gerado pelo console pode ser vista na Figura 7.3 - Retorno do comando de geração e busca dos dados. Figura 7.3.

```
Gerando dados...
Primeiras 24 horas de processamento.
Processamento estimado em 43 horas|
Gerando dados...
Dados gerados em 46horas, 32 minutos, 17 segundos
Importando dados para o banco de dadso
Dados importados em 52horas, 01 minutos, 36 segundos
Pressione enter para cancelar o processo de busca de dados... /
Iniciando busca por dados...
66.780 dados selecionados pelo objeto View em 16 minutos, 39 segundos
```

Figura 7.3 - Retorno do comando de geração e busca dos dados.

Essa busca demonstra que o formato de dados sugerido, embora gere um grande número de registros e um volume de dados ainda maior, é passível de aplicação de recortes de interesse e exportações. Evidencia-se, assim, a possibilidade de uso dos dados para finalidades de planejamento da operação, através da realização de seleções de dados históricos e de relevância.

7.3.3 - TESTE DE SIMULAÇÃO DE FORNECIMENTO DE DADOS

No prisma de relevância deste trabalho, é demonstrada a possibilidade de uso dos dados gerados para o planejamento dinâmico da operação dos transportes, através da detecção e resposta instantânea a fatos que perturbem a normalidade da operação.

Assim como as aplicações apresentadas por Zou *et al.* (2013), inúmeras outras formas de operar os transportes públicos rodoviários podem ser desenvolvidas com tecnologias disruptivas. Entretanto, as formas responsivas de gerir os transportes públicos carecem do fornecimento de dados online.

A fim de validar a possibilidade do uso dos dados transmitidos de maneira online pelo Módulo de Transmissão de Dados (MTD), um segundo software foi desenvolvido, conforme se pode visualizar na Figura 7.4. Este segundo programa, com interface visual, tem como finalidade a visualização dos dados transmitidos via MTD e a apresentação da síntese do mesmo.

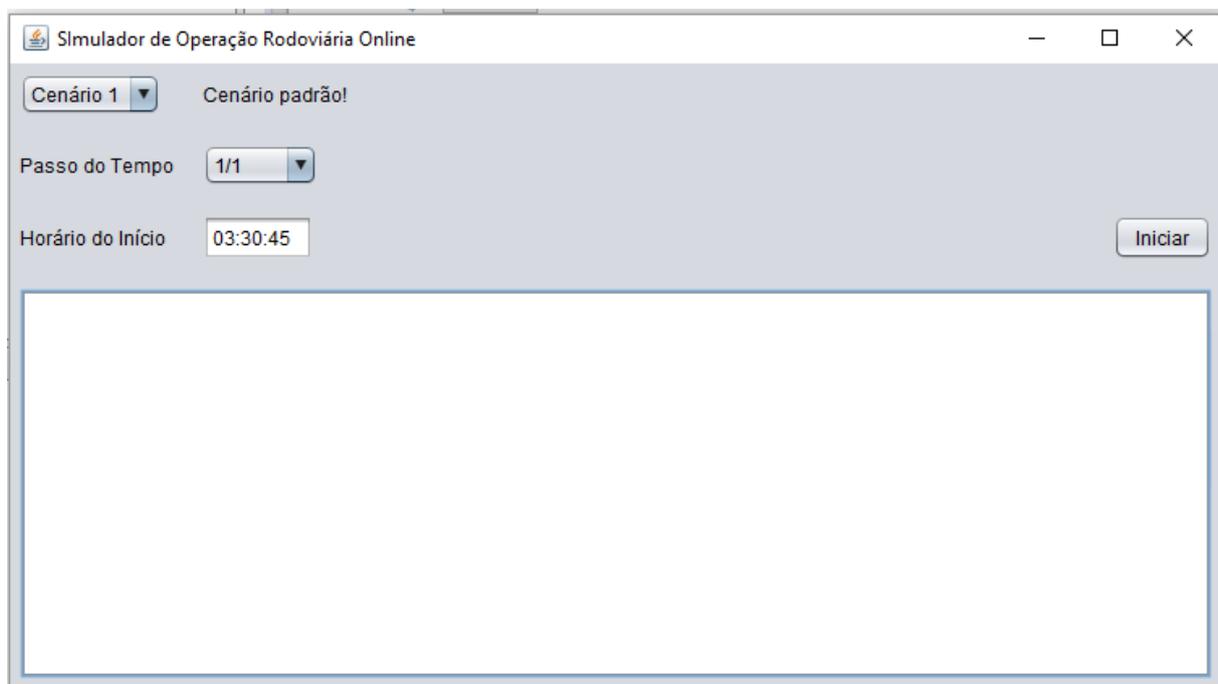


Figura 7.4 - Programa para simulação de Operação Rodoviária Online

Tendo a Figura 7.4 como referência, a caixa de seleção intitulada “Cenário 1” permite a seleção do cenário a ser testado para a simulação. A segunda caixa de seleção, chamada “Passo do tempo” indica que se a simulação ocorrerá em tempo real ou acelerado, acrescentando a conveniência de verificar as simulações de ocorrência de um dia todo em velocidades mais rápidas.

A caixa de texto com o nome “Horário do início” serve para que seja inserido o horário de início da simulação, evitando que o programa tente ler, desnecessariamente, horários fora daqueles de funcionamento. Por sua vez, o botão “Iniciar” começa o processo de simulação. Finalmente, a caixa de texto no canto inferior da tela apresenta os resultados sintéticos da simulação realizada.

O funcionamento do programa é, especialmente, simples. O programa apresentado no subitem anterior, foi modificado de forma a gerar dados aleatórios quanto à entrada de passageiros nos veículos, de forma que os somatórios desses dados fossem condizentes com a tabela de demanda por viagens do modo público demonstrada no PDTU/DF.

Os dados referentes ao posicionamento do veículo foram emulados para condizer com as linhas utilizadas para aquele estudo, eventualmente, tendo algumas linhas atuais inseridas no modelo para complementação, sem prejuízo da visualização proposta. As estampas de tempo geradas pelo programa foram planejadas para condizer com aquelas previstas na tabela horária, com um limite de funcionamento das 06h00m às 00h00m.

