



GERAÇÃO DINÂMICA DE PADRÕES DE MOVIMENTAÇÃO RODOVIÁRIA DE  
CARGAS E VEÍCULOS BASEADA EM IDENTIFICAÇÃO POR  
RADIOFREQUÊNCIA

Alexandre Rojas

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Transportes, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Transportes.

Orientador: Paulo Cezar Martins Ribeiro

Rio de Janeiro  
Dezembro de 2013

GERAÇÃO DINÂMICA DE PADRÕES DE MOVIMENTAÇÃO RODOVIÁRIA DE  
CARGAS E VEÍCULOS BASEADA EM IDENTIFICAÇÃO POR  
RADIOFREQUÊNCIA

Alexandre Rojas

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ  
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA  
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM  
CIÊNCIAS EM ENGENHARIA DE TRANSPORTES.

Examinada por:

---

Prof. Paulo Cezar Martins Ribeiro, Ph.D.

---

Prof. Carlos David Nassi, D.Sc.

---

Prof. Felipe Maia Galvão França, Ph.D.

---

Prof. Werner Kraus Junior, D.Sc.

---

Prof<sup>a</sup>. Regina Serrão Lanzillotti, D.Sc.

---

Prof<sup>a</sup> Rosa Maria Esteves Moreira da Costa, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

DEZEMBRO DE 2013

Rojas, Alexandre

Geração Dinâmica de Padrões de Movimentação Rodoviária de Cargas e Veículos baseada em Identificação por Radiofrequência / Alexandre Rojas – Rio de Janeiro: UFRJ/COPPE, 2013.

XIII, 159 p.: il. ; 29,7 cm.

Orientador: Paulo Cezar Martins Ribeiro

Tese (Doutorado) – UFRJ/ COPPE/ Programa de Engenharia de Transportes, 2013.

Referências Bibliográficas: p. 146-159.

1. Sistemas Inteligentes de Transportes - ITS 2. Padrões Frequentes de Movimentação 3 Planejamento dos Transportes 4. Sistemas de Informação 5. Mineração de Dados 6. Ontologia  
I. Ribeiro, Paulo Cezar Martins. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Transportes. III. Título.

## AGRADECIMENTOS

Ao Professor Paulo Cezar por seus ensinamentos, dedicação, amizade e confiança, elementos sem os quais nenhum trabalho seria possível;

Ao Professor Carlos Nassi e demais professores do PET pela transferência de conhecimento e amizade durante essa jornada;

Ao Professor Amaranto Lopes Pereira (*in memoriam*) pelo modelo, exemplo e excelência no ensino;

Aos Professores Felipe França, Werner Krauss e Rosa Costa, membros da Banca, por sua contribuição e participação;

Às Professoras Neide Santos e Regina Lanzillotti por sua ajuda em todos os momentos;

Ao Professor Antonio Carlos de Azevedo Ritto por sua amizade e incentivo;

Aos alunos da UERJ, Alexandre Neves Ferreira, Grace Rodrigues da Conceição, Jean Pacheco Dias, Luciano Silva Barroso e Edilon de Andrade, por sua ajuda na implementação dos modelos concebidos nesta Tese;

Aos colegas do PET, Antonio Gusmão, Bianca Côrtes e João Pereira, por seu companheirismo;

Ao Professor e colega Marcelo Schots por seu incentivo e ajuda na elaboração desta Tese;

À Jane e Helena da secretaria do PET por sua estimável ajuda em toda vida acadêmica.

À Marluce, companheira inseparável, por seu apoio e afeto.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.).

GERAÇÃO DINÂMICA DE PADRÕES DE MOVIMENTAÇÃO RODOVIÁRIA DE  
CARGAS E VEÍCULOS BASEADA EM IDENTIFICAÇÃO POR  
RADIOFREQUÊNCIA

Alexandre Rojas

Dezembro/2013

Orientador: Paulo Cezar Martins Ribeiro

Programa: Engenharia de Transportes

Este trabalho desenvolve um método teórico fundamentado nos princípios dos Sistemas Inteligentes de Transporte (ITS) com o objetivo de determinar padrões de movimentação de veículos e cargas a partir de dados coletados por sistemas nacionais baseados em identificação por radiofrequência. Composto por três modelos, sendo dois para aquisição e análise exploratória dos dados e um conceitual para representação do conhecimento em relação à movimentação do tráfego. O primeiro identifica os padrões frequentes de movimentação de carga em rodovias com o desenvolvimento de aplicativo computacional que disponibiliza relatórios estatísticos descritivos, validado a partir de resultados compatíveis com o tráfego de veículos na via. O segundo estabelece padrões de movimentação de veículos em região urbana e foram usadas técnicas de mineração de dados. O terceiro, desenvolvido em sua forma conceitual, usa paradigmas da Ontologia para representar o conhecimento para padrões de movimentação considerando aspectos relativos ao tráfego e à localização do veículo.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

DYNAMIC GENERATION OF FREQUENT PATTERN VEHICLE AND FREIGHT  
MOVEMENT OVER BRAZILAN ROAD NETWORK BASED ON RFID DATA

Alexandre Rojas

December/2013

Advisor: Paulo Cezar Martins Ribeiro

Department: Transportation

This thesis develops theoretical method based on the principles of Intelligent Transportation Systems (ITS) in order to determine movement patterns of vehicles and loads from data collected by national systems based on RFID. The method consists of three models, two models for acquisition and exploratory data analysis, and a conceptual model for knowledge representation in relation to the movement of traffic. The first model identifies the frequent patterns of cargo on highways with the development of computer application that provides descriptive statistical reports, and validated based on results compatible with vehicle traffic on the road. The second model aims to identify patterns of movement of vehicles in urban areas and techniques were used for data mining. The selected algorithm to identify the pattern of movement was the "nearest neighbor" and the results compared with previous studies obtaining compatible values. The third model, developed in its conceptual form, using paradigms of ontology, aims to represent knowledge in the field of Transportation Engineering for movement patterns considering aspects of the context of traffic and vehicle location.

## SUMÁRIO

Capítulo 1 - Introdução.....	1
1.1 Objetivo.....	2
1.2 Características do Método.....	2
1.3 Contribuições.....	3
1.4 Premissas.....	4
1.5 Relevância.....	4
1.6 Estrutura do trabalho.....	5
Capítulo 2 - Planejamento dos Transportes.....	6
2.1 Modelagem do tráfego e transporte.....	6
2.2 Modelo Convencional ou das Quatro Etapas.....	6
2.2.1 Coleta de Dados para o Planejamento.....	7
2.2.2 Primeira Etapa: Geração de Viagens.....	23
2.2.3 Segunda Etapa: Distribuição de Viagens.....	25
2.2.4 Terceira Etapa: Distribuição Modal.....	28
2.2.5 Quarta Etapa: Alocação de Tráfego.....	29
2.2.5.1 Abordagens principais.....	30
2.2.5.2 Alocação Dinâmica do Tráfego.....	32
2.3 Métodos para previsão da demanda futura.....	36
2.3.1 Métodos baseados em fatores de crescimento.....	36
2.3.2 Métodos baseados em atividades.....	37
Capítulo 3 - Modelos adotados no contexto do Planejamento de Transporte de Carga....	38
3.1 Conceitos básicos.....	38
3.2 Modelos adotados no Brasil.....	39
3.2.1 Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro - PDTU.....	40
3.2.2 Plano Nacional de Logística e Transportes - PNLT.....	45
3.3 Modelos adotados no Exterior.....	51
3.3.1 Estados Unidos da América.....	51
3.3.1.1 Considerações preliminares.....	51
3.3.1.2 Considerações sobre aspectos geográficos.....	52
3.3.1.3 Considerações sobre a construção da Matriz OD.....	53
3.3.2 União Europeia.....	55
3.4 Análise comparativa dos Modelos de Transporte de Carga.....	55
3.4.1 Modelos Norte-Americanos.....	55
3.4.2 Modelos Europeus.....	58
3.5 Considerações Finais.....	59

Capítulo 4 – Aspectos da Tecnologia da Informação aplicáveis ao desenvolvimento dos Modelos .....	60
4.1 Levantamento das Bases de Dados Nacionais.....	61
4.1.1 Premissas básicas .....	61
4.1.2 Sistema relacionado à Identificação Automática de Veículos.....	62
4.1.3 Sistemas relacionados ao transporte de carga .....	66
4.2 Considerações sobre a integração dos sistemas: .....	70
4.3 Modelagem de Domínio para determinação da movimentação do veículo ou da carga .....	70
4.4 Tecnologias da Informação da Abordagem Sistema de Informações .....	73
4.5 Tecnologias da Informação da Abordagem Mineração de Dados.....	75
4.5.1 Considerações Preliminares .....	75
4.5.2 Etapas do Processo de KDD.....	77
4.5.2.1 Seleção de Dados .....	78
4.5.2.2 Pré-processamento .....	78
4.5.2.3 Mineração de Dados.....	80
4.6 Modelagem ‘Representação do Conhecimento’ .....	84
4.6.1 Sistemas de Informação Sensíveis a Contexto .....	86
4.6.2 Ontologia – Contexto histórico .....	87
4.6.3 Ontologia- Descrição .....	88
4.7 Sistemas Simuladores ou Geradores de Dados para Transporte .....	91
4.7.1 <i>Softwares</i> Simuladores de Tráfego.....	91
4.7.2 Critérios Usados para a Seleção:.....	93
4.7.3 <i>Softwares</i> analisados .....	94
Capítulo 5 - Método Proposto .....	99
5.1 Descrição do Método.....	99
5.2 Considerações Preliminares.....	101
5.3 Abordagem – Sistema de Informações.....	103
5.3.1 Considerações Preliminares .....	103
5.3.2 Estrutura dos Dados .....	104
5.3.3 Localização das Antenas .....	105
5.3.4 Linguagens Utilizadas no Desenvolvimento do Modelo.....	105
5.3.5 Arquitetura do Sistema.....	107
5.3.6 Modelo de Requisitos.....	108
5.3.7 Principais Atores do Sistema.....	109
5.3.8 Casos de Uso.....	110
5.3.9 Modelagem dos Dados .....	111
5.3.10 Validação do Modelo .....	113
5.3.10.1 Parte 1 – considerando apenas o deslocamento na via .....	113
5.3.10.2 Parte 2 – considerando as origens e os destinos do CT-e.....	116
5.4 Abordagem – Mineração de Dados.....	118

5.4.1	Objetivo da Abordagem .....	118
5.4.2	Validação do Modelo .....	122
5.4.2.1	Configurações das ruas .....	122
5.4.2.2	Configurações dos veículos.....	123
5.4.2.3	Simulação.....	124
5.4.2.4	Considerações Finais.....	133
5.5	Abordagem Representação do Conhecimento.....	135
5.5.1	Considerações Iniciais.....	135
5.5.2	Diagrama de Contexto.....	136
5.5.3	Ontologia Tempo .....	136
5.5.4	Ontologia EventoTemporal .....	138
5.5.5	Ontologia Veículo .....	139
5.5.6	Ontologia Localização .....	139
5.5.7	Ontologia Carga .....	141
5.5.8	Ontologia Rodovia .....	141
5.5.9	Ontologia ODCM.....	141
5.6	Apresentação em Congressos Internacionais .....	142
5.7	Considerações Finais.....	142
Capítulo 6 – Conclusões e Recomendações .....		143
BIBLIOGRAFIA .....		146

## LISTA DE FIGURAS

Figura 2.1 Modelo das Quatro Etapas -baseado em (MCNALLY, 2007) .....	7
Figura 2.2 Componentes da área de estudo .....	8
Figura 2.3 Volume de Trafego – Linha de Desejo .....	9
Figura 2.4 Laço (loop) magnético – foto obtida de (DNIT, n/a) .....	15
Figura 2.5 Câmera de captura OCR – Fonte ITS International march/april 2013 .....	16
Figura 2.6- Veiculo Aéreo Não Tripulado – Fonte ITS International fev 2013 .....	20
Figura 2.7 Matriz de viagens – baseado em (CAMPOS, N.a.) .....	24
Figura 2.8 – Matriz OD dinâmica – baseada em (GUOZHEN, LINDONG, et al., 2011).....	32
Figura 3.1 Linha de Contorno RMRJ obtido em (SECTTRAN, 2002).....	41
Figura 3.2 Zoneamento PDTU Fonte (SECTTRAN, 2002).....	42
Figura 3.3 Rodovias consideradas no PDTU Carga Fonte (SECTTRAN, 2002).....	43
Figura 3.4 Principais Agregações Fonte (SECTTRAN, 2002) .....	44
Figura 3.5 Processo de Planejamento de Transportes usado no PNLT (2007).....	46
Figura 3.6 Zonas de Transporte do PNLT Fonte: PNL T(2007) .....	47
Figura 3.7 Vetores Logísticos Fonte (PNLT, 2007).....	48
Figura 3.8 Carregamentos Multimodais Fonte: (PNLT, 2007).....	49
Figura 3.9 Principais Produtos do FAF 3 Fonte (FHWA, 2011) .....	51
Figura 3.10 Geografia do FAF3 Fonte (FHWA, 2011) .....	52
Figura 3.11 Fluxograma da Construção da Matriz OD Fonte FHWA (2011) .....	53
Figura 4.1 - Visão geral do SINIAV Fonte Baseado em (CONTRAN, 2006).....	63
Figura 4.2 Funcionalidades Básicas do Brasil-ID .....	65
Figura 4.3 – Principais Corredores do Brasil-ID Projeto Piloto (VON BRAUN, 2010).....	66
Figura 4.4 - Exemplo de Trajetórias de Objetos entre Pontos de Coleta de Dados .....	71
Figura 4.5 Exemplo de Dígrafo Linear.....	72
Figura. 4.6 Etapas Operacionais do Processo de KDD.....	76
Figura 5.1 Principais Componentes.....	98
Figura 5.2. Projeto SINIAV -Monitoramento dos Veículos .....	99
Figura 5.3 Concepção Lógica para aquisição de dados .....	100
Figura 5.4 – Passos para Configuração do Sistema .....	106
Figura 5.5 –Geração da Matriz ODCM .....	107
Figura 5.6 – Responsabilidades dos atores do sistema .....	109
Figura 5.7. Diagrama do Modelo de Entidade-Relacionamento (MER) .....	110
Figura 5.8. Tela Principal - Perfil Gerador de Tráfego.....	111
Figura 5.9. Tela de gerenciamento de portais.....	112
Figura 5.10. Gráfico Tipo de Cargas por Semana .....	113
Figura 5.11. Relatório de Cargas .....	114
Figura 5.12. Tela Principal - Perfil Administrador .....	115

Figura 5.13 . Tela de Geração do Relatório ODRCM .....	116
Figura 5.14 - Layout da rede usado como base para simulação. ....	122
Figura 5.15 - Novo layout da rede de simulação. ....	123
Figura 5.16- Simulação das antenas de RFID. ....	124
Figura 5.17 - Simulação em andamento. ....	125
Figura 5.18 - Weka Explorer .....	126
Figura 5.19- Tela de escolha do algoritmo .....	126
Figura 5.20 – Relatório- Todos os Veículos .....	127
Figura 5.21 - Utilização do filtro Discretize .....	128
Figura 5.22 –Relatório com agrupamento por faixas de velocidade.....	129
Figura 5.23 -Relatório com dados da categoria caminhão.....	130
Figura 5.24 - Relatório excluindo os registros do tipo taxi.....	131
Figura 5.25 Diagrama de Contexto.....	134
Figura 5.26 Ilustração da Ontologia Tempo .....	135
Figura 5.27 Ontologia EventoTemporal .....	136
Figura 5.28 Ontologia Veiculo .....	137
Figura 5.29 Ontologia Localização .....	138
Figura 5.30 Ontologia Carga .....	139

## LISTA DE QUADROS

Quadro 2.1 Vantagens e Desvantagens dos Tipos de Pesquisa .....	10
Quadro 2.2 Vantagens e Desvantagens das Tecnologias de equipamentos de detecção de veículos .....	16
Quadro 3.1 Principais produtos analisados.....	44
Quadro 3.2 Comparativo Políticas e Necessidades de Análise x Classes de Modelo.....	57
Quadro 4.1 Mapa de memória da <i>Tag</i> do SINIAV (CONTRAN, 2006).....	65
Quadro 4.2 Estrutura dos dados do CT-e (parcial) Fonte (CT-e) .....	68
Quadro 4.3 Modelo de Dados Organizados.....	72
Quadro 5.1 Atores e responsabilidades .....	110
Quadro 5.2 Principais cargas transportadas (fonte PDTU).....	124
Quadro 5.3 Erro relativo por MPR do modelo ODAVL .....	134
Quadro 5.4 Comparação dos modelos ODAVL e Modelo Proposto.....	135

## LISTA DE SIGLAS

ANTT	Agencia Nacional de Transportes Terrestres
API	Application Programming Interface
CENTRAN	Centro de Excelência em Engenharia de Transportes
CNT	Confederação Nacional dos Transportes
CONTRAN	Conselho Nacional de Trânsito
DANFE	Documento Auxiliar da Nota Fiscal Eletrônica
DENATRAN	Departamento Nacional de Trânsito
DETRAN/RJ	Departamento de Transito do Rio de Janeiro
DNIT	Departamento Nacional de Infraestrutura do Transporte
ENCAT	Encontro Nacional de Coordenadores e Administradores Tributários
FAF	Freight Analisys Framework
GIS	Geographic Information Systems
GPL	General Public License
HQL	Hibernate Query Language
IDE	Integrated Development Environment
ITS	Intelligent Transportation Systems
J2EE	Java 2 Platform, Enterprise Edition
JVM	Java Virtual Machine
KDD	knowledge-discovery in databases
MER	Modelo de Entidade e Relacionamento
MXML	Macromedia XML
NHCRP	National Highway Cooperative Research Program
PDF	Portable Document Format
PDTU	Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro
PNLT	Plano Nacional de Logística e Transportes
RFID	Radio Frequency Identification
RIA	Rich Internet Application
RMRJ	Região Metropolitana do Rio de Janeiro
RUP	Rational Unified Process
SCM	supply chain management
SDK	Software Developement Kit
SECEX	Secretaria de Comércio Exterior
SECTRANS	Secretaria de Estado de Transportes
SERPRO	Serviço Federal de Processamento de Dados
SINIAV	Sistema Nacional de Identificação Automática de Veículos
SQL	Structured Query Language
UML	Unified Modeling Language
UTMS	Urban Transportation Model System
Weka	Waikato Environment for Knowledge Analysis
XML	Extensible Markup Language

## Capítulo 1 - Introdução

O acelerado processo de crescimento da economia brasileira observado nas últimas décadas não vem sendo acompanhado pelo aumento proporcional da oferta de serviços de transportes, principalmente no que se refere ao setor rodoviário (CNT, 2012) no qual a infraestrutura existente é incapaz de suprir as necessidades vigentes. Para que se tenha um planejamento efetivo dos Sistemas de Transporte, devem-se identificar os aspectos atuais de mobilidade de pessoas, veículos e mercadorias e realizar projeções para saber o comportamento da demanda futura (ORTÚZAR e WILLUMSEN, 2011). Desta forma, é possível estabelecer premissas para avaliar as alternativas de investimento nos sistemas de transporte ou na malha viária.

Os métodos para levantamento de dados convencionais nem sempre permitem a percepção de todos os detalhes e implicações da situação atual e sua projeção no horizonte de médio e longo prazo. Outro aspecto limitador relaciona-se aos custos dessas pesquisas, que representam valores expressivos e consomem uma importante parcela do orçamento para o Planejamento de Transportes.