Uma segunda geração foi realizada, aproveitando todas as condições apresentadas anteriormente, exceto pelo fato de que os ônibus que circularam pelas vias das Regiões Administrativas de Águas Claras, Taguatinga, Ceilândia, Samambaia e Setor Policial Sul, experimentaram um atraso artificial em seu programa horário, simulando, assim, um atraso causado por problemas viários.

Em última instância, uma terceira geração de dados foi realizada, aproveitando todas as condições da primeira geração, com a exceção de que os dados de alguns veículos não foram gerados durante um determinado período de tempo, simulando assim a falta de comunicação temporária de alguns carros.

Os três cenários formados - “pleno funcionamento”, “trânsito” e “falta de comunicação” - foram escolhidos por tratar das situações mais básicas enfrentadas na gestão dinâmica dos transportes públicos.

Quanto ao funcionamento do programa, quando o botão “Iniciar” é clicado, este gera uma estampa de tempo, no padrão Unix e, a partir dela, é realizada uma busca pelos arquivos de todos os ônibus disponíveis no diretório programado para aquele horário. Dessa maneira, ele simula o recebimento das informações online, como se elas tivessem sido transmitidas naquele instante.

Ao final do processamento desses arquivos, o programa muda seu relógio interno para 30 segundos após a primeira leitura, gerando uma nova estampa de tempo e fazendo a busca no diretório por arquivos dos ônibus que tenham essa estampa de tempo. Isso ocorre de maneira cíclica, até que o final do dia de operação tenha terminado e, por consequência, os arquivos gerados.

No primeiro cenário simulado é possível verificar que foram lidos de maneira constante e sem mensagens de erro. Na tela do programa foram apresentadas as métricas das viagens que virtualmente estariam ocorrendo naquele momento. Os resultados podem ser observados na Figura 7.5.

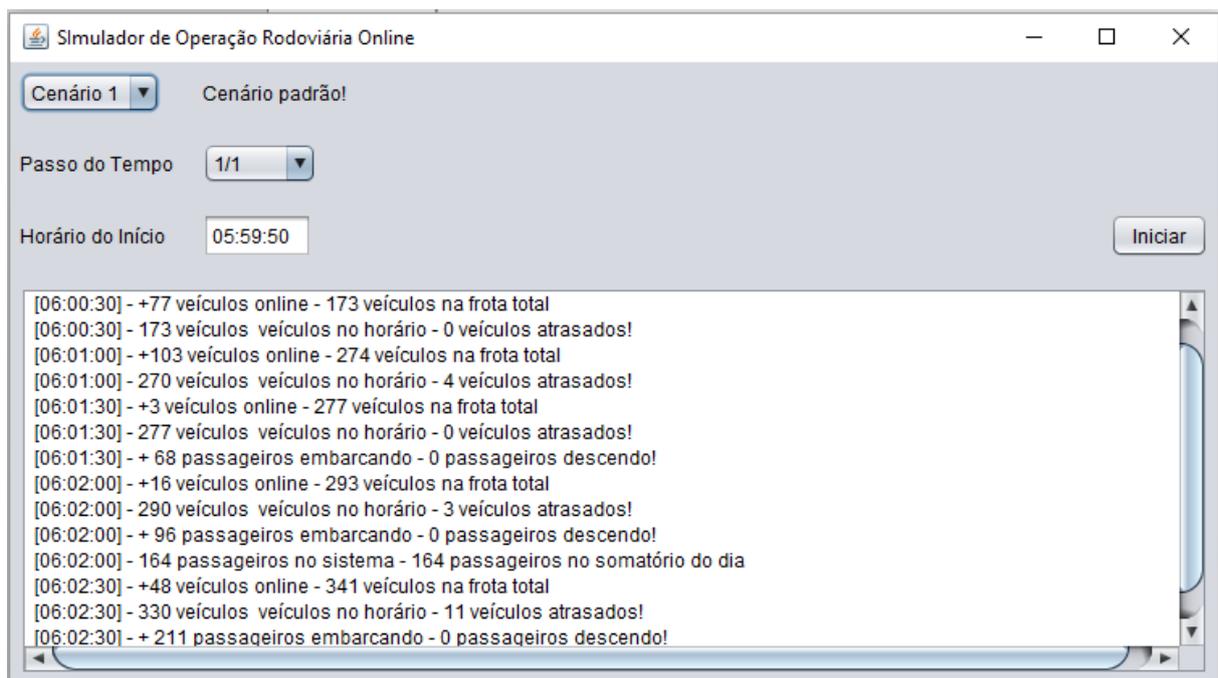


Figura 7.5 - Resultado da simulação para o cenário padrão.

Na simulação do segundo cenário, ao ler os dados do período em que havia uma retenção nas vias de Águas Claras, Taguatinga, Ceilândia, Samambaia e Setor Policial Sul, o programa passou a apresentar mensagens de atenção para que o operador tomasse as atitudes necessárias.

O programa continha instruções para ignorar atrasos inferiores a 3 minutos, considerando estes como atrasos normais e aceitáveis.

Os avisos, que podem ser vistos na Figura 7.6, indicam as linhas com ocorrência de atrasos. Interessante observar que o sistema continuou o reporte de atrasos mesmo para os veículos que não estavam mais nas zonas congestionadas, mas que haviam tido efeitos de retardo graças à mesma.

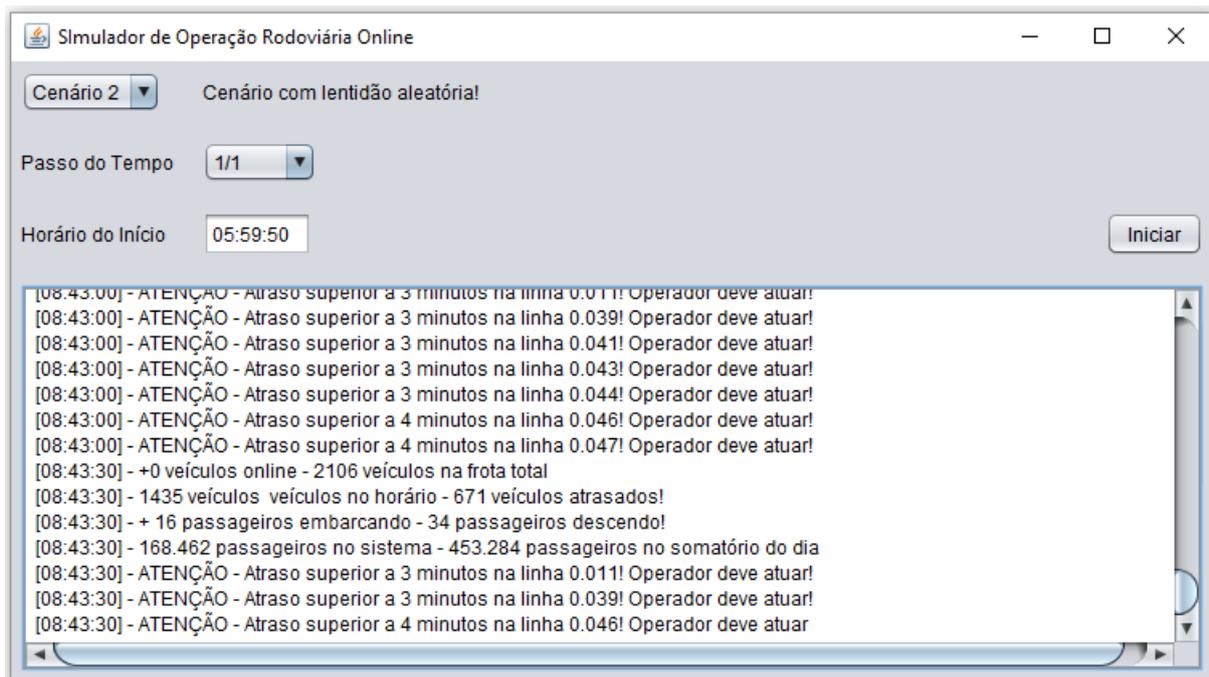


Figura 7.6 - Resultado da simulação para o segundo cenário

Os apontamentos apresentados pelo programa, embora bastante rudimentares, são indício categórico que o formato de dados híbrido, em sua versão que é coletada e transmitida via internet, supre as necessidades mínimas de sistemas DTPO e demais formas dinâmicas de operação.

Evidencia-se, ainda que possuir, de maneira dinâmica, a quantidade de passageiros presentes no sistema, permite ao operador a tomada de decisões táticas imediatas, como a injeção de mais veículos numa determinada linha. Embora o programa apresente dados sintéticos, outras formas de desagregação podem ser tentadas em análises por programas mais elaborados, permitindo a visualização do volume de passageiros numa determinada linha ou, até mesmo num carro específico.

Por último, foram simuladas as perdas de transmissão de dados pelo programa, replicando um cenário no qual há uma falha de sinal concentrada numa região ou um problema generalizado na transmissão realizada por um conjunto de equipamentos.

O sistema foi calibrado de maneira a somente acusar problemas na transmissão de dados depois de 10 ciclos (5 minutos), sem conseguir ler dados de um determinado conjunto de veículos. Como resultado, o sistema passou a apresentar informações sobre a perda de sinal com os veículos acometidos pelo problema, conforme Figura 7.7.

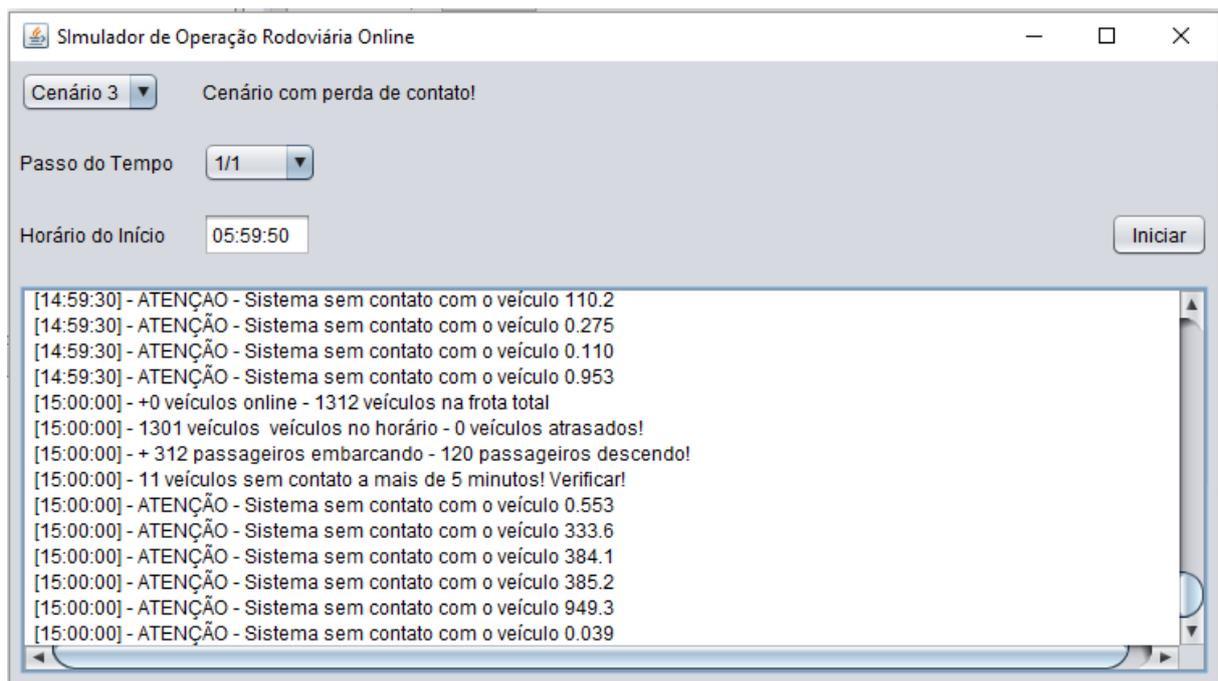


Figura 7.7 - Resultado da simulação para o terceiro cenário

Embora o sistema indique que os dados são um problema de conexão com a linha toda, o que pode não ser verdade, uma vez que é possível haver mais de um veículo operando uma determinada linha simultaneamente. Essa resposta ocorreu graças às características sintéticas dos dados apresentados, sendo completamente possível trabalhá-los para que representem dados por veículo.