Em resposta a essas necessidades, os planejadores de Sistemas de Transporte contam com um elenco de tecnologias integrantes dos Sistemas Inteligentes de Transportes ou *Intelligent Transportation Systems* – ITS (LEUNG *et al.* 2011), possibilitando o uso de sistemas baseados em tecnologia de identificação por radiofrequência (RFID – *Radio Frequency Identification*) para aquisição e tratamento de dados de rastreamento de cargas e veículos.

Conceitua-se ITS como a aplicação de técnicas de informação e comunicação para a solução dos problemas da Engenharia de Transportes (IRF, 2012) e RFID como um termo genérico para descrever sistemas de identificação gravados em um componente eletrônico, passíveis de serem “lidos” por equipamentos, que atualmente são utilizados para identificação automática de veículos, de mercadorias, animais ou mesmo uma pessoa.

Do tratamento do binômio – identificação automática de veículos e rastreamento de cargas – certamente emergirão novas propriedades aplicáveis à área de transporte, o que possibilitará aprimorar o Planejamento do Transporte.

## 1.1 Objetivo

O objetivo desta tese é estabelecer um método automatizado e fundamentado nos princípios dos Sistemas Inteligentes de Transporte ITS (IRF,2012) para identificar os padrões de movimentação de cargas e de veículos a partir de dados coletados por sistemas nacionais de identificação automática de veículos e de transporte de cargas.

O método se propõe a reduzir custos na aquisição de dados da movimentação de veículos e de carga, aumentar a frequência da coleta, ampliar o elenco de dados coletados e tratá-los de forma automatizada, possibilitando que sejam aprimorados os processos para o gerenciamento do tráfego, planejamento da infraestrutura e gestão da logística, a partir da identificação permanente e contínua das principais origens e destinos, rotas utilizadas, tipos de veículos, bem como das cargas transportadas.

Para atingir os objetivos, pretende-se consolidar os dados dos seguintes sistemas: Sistema Nacional de Identificação de Veículos – SINIAV, Brasil-ID para identificação automática de mercadorias e para rastreamento de cargas, ambos com tecnologia de identificação por radiofrequência (RFID), e o sistema Nota Fiscal Eletrônica para os serviços de transporte CT-e (conhecimento do transporte eletrônico).

## 1.2 Características do Método

O método é composto por três modelos, sendo dois computacionais e um conceitual, voltados ao planejamento e à gestão do transporte e com as seguintes características básicas:

- Alto grau de expressividade e formalidade quanto à representação de conceitos e relações envolvidos em cenários de movimentação de carga rodoviária;
- Diversidade de informações relativas aos veículos e carga em movimento, e a infraestrutura onde se inserem, com o intuito de atender aos vários domínios da aplicação, entre eles o planejamento do transporte de carga, gestão do transporte urbano e gestão logística;
- Conformidade com padrões de representação de informação para facilitar o intercâmbio, o reuso e o compartilhamento de informações entre aplicações de tráfego rodoviário.

O primeiro modelo objetiva identificar os padrões frequentes de movimentação de carga em rodovias e com o desenvolvimento de aplicativo computacional que

consolida dados dos sistemas acima relacionados e disponibiliza relatórios estatísticos. A notação utilizada para modelar os artefatos da fase de análise e projeto será a UML (*Unified Modeling Language*). A aplicação seguirá o paradigma RIA (*Rich Internet Application*), que visa a juntar a usabilidade das aplicações Desktop com a flexibilidade de uma aplicação Web. As linguagens previstas para o desenvolvimento são, entre outras, o Adobe Flex para Apresentação, o Java para o Controle e o acesso aos dados pela linguagem HQL. A aplicação deverá usar um misto das arquiteturas Cliente-Servidor e Três Camadas

O segundo modelo será desenvolvido usando-se *softwares* livres e tem por objetivo identificar e prever padrões de movimentação de veículos em região urbana. Nesse modelo, serão usadas técnicas de mineração de dados para descoberta de conhecimento em bases de dados. Serão selecionados e modelados os *softwares* para simulação de tráfego, captura da passagem dos veículos pelas antenas e descoberta de conhecimento em bases de dados.

O terceiro modelo será desenvolvido em sua forma conceitual, usando os paradigmas da Ontologia e com maior aplicação na gestão Logística, objetiva representar o conhecimento para o planejamento do transporte de carga, segundo aspectos relativos à “computação sensível a contexto”. Neste modelo, será utilizada a abordagem proposta por Noy e McGuinness (2001) quanto à representação do modelo.

### 1.3 Contribuições

Como contribuições, as questões de projeto que possibilitem a identificação dinâmica de padrões de movimentação de cargas e veículos pela malha rodoviária com o desenvolvimento de um modelo computacional baseado no reuso de dados dos sistemas nacionais para monitoração de veículos.

Os aspectos inovadores desta pesquisa em relação aos estudos anteriores são:

- Possibilitar a identificação de padrões frequentes de Origem-Destino-Rota-Carga-Modo para cargas e veículos em circulação pela malha rodoviária de forma dinâmica;
- Reutilizar dados obtidos por sistemas nacionais, reduzindo o custo do levantamento de dados primários;

- Possibilitar a percepção de tendências e mudanças devido à criação de uma base de dados permanente;
- Identificar propriedades emergentes da movimentação de veículos.

## 1.4 Premissas

A primeira premissa considera que a modelagem do transporte de carga é influenciada pela complexidade e pelo custo das pesquisas de campo, limitada pela quantidade e pela qualidade dos dados.

A segunda assume que o levantamento de dados é necessário à construção das matrizes de transporte de carga que não contemplam, integralmente, aspectos de mudanças de cenário da economia nacional ou internacional.

A terceira sustenta a possibilidade de estruturação de um método para geração de padrões de movimentação de cargas e veículos em diversas formas de agregação com o uso de dados obtidos de sistemas nacionais para controle tributário e fiscal.

## 1.5 Relevância

Na literatura, emergem estudos que utilizam dados extraídos por sistemas RFID com diversas finalidades em planejamento de transportes, gerenciamento da infraestrutura, estudos de congestionamentos, logística, segurança, entre outros. Uma das áreas de estudos, aplicados ao planejamento de transporte usualmente trata da identificação e geração de padrões de comportamentos para movimentação de veículos e pessoas. A identificação dos padrões frequentes da movimentação de cargas pelas rodovias tem menor destaque para aplicações governamentais, focando seu desenvolvimento sob a ótica de aplicações logísticas.

Os estudos de movimentação de carga em território brasileiro são realizados periodicamente mediante pesquisas nas rodovias e ficam limitados devido ao elevado custo. Retratam aspectos pontuais do transporte de bens e mercadorias, conforme indica estudo elaborado pelo *Transportation Research Board* (TRB, 2013), que ressalta a importância de se desenvolverem ferramentas computacionais para coletar e tratar de dados de transporte continuamente e a necessidade de existir interoperabilidade de sistemas de informações em nível nacional.

## 1.6 Estrutura do trabalho

Esta tese está estruturada em seis capítulos.

No Capítulo 1, identificou-se o problema sobre o planejamento do transporte de carga, definiram-se o objetivo e as premissas e justificou-se o estudo desenvolvido com delimitação de sua estrutura. Também foram descritos resumidamente os modelos que compõem o método e as tecnologias empregadas em sua elaboração.

No Capítulo 2, discutem-se os conceitos da Engenharia do Transporte aplicáveis ao planejamento de transportes, caracterizando-se seus principais componentes visando a identificar a Matriz Origem-Destino (OD) e o tratamento das informações integrantes do planejamento dos transportes.

No Capítulo 3, apresenta-se o estado da arte dos Modelos de Planejamento do Transporte no Brasil e no exterior, como tentativa de obtenção de subsídios para a construção dos modelos propostos.

O Capítulo 4 discute e apresenta brevemente os conceitos de Engenharia de *Software* aplicáveis ao estudo em questão. Também se discutem Sistemas Inteligentes de Transporte com suas formas de representação, a descoberta de padrões em bases de dados (*data mining KDD*) e a representação do conhecimento com uso de Ontologia, visando a compor o referencial teórico de sustentação ao entendimento e ao desenvolvimento dos modelos propostos.

O Capítulo 5 apresenta o desenvolvimento do Método, por meio dos três modelos que o compõem, com a aplicação das técnicas apresentadas no Capítulo 4, bem como os aspectos para a validação das diversas abordagens. O Modelo Conceitual para o desenvolvimento dos Sistemas de Identificação Automática Origem-Destino-Modo-Carga encontra-se disponível em

<https://skydrive.live.com/redir?resid=F27A4F06F5CFE1C9!432>

No Capítulo 6, são apresentadas as conclusões e recomendações da pesquisa.

## Capítulo 2 - Planejamento dos Transportes

Esta etapa, que corresponde à revisão bibliográfica, se destina a apresentar a fundamentação teórica para caracterizar os principais componentes do Planejamento dos Transportes utilizados nesta tese.

### 2.1 Modelagem do tráfego e transporte

Segundo a classificação de Hensher e Button (2001), três principais categorias de modelos são utilizadas em planejamento de transportes:

- (a) Modelos convencionais empíricos – são os mais utilizados em Transportes, conhecidos como “Modelo de Quatro Etapas” ou “Modelo Sequencial”;
- (b) Modelos comportamentais – utilizam fatores de motivação dos usuários;
- (c) Modelos atitudinais – procuram captar as reações dos usuários não compreendidas pelos modelos de cunho quantitativo.

Alguns aspectos relativos à reengenharia dos processos de planejamento de transporte com o uso massivo, tanto de dados relativos ao uso de veículos quanto da infraestrutura viária com dispositivos computacionais integrados, foram analisados por Haode *et al.* (2008) e caracterizados pela possibilidade de capturar aspectos multidisciplinares que envolvam os dados dos veículos, tráfego, meio ambiente, condições da infraestrutura, entre outros. Os principais modelos de planejamento de transportes são apresentados a seguir.

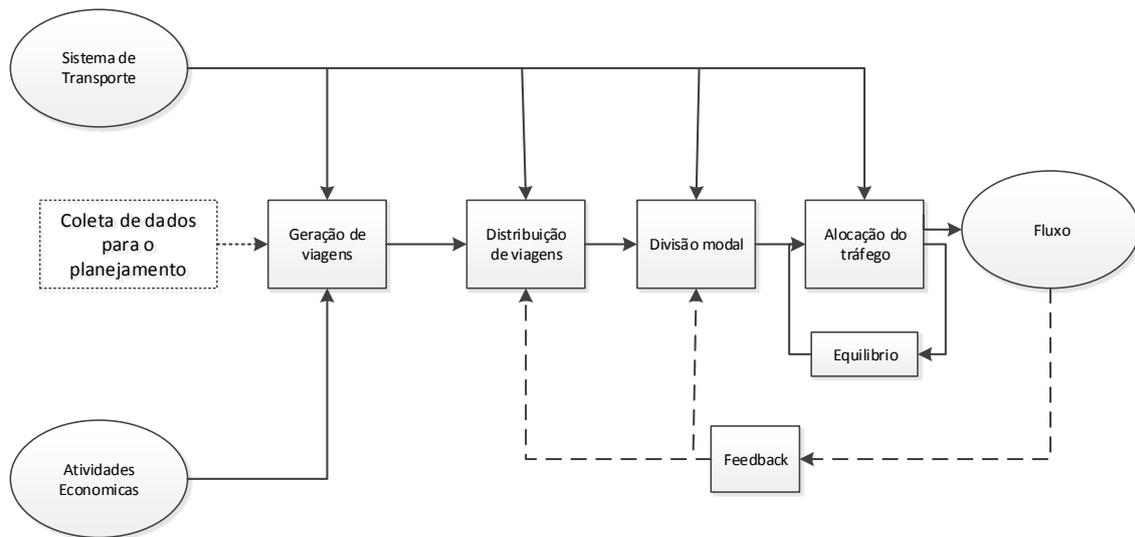
### 2.2 Modelo Convencional ou das Quatro Etapas

Um dos métodos mais utilizados para prever o fluxo de viagens em uma determinada área de estudo é dividir o planejamento do transporte em quatro passos ou etapas. Esse procedimento possui várias designações, todas com o mesmo significado: Modelo Convencional Empírico, Modelo Clássico de Transporte, UTMS (*Urban Transportation Model System*), Modelo de Quatro Etapas ou ainda Modelo Sequencial (HENSHER, 2001). As etapas que compõem o modelo, apresentadas na Figura 2.1, resumidamente são: geração de viagens, distribuição de viagens, divisão modal e alocação do tráfego.

O modelo fundamenta-se na estimativa dos fluxos de viagem com base nas características de comportamento dos usuários do sistema de transportes, sendo este definido a partir de informações do conjunto de indivíduos presentes em locais previamente definidos.

As principais vantagens dessa abordagem são as diferentes formas possíveis de determinar cada passo, incluindo meios que não exijam informações complexas e precisas, o que a torna uma opção muito atrativa. Outra característica interessante é a possibilidade de utilizar as diversas partes de um modelo desenvolvido para uma região em outras com características semelhantes, sem a necessidade de muitos ajustes. Esse tipo de análise tem seus resultados normalmente baseados em hipóteses simplificadoras, o que acaba por gerar resultados menos precisos e requer muitas vezes um processo de calibração ou validação estatística.

A Figura 2.1 mostra as etapas e suas relações.



**Figura 2.1:** Modelo das Quatro Etapas – baseado em MCNALLY (2007)

A seguir, são descritas as etapas desse modelo e suas principais características.

### 2.2.1 Coleta de Dados para o Planejamento

Para o Planejamento dos Sistemas de Transportes são necessárias informações a respeito deste sistema (rede viária, demanda por viagens atraídas e produzidas, pesquisa Origem e Destino, frota), coleta de variáveis sócio-econômicas (população, densidade

populacional, renda, empregos, escolas, dentre outras) e estudos sobre a legislação e o uso e ocupação do solo da região em análise. (Hensher e Button, 2000).

Os dados coletados auxiliam na análise dos Sistemas de Transportes e na formulação de modelos matemáticos que permitem prever o comportamento futuro da demanda por transportes (Hensher e Button, 2000), assim se torna possível programar/planejar o seu desenvolvimento.

As “Pesquisas de Origem e Destino” (OD) são importantes fontes de informação para o Planejamento dos Transportes que podem ser realizadas de várias maneiras, visando sempre coletar dados que possibilitem a planificação dos Sistemas de Transportes a partir determinação dos pontos iniciais e finais dos deslocamentos e a obtenção de informações sobre os motivos das viagens (emprego, lazer, compras), horários das viagens, número de veículos particulares, uso de transporte público, número de moradores, renda média, entre outras para o transporte de carga a pesquisa envolve o tipo de carga, origem, destino, pontos de parada e rota empregada. O objetivo da coleta de dados é definir o padrão de viagens e uso do solo na área de estudo e fazer um diagnóstico sobre o sistema de transporte existente que engloba todos os movimentos : internos ,externos e internos-externos na área de estudo e os horários em que os mesmos acontecem. Para áreas rurais observa-se a sazonalidade das cargas, tipo de transporte utilizado, entre outros.

Os padrões de viagem são representados por uma matriz Origem-Destino (OD). A pesquisa OD, usualmente adotada para identificar a movimentação de pessoas e veículos em área urbana é um sistema de obtenção de informações, normalmente realizada em forma de entrevistas pessoais, em que se questionam as viagens realizadas pelos entrevistados em um dia típico. Contém as origens, os motivos, os horários, os modos de transporte utilizados, o tempo gasto e finalmente os destinos. Podem-se obter informações abrangentes ou não, dependendo dos objetivos e recursos associados à pesquisa. Esse método tem como vantagem o levantamento dos dados das viagens estarem o mais próximo possível da realidade. Podem-se definir apenas os pontos de início e término das viagens ou também pontos intermediários e, em alguns casos, até mesmo a trajetória completa, embora para isso sejam normalmente necessários métodos auxiliares.

Alguns autores retratam que a obtenção de uma matriz OD constitui processo moroso e que despende de significativa quantidade de recursos financeiros. Além do

mais, por questões práticas e econômicas, raramente é possível obter matrizes para longos períodos de tempo, o que constitui uma restrição importante para a tomada de decisão em Engenharia de Tráfego, bem como para aplicação das técnicas de simulação de tráfego (HENSHER, 2001).

Para contornar os aspectos da complexidade e de custo no levantamento dos dados, os engenheiros de transporte utilizam a alternativa para a obtenção de matrizes OD a partir de contagens de fluxo. Em função dessa prática, vários modelos que possibilitam a estimação da matriz de viagens foram desenvolvidos e adotados – denominados sintéticos, por sintetizarem a quantidade de informação (ORTÚZAR e WILLUMSEN, 2011). O modelo sintético pode ser dividido em: Estático e Dinâmico. O primeiro não leva em consideração a variação temporal da demanda. Por sua vez, o segundo é capaz de representar as viagens em função da variação temporal da demanda ao longo do dia. Por ser mais fácil a sua modelagem e permitir a transição para um modelo dinâmico, os processos estáticos são mais utilizados.

Algumas considerações sobre o levantamento de campo:

- Os modelos convencionais de Planejamento de Transportes usualmente utilizados apresentam elevados custos para a estimativa de demanda (WILLUMSEN, 1978a). A maioria desses modelos está relacionada com a coleta e a análise de entrevistas domiciliares e de campo, que nem sempre são viáveis;
- Para o planejamento do transporte urbano, algumas alternativas para realização da pesquisa automatizada da matriz OD de passageiros vêm sendo adotadas, porém, na área de transporte rodoviário de cargas, outros componentes adicionam elementos que dificultam o levantamento automático de dados;
- Para o rastreamento de veículos e mercadorias, merece destaque os avanços nos modelos de planejamento baseados em *Geographic Information Systems – GIS* (ZHANG e WEI, 2009; XIAOLIN, 2004). Neste sentido, a integração de sistemas promovidos pelos Sistemas Inteligentes de Transportes se apresenta como uma alternativa viável para aquisição e tratamento de dados de rastreamento de cargas e veículos (KUTZ, 2011), notadamente em países com dimensões como a do Brasil.

A coleta de dados pode ser realizada de forma manual ou automática.

- Forma manual

Pesquisa domiciliar: Para essa pesquisa, é determinada a amostra em função dos domicílios a serem visitados e o questionário é aplicado de forma controlada, retratando o deslocamento dos moradores.

Pesquisa por envio de questionário: Nessa pesquisa, os usuários de uma via são identificados junto ao Departamento de Trânsito local para obtenção do endereço de seu proprietário e sendo enviado questionário ou mensagem eletrônica sobre a origem e destino do deslocamento.

Entrevista com o motorista: Usualmente aplicado em postos de pedágio, onde um pesquisador consulta o motorista sobre os motivos do deslocamento.

As vantagens e desvantagens de cada método estão resumidas no Quadro 2.1

**Quadro 2.1:** Vantagens e Desvantagens dos Tipos de Pesquisa

continua

Pesquisa Domiciliar	
Vantagens	Desvantagens
Não interfere no tráfego.	Restrição das pessoas em divulgar dados pessoais.
Permite controle estatístico das respostas.	
Possibilita que o entrevistador esclareça dúvidas do entrevistado.	
É o processo mais utilizado no Brasil (por exemplo, Metrô SP e PDTU).	
É o mais confiável dos processos.	
Aplica-se ao deslocamento de pessoas e veículos.	