7.4 - TÓPICOS CONCLUSIVOS

Neste capítulo foi apresentada a validação dos dados emulados gerados a partir dos requisitos estabelecidos neste trabalho. Verificou-se que, embora o volume de dados seja consideravelmente grande, ultrapassando a casa dos 6,5 bilhões de registros no final do primeiro ano, o formato indicado de exportação por esta dissertação se mostrou bastante eficaz.

Adicionalmente, demonstrou-se que a pesquisa nesses dados pode ser realizada num tempo razoável.

Deve ser pesado, dentro dessa esfera, que resultados melhores no quesito podem ser obtidos através de operações de otimização do hardware, do software e de requisições de pesquisa.

Na segunda validação realizada, quanto à pertinência dos dados gerados para a operação dinâmica dos transportes públicos, os dados se mostraram bastante promissores fornecendo informações de gestão bastante valiosas.

O formato de funcionamento do sistema replicado foi similar àquele encontrado em servidores de transferência de arquivos, no qual o MTD faria o envio do arquivo gerado para um diretório na internet, que, por sua vez, seria consumido pelo sistema de maneira volátil, ou seja, sem deixar registros gravados.

A volatilidade dos dados no cenário proposto é justificada pelo fato de que, se o sistema está transmitindo dados através do módulo específico, ele também está salvando dados no Módulo de Memória, com uma precisão três vezes maior. Dessa feita, o armazenamento perene dos dados coletados via internet pode ser considerada desnecessária.

8 - CONCLUSÕES E RECOMENDAÇÕES

8.1 - CONSIDERAÇÕES INICIAIS

Este trabalho apresentou uma abordagem acerca da lacuna presente na intersecção dos métodos de coleta automatizada existentes. Abordou, ainda, a ausência de uma padronização clara para as situações nas quais são necessárias as coletas de dados envolvendo mais de um dos métodos existentes, gerando um universo de dados incompatíveis entre si.

A tônica dos assuntos discutidos foi permeada pelas necessidades latentes para as formas de Planejamento da Operação dos Transportes Públicos Urbanos rodoviários, especialmente, para os formatos de gestão mais modernos, que carecem de um volume de dados consideravelmente grande e com pluralidade de informações.

Desta forma, a discussão de requisitos e padrões para uma coleta automatizada de dados híbridos ganha relevância e pertinência, especialmente num momento em que as cidades e seus gestores tendem a abordagens dinâmicas, modernas e transparentes sobre mobilidade.

Em que pese o desenvolvimento da obra, tanto pela fundamentação teórica apresentada - abordando os tópicos dominantes de ITS para que as coletas de dados híbridos fossem conceituadas e posicionadas em relação aos demais métodos - quanto pela conceituação da operação dos transportes públicos rodoviários - seus insumos e seus métodos modernos, pode-se afirmar que os objetivos propostos no capítulo inicial foram completamente satisfeitos.

Mediante o embasamento teórico que pavimentou as discussões ora apresentadas, foi possível trazer à luz a confluência de ambos os assuntos, permitindo que fossem elencadas variáveis mínimas para que um método de coleta automatizado capture dados híbridos, haja vista a gama de métodos clássicos já consagrados. Essa definição foi possível através de uma extensa análise da bibliografia existente.

Uma vez determinadas as variáveis de relevância, foram identificados os requisitos aplicáveis aos diferentes tipos de equipamentos de coleta, sejam eles módulos ou sensores, com o intuito de prover dados de maneira compatível com as necessidades aventadas.

Através dos resultados obtidos nos testes de validação aplicados, nos quais foram simulados os dados gerados por uma rede de equipamentos de coleta automatizada de dados híbridos, observa-se que o presente trabalho trouxe grandes contribuições nos seguintes enfoques:

(a) Exposição quanto aos equipamentos de ITS que tratam da coleta automatizada de dados e sua conexão com o Planejamento da Operação dos Transportes Públicos Urbanos, através da avaliação da literatura;

(b) Indicação da recorrência de dados coletados para pesquisas que visem a elaboração de planejamentos e estudos para a operação dos transportes públicos, inclusive quanto aos aspectos de qualidade do serviço prestado; e, ainda,

(c) Proposição de requisitos mínimos a serem atendidos pelos equipamentos de coleta automatizada construídos com o enfoque no fornecimento de dados para o planejamento e o Planejamento da Operação dos Transportes Públicos Urbanos.

8.2 - CONCLUSÕES

A condução desta dissertação permitiu que fosse concluído o quanto segue:

(1) A padronização mínima dos dados a serem coletados, assim como a apresentação de requisitos mínimos para a arquitetura dos módulos e sensores associados, traz clareza ao processo de determinação de tecnologias de ITS, com a finalidade de modernização dos processos de Planejamento da Operação dos Transportes Públicos Urbanos rodoviários. A presente obra pode ser usada como base de processos de estudos técnicos e licitatórios com esse intuito;

(2) Em que se pese os dados coletados dentro do padrão proposto, os dados coletados podem ser mais representativos no enfoque do comportamento de viagem do que aqueles dados que são costumeiramente coletados, principalmente quando devidamente correlacionados por métodos estatísticos;

(3) Dentro da ótica acadêmica, a determinação dos requisitos para a coleta automatizada de dados híbridos pode servir como diretriz para a realização de pesquisas dessa seara, permitindo a geração de dados intercambiáveis e comparáveis entre si, mesmo que de localizações diferentes;

(4) Adicionalmente, o setor industrial pode se valer, também, desta fonte como pré-requisito básico para a criação de equipamentos e de tecnologias para a coleta automatizada de dados. Ainda, para o desenvolvimento de aplicações que utilizem os dados gerados. Espera-se que, com a uniformização dos dados brutos, o processo de desenvolvimento de softwares e aplicações ganhe em assertividade e economia;

(5) Os três formatos de dados apresentados pelo trabalho demonstraram características de relevância, tanto para a formação de *big data*, nos quesitos variedade, valor, volume e veracidade. O quesito de velocidade, por depender significativamente de elementos que excedem o contexto do trabalho, não foi verificado. Ainda com testes rasos e rudimentares, foi possível demonstrar grande aplicabilidade nos dados gerados, abrindo uma imensa gama de possibilidades de aplicações.

8.3 - LIMITAÇÕES DA PESQUISA

A presente obra apresenta algumas limitações, elencadas a seguir:

(1) Embora tenham sido apresentados minimamente, os quesitos inerentes aos componentes eletrônicos, sejam aqui considerados módulos ou sensores, trabalhos mais aprofundados nesse tema precisam ser desenvolvidos a fim de exemplificar e testar mais amplamente a arquitetura e a sinergia dos equipamentos a serem utilizados;

(2) O mesmo pode ser observado no que tange os testes realizados com os softwares desenvolvidos. A característica minimalista dos programas desenvolvidos para a realização deste trabalho serviu apenas para a validação dos dados simulados. Entretanto nuances mais profundas de usabilidade, bem como melhores resultados na capacidade de pesquisa e de seleção de dados, podem ser reveladas por estruturas melhor elaboradas;

(3) O trabalho não cobre outros modais importantes dos transportes públicos, como modais metroferroviários e aqueles realizados como particulares. Entendendo os transportes públicos como uma rede, o planejamento da operação precisa contar com dados de todos os atores em igualdade de precisão e de frequência.

8.4 - RECOMENDAÇÕES PARA PESQUISAS FUTURAS

Pesquisas futuras derivadas do presente trabalho podem considerar os tópicos a seguir:

- Utilizando a estrutura mínima demonstrada, sugere-se o desenvolvimento de algumas unidades do equipamento para a coleta dos dados e, ainda, a implantação deste em veículos de uma determinada linha, com o intuito de verificar a aplicabilidade do método em situações reais de uso;
- A investigação das possíveis diferenças na expansão das matrizes de origem e destino usando dados contínuos provindo de equipamentos AHDC *versus* a expansão realizada

tendo como base pesquisas de sobe e desce pontuais poderiam validar a potência dos dados coletados;

- Tendo em vista os dados reais coletados por equipamentos de AHDC, podem ser estudados os possíveis ajustes detalhados da plataforma computacional de suporte para esse tipo de coleta, como a determinação do banco de dados baseado em documentos que apresente as melhores e mais rápidas respostas para o procedimento de busca e armazenamento.

- Um trabalho pode ser elaborado analisando as variações sazonais no volume de passageiros tentando relacionar com a com os fatores ambientais e de qualidade coletados pelo padrão, através de correlação estatística.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alsger, A., Assemi, B., Mesbah, M., e Ferreira, L. (2016) Validating and improving public transport origin-destination estimation algorithm using smart card fare data. *Transportation Research Part C: Emerging Technologies*, 68, 490–506. doi:10.1016/j.trc.2016.05.004
- Alvarez-Garcia, J. A., Ortega, J. A., Gonzalez-Abril, L., e Velasco, F. (2010) Trip destination prediction based on past GPS log using a Hidden Markov Model. *Expert Systems with Applications*, 37(12), 8166–8171. doi:10.1016/j.eswa.2010.05.070
- Antunes, E. M. (2009) *Avaliação da qualidade do transporte público por ônibus sob o ponto de vista do usuário em cidades médias paranaenses*. Universidade Estadual do Paraná.
- Baroni, T. H. F., Berardinelli, L. M. A., e Silva, A. H. (2016) Virtual Ticket para Acesso às Estações do Metrô-DF. *XXII Semana de Tecnologia Metroferroviária* (p. 1–23). AEAMESP - Associação dos Engenheiros e Arquitetos de Metrô de São Paulo, São Paulo.
- Baroni, T. H. F., e Souza, Y. T. de. (2015) Impacto da regulação da coleta de dados nos métodos de coleta automatizada e no planejamento dinâmico dos transportes. *XXIX Congresso Nacional de Pesquisa em Transporte da ANPET* (p. 2332–2341). ANPET, Ouro Preto. Obtido de <http://www.anpet.org.br/xxixanpet/anais/documents/AC789.pdf>
- Ben-Akiva, M. (1987) Methods to Combine Different Data Sources and Estimate Origin-Destination Matrices. *Transportation and Traffic Theory*.
- Ben-Akiva, M., Macke, P. P., e Hsu, P. S. (1985) Alternative Methods to Estimate Route-Level Trip Tables and Expand on Board Surveys. *Transportation Research Record*, 1037, 11.
- Ben-Akiva, M., e Morikawa, T. (1989) Data Fusion Methods and Their Applications to Origin-Destination Trips Tables. *Transportation and Traffic Theory*.
- Bertozzi, P. P., e Lima Júnior, O. F. (1998) A qualidade no serviço de transporte público sob as óticas do usuário, do operador e do órgão gestor. *Revista dos Transportes Públicos*.
- Bohte, W., e Maat, K. (2009) Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*, 17(3), 285–297. doi:10.1016/j.trc.2008.11.004
- Brisken, A. F., Anderson, R. E., Frey, R. L., e Lewis, J. R. (1979) Land mobile communications and position fixing using satellites. *IEEE Transactions on Vehicular Technology*, 28(3), 153–170. doi:10.1109/T-VT.1979.23787
- Bruton, M. J. (1979) *Introdução ao Planejamento dos Transportes*. (1ª.). Interciência/USP, Rio de Janeiro.
- Bubicz, M. E., e Sellito, M. A. (2009) Qualidade em serviço de transporte de passageiros: um estudo de caso no sistema urbano de Porto Alegre. *Revista Produção Online*.
- Câmara, G., e Monteiro, A. M. V. (2001) Conceitos básicos em ciência da geoinformação. G.