**Quadro 2.1:** Vantagens e Desvantagens dos Tipos de Pesquisa

conclusão

Vantagens	Desvantagens
<b>Envio de Questionário</b>	
Esse método é mais seguro porque o tráfego não está parado, em oposição com os outros métodos.	Não há contatos pessoais no levantamento de dados, nem há oportunidade para esclarecer dúvidas ou explicar aspectos da pesquisa.
O tamanho da equipe é menor que o dos outros métodos. O trabalho é realizado em escritório.	É fundamental que os questionários sejam enviados dentro de um curto período de tempo após as placas serem observadas (1-2 dias é normalmente o máximo).
O questionário ou <i>e-mail</i> pode ser mais extenso do que as entrevistas em termos de número de perguntas (especialmente sobre aspectos socioeconômicos e familiares).	O método é essencialmente um levantamento por correspondência ou <i>e-mail</i> , por isso é provável que o índice de ausência de resposta seja relativamente alto e há pouca oportunidade para conduzir <i>follow-up</i> .
Pode ser aplicado em locais mesmo com alto volume de tráfego. Os impactos no trânsito são menores do que para os métodos de entrevista.	Pouco adaptado à cultura brasileira.
<b>Entrevista com o motorista</b>	
Esse método é geralmente menos oneroso do que outros métodos.	A abordagem demanda a consulta prévia às autoridades de trânsito locais.
Pode-se planejar a distribuição dos formulários para determinados grupos de veículos.	Atrasos de trânsito ocorrem especialmente em instalações de alto volume e durante períodos de pico de tráfego. A entrevista deve ser conduzida rapidamente.
A taxa de resposta é maior que a de outros métodos, de modo que o potencial para o sucesso da pesquisa é grande.	É necessária a autorização das autoridades locais para a realização da pesquisa.
Pode-se selecionar tipos de amostra em cada local para satisfazer os padrões de análise estatística.	O método é o menos seguro dos métodos
Os dados estão disponíveis muito mais cedo do que para os outros métodos, que dependem de pesquisas com devolução pelos correios.	Devido ao potencial para os atrasos, a entrevista deve ser extremamente curta.

Outro grupo considera o levantamento automático de dados para obtenção da contagem volumétrica, que é integralmente realizada por meio de sensores distribuídos em local da rodovia ou do corredor de tráfego. Essa contagem volumétrica aplica-se principalmente ao método para determinação da Matriz OD, descrito por Cascetta (1984), em que o número de viagens entre pares OD pode ser estimado a partir de contagens volumétricas obtidas em diversos locais da rede viária. A vantagem desse método é que os volumes de tráfego podem ser obtidos de maneira relativamente simples e com menor custo a partir de sensores permanentes, no caso de redes monitoradas por uma central de controle de tráfego, ou por um dos inúmeros processos eletrônicos disponíveis ou mesmo contagens manuais, no caso de redes não monitoradas. Guozhen *et al.* (2011) cita que, dentro desse grupo, várias novas tecnologias vêm se destacando a partir desenvolvimento de Sistemas de Informação Georreferenciados – GIS, que permitiram novas possibilidades para estimativa da matriz OD.

- Forma automática

Segundo Queiroga e Bullock (1998), os estudos relativos ao levantamento automático de dados de viagens podem ser classificados em dois grupos: *Roadside Techniques* e *Vehicle Techniques*. As *Roadside Techniques* são baseadas no uso de detectores fisicamente localizados ao longo de rotas de interesse, em intervalos pré-definidos, que obtêm os dados de tempos de viagem de veículos que atravessam a rota, registrando o tempo em que passam pelos pontos de verificação. Os sistemas *Automatic Vehicle Identification* – AVI (em português, Identificação Automática de Veículos – IAV), baseados em identificação de placas, são exemplos de *Roadside Techniques*.

Os sistemas designados *Vehicle Techniques* são baseados no uso de detectores embarcados nos veículos. Os sistemas AVI, que preveem o uso de *transponders* ou *tags* instalados nos veículos com leitores localizados ao longo da via e uma rede de comunicação para transmitir as informações a uma central de operação, estão incluídos nessa categoria. O *Probe Car* é um veículo sonda com equipamento de posicionamento embarcado que se movimenta ao longo das vias, coletando dados de interesse (posição e tempo principalmente), funciona como um detector móvel, técnica bastante comum do grupo das *Vehicle Techniques*.

- Tecnologia na Infraestrutura (*Roadside Techniques*)

Resumidamente, as tecnologias para detecção de veículos dentro do grupo *Roadside Techniques* que possibilitam a identificação do volume do tráfego são (FHWA, 2007; DNIT,2006; CONSTANTINOS *et al.*, 2011):

- Tubos pneumáticos;
- Laços Indutivos;
- Magnetometro (*Two-axis fluxgate*);
- Piezoelétricos;
- Radar de Microondas;
- Infravermelhos Ativos (*radar laser*);
- Infravermelho Passivos;
- Detectores Ultrassônicos;
- Detectores Acústicos;
- Processadores de Imagens de Vídeo.

A seguir, apresentam-se as características das principais tecnologias de detecção de veículos:

a) Tubos pneumáticos

Os sensores de tubo pneumático enviam um aumento de pressão de ar ao longo de um tubo de borracha quando os pneus de um veículo passam por cima do tubo. O pulso de pressão fecha um interruptor do ar, produzindo um sinal elétrico que é transmitido para um contador ou *software* de análise. O sensor de tubo pneumático estrada é portátil, usando chumbo-ácido, gel, ou outras baterias recarregáveis como fonte de energia.

b) Sensores de laço indutivo (*loop*)

O equipamento fica enterrado na via e é composto de uma espiral com fio de cobre, usualmente 3 ou mais espiras, no qual é injetada sinais elétricos cujas frequências variam de 10 KHz a 50 KHz e superior. As funções de *loop* funcionam como um elemento indutivo em conjunto com a unidade de controle eletrônico. Quando um veículo passa sobre o *loop*, a indutância é diminuída. A indutância diminuiu ou aumenta

a frequência de oscilação e faz com que a unidade eletrônica possa enviar um pulso ao controlador, indicando a presença ou passagem de um veículo.

Os dados fornecidos pelos detectores convencionais indutivos são passagem de veículos, presença, contagem e ocupação. Apesar de *loops* não poderem medir diretamente a velocidade, essa grandeza pode ser determinada usando um artifício para a medição de duas velocidades de ciclo ou de um detector de *loop* simples e um algoritmo cujos insumos são comprimento do laço, duração média dos veículos ao longo do tempo sobre o detector e número de veículos contados. A classificação dos veículos é suportada por versões mais recentes do laço indutivo, contendo unidades de eletrônicos que excitam o laço do fio nas frequências mais altas que identificam partes específicas de metal sob o veículo. O laço indutivo é o tipo mais utilizado na contagem volumétrica utilizada por inúmeros municípios e pelo Departamento Nacional de Infraestrutura do Transporte – DNIT (n/a), aplicado para o Plano Nacional de Contagem de Trânsito.

O sistema, no modelo adotado pelo DNIT, Figura 2.4, consiste de 4 (quatro) subconjuntos: abrigo, linha de transmissão, laço e detectores. A Figura 2.4 mostra um laço instalado em uma rodovia.



**Figura 2.4** : Laço (*loop*) magnético – foto obtida de DNIT(n/a)

- Magnetômetro (*Two-axis fluxgate*)

Os sensores magnéticos são dispositivos passivos que indicam a presença de um objeto metálico, detectando a perturbação (conhecida como uma anomalia magnética) no campo magnético da terra criado pelo objeto.

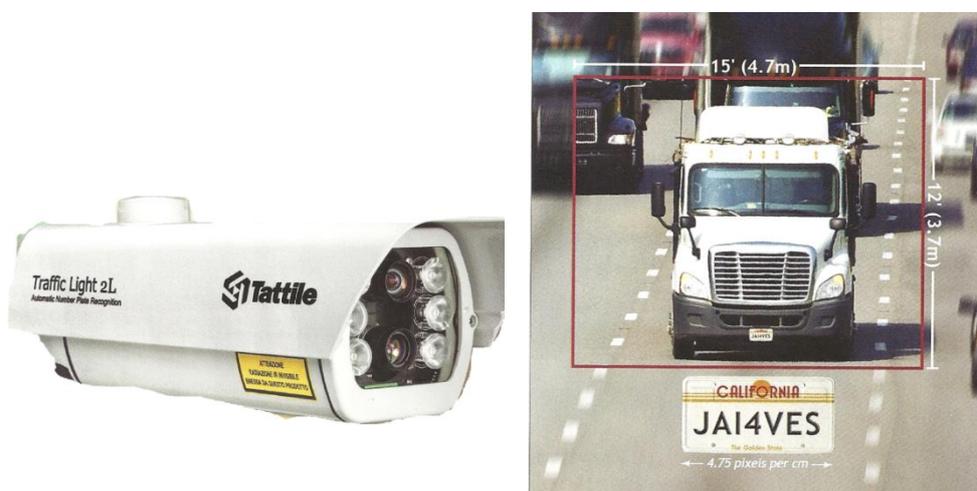
- Sensores Piezoelétricos

Materiais piezoelétricos podem gerar uma tensão quando submetidos ao impacto mecânico ou vibração. A tensão medida é proporcional à força ou ao peso do veículo. A magnitude do efeito piezoelétrico depende da direção da força em relação aos eixos do cristal. Sensores piezoelétricos são utilizados para classificar os veículos por contagem de eixo e espaçamento de eixo e para medir o peso do veículo e velocidade (o último, quando vários sensores são implantados). Eles são frequentemente usados como parte de pesagem em movimento-sistemas.

Sob a ótica das *Roadside Technologies*, outro grupo que possibilita, além da contagem de tráfego, monitorar o deslocamento dos veículos, fornecendo dados para os sistemas ATMS – *advanced traffic management system*, ATIS – *advanced traveler information system* e APTS – *advanced public transportation system* são as que possibilitam a identificação automática de veículos. Essas tecnologias são consideradas como sensores ponto a ponto (CONSTANTINOS *et al.*, 2011). O principal exemplo dessa tecnologia é identificação por reconhecimento digital de imagem.

- Identificação por OCR (*optical character recognition*)

Com essa tecnologia, uma câmera captura a imagem da placa do veículo e, por meio de *software* específico, identifica seus caracteres. Segundo os principais fornecedores dessa tecnologia, o sistema – câmeras e *softwares* – é capaz de ler sob qualquer condição de clima e determinar as dimensões do veículo, classificando-os por tipo (carro, caminhão entre outros) (Figura 2.5).



**Figura 2.5:** Câmera de captura OCR (ITSINTERNATIONAL, 2013)

Essa tecnologia vem se popularizando nos últimos tempos devido a sua redução de custo e aumento de confiabilidade.

O Quadro 2.2 apresenta um resumo das principais vantagens e desvantagens de cada tecnologia.

**Quadro 2.2** Vantagens e Desvantagens das Tecnologias de equipamentos de detecção de veículos  
(continua)

Vantagens	Desvantagens
<b>Tubos pneumáticos</b>	
Fácil Instalação. Baixo custo. Tecnologia bastante conhecida.	Sofre com desgaste do tráfego. Sujeita a atos de vandalismo. Pouca precisão em veículos de vários eixos Suscetível à temperatura ambiente.
<b>Laço ou <i>loop</i> indutivo</b>	
<i>Design</i> flexível para satisfazer a grande variedade de aplicações. Tecnologia madura bem compreendida Grande base de experiência. Fornece parâmetros de tráfego básicos (ex.: volume, presença, ocupação, velocidade etc). Insensível a intempéries como chuva e neblina. Oferecem melhor precisão para dados de contagem em comparação com outras técnicas comumente usadas. Padrão comum para a obtenção de medições precisas de ocupação. Os modelos atuais usando tecnologia de alta frequência possibilitam fornecer dados de classificação do veículo. Padrão utilizado pelo DNIT e diversas Prefeituras.	A instalação requer cortar pavimento. Diminui a vida útil do pavimento. Instalação e manutenção exigem fechamento de pista. Equipamentos na via sujeitos ao <i>stress</i> do tráfego e da temperatura. Múltiplos detectores geralmente tornam necessários equipamentos complementares para monitorar um local. A precisão pode diminuir quando o desenho exigir a detecção de uma grande variedade de classes de veículos. A manutenção do pavimento pode danificar o <i>loop</i> .

**Quadro 2.2** Vantagens e Desvantagens das Tecnologias de equipamentos de detecção de veículos

(continuação)

Magnetômetro ( <i>Two-axis fluxgate magnetometer</i> )	
Menos suscetíveis do que os <i>loops</i> indutivos aos estresses do trânsito.	A instalação requer cortar pavimento.
Insensível a intempéries como chuva e neblina.	Diminui a vida do pavimento.
Alguns modelos podem transmitir dados através de ligação sem fio via Radiofrequencia – RF.	Instalação e manutenção exigem fechamento de pista.
	Modelos com zonas de detecção pequenas requerem múltiplas unidades para detecção de todas as pistas da via.
Magnéticos (indução ou <i>search coil magnetometer</i> )	
Podem ser usados em locais onde outras tecnologias não são viáveis (por exemplo, tabuleiro de ponte).	A instalação requer cortar o pavimento.
Alguns modelos são instalados em estrada sem necessidade de cortes de pavimento.	Não detecta veículos parados, exceto se complementada por outros equipamentos.
Insensível a intempéries como chuva e neblina.	
Menos suscetíveis a estresses de trânsito.	
Micro-ondas ( <i>microwave radar</i> )	
Normalmente insensível aos aspectos climáticos em curtas distâncias, características das aplicações de gerenciamento de tráfego.	Esta tecnologia não identifica veículos parados.
Medição direta da velocidade	
Viável à medição de múltiplas pistas de rolamento.	
Infravermelho Ativo ( <i>radar laser</i> )	
Medição apurada da velocidade, posição e tipo do veículo.	A operação poderá ser afetada quando a visibilidade é inferior a 6 metros.
Viável à medição de múltiplas pistas de rolamento.	Instalação e manutenção demandam o fechamento da via.
Infravermelho Passivo	
Medição de velocidade multizonas.	A operação poderá ser afetada com baixa visibilidade
	Alguns modelos não possuem recursos para detecção da presença de veículos

**Quadro 2.2:** Vantagens e Desvantagens das Tecnologias de equipamentos de detecção de veículos  
(conclusão)

Processadores Ultrassônicos	
Viável medir em múltiplas pistas. Capaz de identificar sobre carga em veículos. Largamente utilizado no Japão.	As condições ambientais podem influir na precisão das medidas. Menor precisão para veículos trafegando em baixa velocidade.
Processadores Acústicos	
Detecção passiva. Viável medir em múltiplas pistas.	As condições ambientais podem influir na precisão das medidas. Alguns modelos não são recomendáveis para tráfego de baixa velocidade ou em regime de para-anda ( <i>stop and go</i> ).
Processadores de Imagens de Vídeo	
Permite a identificação da placa veicular com uso de OCR. Detecção de múltiplas pistas ou zonas. Capaz de medir diversas grandezas do tráfego. Aumenta sua capacidade com a integração de diversas câmeras Câmeras de vigilância podem ser usadas para a detecção de incidentes na via.	A instalação e a manutenção da câmera exigem o fechamento da via. As condições ambientais podem influir na precisão das medidas. Para a detecção noturna, são necessários equipamentos especiais.

Outro grupo de tecnologias não se enquadra perfeitamente dentro das tecnologias *Roadside Techniques*, são derivadas de observação aérea realizada por aviões, helicópteros ou, mais recentemente, veículos não tripulados.

- Observação aérea

O custo e as limitações de clima para a decolagem das aeronaves fazem seu uso limitado. Uma alternativa de menor custo de aquisição dos equipamentos e custo operacional são os veículos não tripulados.



**Figura 2.6:** Veículo Aéreo Não Tripulado (ITSINTERNATIONAL, 2013)

Segundo especialistas internacionais (ITSINTERNATIONAL, 2013), a liberação em 2012 pelo FAA (*Federal Aviation Administration*) para o uso civil de veículos não tripulados VANT – Veículo Aéreo não Tripulado (Figura 2.6), conhecidos por sua forma em Inglês, UAV (*unmanned aerial vehicles*) – a partir do final de 2015, deverá proporcionar uma nova dimensão do controle do tráfego. Chris Tindall (ITSINTERNATIONAL, 2013) cita que: ‘Uma nova fronteira no controle do tráfego rodoviário se apresenta com o uso de veículos aéreos não tripulados’ (tradução livre)

- Tecnologia Embarcada (*Vehicle Technology*)

As tecnologias integrantes do grupo *Vehicle Technology* caracterizam-se por necessitar de equipamento instalado no veículo e que possa ser identificado por sistema instalado da infraestrutura da via.

Os principais exemplos dessa tecnologia são:

- Identificação por RFID (*radio frequency identification*)

RFID é um termo genérico usado para descrever sistemas que transmitem a identidade (normalmente um número serial único) de um objeto ou pessoa. Existem três elementos básicos neste tipo de sistema:

- *Tag* (também chamada de etiqueta ou *transponder*), é composta de um *chip* semicondutor, uma antena e por vezes uma bateria;
- Leitor (também chamado *interrogator*), é composto de uma antena, um módulo eletrônico de radiofrequência RF e um módulo controlador eletrônico;

- Controlador (também chamado de *host* ou *controller*), normalmente é um computador pessoal ou estação de trabalho, executando um *software* chamado de *middleware*.

A função básica de uma *tag* RFID é armazenar dados e transmitir dados ao leitor.

De forma resumida, pode-se dizer que uma *tag* consiste de um *chip* eletrônico e uma antena encapsulada em um pacote que servirá de proteção e molde. Geralmente o *chip* contém uma memória onde informações podem ser armazenadas, lidas e, em alguns casos, reescritas. Algumas *tags* podem conter pequenas baterias, o que permite classificá-las em dois grupos: *tags* ativas e *tags* passivas. Existe um terceiro grupo designado por semi-ativas.

- Tags Ativas

*Tags* RFID são ditas ativas se contêm em sua construção uma fonte de energia como uma bateria. Quando a *tag* precisa transmitir dados para o leitor, usa sua própria fonte de energia para transmitir. Essa é a razão pela qual, em geral, as *tags* ativas podem se comunicar com leitoras menos potentes e podem transmitir informações a distâncias maiores. Além disso, esse tipo de *tag* tem, tipicamente, memórias maiores, superiores a 128 Kb e componentes maiores, mais complexos, tornando-o um tipo mais caro se comparado com o tipo passivo.

Os primeiros sistemas RFID eram ativos e ainda hoje há muitas aplicações que requerem *tags* ativas.

As principais características dos sistemas de RFID ativos são:

- Grandes distâncias de leitura (até 300 metros);
- Possuem dimensões elevadas, se comparadas com as de uma *tag* passiva (devido à presença da bateria);
- Melhor utilização dos *tags* pelos sensores;
- Alta velocidade de resposta (objetos podem se mover a até 220 Km/h);
- Tempo de vida útil determinado (entre um e dez anos);
- Custo mais elevado que o de *tags* passivas.

○ *Tags* Passivas

*Tags* RFID passivas não têm fonte de energia interna e obtêm força para transmitir seus dados da indução gerada pelo sinal emitido da leitora, também têm memórias internas menores, da ordem de alguns *bytes*. Como resultado, elas tendem a ser menores e a ter um custo de produção também menor que o das *tags* ativas. A distância efetiva de alcance para transmissão é menor (por vezes, menos que um metro), demandando assim equipamentos leitores mais poderosos e também mais caros.

As principais características das *tags* passivos são:

- Distância de leitura de até 10 metros;
- Possuem dimensões reduzidas, podendo ter a espessura de uma folha de papel;
- Vida útil teoricamente infinita;
- Mais suscetíveis a interferências eletromagnéticas;
- Baixo custo.

Existe ainda um tipo semi-ativo que conjuga as características de *tags* passivas e ativas e que possui baterias ligadas a ela, mas essas baterias permanecem em *stand by* aguardando um pulso para alimentar partes de sua estrutura eletrônica.

A *tag* e o leitor se comunicam por meio de ondas de rádio. Quando uma *tag* entra na zona de leitura de um dispositivo leitor, é ativada pela exposição ao campo eletromagnético e recebe um sinal para transmitir seus dados armazenados. Esses dados podem ser de natureza diversa, mas podem simplificar, transmitindo, conforme um determinado protocolo, um conjunto de *bytes* cujo significado apenas compete à aplicação, podendo ser, números de série, instruções de configuração, números identificadores de veículos, matrículas de funcionários, entre outros. Uma vez que os sinais da *tag* tenham sido recebidos pelo leitor, ela os decodifica e os repassa para a aplicação controladora via canal e protocolo conhecidos (por exemplo: RS232, RS485 ou TCP/IP).

Um sistema de controle desse tipo pode ter diversos leitores espalhados e, ao mesmo tempo, ter um único controlador, ligando toda essa rede de leitores. Da mesma forma, um único leitor pode se comunicar com diversas etiquetas eletrônicas ao mesmo tempo.

- *Bluetooth* e telefones celulares

Recentemente, pesquisadores passaram a utilizar os sinais emitidos pelos veículos equipados com *bluetooth*, bem como sinais de radio *wi-fi* emitidos por telefones celulares do motorista ou de passageiros visando à identificação da movimentação desses veículos.

A vantagem do uso dessas tecnologias é que ocorre puramente um ‘*ping*’ – resposta do equipamento *bluetooth* ou celular a uma interrogação de outro equipamento, informando o seu *MAC address* – sem necessidade de ‘colaboração’ do motorista. O endereço *MAC (Media Access Control)* é um endereço físico associado à interface de comunicação que conecta um dispositivo à rede e possui um endereço “único”, não havendo duas portas com a mesma numeração. Um aspecto relevante é que o equipamento que interroga possui alcance limitado, estabelecendo limites para a localização dos veículos.

Considerando que essa comunicação veículo-infraestrutura não ocorre para todos os veículos, o uso dessa tecnologia torna-se mais apropriado para identificação da origem destino, ou rotas adotadas, mas não para controle de tráfego.

Um exemplo recente é relatado por Zhengyu Duan *et al.* (2011), que apresentaram uma pesquisa realizada na cidade de Shanghai, com 10 milhões de usuários de telefones celulares para identificação da movimentação de pessoas. Segundo os autores, possibilitou, entre outros, capturar os padrões de movimentação associados às variações sazonais, à identificação do deslocamento de trabalho e não-trabalho e aos modais utilizados.

- Equipamentos de posicionamento global

Uma tecnologia que vem sendo utilizada notadamente na área logística é baseada em dados fornecidos por sistemas de posicionamento global.

Na área de controle de tráfego, um exemplo do uso dessa tecnologia está associado à monitoração de uma frota de veículos ou a um ‘*probe car*’ – veículo que se desloca pela via, fornecendo informações sobre o tráfego por meio de sistemas *GSM Global System for Mobile Communications*, ou Sistema Global para Comunicações Móveis.

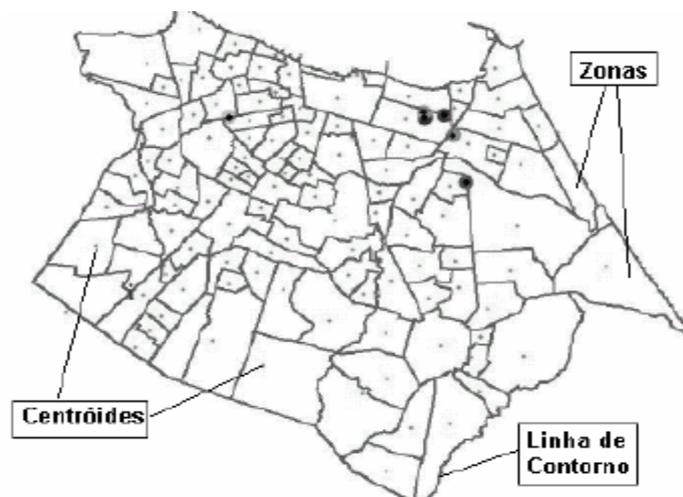
## 2.2.2 Primeira Etapa: Geração de Viagens

O primeiro passo para a modelagem considera os motivos que levam os indivíduos a realizar viagens e é dividido em duas partes: produção de viagens e atração de viagens. Em geral, são associadas às origens e aos destinos, respectivamente. Ortúzar e Willunsen (2011) conceituam essa etapa como sendo “o número total de viagens geradas de uma origem e atraídas para um destino de cada zona da área de estudo”, para o qual é enfatizada a importância da homogeneidade dessas zonas, e complementa que o número de viagens ocorrerá em função das características das zonas de tráfego.

Na modelagem do sistema é considerada a representação discreta do espaço contínuo e o sistema de transporte é representado por um grafo com arcos e nós (trechos de via que conectam pontos do sistema viário, representados pelos nós) (Figura 2.2).

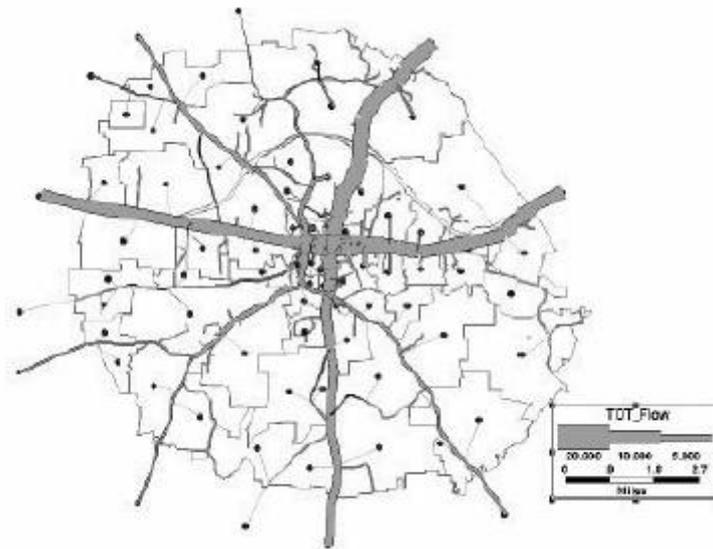
Os Planos de Transportes para zonas urbanas devem definir uma área na qual serão coletadas as informações necessárias à elaboração dos projetos. A fronteira que separa essa área das demais é denominada “Cordão Externo” ou “Linha de Contorno”. A área interna, delimitada por essa linha, deverá ser subdividida em áreas menores, chamadas de “Zonas de Tráfego”. Cada zona possui um centro de gravidade denominado “centroide”, no qual está concentrada a maior parte das atividades realizadas na zona.

Os “centroides” das zonas de tráfego são ligados à rede por conectores que representam os pontos de acesso ou egresso ao sistema de transporte para as viagens com origem ou destino naquela zona (Figura 2.2).



**Figura 2.2:** Componentes da área de estudo

De posse dos dados de origem e destino das viagens, será possível traçar diagramas que indiquem a intensidade da demanda. Esses diagramas são denominados “Linhas de Desejo” que, traçados em escala, possibilitam visualizar o volume da demanda de acordo com o sentido de seu deslocamento (Figura 2.3).



**Figura 2.3:** Volume de Tráfego – Linha de Desejo

Zona de Tráfego é a área de estudo em que se deseja avaliar as consequências das mudanças, portanto, vinculada ao contexto do estudo. Nas várias cidades brasileiras que passaram por estudos de planejamento de tráfego foram estabelecidas zonas de tráfego (ZT), quando da montagem de suas matrizes OD, a partir da agregação de setores censitários do IBGE. (ORTÚZAR e WILLUMSEN, 2011)

Polos Geradores de Tráfego (PGT) são definidos como:

“empreendimentos de grande porte que atraem ou produzem grande número de viagens, causando reflexos negativos na circulação viária em seu entorno imediato e, em certos casos, prejudicando a acessibilidade de toda a região, além de agravar as condições de segurança de veículos e pedestres (DENATRAN, 2001)”.

Como exemplos de PGT, constam, entre outros estabelecimentos: shopping centers, supermercados, hotéis, centro de convenções, teatros, escolas, portos e aeroportos.

### 2.2.3 Segunda Etapa: Distribuição de Viagens

O segundo passo consiste em definir cada viagem ou conjunto de viagens, associando suas origens a seus destinos, normalmente atribuindo um número de viagens para cada zona de tráfego. Gera-se uma matriz quadrada de dimensão “n” que representa o número de zonas de tráfego analisadas que contém, em cada célula, o número de viagens entre cada par ordenado de zonas, além das viagens internas de cada zona, ou seja, as viagens com origem e destino na mesma zona. Ou seja, define-se uma matriz conforme o Quadro 2.3

**Quadro 2.3:** Matriz de viagens

I/J	Z <sub>1</sub>	Z <sub>2</sub>	.....	Z <sub>n-1</sub>	Z <sub>n</sub>	Produção
Z <sub>1</sub>	T <sub>11</sub>	T <sub>12</sub>	.....	T <sub>n-1</sub>	T <sub>n</sub>	P <sub>1</sub>
Z <sub>2</sub>	T <sub>21</sub>	T <sub>22</sub>	.....	T <sub>2n-1</sub>	T <sub>2n</sub>	P <sub>2</sub>
.....	.....		.....	.....	.....	.....
Z <sub>n-1</sub>	T <sub>n-1,1</sub>	T <sub>n-1,2</sub>	.....	T <sub>n-1,n-1</sub>	T <sub>n,n</sub>	P <sub>n-1</sub>
Z <sub>n</sub>	T <sub>n</sub>	T <sub>n,2</sub>	.....	T <sub>n.n-1</sub>	T <sub>n,n</sub>	P <sub>n</sub>
Atração	A <sub>1</sub>	A <sub>2</sub>	.....	Nn <sub>-1</sub>	A <sub>n</sub>	

De forma geral, a distribuição de viagens é realizada pela potencialidade de cada zona gerar viagens, na atratividade das zonas de destino e na distância ou custo de transporte entre cada par de zonas de origem e destino.

O modelo mais comumente usado para a projeção da matriz OD é o chamado modelo gravitacional, analogamente à lei de atração de massas (ORTÚZAR e WILLUMSEN, 2011). Esse modelo possui algumas possíveis variações, mas é basicamente constituído de uma equação que define o número de viagens entre cada par de zonas, proporcionalmente ao número de viagens produzidas e atraídas para cada uma delas em relação ao total de viagens. Considera-se também o custo operacional (distância, tempo ou um conjunto de variáveis) entre as zonas, além de alguns fatores de ajuste que variam de acordo com a região analisada.

A distribuição de viagens tem por objetivo estimar o número de viagens ( $T_{ij}$ ) entre as ZT ( $i$  origens e  $j$  destinos), criando uma matriz de Origem Destino (OD) com base nos resultados obtidos na etapa de geração de viagens. Nessa etapa, pode-se utilizar o método de fator de crescimento, fator socioeconômico, custo generalizado ou modelo gravitacional (SOUZA e D'AGOSTO, 2013), sendo comum o uso do modelo gravitacional, vinculado ou não, conforme mostra a equação (1), de modo que a atração e a produção possam ser representadas pela quantidade de carga, pelo número de viagens de veículos ou, ainda, por valor monetário da carga. Como fator de fricção, utiliza-se, predominantemente, a distância entre as zonas de tráfego, podendo-se ainda utilizar o tempo de viagem ou seu custo.

$$T_{i,j} = K_i \cdot L_j \frac{P_i \cdot A_j}{d_{i,j}^a} \quad (1)$$

Onde:

$T_{ij}$ : fluxo de carga entre a ZT de origem  $i$  e de destino  $j$ ;

$K_i$  e  $L_j$ : fatores de ajuste;

$P_i$ : produção de carga a ZT de origem  $i$ ;

$A_j$ : atração de carga para a ZT de destino  $j$ ;

$d_{ij}$ : fator de fricção;

$a$ : constante de calibração.

- Estimação da Matriz OD

O primeiro método consiste em estimar o número de viagens de cada célula da matriz diretamente a partir da contagem volumétrica realizada por meio de entrevistas domiciliares, entrevistas em locais da própria rede viária, monitoramento de veículos, ou alguma combinação dos procedimentos mencionados denominado de sintético. Segundo Rakha *et al.* (2005), ao longo das últimas quatro décadas, esse modelo foi pesquisado com o foco dirigido principalmente na resolução de algumas inconsistências: obtenção de múltiplos resultados, falta de continuidade do fluxo em arcos contíguos e desconhecimento das rotas de viagem. O modelo gravitacional e o modelo de oportunidades intervenientes descritos por Ortúzar e Willunsen (2011) são alguns exemplos de modelos sintéticos empregados para estimar a distribuição de viagens entre pares OD.

O segundo método baseia-se na definição de um modelo ou conjunto de modelos analíticos pelos quais as viagens são estimadas sinteticamente a partir de informações relativas ao padrão de uso do solo e em função de medidas de impedância à realização de viagens entre diferentes pares OD.

Os trabalhos de Willumsen (1978a), Van Zuylen e Willumsen (1982), Bell (1983), Cascetta (1984), Carey e Relvelli (1986) e Nielsen (1993) são os pioneiros quanto aos estudos relativos aos métodos para estimar a matriz OD. Outros pesquisadores, como Nguyen (1977), Gur (1978), Gur *et al.* (1980) e Fisk (1988) apresentaram contribuições importantes quanto à tentativa de definir as rotas de viagem.

De modo geral, esses modelos trabalham com fluxos veiculares, mas nada impede que outros fluxos (por exemplo: de carga, de pedestres, de passageiros do transporte público, entre outros) sejam considerados, desde que respeitadas as características de cada um deles.

A ideia consiste basicamente em uma técnica inversa ao processo de alocação do fluxo de tráfego à rede de transportes. Ou seja, quando uma matriz OD conhecida é alocada à rede viária, são obtidos os fluxos em cada arco constituinte, portanto, a técnica inversa seria obter a matriz OD que corresponda ao conjunto de fluxo observado em cada arco, de acordo com o método de estimação empregado. No entanto, a alocação de matrizes distintas pode retornar iguais conjuntos de fluxo a um mesmo conjunto que poderá gerar diferentes combinações para a matriz OD (WILLUMSEN, 1981). Esse seria o principal foco dos estudos que tiveram início na segunda metade da década de 1970, propor um modelo capaz de estimar a matriz OD mais viável que corresponda às contagens de tráfego. Esses estudos consideram exclusivamente a movimentação dos veículos não efetuando o cotejamento com os dados da carga, da origem e do destino, entre outros.

Outra forma de calcular as matrizes OD envolve o conceito de padrões frequentes (ou, do inglês, *Frequent Pattern*) inicialmente proposto por Agrawal *et al.*, (1993) como um processo de mineração de dados para estabelecer associações entre seus itens. Por meio do estudo do conjunto de dados em sequência, podem-se aferir algumas regras que são chamadas de padrões de sequência. As técnicas e modelos de geração de tempos de viagem e velocidades, a partir de dados geoposicionados, correspondem a aplicações crescentes para finalidades de transporte e logística.

Estudos relativos à geração de padrões frequentes da movimentação em rodovias que tratam do monitoramento, baseado em dados de referência sobre viagens prévias agrupam os dados em:

- Caso 1 – do mesmo veículo, na mesma rodovia a ser monitorada;
- Caso 2 – de vários veículos, na mesma rodovia a ser monitorada;
- Caso 3 – vários veículos, em rodovias de mesma classe rodoviária da rodovia a ser monitorada.

Nessa linha de pesquisa, pode-se citar o trabalho de Bolla e Davoli (2000), que apresentaram um modelo para estimar o tráfego, usando um algoritmo que permite calcular os parâmetros de tráfego, com base na localização do telefone móvel. Os pesquisadores Calabrese *et al.* (2011) desenvolveram um algoritmo para monitoramento de veículos e pedestres em tempo real em área urbana, utilizando dados agregados de telefonia móvel. Cayford e Johnson (2003) analisaram os principais parâmetros a serem considerados para geração de dados de tráfego, ou seja, a precisão, a frequência de medição e o número de localizações necessárias para se obter descrições precisas de tráfego. Várias empresas em todo o mundo, entre outras a ITIS Holdings (Inglaterra), Delcan (Canadá), CellInt (Israel), AirSage e IntelliOne (EUA), começaram a desenvolver aplicações comerciais de telefone celular, baseadas em monitoramento de tráfego (CALABRESE *et al.*, 2011).

#### 2.2.4 Terceira Etapa: Distribuição Modal

A divisão modal pode ser definida como a divisão proporcional das viagens realizadas pelas pessoas, entre os diferentes modos de transportes: público e privado. Ou ainda entre diferentes modalidades de transportes: ônibus, Veículo Leve sobre Trilhos – VLT, metrô, trem de subúrbio, entre outros. Segundo Ortúzar e Willumsen (2011), os fatores que influenciam a escolha do modo de transportes estão divididos em três grupos: características do usuário (renda, estrutura domiciliar, posse de veículos, entre outros), características da viagem (propósito da viagem, hora do dia em que a viagem é realizada, entre outros) e características do Sistema de Transportes (tempo de viagem, custo, condição de conforto, entre outros). Esses modelos são de aplicação restrita e de difícil generalização. A partir disso, são definidas as funções de utilidade para cada modo. Com base em tais características e informações socioeconômicas, são determinadas segundo as escolhas de cada tipo de usuário.

Existem basicamente dois tipos de modelos de divisão modal:

- Determinísticos: estabelecem a proporção de viagens por cada modo, utilizando métodos quantitativos simples;
- Probabilísticos: utilizam a probabilidade de escolha de cada modo para determinar a percentagem de viagens para seus respectivos modos. Os modelos mais utilizados são: Logit Binomial, Logit Multinomial e Logit Hierárquico ou Aninhado.

### 2.2.5 Quarta Etapa: Alocação de Tráfego

No último passo, as viagens definidas nas matrizes OD são então associadas às suas respectivas rotas da rede viária contendo todas as vias alternativas por onde os veículos podem trafegar.

Para cada modo de transporte, considera-se uma rede viária diferente, já que exige uma estrutura viária particular. No último passo, as viagens definidas nas matrizes OD são então associadas às suas respectivas rotas da rede viária contendo todas as vias alternativas por onde os veículos possam trafegar.

A alocação tem por finalidade estabelecer os volumes de tráfego que irão utilizar os diferentes trechos da rede. A última etapa do modelo, a alocação de tráfego ou alocação de viagens, considera aspectos da oferta tais como: as vias ou a rede, o equilíbrio oferta-demanda e o equilíbrio da rede com a oferta de veículos. Primordialmente aloca as viagens de veículos (automóveis individuais e transporte coletivo público) entre os pares origem-destino para o sistema de vias.