- Câmara, C. Davis, & A. M. V. Monteiro (Eds), *Introdução à Ciência da Geoinformação* (1ª., p. 1–35). INPE, São José dos Campos. Obtido de <http://www.dpi.inpe.br/gilberto/livro/introd/index.html>
- Cardoso, B. C. (2006) *Qualidade de serviço no setor de transportes sob a ótica da Teoria dos Topoi*. Universidade Federal do Rio de Janeiro.
- Cascetta, E., Inaudi, D., e Marquis, G. (1993) Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts. *Transportation Science*, 27(4), 363–373. doi:10.1287/trsc.27.4.363
- Cintra, M. (2014) Os custos dos congestionamentos na cidade de São Paulo. Obtido de <http://bibliotecadigital.fgv.br/dspace/handle/10438/11576>
- Corrêa, E. ([s.d.]) Introdução ao formato JSON. Obtido 5 de março de 2017, de <http://www.devmedia.com.br/introducao-ao-formato-json/25275>
- Couto, D. M. (2011) *Regulação e Controle Operacional no Transporte Coletivo Urbano: Estudo de Caso no Município de Belo Horizonte/MG*. Universidade Federal de Minas Gerais.
- Cui, A. (2006) Bus passenger Origin-Destination Matrix estimation using Automated Data Collection systems. Obtido de <http://dspace.mit.edu/handle/1721.1/37970>
- Deibel, L. E., e Zumwalt, B. (1984) Modular approach to on-board automatic data collection systems., (Transportation Research Board, National Research Council), 123. Obtido de <http://trid.trb.org/view.aspx?id=217011>
- Dell’Olio, L., Ibeas, A., e Cecin, P. (2011) The quality of service desired by public transport users. *Transport Policy*, 18(1), 217–227. doi:10.1016/j.tranpol.2010.08.005
- DFTrans - Transporte Urbano do Distrito Federal. (2016) *Total de frotas cadastradas por tipo de ônibus*. Brasília.
- Du, J., e Aultman-Hall, L. (2007) Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues. *Transportation Research Part A: Policy and Practice*, 41(3), 220–232. doi:10.1016/j.tra.2006.05.001
- Eboli, L., e Mazzulla, G. (2008) A Stated Preference Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology*, 31(5), 509–523. doi:10.1080/03081060802364471
- Eren, H., Makinist, S., Akin, E., e Yilmaz, A. (2012) Estimating driving behavior by a smartphone. *2012 IEEE Intelligent Vehicles Symposium* (p. 234–239). IEEE. doi:10.1109/IVS.2012.6232298
- Faouzi, N.-E. El, Leung, H., e Kurian, A. (2011) Data fusion in intelligent transportation systems: Progress and challenges – A survey. *Special Issue on Intelligent Transportation Systems*, 12(1), 4–10. doi:10.1016/j.inffus.2010.06.001
- Fernández, L. E., de Cea, J., e Malbran, R. H. (2008) Demand responsive urban public transport system design: Methodology and application. *Transportation Research Part A: Policy and*

Practice, 42(7), 951–972. doi:10.1016/j.tra.2007.12.008

- Ferraz, A. C. P., e Torres, I. G. E. (2004) *Transporte Público Urbano*. Rima, São Paulo.
- Gil, A. C. (2007) *Como elaborar projetos de pesquisa*. (4. ed.). Editora Atlas, São Paulo.
- Gomide, A. Á. (2003) *Transporte urbano e inclusão social: elementos para políticas públicas* (No. 960). Brasília.
- Hoel, L. A., Garber, N. J., e Sadek, A. W. (2011) *Engenharia de Infraestrutura de Transportes*. (1ª.). CENGAGE Learning, São Paulo.
- Huang, A., e Levinson, D. (2015) Axis of travel: Modeling non-work destination choice with GPS data. *Transportation Research Part C: Emerging Technologies*, 58, 208–223. doi:10.1016/j.trc.2015.03.022
- Jinhua, Z. (2004) The planning and analysis implications of automatic data collection systems: Rail transit OD matrix inference and path choice modelling examples. *Thesis of Master, M.S. thesis, MIT, USA*. Obtido de <http://18.7.29.232/handle/1721.1/28752>
- Kneib, E. C. (2008) *Subcentros urbanos: contribuição conceitual e metodológica à sua definição e identificação para planejamento de transportes*. Universidade de Brasília, Brasília. Obtido de http://repositorio.unb.br/bitstream/10482/1765/1/2008_ErikaCristineKneib.pdf
- Komninos, N., Schaffers, H., e Pallot, M. (2011) Developing a Policy Roadmap for Smart Cities and the Future Internet. *eChallenges e2011*, 1–8. doi:10.1109/MCOM.2013.6525605
- Li, Y., Xue, F., Feng, L., e Qu, Z. (2016) A driving behavior detection system based on a smartphone's built-in sensor. *International Journal of Communication Systems*, 23(5), 633–652. doi:10.1002/dac.3178
- Lima Júnior, O. F. (1995) *Qualidade em serviços de transportes: conceitualização e procedimentos para diagnóstico*. São Paulo.
- Lima Júnior, O. F., e Gualda, N. D. F. (1995) Qualidade em serviços de transportes: conceitualização e procedimentos para diagnóstico. *Congresso da Associação Nacional de Pesquisa e Ensino em Transportes* (p. 668–679). ANPET IX, São Carlos.
- Lobo, R. (2015) Os 8 maiores sistemas de transporte do Brasil. *Via Trolebus*. Obtido 31 de dezembro de 2016, de <http://viatrolebus.com.br/2015/03/os-8-maiores-sistemas-de-transporte-do-brasil/>
- Ma, X., Wang, Y., Chen, F., e Liu, J. (2012) Transit smart card data mining for passenger origin information extraction. *Journal of Zhejiang University SCIENCE C*, 13(10), 750–760. doi:10.1631/jzus.C12a0049
- Martins, W. T. (2015) *Avaliação Da Qualidade Do Transporte Público a Partir Da Definição De Serviço Adequado*. Universidade de Brasília. Obtido de <http://repositorio.unb.br/handle/10482/18393>
- Mattos, A. N. de. (2004) *Telemetria e conceitos relacionados*. Artigo de Discussão.