Wardrop (1952) identificou dois princípios passíveis para prever rotas tomadas por motoristas entre um par OD: (1) os tempos de viagens de todas as rotas usadas, iguais ou menores do que aqueles que seriam experimentados por um único veículo em uma rota não utilizada, e (2) o tempo médio de viagem de todos os motoristas como sendo mínimo, o que implica que o total das horas por veículo agregadas gastas em viagens seja mínimo.

O primeiro princípio considera apenas a visão do tempo médio total do sistema, equivalente ao custo médio da teoria econômica, entendido como se os motoristas calculassem somente seus próprios tempos de viagens na escolha das rotas. O segundo, por outro lado, é mais abrangente, uma vez que a escolha das rotas pelos motoristas é

determinada a partir do tempo marginal do sistema, equivalente ao custo marginal da teoria econômica, considerando os tempos do sistema como um todo e não individualmente como o anterior.

A classificação dos métodos de alocação de tráfego, segundo Ortúzar e Willumsen (2011), baseia-se na resposta de duas perguntas: (i) se a restrição da capacidade está incluída na alocação e (ii) se os efeitos estocásticos estão incluídos. A partir do cruzamento das respostas dessas questões têm-se os métodos de alocação.

### 2.2.5.1 Abordagens principais

Complementam Nicholas e Lester (2009) que a alocação de tráfego possui as seguintes abordagens principais:

*1. Atribuição Tudo ou Nada:* Nesse método, todo o fluxo de um par OD é alocado ao caminho mais curto que liga a origem e o destino (dois centroides). Trata-se de um modelo não-realista, pois somente um caminho é utilizado entre todo par OD, mesmo que existam outros caminhos com custo ou tempo de viagem iguais ou próximos. Não considera também problemas de capacidade ou o nível de congestionamento nos arcos e o tempo de viagem é um parâmetro constante não variando com o fluxo no arco;

*2. Atribuição Estocástica:* Distribui o fluxo entre os múltiplos caminhos dos pares OD, estabelecendo que a proporção do fluxo nos diferentes caminhos é igual à probabilidade de escolha de cada um deles. Tal lógica é implementada comparando-se o menor tempo de viagem de um caminho com os tempos de viagens dos demais. Esse método também não considera uma análise de volume de tráfego, portanto, não tende à estabilidade;

*3. Atribuição Incremental:* Processo no qual os volumes de tráfego são alocados por passos. A cada passo, uma proporção da demanda total é atribuída baseada no método da atribuição “Tudo ou Nada”. Para cada passo seguinte, os tempos são recalculados, baseando-se nos volumes dos trechos. Quando ocorrem muitos incrementos, pode-se pensar que se trata de uma atribuição de equilíbrio, mas não é, pois se observam inconsistências nos volumes dos trechos e nos tempos de viagens;

*4. Restrição de Capacidade:* Esse método procura se aproximar de uma solução de equilíbrio por meio de interação, efetuando uma alocação “Tudo ou Nada” para carregamento de tráfego e recalculando os tempos de viagens por meio de uma função de congestionamento que reflete a capacidade nos arcos. Esse método não converge

para uma solução de equilíbrio e tem um problema adicional, pois os resultados dependem muito do número específico de iterações efetuadas. O número de interações pode influenciar os resultados substancialmente;

5. *Equilíbrio de Usuários*: É um processo iterativo para se encontrar uma solução convergente. Considera-se que todos os viajantes têm perfeita informação de todas as possibilidades de escolhas na rede e que todos escolhem as rotas que minimizam seus tempos de viagens e custos, o que não corresponde com a realidade;

6. *Equilíbrio de Usuários Estocástico*: É uma generalização do modelo anterior que assume que os viajantes não têm um conhecimento perfeito da rede e eles atribuem aleatoriamente custos diferentes em diferentes caminhos. Produzem, assim, resultados mais realísticos;

7. *Sistemas Ótimos*: Calcula uma atribuição que minimiza o tempo total na rede. Assim, o sistema ótimo consiste em um modelo que reduz o congestionamento da rede, indicando a todos os usuários quais rotas devem usar. Apesar de difícil aplicação em casos reais, é usado para cenários de Sistemas Inteligentes de Transportes.

Os objetivos da alocação de tráfego, de acordo com Bruton (1979), são:

- 1) Verificar as deficiências do atual Sistema de Transportes, atribuindo-se as estimativas de viagens futuras à rede viária existente;
- 2) Avaliar os efeitos de pequenos melhoramentos e extensões do Sistema de Transportes existente, atribuindo-se as estimativas das viagens futuras à rede de transportes modificada/melhorada;
- 3) Determinar prioridades de construções (viadutos, vias, entre outras) no Sistema de Transportes;
- 4) Testar várias propostas de Sistemas de Transportes.

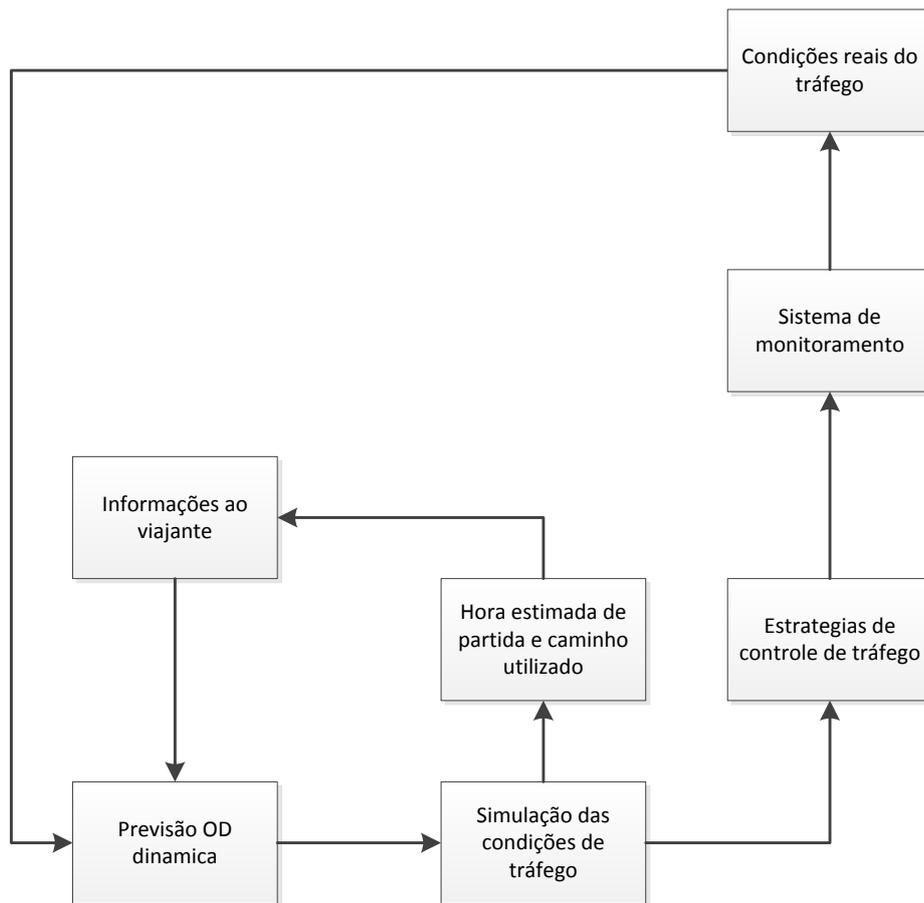
A má qualidade e a não integração dos dados são os maiores problemas enfrentados pelos planejadores. Os dados precisam ser armazenados, acessados, filtrados, organizados, adequados, integrados e analisados para que gerem um resultado que apoie a tomada de decisões pelos responsáveis pelo transporte. Para atender a essas necessidades, é necessário um conjunto de ferramentas e técnicas que estão contempladas em BI – *Business Intelligence*, que é o termo utilizado para designar a utilização de técnicas computacionais durante a coleta, organização, análise e

monitoramento de informações que irão oferecer suporte ao processo de tomada de decisão de gestores.

### 2.2.5.2 Alocação Dinâmica do Tráfego

Os primeiros trabalhos para estimar matrizes dinâmicas OD foram elaborados por Cremer e Keller (1981), que identificaram aspectos da determinação da origem-destino a partir da contagem nos cruzamentos dos fluxos. Bell (1983) introduziu parâmetros adicionais nos tempos de viagem para reavaliar a estimativa do tempo de viagem e propôs uma formulação baseada na aplicação do Método dos Mínimos Quadrados Ponderado. Willumsem (1984) usou contagem de tráfego para determinar a matriz OD dependente do tempo. Cascetta *et al.* (1993) propuseram o uso do Método dos Mínimos Quadrados Generalizado na malha rodoviária. Chang e Wu (1994) formularam o problema como uma extensão do filtro de Kalman com o vetor de estado definido como uma combinação de parcelas e parâmetros de alocação do tráfego. O filtro de Kalman produz estimativas dos valores reais de grandezas medidas e valores associados, predizendo um valor, estimando a incerteza do valor predito e calculando uma média ponderada entre o valor predito e o valor medido. O peso maior é dado ao valor de menor incerteza. As estimativas geradas pelo método tendem a permanecer mais próximas dos valores reais que as medidas originais, pois a média ponderada apresenta uma melhor estimativa de incerteza que ambos os valores utilizados no seu cálculo. Bierklaire e Crittin (2004) propuseram um novo modelo de estimação e previsão usando tabelas de contingência.

A Matriz Dinâmica OD, também chamada de tempo-dependente, é um insumo essencial para os modelos de atribuição de tráfego dinâmico (DTA – *Dynamic Traffic Assignment*), sendo usada, principalmente, para descrever e prever o tempo de viagem em função do fluxo de tráfego na rede conhecidos como ATIS (*Automatic Traveler Information System*), assim como estabelecer estratégias de controle e gestão para sistemas de tráfego ATMS (*Advanced Traffic Management Systems*). A Figura 2.7 descreve a interação entre a matriz OD dinâmica e os sistemas ATIS e ATMS.



**Figura 2.8** – Matriz OD dinâmica – baseada em (GUOZHEN *et al.* 2011)

A matriz OD dinâmica é determinada em sistema de simulação para estimar a condição de tráfego de rede. Utilizando as condições de tráfego da rede, o sistema ATIS fornece informações de orientação para os usuários e gera no sistema ATMS às estratégias de controle.

Uma das modalidades para estabelecimento do processo de estimação, também bastante utilizado em pesquisas sob o enfoque de séries temporais, consiste no emprego de Modelos de Espaço de Estados e o filtro dos mínimos quadrados generalizados ou filtro GLS. Esse filtro permite o tratamento de modelos com erros de observação autocorrelacionados de forma mais simples do que utilizando o filtro de Kalman e, além disso, possibilita a modelagem conjunta de várias subpopulações ou domínios sob restrições de *benchmark* obtidas a partir da mesma pesquisa. Isso não só permite manter a coerência entre as estimativas obtidas pelo método, e estimativas agregadas baseadas no planejamento da amostra, como ajudam na proteção contra possíveis erros de especificação dos modelos.

Guozhen *et al.* (2011) citam que o método GLS (*Generalized Least Squares*) apresenta-se como um dos mais utilizados para estimar a matriz OD dinâmica. Nesse trabalho, os autores usam o Método GLS e relevam o uso do histórico do tráfego adotando apenas os dados em tempo real. Essa abordagem será mais bem analisada por se relacionar com o nosso propósito.

A seguinte notação será usada para representar as variáveis da formulação da matriz OD dinâmica:

I	Zona de origem	$x_{(i,j,\tau)}^{AVL}$	Número total de veículos equipados com AVL com partida em i e destino j durante o intervalo $\tau$
J	Zona de destino	$x_{(i,\tau)}^{AVL}$	Número total de veículos equipado com AVL
I	<i>Link</i> com as medidas de tráfego	$\xi_{(i,j,\tau)}$	Diferença entre $x_{(i,j,\tau)}/x_{(i,\tau)}$ e $x_{(i,j,\tau)}^{AVL}/x_{(i,\tau)}^{AVL}$
T	Intervalo de tempo de partida	$\gamma_{(l,t)}$	Número de veículos no <i>link</i> l durante a observação t
T	Intervalo de tempo de observação	$\gamma_{(l,t)}^{sim}$	Número de veículos no link l durante o intervalo de observação t para o DTA usando a matriz OD estimada
$x_{(i,j,\tau)}$	Volume de tráfego com partida em i e destino j durante o intervalo $\tau$	$\eta_{(l,t)}$	Diferença entre $\gamma_{(l,t)}$ e o fluxo estimado no <i>link</i>
$\hat{x}_{(i,j,\tau)}$	Demanda estimada partindo de i e destino j durante o intervalo $\tau$	$\alpha_{(l,t)(i,j,\tau)}$	Proporção do fluxo no <i>link</i> que descreve a proporção do fluxo com origem i e destino j, partindo as viagens durante o intervalo $\zeta$ , medido no link l durante o período de

observação t.

$x_{(i,j,\tau)}^{his}$	Dados históricos de $x_{(i,j,\tau)}$	$\hat{\alpha}_{(l,t)(i,j,\tau)}$	Valor estimado de $\alpha_{(l,t)(i,j,\tau)}$ obtido do modelo DTA
$\varepsilon_{(i,j,\tau)}$	Diferença entre $\hat{x}_{(i,j,\tau)}$ e $x_{(i,j,\tau)}^{his}$	X	Matriz OD dinâmica consistindo de $x_{(i,j,\tau)}$
$x_{(i,\tau)}$	Número total de veículos partindo de i no intervalo $\tau$	$\dot{X}$	Objetivo da matriz OD, tal que atenda à OD
$\dot{Y}$	Vetor que mede os fluxos nos links consistindo de $Y_{(l,t)}$	$\bar{Y}$	Estimativa de Y

A abordagem de Guozhen *et al.* (2011) pode ser representada por (2):

$$\min F(X, Y) = W_1 \cdot F_1(X, \hat{Y}) + W_2 \cdot F_2(X, \bar{Y}) \quad (2)$$

Onde  $W_1 \cdot F_1(X, \hat{Y})$  e  $W_2 \cdot F_2(X, \bar{Y})$  representam, respectivamente, o desvio entre a matriz OD estimada e seu objetivo e o desvio entre o volume estimado e o volume medido.  $W_1$  e  $W_2$  são os pesos de  $F_1$  e  $F_2$ . O objetivo é fazer o valor da função  $F(X, Y)$  mínimo.

Muitos modelos para estimar a matriz OD usam a contagem do fluxo como parte integrante, o que na prática introduz erro de medição. Partindo do pressuposto de ignorar o erro de medição do fluxo de ligação, a relação entre o fluxo no *link* e os pares OD é medida da seguinte forma (3):

$$Y_{(l,f)} = \sum_{i,j,\tau} \alpha_{(l,t)(i,j,\tau)} \cdot x_{(i,j,\tau)} \quad (3)$$

A proporção de fluxo (l, t) (i, j) reflete a relação entre fluxo do *link* e os pares OD em tempo real, que não seria possível obter com precisão. Assim, a fórmula (3) pode ser reformulada conforme mostrado a seguir:

$$Y_{(l,f)} = \sum_{i,j,\tau} \hat{\alpha}_{(l,t)(i,j,\tau)} \cdot x_{(i,j,\tau)} + \eta_{(l,t)} \quad (4)$$

Usualmente os pesquisadores utilizam a matriz OD existente como parte para os cálculos da estimação da matriz OD dos próximos períodos conforme formula (5):

$$x_{(i,j,\tau)} = x_{(i,j,\tau)}^{his} + \varepsilon_{(i,j,\tau)} \quad (5)$$

$\varepsilon_{(i,j,\tau)}$  reflete a diferença entre a matriz OD real e os dados históricos. Assim,  $F_1$  pode ser formulado como:

$$F_1 = \sum_{i,j} [X_{(i,j,\tau)} - x_{(i,j,\tau)}^{his}]^2 \quad (6)$$

Esses resultados serão utilizados adiante para aferição do modelo proposto onde foi adotado o algoritmo do “vizinho mais próximo” para determinação dos padrões frequentes de movimentação dos veículos.

Em um horizonte mais recente, outro processo para alocação na malha rodoviária é a utilização de técnicas de *Business Intelligence* – BI para o estabelecimento de padrões frequentes de deslocamento.

## 2.3 Métodos para previsão da demanda futura

### 2.3.1 Métodos baseados em fatores de crescimento

Alguns métodos para estimar as demandas futuras baseiam-se nos dados atualizados por fatores de crescimento definidos para as zonas ou setores em questão. Como as viagens possuem origens e destinos, é comum ter dois fatores de crescimento associados a uma única viagem, permitindo diferentes soluções, como encontrar a média entre os dois fatores ou estabelecer uma proporção expressiva por processos interativos. Entretanto, a utilização de fatores de crescimento é bem mais limitada no caso de ocupações em novas zonas (ORTÚZAR e WILLUMSEN, 2011), já que um crescimento que parte de valores próximos de zero não permite um cálculo razoável de um fator de crescimento percentual.

### 2.3.2 Métodos baseados em atividades

Outra abordagem para estimar os volumes de viagem, sob a ótica de fatores diversos em relação às abordagens convencionais, consiste em considerar padrões de comportamento familiar com diferentes características de suas viagens diárias com base nas atividades (KUTZ, 2011). O método leva em conta fatores como necessidades biológicas, atividades interpessoais, horários de abertura e fechamento dos locais de destino e a sequência de atividades ao longo do dia. Esses modelos possuem uma maior complexidade, porém, oferecem uma boa alternativa para analisar o impacto de mudanças pontuais, como a inauguração de um novo polo de atração de viagens ou a mudança dos horários de funcionamento de polos já existentes.

## Capítulo 3 - Modelos adotados no contexto do Planejamento de Transporte de Carga

Este Capítulo destina-se a consolidar os principais modelos para o Planejamento do Transporte de Carga utilizados no Brasil, nos Estados Unidos e na União Europeia com uma análise crítica relacionada aos modelos.

### 3.1 Conceitos básicos

Ao longo das últimas décadas, a modelagem de transporte de carga foi estabelecida a partir de uma variedade de perspectivas (CAMBRIDGE SYSTEMATICS INC., 1997) das quais se destacam:

- Modelo das Quatro Etapas. Souza e D'Agosto (2013) citam que o modelo das quatro etapas apresenta-se como o mais utilizado para vários níveis de planejamento:

“o modelo é versátil, pode ser empregado de forma integral ou parcial, considerando fluxos de unidade de massa ou de veículos, no caso destes últimos diretamente aplicáveis ao dimensionamento de infraestrutura de tráfego, em todos os níveis de abrangência geográfica (de municípios a países), sem limitação do tipo de carga que pode ser transportada por qualquer dos cinco modos de transporte ou suas combinações, configurando ferramenta de planejamento estratégico e apoio ao desenvolvimento socioeconômico do país”

- Para o atendimento à Cadeia de Suprimento *Supply Chain Management*, vários modelos foram desenvolvidos com o objetivo de melhorar a eficiência das operações de negócios para as empresas individuais ou grupos de empresas (MEIXELL e GARGEYA, 2005; CRAINIC e LAPORTE, 1997);
- Outros modelos para a tomada de decisão do setor público, com diferentes abordagens e para fins específicos, objetivam avaliar as melhorias de infraestrutura, as alternativas de uso da terra e as políticas de transporte (TRB, 2008);

Essas abordagens foram, em grande parte, modeladas sobre os conceitos relativos aos métodos desenvolvidos para a previsão de demanda de passageiros (PENDYALA *et al.*, 2000).

Segundo Roorda *et al.* (2010), mais recentemente, um pequeno número de modelos híbridos tem sido desenvolvidos para tentar incorporar elementos comportamentais da gestão da cadeia de suprimentos e de decisão do setor público.

Os modelos híbridos consideram que:

- i. Existem diversos atores envolvidos na produção e distribuição de bens, nenhum dos quais pode ter controle total ou mesmo conhecimento de todas as decisões tomadas ao longo da cadeia de abastecimento;
- ii. As interações entre as empresas são diversas. Cadeias de abastecimento cada vez mais bem-sucedidas envolvem alianças de longo prazo entre fornecedores, fabricantes, varejistas, operadoras e empresas de logística entre outros;
- iii. Os modelos de negócio estão mudando ao longo do tempo, incluindo mudanças na operação do negócio para níveis mais baixos de estoque, apenas na entrega *just in time*, e um ambiente empresarial cada vez mais impulsionado por encomendas de clientes. A globalização do comércio também aumentou, impulsionado por reduções no custo de transporte, diferenças de recursos, salários e custos de produção.

### 3.2 Modelos adotados no Brasil

No Brasil, alguns estudos se destacam para o planejamento do transporte. A grande maioria baseados no Modelo das Quatro Etapas, complementados por estudos estatísticos para a projeção da demanda em cenário futuro. No cenário estadual, o PDTU – Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro - PDTU/RMRJ (2002), com dados iniciados em 2004, e atualmente em fase de revisão pela SECTRAN – Secretaria Estadual de Transportes, apresenta o cenário para o transporte de passageiros e de carga. O Plano Nacional de Logística e Transportes – PNLT (2007) foi elaborado pelo Ministério dos Transportes em colaboração com o Ministério da Defesa por meio do CENTRAN – Centro de Excelência em Engenharia de Transportes que objetiva:

“formalizar e perenizar instrumentos de análise, sob a ótica logística, para dar suporte ao planejamento de médio e longo prazo de intervenções públicas e privadas na infraestrutura e na organização dos transportes, de modo a que o setor possa contribuir para a consecução das metas rumo ao desenvolvimento sustentado”.

Outro estudo de âmbito nacional para a área de transporte de carga foi elaborado pela Confederação do Transporte com o Plano CNT de Transporte e Logística 2011 (CNT, 2011), tendo como meta enumerar e sequenciar projetos prioritários sob a ótica sistêmica de integração da infraestrutura. Na seleção de projetos dos Eixos Estruturantes, foram priorizadas: a intermodalidade, a conexão com os países da América Latina, a acessibilidade aos pontos de exportação da economia brasileira e a integração entre as zonas de produção e de consumo interno. Em relação às Regiões Metropolitanas, as principais considerações foram: a mobilidade e a acessibilidade urbanas, a integração tanto dos municípios como dos modais de transporte urbano e o nível de serviço das infraestruturas.

### 3.2.1 Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro - PDTU

Esse estudo teve seu início em março de 2002, sendo executado pelo consórcio LOGIT-OFFICINA-JGP com a participação de técnicos da Companhia Estadual de Engenharia de Transportes e Logística – CENTRAL e com a colaboração da Secretaria Estadual de Transportes – SECTRAN e de outros órgãos municipais e estaduais (SECTRANS, 2002).

Dois aspectos foram estudados: transporte de passageiros e transporte de carga para a Região Metropolitana do Rio de Janeiro (RMRJ).

Os objetivos gerais do transporte de passageiros foram:

- O diagnóstico da situação atual da demanda e oferta dos transportes da RMRJ;
- A formulação de alternativas que visem à racionalização do sistema de transporte da RMRJ, com especial ênfase à política de integração intermodal física e tarifária;
- A formulação de uma política de investimentos em infraestrutura viária e nos transportes coletivos;

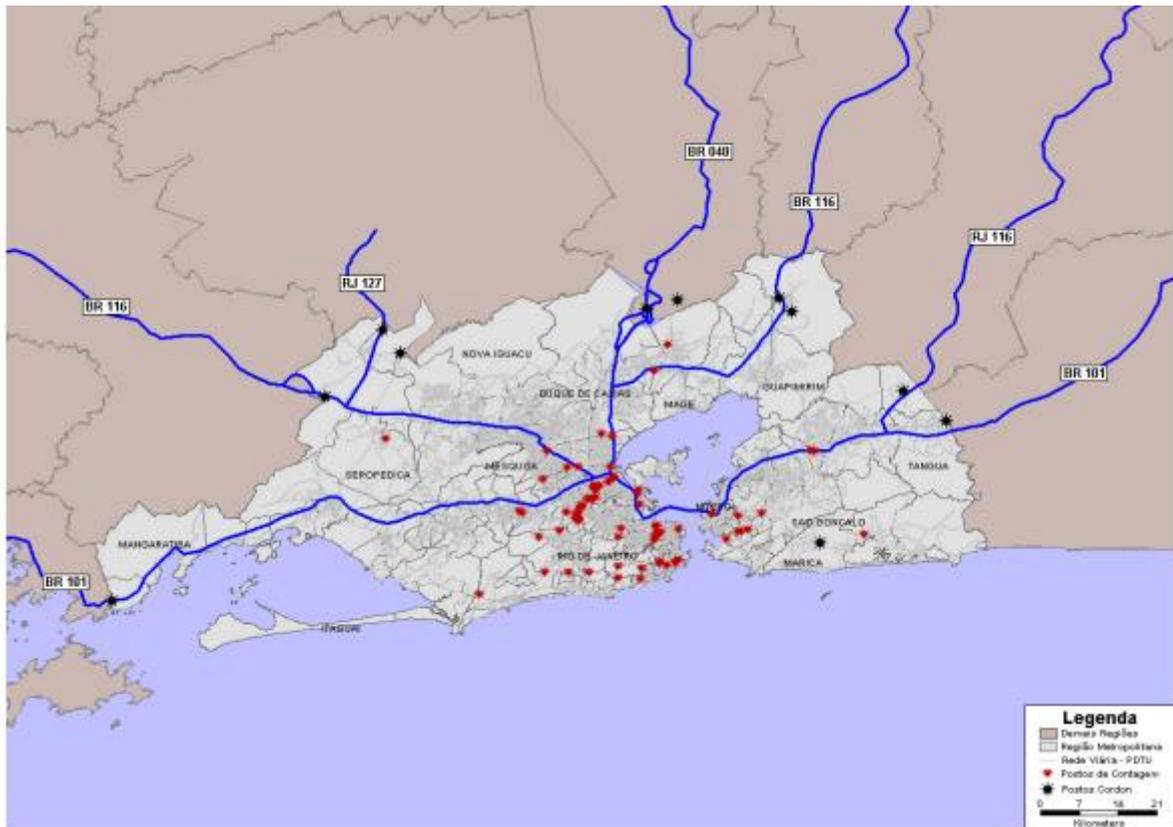
- O fornecimento de um instrumental que permita implementar um processo permanente de planejamento.

Para o transporte de carga, foi elaborada uma extensão do PDTU, visando a obter:

“uma avaliação sobre os Portos de Sepetiba, Rio de Janeiro e Angra dos Reis, sobre os aeroportos do Rio (Galeão), Cabo Frio, Campos e Macaé, e um Plano Diretor de Desenvolvimento, Promoção e Integração Logística para a área de influência do Corredor Centro-Sudeste e, por consequência, do sistema portuário e aeroportuário do Estado do Rio de Janeiro e de sua hinterlândia”

Além dos levantamentos de dados pelas fontes secundárias (junto às fontes operadoras de transporte de carga aeroviária, ferroviária e hidroviária), foi realizada Pesquisa Origem-Destino com os transportadores rodoviários de carga, ao longo dos postos de pesquisa localizados no cordão externo da área de estudo definidos no PDTU

A Pesquisa Origem-Destino de Carga foi realizada em 12 postos na linha de contorno da RMRJ e foi complementada com contagens volumétricas. A localização dos postos de leitura é mostrada na Figura 3.1, que mostra também a localização das rodovias pesquisadas e os respectivos postos de coleta.

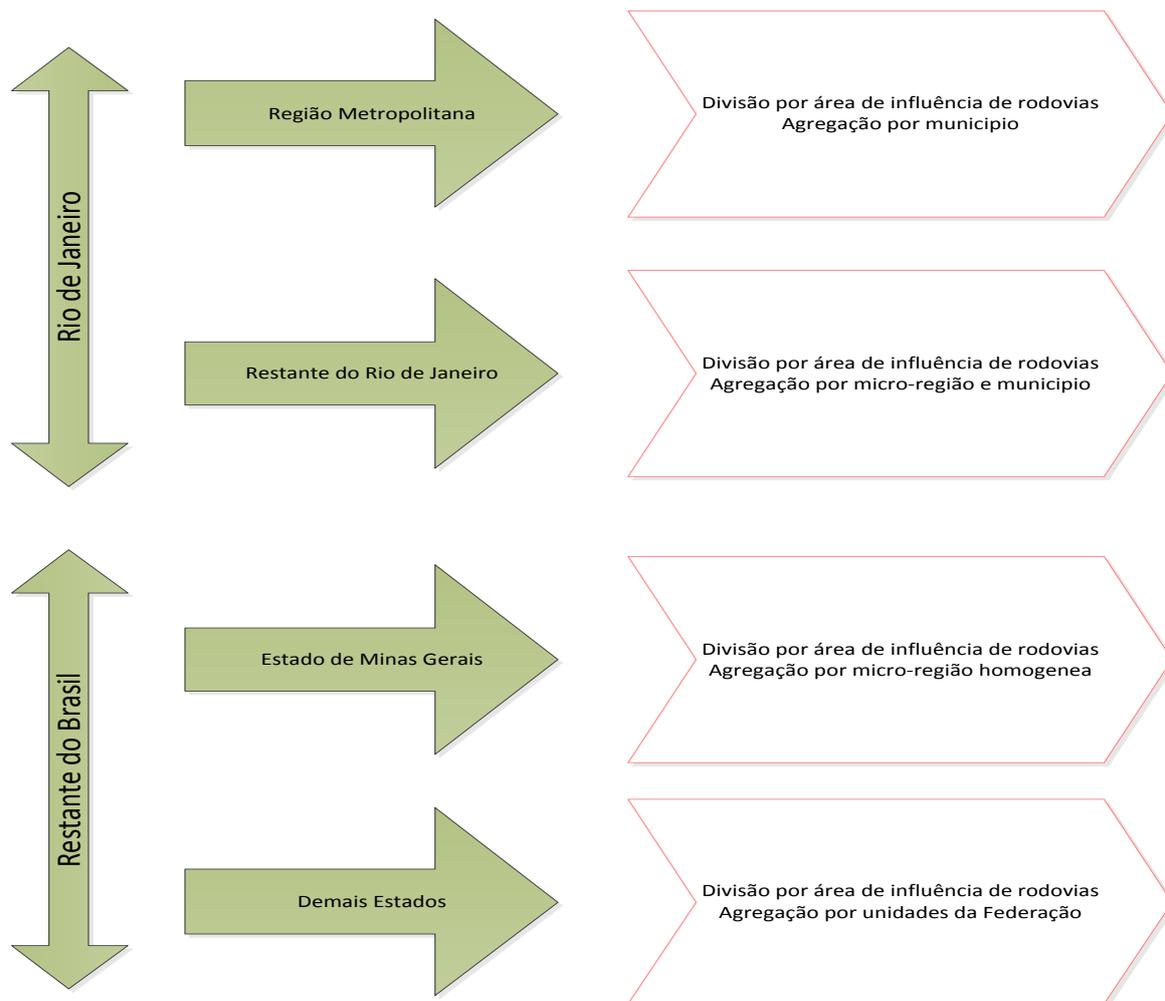


**Figura 3.1:** Linha de Contorno RMRJ (SECTRANS, 2002)

Para a construção da Matriz OD, foi aplicado um questionário contendo os seguintes itens: Localização da medição, Data, hora e sentido; Propriedade do veículo; Tipo de carroceria; Tipo e peso da carga; Origem da viagem; Destino da viagem; Motivo da viagem; Fluxo de passagem (identificação de eventuais paradas na RMRJ); Local de estacionamento, carga e descarga.

A partir da Pesquisa OD de carga e de informações em fontes secundárias, tais como ANTT, SECEX, inclusive mantendo contato com associações de classe e sindicatos, foi desenvolvida a terceira etapa dos estudos, que tratou da elaboração das matrizes de carga para os produtos julgados relevantes para o Estado do Rio de Janeiro, considerando as seguintes etapas: Descrição das fontes de dados; Definição do zoneamento; Processamento da pesquisa de campo; Escolha dos produtos relevantes; Elaboração das matrizes OD de carga para o ano base (2003); Projeção das matrizes OD para os anos horizonte (2008 e 2013).

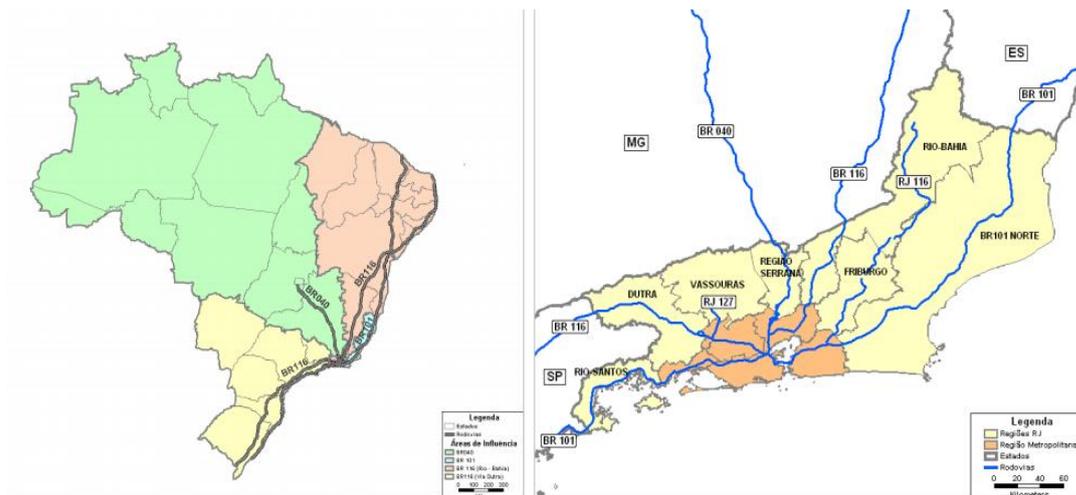
O zoneamento adotado para as matrizes estaduais de carga considera nível de detalhamento compatível com os estudos regionais, conforme Figura 3.2:



**Figura 3.2:** Zoneamento PDTU (SECTTRANS, 2002).

Essa Figura mostra que o estabelecimento do zoneamento considerou os principais centros produtores e atratores para a carga do Rio de Janeiro e para o restante do Brasil. Dentro do Estado do Rio de Janeiro, foram consideradas divisões por área de influência das rodovias por município, enquanto na região metropolitana da Cidade do Rio de Janeiro por micro-região. No restante do Brasil, o Estado de Minas Gerais teve papel preponderante, fazendo-se a agregação por micro-região homogênea. Para os demais Estados da Federação, considerou-se a área de influência das rodovias.

A Figura 3.3 mostra os principais eixos rodoviários considerados, tendo por origem o Estado do Rio de Janeiro.



**Figura 3.3:** Rodovias consideradas no PDTU Carga (SECTRANS, 2002)

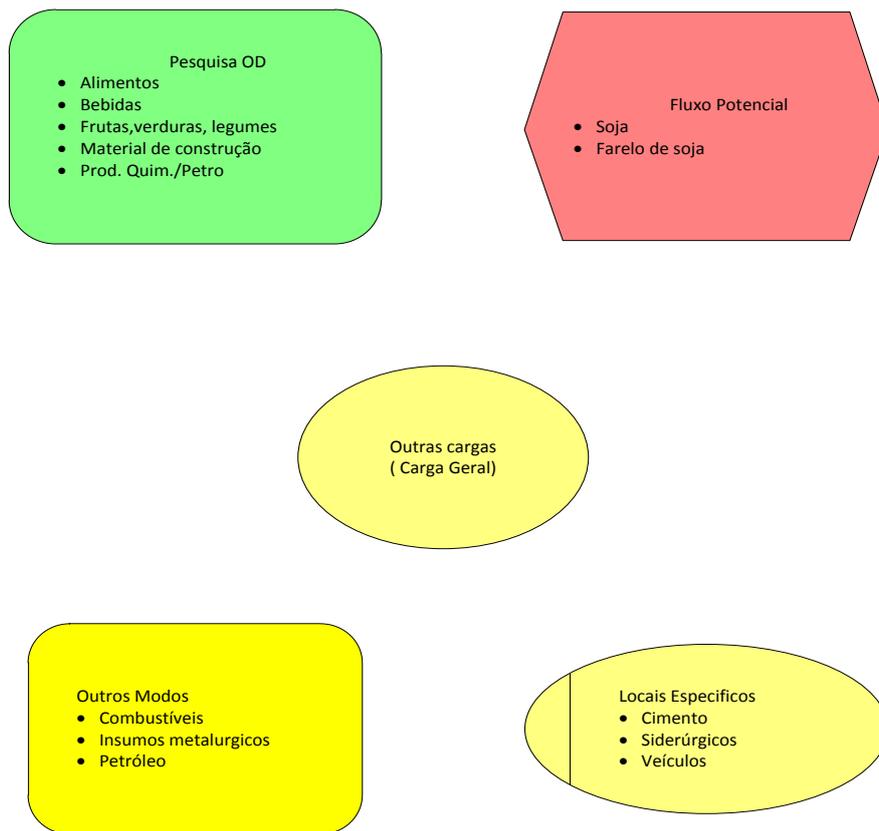
O levantamento dos principais produtos ofertados e demandados nas regiões de influência no corredor de transportes centro-sudeste constatou a ocorrência de produtos agrícolas, relacionados à pecuária, de produção industrial, relativos à importação.

Para o Estado do Rio de Janeiro, foram selecionados os produtos constantes do Quadro 3.1:

**Quadro 3.1:** Principais produtos analisados

Sector de Atividade	Produtos
Agricultura	Café, Feijão, Frutas, Soja.
Pecuária	Carne Bovina, Carne Suína Aves.
Produção Industrial	Farelo de Soja, Gesso, Minério, Aço, Veículos, Açúcar, Carvão, Celulose, Granito.
Importação	Trigo, Arroz, Malte, Fertilizantes.

Com base nas pesquisas realizadas e nos critérios estabelecidos, foram realizadas agregações e desagregações de forma a se obter um número de matrizes representativas dos principais fluxos do Estado. Essas agregações estão apresentadas na Figura 3.4:



**Figura 3.4:** Principais Agregações (SECTRANS, 2002)

A Figura 3.4 apresenta as principais fases para estimativa de matrizes com base em informações de fontes secundárias que estabeleceram a caracterização da Cadeia Logística, a montagem do Balanço de Oferta e Demanda e as Matrizes OD para os quais foram usados o Modelo Gravitacional.

Da análise das produções e atrações apresentadas, identificam-se os produtos relevantes com maior impacto nos transportes do Estado. O resultado da pesquisa foi representado por matrizes tridimensionais  $6 \times 10 \times 10$  relativas respectivamente aos produtos, origens e destinos.