- Mayer-schönberger, V., e Cukier, K. (2013) *Big Data - Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana*. (1º ed). Campus, São Paulo.
- Medrano, R. M. A. (2016) *O Modelo Intencional de Transporte: Contribuições da Ontologia de Bunge para formalização da Teoria de Comportamento em Transporte*. Universidade de Brasília.
- Metrô/DF - Companhia do Metropolitano do Distrito Federal. (2016) Pesquisa de Mobilidade Urbana do DF. Obtido 31 de outubro de 2016, de <http://www.pmudf.com.br/>
- Morais, J. S. (2012) *Proposta de método para avaliação da qualidade do transporte público urbano por ônibus utilizando a Teoria das Representações Sociais*. Universidade de Brasília.
- Moreira-Matias, L., Mendes-Moreira, J., de Sousa, J. F., e Gama, J. (2015) Improving Mass Transit Operations by Using AVL-Based Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 1636–1653. doi:10.1109/TITS.2014.2376772
- Munizaga, M. A., e Palma, C. (2012) Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies*, 24, 9–18. doi:10.1016/j.trc.2012.01.007
- Nachouki, G., e Quafafou, M. (2008) Multi-data source fusion. *Information Fusion*, 9(4), 523–537. doi:10.1016/j.inffus.2007.12.001
- Nam, T., e Pardo, T. A. (2011) Smart city as urban innovation. *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance - ICEGOV '11*, 185. doi:10.1145/2072069.2072100
- Nitsche, P., Widhalm, P., Breuss, S., e Maurer, P. (2012) A Strategy on How to Utilize Smartphones for Automatically Reconstructing Trips in Travel Surveys. *Procedia - Social and Behavioral Sciences*, 48, 1033–1046. doi:10.1016/j.sbspro.2012.06.1080
- Ortúzar S., J. de D., e Willumsen, L. G. (2011) *Modelling Transport*. (Fourth edi.). John Wiley & Sons, Chichester, West Sussex, United Kingdom.
- Parasuraman, A., Zeithaml, V. A., e Berry, L. L. (1988) Servqual: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
- Passos, G. (2015) Internautas questionam exceções em anteprojeto que regula uso de dados pessoais. Obtido 17 de julho de 2015, de <http://www.ebc.com.br/tecnologia/2015/03/internautas-questionam-excecoes-em-anteprojeto-que-regula-uso-de-dados-pessoais>
- Perk, V., e Kamp, N. (2003) Handbook On Automated Data Collection Methods For The National Transit Database. *National Technical Information Service*, (October), 60. Obtido de <http://www.nctr.usf.edu/pdf/473-11.pdf>
- Pires, A. B., Silva, A. C. e, e Vasconcellos, E. A. (1997) *Transporte Humano – Cidades com Qualidade de Vida*. (1ª). ANTP, São Paulo. Obtido de http://files-server.antp.org.br/_5dotSystem/download/dcmDocument/2014/10/03/9AFE933E-903C-

4B31-B2A4-1FB59795FD13.pdf

- PLAMUS, P. de mobilidade urbana sustentável da grande F. (2015) Relatório Final - Consolidação das Propostas e Plano de Implementação.
- Pokorny, J. (2013) NoSQL databases: a step to database scalability in web environment. *International Journal of Web Information Systems*, 9(1), 69–82. doi:10.1108/17440081311316398
- Pontes, B. M. S. (2005) O planejamento dos transportes nas áreas metropolitanas. *Cadernos Metrópole*, (14). Obtido de <http://revistas.pucsp.br/index.php/metropole/article/view/8787>
- Prado, K. C. D., e Santos, P. E. dos. (2014) *Smart Cities: Conceito, Iniciativas E O Cenário Carioca*. Universidade Federal do Rio de Janeiro. Obtido de <http://monografias.poli.ufrj.br/monografias/monopoli10012947.pdf>
- Ramos, M. W. (2013) *Qualidades medida e percebida no sistema de transporte coletivo por ônibus: estudo de caso de Belo Horizonte*. Universidade Federal de Minas Gerais.
- RECIFE. (2015) Prefeitura do Recife inicia Pesquisa de Origem e Destino. Obtido 30 de outubro de 2016, de <http://www2.recife.pe.gov.br/noticias/17/11/2015/prefeitura-do-recife-inicia-pesquisa-de-origem-e-destino>
- Rede Nossa São Paulo. (2016) Pesquisa sobre Mobilidade Urbana é lançada em São Paulo. Obtido 31 de outubro de 2016, de <http://www.nossasaopaulo.org.br/noticias/pesquisa-sobre-mobilidade-urbana-e-lancada-em-sao-paulo>
- Redman, L., Friman, M., Gärling, T., e Hartig, T. (2013) Quality attributes of public transport that attract car users: A research review. *Transport Policy*, 25, 119–127. doi:10.1016/j.tranpol.2012.11.005
- Rodrigues, M. O. (2006) *Avaliação da qualidade do transporte coletivo da cidade de São Carlos*. Universidade de São Paulo.
- SEMOB - Secretaria de Estado de Mobilidade do Distrito Federal. (2010) *Plano Diretor de Transporte Urbano e Mobilidade do Distrito Federal - Relatório Final*. Brasília. Obtido de http://editais.st.df.gov.br/pdtu/final/relatorio_final.pdf
- Valenzuela, M., Ha, V., e Etzioni, O. (2015) Identifying Meaningful Citations. *AAAI Workshop on Scholarly Big Data*, 6. Obtido de <http://ai2-website.s3.amazonaws.com/publications/ValenzuelaHaMeaningfulCitations.pdf>
- Vlassenroot, S., Gillis, D., Bellens, R., e Gautama, S. (2014) The Use of Smartphone Applications in the Collection of Travel Behaviour Data. *International Journal of Intelligent Transportation Systems Research*, 13(1), 17–27. doi:10.1007/s13177-013-0076-6
- Vlassenroot, S., Gillis, D., Bellens, R., e Gautama, S. (2015) The Use of Smartphone Applications in the Collection of Travel Behaviour Data. *International Journal of Intelligent Transportation Systems Research*, 13(1), 17–27. doi:10.1007/s13177-013-0076-6