### 3.2.2 Plano Nacional de Logística e Transportes - PNLT

Datado de 2007, esse Plano foi desenvolvido pelo Ministério dos Transportes juntamente com o Ministério da Defesa.

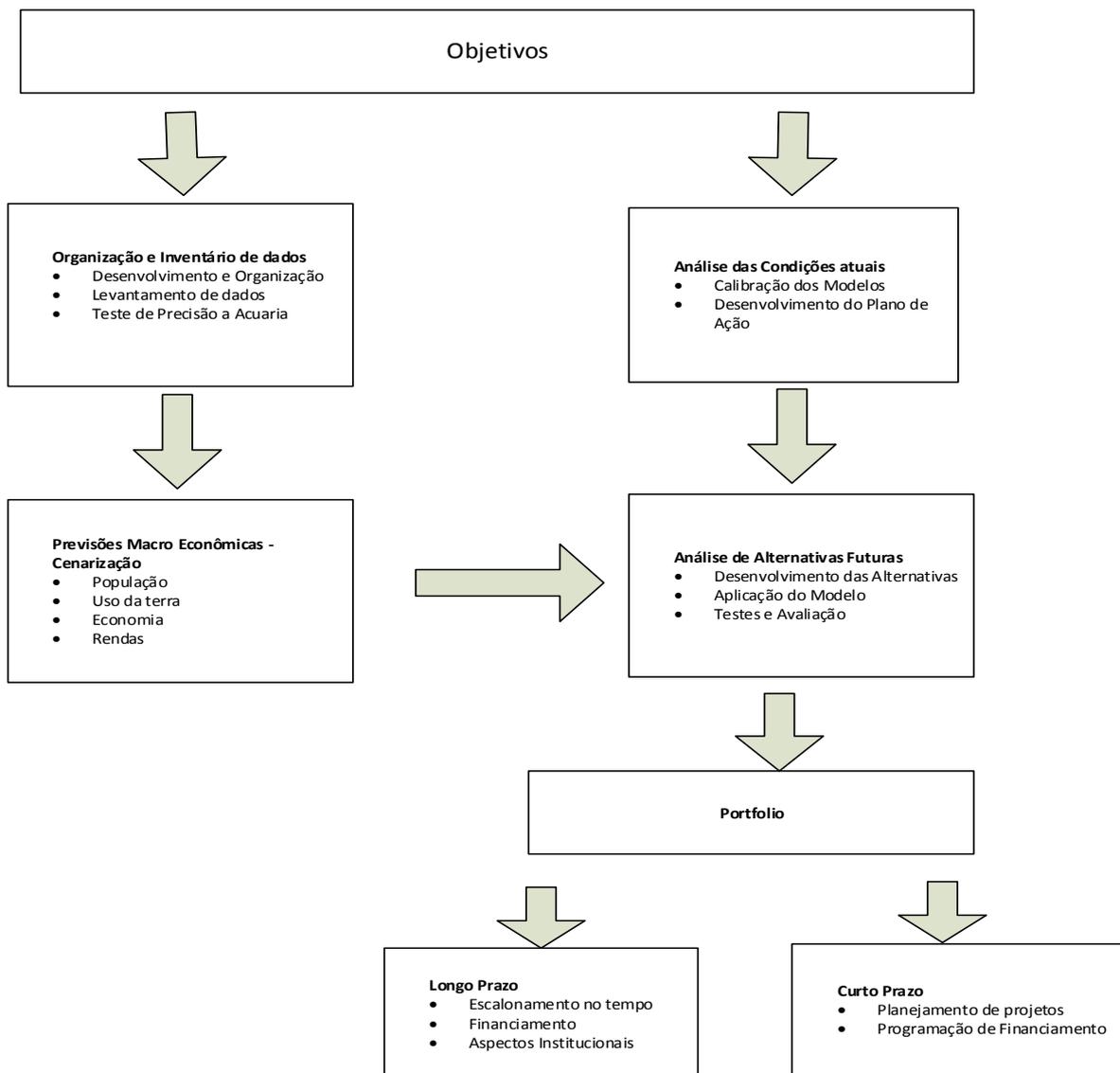
Os objetivos e a abrangência dos Estudos são (MINISTERIO DOS TRANSPORTES, 2007):

“um primeiro e fundamental objetivo do PNLT é a retomada do processo de planejamento no Setor dos Transportes, dotando-o de uma estrutura permanente de gestão, com base em um sistema de informações georreferenciado,

Um segundo objetivo, é a consideração dos custos de toda a cadeia logística que permeia o processo que se estabelece entre as origens e os destinos dos fluxos de transporte...

Um terceiro objetivo é a necessidade de efetiva mudança, com melhor equilíbrio, na atual matriz de transportes de cargas do País, das modalidades ferroviária e aquaviária.”

A abordagem utilizada em seu desenvolvimento retrata o processo clássico no planejamento dos transportes e é ilustrada de forma simplificada na Figura 3.5:



**Figura 3.5:** Processo de Planejamento de Transportes usado no PNLT (2007)

A partir do processo mostrado na Figura 3.5, o PNLT procurou, essencialmente, formalizar e aperfeiçoar a metodologia de planejamento estratégico usada pelo Ministério dos Transportes, objetivando gerar de forma consistente as contribuições requeridas do setor para subsidiar a elaboração dos próximos PPA's – Plano Plurianual, segundo a qual, a modelagem estratégica do transporte foram considerados os principais tipos de cargas, além de passageiros, no nível de agregação espacial dos principais corredores entre microrregiões brasileiras, definidas pelo IBGE.

- **Definição do Zoneamento de Tráfego**

O zoneamento do tráfego utilizado no PNLT (2007) estabelece a divisão da área de estudo em zonas e possibilitou a representação das regiões que apresentam

características homogêneas em relação à demanda por transporte. Os critérios adotados para a definição do zoneamento do estudo levaram em consideração que cada zona de transporte tem como unidade espacial mínima os limites dos municípios.

Em função do nível de detalhamento desejado, as zonas de transporte corresponderam à subdivisão de microrregiões homogêneas definida pelo IBGE, sendo que nas Regiões Centro-Oeste e Norte o zoneamento foi mais detalhado, de forma a permitir uma maior precisão na modelagem de transportes.

O Comércio Exterior foi representado, dependendo do caso, pelas zonas correspondentes aos portos marítimos e postos de fronteira terrestre, ou por zonas correspondentes a agregações de países fronteiriços ou a uma zona exterior correspondente aos demais países com quem o Brasil mantém o seu comércio internacional.

A Figura 3.6 ilustra o zoneamento básico de transportes adotado no PNLT, correspondente às 558 microrregiões homogêneas do IBGE.



**Figura 3.6:** Zonas de Transporte do PNLT ( 2007)

O Zoneamento Básico mostrado na Figura 3.6 possibilitou estabelecer projeções para a oferta e a demanda de 80 tipos de produtos, em cada uma das 558 microrregiões do Brasil. Para o período 2007 – 2011 – 2015 – 2019 – 2023, adotou-se metodologia macroeconômica que consistiu em:

(i) Estruturar um cenário referencial para o ano-base, considerando as características estruturais do sistema econômico atual, sua evolução recente e conhecimentos como os espaços econômicos se inter-relacionam;

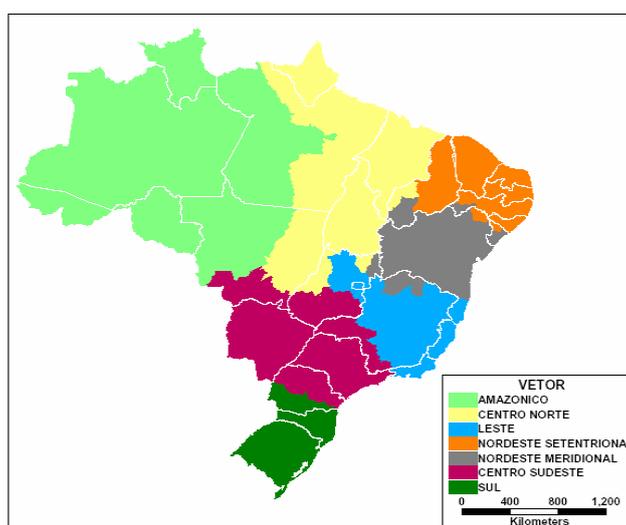
(ii) Aplicar um modelo computável de equilíbrio geral (EFES – *Economic Forecasting Equilibrium System*) ao cenário referencial, permitindo a geração de cenários futuros.

Para o planejamento de transportes, adotou-se o Modelo das Quatro Etapas.

A proposta de organização espacial do País consistiu no agrupamento de microrregiões homogêneas em função da superposição georreferenciada de diversos fatores representativos de suas características:

- Impedâncias ambientais;
- Similaridades socioeconômicas;
- Perspectivas de integração e inter-relacionamento (a antiga noção de “corredores de transporte”);
- Funções de transporte, identificadas a partir da análise de isocustos em relação aos principais portos concentradores de carga do País.

Esses agrupamentos designados “vetores logísticos” são mostrados na Figura 3.7.



**Figura 3.7:** Vetores Logísticos (PNLT, 2007)

Da análise dos Vetores Logísticos apresentados na Figura 4.7 resultaram sete agrupamentos compostos por microrregiões agregadas, segundo os fatores acima referidos:

- ✓ Amazônico, Centro-Norte, Nordeste Setentrional, Nordeste Meridional, Leste, Centro-Sudeste e Sul.

O carregamento do tráfego para os anos 2007, 2011, 2015 e 2023 foi elaborado considerando que o transporte de minério de ferro assume valores elevados, da ordem de 100 milhões de toneladas/ano, enquanto que as demais mercadorias ostentam valores sensivelmente menores.

Adicionalmente são apresentados, para cada ano de análise, os respectivos mapas com o nível de serviço do modal rodoviário, a partir da relação volume/capacidade – V/C de cada trecho considerado, segundo a classificação do Highway Capacity Manual (2000).

O carregamento do transporte multimodal, sem considerar o transporte de minério para o ano 2007, é apresentado na Figura 3.8:



**Figura 3.8:** Carregamentos Multimodais (PNLT, 2007)

A análise dos elementos produzidos e apresentados na Figura 4.8 retrata que a matriz de transporte brasileira depende, preponderantemente, do modo de transporte rodoviário. Em valores aproximados, a rodovia responde pela movimentação de 58%

dos fluxos de carga (se excluído o transporte de minério de ferro, essa participação supera 70%) e por 95% dos passageiros no País. A participação do modo ferroviário atinge 25% e as hidrovias, 13% das cargas, restando aos modos dutoviário e aéreo participações em torno de 3,6% e 0,4%, respectivamente.

### 3.3 Modelos adotados no Exterior

#### 3.3.1 Estados Unidos da América

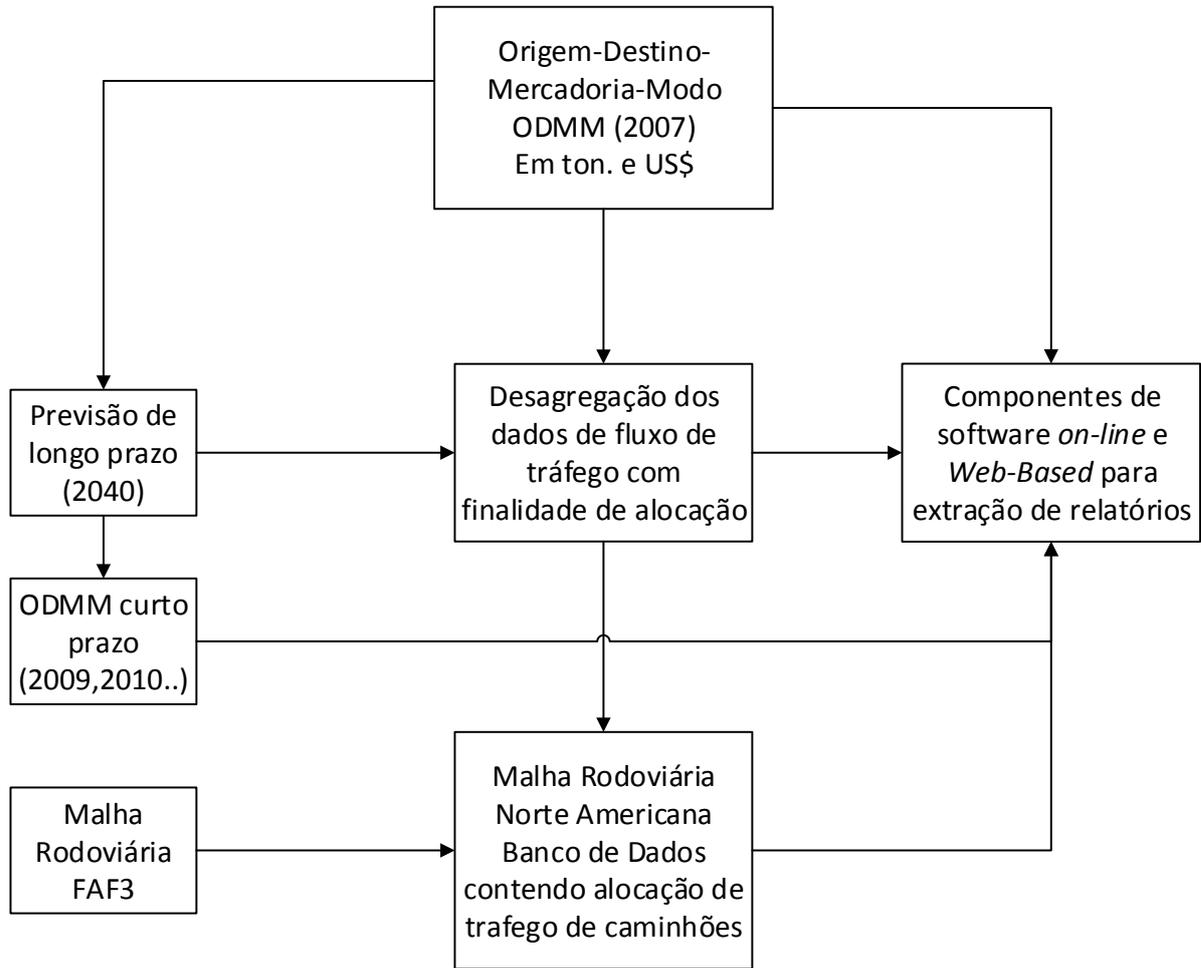
##### 3.3.1.1 Considerações preliminares

O *Freight Analysis Framework* (FAF) é um estudo elaborado para o *Federal Highway Administration*, pertencente ao *U.S Highway Administration*, que integra dados de diversas fontes para criar um cenário compreensível da movimentação de carga pelos Estados Norte Americanos e por suas principais cidades por diversos modos de transporte. O FAF3 é a terceira versão de Banco de Dados de Movimentação de Carga, enquanto o FAF1 foi elaborado considerando estimativas relacionadas ao transporte por caminhão, trem, e fluvial/marítimo para o ano 1998. O FAF2 apresenta um quadro mais completo a partir da Pesquisa de Fluxo de Mercadorias para o ano civil de 2002. A descrição e os critérios de construção do FAF3 (*Network Database and Flow Assignment: 2007 and 2040*) encontram-se no *site* [ops.fhwa.dot.gov/freight/freight\\_analysis/faf/faf3/netwkdbflow/index.htm](http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf3/netwkdbflow/index.htm). Esse estudo inclui um Banco de Dados de movimentação de carga regional e a alocação de tráfego na malha viária americana com resultados representados em toneladas e em dólar americano, baseados em levantamento realizado em 2007 e projetado para o ano 2040, segundo as teorias macroeconômicas para o crescimento da economia norte-americana (FHWA, 2011). As principais dimensões da matriz de fluxo são:

- Local de embarque – Origem (O);
- Local de destino – Destino (D);
- Classe do bem transportado – “*Commodity*” (C);
- Modo de transporte – (M).

No total, a Matriz de Fluxo apresenta quatro dimensões 131 x 131 x 43 x 7 para Origem, Destino, “*Commodity*” e Meio de Transporte, respectivamente. A Figura 3.9

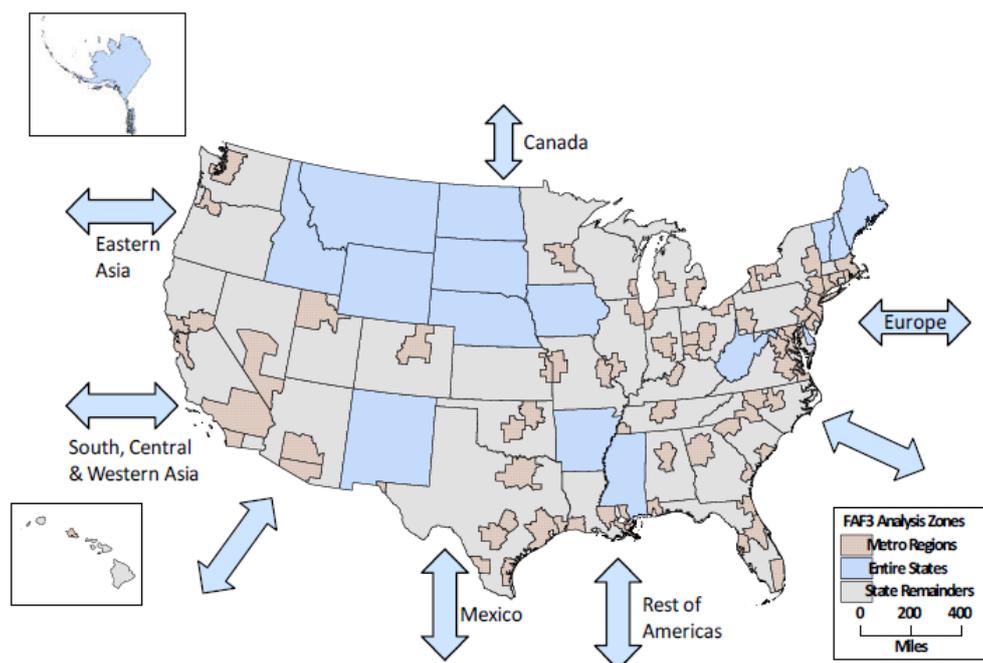
ilustra as principais etapas e relações entre os diversos produtos do FAF3, iniciando com os dados de movimentação de carga do ano de 2007.



**Figura 3.9:** Principais Produtos do FAF3 (FHWA, 2011)

### 3.3.1.2 Considerações sobre aspectos geográficos

A principal fonte de informações do FAF3 foi a pesquisa CFS 2007, coletadas pelo Censo e pelo Departamento de Transportes dos EUA, em que foram estabelecidas 123 zonas de tráfego domésticas e 8 regiões de comércio internacional, conforme mostrado na Figura 3.10:



**Figura 3.10:** Geografia do FAF3 (FHWA, 2011)

Nessa figura, é possível observar, nos três conjuntos de regiões em destaque, 74 principais regiões metropolitanas, 33 remanescentes fora dessas metropolitanas e 16 regiões identificadas como estados.