- Vuchic, V. R. (2005) *Urban Transit: Operations, Planning and Economics*. John Wiley & Sons Inc, Nova Iorque.
- Wang, W., Attanucci, J. P., e Wilson, N. H. M. (2011) Bus Passenger Origin-Destination Estimation and Related Analyses Using Automated Data Collection Systems. *JOURNAL OF PUBLIC TRANSPORTATION*, 14(4), 131–150. Obtido de <http://scholarcommons.usf.edu/jpt/vol14/iss4/7/>
- Wang, W. L., Lo, S. M., e Liu, S. B. (2015) Aggregated Metro Trip Patterns in Urban Areas of Hong Kong: Evidence from Automatic Fare Collection Records. *Journal of Urban Planning and Development*, 141(3), 5014018. doi:10.1061/(ASCE)UP.1943-5444.0000225
- Williams, B. (2008) *Intelligent Transport Systems Standards*. Transportation. Artech House, Norwood.
- Willumsen, L. G. (1978) *Estimation of an O-D matrix from traffic counts - A Review* (No. 99). *Institute of Transport Studies*. Obtido de <http://eprints.whiterose.ac.uk/2415>
- Zhao, J., e Rahbee, A. (2007) Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems. *Computer-Aided Civil and Infrastructure Engineering* 22, 22, 376–387.
- Zou, M., Chen, X. (Michael), Yu, H., Tong, Y., Huang, Z., Li, M., e Zou, H. (2013) Dynamic Transportation Planning and Operations: Concept, Framework and Applications in China. *Procedia - Social and Behavioral Sciences*, 96(Cictp), 2332–2343. doi:10.1016/j.sbspro.2013.08.262

ANEXO I – TRANSCRIÇÃO DO FORMULÁRIO DA PESQUISA APLICADA

 <p>Comportamento em Transportes e Novas Tecnologias</p>	Impressão de pesquisa <i>online</i>	
	Formulário de aplicação <i>offline</i>	
PESQUISA SOBRE O CONFORTO NOS TRANSPORTES PÚBLICOS		
PLATAFORMA SURVEY - www.srvey.co		

PESQUISA FINALIZADA EM 19/03/2016

Nome:		Sexo:	
Idade:	RA que reside:		
É usuário de Transportes Públicos? S / N	Possui veículo? S / N		

Pensando nas suas experiências GERAIS com os transportes públicos, quanto ao ponto de ônibus classifique as afirmações abaixo.					
	1	2	3	4	5
O caminho até o ponto de ônibus é seguro.					
O caminho até o ponto de ônibus tem calçadas em bom estado.					
Sinto-me tranquilo (tranquila) em acessar o ponto de ônibus a qualquer horário.					
O ponto de ônibus é seguro.					
O ponto de ônibus é confortável.					
O ponto de ônibus é coberto.					
No ponto de ônibus existem informações suficientes sobre horários e rotas que param ali.					
Fora do horário de pico há bancos suficientes.					
Não me sinto desconfortável em esperar mais que 10 minutos na parada.					

Para o entrevistador: 1 - Concordo Totalmente / 2 - Concordo Parcialmente / 3 - Indiferente / 4 - Discordo Parcialmente / 5 - Discordo Totalmente

Classifique os itens abaixo do mais importante (1) para o menos importante (6)						
	1º	2º	3º	4º	5º	6º
Caminho seguro até a parada						
Caminho confortável até a parada						
Parada segura						
Parada confortável						
Esperar pouco tempo na parada						
Ter informações atualizadas sobre horários e linhas						

Para o entrevistador: Apenas uma classificação por linha!



Impressão de pesquisa *online*

Formulário de aplicação *offline*

PESQUISA SOBRE O CONFORTO NOS TRANSPORTES PÚBLICOS



PLATAFORMA SURVEY - www.srvey.co

PESQUISA FINALIZADA EM 19/03/2016

Pensando nas suas experiências GERAIS com os transportes públicos, classifique as afirmações abaixo.

	1	2	2	3	4	5
Os veículos cumprem seus horários.						
Os veículos estão em bom estado de conservação.						
Sinto-me seguro (segura) em relação a acidentes de trânsito enquanto estou no veículo.						
Sinto-me seguro (segura) em relação a assaltos/violência enquanto estou no veículo.						
O motorista dirige de maneira não agressiva e respeitando as leis de trânsito.						
Sinto-me confortável com a quantidade de passageiros nos ônibus.						
Sinto que posso decidir não embarcar em um ônibus muito cheio sem prejudicar meus horários.						
A qualidade do ar dentro do ônibus me parece boa.						
A temperatura dentro dos ônibus é sempre agradável.						
A luminosidade dentro dos ônibus me parece boa.						
A quantidade de ruído dentro dos ônibus me parece agradável.						
Não me incomoda em circular, por curtos trajetos, empé.						
Os funcionários da empresa de transportes são sempre cordiais e prestativos.						

Para o entrevistador: 1 - Concordo Totalmente / 2 - Concordo Parcialmente / 3 - Indiferente / 4 - Discordo Parcialmente / 5 - Discordo Totalmente

Classifique os itens abaixo do mais importante (1) para o menos importante (13)

	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º	11º	12º	13º
Viajar sentado													
Temperatura dentro do ônibus													
Quantidade de pessoas de pé													
Iluminação no ônibus													
Qualidade do ar (cheiro ou fumaça) dentro do veículo													
Comportamento do motorista (dirigindo)													
Polidez do motorista/cobrador													
Quantidade de barulho dentro do veículo													
Tempo de viagem													
Conservação do veículo													
Idade do veículo													
Confiança de chegar no horário													
Confiança que os ônibus irão transitar todos os dias													

Para o entrevistador: Apenas uma classificação por linha!