### 3.3.1.3 Considerações sobre a construção da Matriz OD

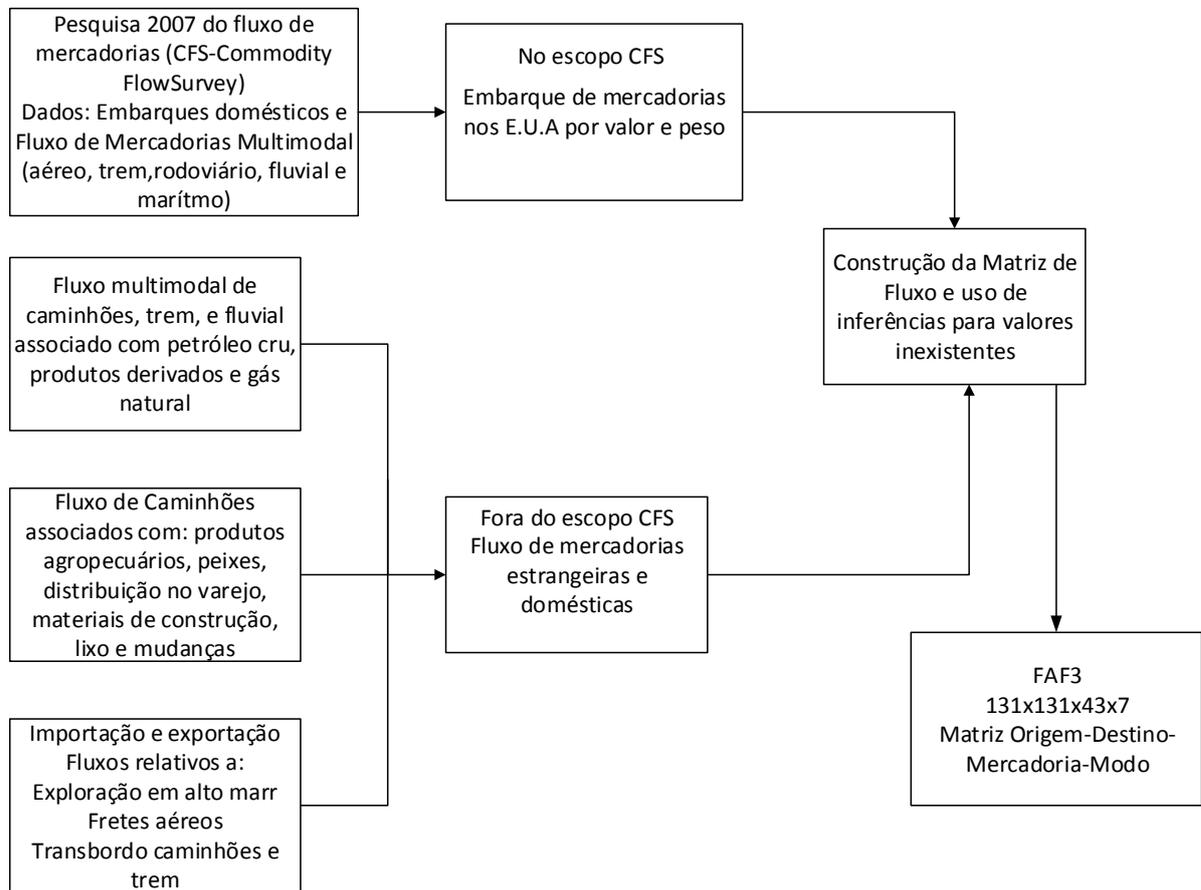
No desenvolvimento desse trabalho, os autores utilizaram um conceito análogo ao de Zonas de Tráfego, porém, relacionado ao transporte de carga e denominado de *Freight Analysis Zone* (FAZ) ou Zonas de Tráfego de Frete para cada tipo de mercadoria selecionadas, segundo o *Standard Classification of Transported Goods* (SCTG).

Para tanto, os códigos de cinco dígitos das Zonas Postais (ZBP) do Censo dos EUA foram usados como base de unidades espaciais (Bsus). A análise de questões importantes no projeto foi elaborada em três cenários: área metropolitana, atividade econômica e rede de estradas, sendo os cenários construídos, utilizando um sistema de informação geográfica para um número predefinido de zonas.

Os resultados foram tratados para a eventual perda de informações e posteriormente avaliados por meio de três critérios de desempenho: distância mínima, correlação entre as atividades econômicas e agregação por zonas. Os resultados obtidos

indicaram um bom desempenho para a metodologia adotada, especialmente, a área metropolitana baseada em FAZ.

O fluxograma com as principais etapas para a construção da matriz OD é apresentado na Figura 3.11:



**Figura 3.11** Fluxograma da Construção da Matriz OD (FHWA, 2011)

Pela análise dos procedimentos retratados no fluxograma, constata-se um esforço significativo em compatibilizar a estrutura da pesquisa CFS com o levantamento de dados considerados inconsistentes. Nesta depuração de dados, três aspectos destacam-se: o primeiro é relativo à pesquisa CFS, quando foram distribuídos formulários para as principais empresas norte-americanas questionando sobre aspectos da movimentação de carga, uma vez que esses dados foram apresentados de forma agregada para um determinado elenco de produtos. Em uma segunda etapa, o preenchimento das células para os quais não existiu informação e, o terceiro aspecto, o tratamento das mercadorias fora do escopo do CFS.

### 3.3.2 União Europeia

Para a União Europeia – UE, o órgão de pesquisa TRKC – *Transport Research Knowledge Center* elaborou e apresentou em 2010 a versão do *Thematic Research Summary (TRS) on Freight Transport*, que consolida os projetos europeus para o transporte de carga (TRKC, 2010), evidenciando a política da UE de apoio à intermodalidade, tal como o projeto Marco Polo de fornecimento de fundos aos operadores de transporte para realizar uma mudança do modo rodoviário para o transporte ferroviário e marítimo. Outra política de desenvolvimento principal diz respeito à promoção da Trans-Redes europeias que abrangem todas as modalidades e incluem as autoestradas.

Um dos objetivos da política atual é definir um quadro de tarifação diferenciada para melhorar a eficiência ambiental do transporte rodoviário de carga.

Alguns projetos cobertos pelo relatório TRKC (2010) são: ALSO DANUBE; ASAP; BRAVO; CESAR II; iBoS; IP;ITIP; ParcelCall; REALISE; REMARCC II; SIMTAG;SPIN-TN; TRAPIST; TROP. BESTLOG; D2D; FIRE; GAUSS; GIFTS; RECORDIT; SPIN; TASKU (FI); *Supply chain resilience* (FL0123) (UK).

## 3.4 Análise comparativa dos Modelos de Transporte de Carga

Analisando-se a bibliografia relativa ao Planejamento do Transporte, constata-se que, mesmo no cenário internacional, o modelo das quatro etapas é largamente utilizado, variando os modelos de previsão da demanda de transporte de carga.

Dentre os trabalhos analisados, dois aparecem em destaque:

### 3.4.1 Modelos Norte-Americanos

O trabalho elaborado por Chow *et al.* (2010) apresenta aspectos comparativos relacionados ao modelo dos Estados Unidos da América e da Europa, contudo, estabelece um foco primário nos Estados Unidos, do qual ressalta o trabalho do *National Highway Cooperative Research Program* (NCHRP 606, 2008). Em resumo, suas conclusões são descritas a seguir:

O Relatório 606 do NCHRP estabelece cinco classes de modelos para o planejamento do transporte de carga, definidos e categorizados por seus componentes:

- Classe A: Fatores de Crescimento – O fluxo de cada *link* é obtido aplicando-se taxas de crescimento ao volume de transporte de cargas. Esse método é bastante simples e usado principalmente em previsões de curto prazo;
- Classe B: *Factoring Method* – Considera os padrões de viagens dos veículos comerciais e usa dois recursos adicionais além do modelo Classe A, a divisão modal e a alocação de tráfego junto com a distribuição OD;
- Classe C: Modelos de Caminhões – Gera um modelo agregado de viagens de caminhões e faz a alocação na malha rodoviária;
- Classe D: Modelo das Quatro etapas – Já tratados anteriormente, porém, adaptados para área de carga. Esse modelo é mais adotado pelos órgãos governamentais de planejamento de transporte. As alterações nas etapas do modelo são:
  - ✓ Produção e Atração: Nessa etapa, são determinadas as quantidades de bens transportados de várias origens para vários destinos. A dimensão é estabelecida em peso por toneladas. Em um estágio intermediário do modelo de produção/atração, a dimensão pode ser o valor monetário;
  - ✓ Distribuição: O fluxo de bens transportados entre as diversas origens e respectivos destinos é calculado. A unidade utilizada é toneladas;
  - ✓ Divisão modal: Nessa etapa, considera-se a divisão modal por transporte rodoviário, ferroviário, fluvial ou multimodal, entre outros;
  - ✓ Alocação: Após a conversão do fluxo em toneladas para unidades de veículos, eles podem ser alocados à malha rodoviária.
- Classe E: Modelos de Atividade Econômica – Consideram os modelos econômicos do uso do solo que incorporam aspectos de custo dos sistemas de transporte.

As diversas lacunas entre as necessidades analíticas dos planejadores de transporte e os modelos citados anteriormente são mencionadas no Relatório NCHRP 606 e apresentadas no Quadro 3.1. Na elaboração do quadro, utilizou-se a notação indicada pela letra P, se o item é considerado prioritário, ou pela letra S se é considerado secundário ou por um traço (–) se o item não é considerado.

**Quadro 3.2** Comparativo Políticas e Necessidades de Análise x Classes de Modelo

Políticas e necessidades de análise	Classes de Modelo				
	A	B	C	D	E
Análise da divisão modal	-	S	-	P	P
Análise da movimentação de carga	-	P	-	P	P
Análise de <i>performance</i> dos programas de avaliação	-	-	-	-	-
Análise dos engarrafamentos	-	-	S	S	S
Análises setoriais	P	S	P	P	P
Corredores de comércio e de fronteiras internacionais	-	-	-	-	-
Desenvolvimento de projetos com inovação tecnológica	P	S	S	S	S
Estudos de Políticas Públicas	-	-	-	-	-
Gerenciamento de pavimentos, pontes e segurança	P	S	P	P	P
Operações, acidentes, roubos de carga, tamanho caminhões, excesso de peso	-	-	-	-	-
Planejamento de acesso aos terminais	-	S	-	S	P
Planejamento de Transportes governamentais	-	P	P	P	P
Planejamento Ferroviário	-	S	-	P	P
Planejamento macroeconômico	P	S	P	P	P
Previsão e análise do fluxo de caminhões	-	S	P	P	P

*Nota: Tipo de análise: (P) prioritária; (S) secundária.*

Comparando-se as políticas e as necessidades de análise identificadas no relatório NCHRP 606 com os requisitos dos modelos, é possível constatar que diversos aspectos não estão contemplados (ou superficialmente tratados) nos modelos de planejamento de transporte de carga dos quais podem-se citar: o impacto das Políticas Públicas, os corredores de transporte, os aspectos de operações e segurança e a medição de desempenho.

Essas insuficiências vêm ao encontro das conclusões para esses modelos apresentadas por Liedtke e Scepperle (2004), Hesse e Rodrigues (2004), Friesz e Holguín-Veras (2005), entre outros. Hensher e Figliozzi (2007) enfatizam as

inadequações entre o Modelo das Quatro Etapas e a tecnologia do século XXI, concluindo ser crucial “para o transporte de carga modelos que contemplem as relações das cadeias de suprimento e das restrições da logística” (tradução livre). Essencialmente, os modelos de planejamento de carga, na prática, dependem das abordagens em que são agregados os aspectos relativos ao comportamento econômico, quer sejam segundo a ótica das empresas ou dos gestores políticos. Além disso, os dois principais tipos de modelos, baseados na movimentação do veículo ou da mercadoria, apresentam falhas. Modelos baseados em veículo, tais como os das Quatro Etapas, não conseguem modelar o subjacente comportamento econômico, tais como fluxos de mercadorias, a partir do qual a demanda é realmente derivada. Modelos baseados em mercadorias, não contemplam todos os aspectos das movimentações dos veículos, especialmente em ambientes urbanos (HOLGUIN-VERAS *et al.*, 2004).

Independentemente do método utilizado, gera-se, enfim, o fluxo de viagens em cada via. Pode-se dessa forma quantificar a demanda por transporte na região analisada e verificar os impactos nessa demanda provocados por fatores externos, como a reestruturação da rede viária ou o crescimento populacional, contanto que sejam executados os quatro passos mencionados, utilizando os dados anteriores e posteriores às mudanças a serem avaliadas.

### 3.4.2 Modelos Europeus

O trabalho desenvolvido pelos pesquisadores Jong *et al.* (2004) apresenta uma ampla revisão dos atuais modelos utilizados para a União Europeia, que serão detalhados adiante.

Um significativo elenco de países integrantes da Comunidade Europeia desenvolveu modelos para previsão do transporte de carga e de seus respectivos fluxos. Dentre esses modelos, pode-se citar:

- O modelo para a Europa SCENES (*Scenarios for European transport*) e NEAC;
- O modelo belga Walloon *region freight model system* (WFTM), que usa o *software* de alocação multimodal NODUS;
- O modelo nacional italiano, que usa o modelo de frete *input-output*;

- Os modelos holandeses TEM (*Transport Economics Model*) e SMILE (*Strategic Model for Integrated Logistic Evaluation*). Sobre esses modelos, as versões anteriores utilizavam métodos *input-output*, porém, as versões atuais possuem módulos para localização de Centros de Distribuição;
- Os modelos usados no Reino Unido para previsão da movimentação de carga baseados no modelo STEMM, Strategic European Multimodal Modelling Project;
- SAMGODS – Modelo nacional sueco para transporte de carga;
- NEMO – Modelo nacional norueguês para transporte de carga. Tanto o modelo SAMGODS quanto o NEMO utilizam o *software* STAN para realizar a alocação do transporte.

### 3.5 Considerações Finais

Observa-se que o modelo das quatro etapas foi adotado nos países analisados quando a abordagem foi relacionada com planejamento nacional de transporte. Diversas críticas são formuladas ao modelo em virtude de diversos aspectos não estarem contemplados (ou superficialmente tratados) nos modelos de planejamento de transporte de carga como, por exemplo, o impacto das Políticas Públicas, os corredores de transporte, os aspectos de operações e segurança e a medição de desempenho.

Essas deficiências serão abordadas adiante no modelo de aquisição e tratamento de dados para o planejamento do transporte de carga tratado nesta tese.

## Capítulo 4 – Aspectos da Tecnologia da Informação aplicáveis ao desenvolvimento dos Modelos

Este capítulo, voltado aos aspectos de Engenharia de *Software*, complementa o referencial teórico anterior, que evidenciou o uso intensivo do Modelo das Quatro Etapas e a ocorrência de lacuna entre os pressupostos utilizados nos planejamentos e as atuais necessidades de informação. E também se propõe a discutir e analisar os principais componentes necessários ao desenvolvimento dos modelos propostos.

Alguns aspectos críticos são retratados diante da tecnologia atualmente disponível, dentre os quais se podem citar:

- O planejamento nacional de transporte de carga ainda é fundamentado em pesquisa de campo realizada nas principais rodovias, o que evidencia a necessidade de se desenvolverem sistemas capazes de agregar informações a baixo custo para aquisição e tratamento de dados. Caso análogo é retratado pelo relatório do *Transportation Research Board* (TRB, 2013) para os Estados Unidos da América, que menciona em *Background*:

‘Integrar aspectos relacionados a movimentação de cargas em processos de planejamento de transporte tornou-se cada vez mais difícil no momento em que essas influências apresentam-se de modo crítico para a nossa capacidade de previsão a longo prazo e para definição das necessidades de infra-estrutura de futuras’. (tradução livre)

- Para o transporte de passageiros, os sistemas de aquisição de dados do tráfego encontram-se em estágio avançado, porém, demandam a instalação de infraestrutura específica. Destaca-se que os sistemas baseados em tecnologia *wi-fi* e *bluetooth*, menos custosos, não são capazes de coletar os dados da movimentação de todos os veículos, o que limita seu uso para a alocação dinâmica do tráfego;
- O desenvolvimento de novas tecnologias de informação e comunicação possibilitou a implantação de inúmeros projetos para identificação da movimentação de veículos e pessoas a partir de aspectos que consideram o

contexto onde ocorre o movimento. Algumas aplicações para a área logística já são documentadas pela literatura especializada.

A seguir são formuladas algumas questões que a pesquisa se propõe a responder:

- Existem sistemas nacionais capazes de gerar automaticamente dados da movimentação de veículos e cargas?
- Como consolidar estes dados para a criação de um planejamento do transporte permanentemente atualizado e com uma maior diversidade de informações?
- Esses sistemas são capazes de tratar os dados com diversos níveis de agregação de forma a alimentar os sistemas ITS, atuais e futuros?

O método para atender aos aspectos da questão da pesquisa deve abordar três tecnologias da informação distintas e complementares em que a informação é coletada pelas antenas dos sistemas nacionais com tecnologia RFID e tratada por *software* específico:

- Abordagem Sistema de Informação – Os relatórios utilizam técnicas estatísticas descritivas;
- Abordagem Descoberta de Conhecimento em Bases de dados – O modelo objetiva a descoberta de novas propriedades;
- A abordagem que considera o contexto em que se insere o veículo ou a carga é tratada em sua forma conceitual, usando os paradigmas da ontologia.

## 4.1 Levantamento das Bases de Dados Nacionais

Em resposta à primeira pergunta, foi realizado o levantamento de sistemas de informação de alcance nacionais que pudessem estabelecer uma base de dados consistente a partir de contato com técnicos das seguintes empresas:

- SERPRO – Serviço Federal de Processamento de Dados;
- DETRAN/RJ – Departamento de Transito do Rio de Janeiro;
- Centro de Pesquisas Avançadas Werner Von Braun;
- Encat – Encontro Nacional de Coordenadores e Administradores Tributários.

### 4.1.1 Premissas básicas

Para atingir aos objetivos, foram adotadas as seguintes premissas básicas:

- Os sistemas pesquisados devem abranger todo o território nacional, ser aplicável aos vários modais de transporte de passageiros e de carga além ter atualização permanente;
- Os dados coletados devem atender aos propósitos de planejamento de curto, médio e longo prazo.

O levantamento dos sistemas, atuais e previstos, com essas características resultou na identificação de três sistemas:

#### 4.1.2 Sistema relacionado à Identificação Automática de Veículos

- Projeto SINIAV – Sistema Nacional de Identificação Automática de Veículos

O primeiro projeto analisado tem por objetivo a Identificação Automática de Veículos (IAV) em território nacional. O SINIAV foi criado pela Resolução 212 do Conselho Nacional de Trânsito – CONTRAN de 13 de novembro de 2006 e modificado pela Portaria 85 de 5 de novembro de 2009, quando foram estabelecidos prazos e normas técnicas a serem atendidos para sua implantação em todo o País. Posteriormente, considerando a necessidade da participação de todos os órgãos do Sistema Nacional de Trânsito, a necessidade de prévia homologação dos equipamentos que irão operar no SINIAV e também da adequação dos sistemas informatizados do DENATRAN, em 19 de Dezembro de 2012, foi publicada a Deliberação Nº 131, que altera o Cronograma de implantação do SINIAV, modificando o início em todo território Nacional, a partir de 01 de janeiro de 2013 e sua conclusão até o dia 30 de junho de 2015.

A tecnologia a ser utilizada no projeto SINIAV foi desenvolvida pelo Centro de Pesquisas Avançadas Wernher von Braun em Campinas SP (CONTRAN, 2006), por delegação do Ministério das Cidades. Esse sistema faz uso da tecnologia por rádiofrequência RFID de 915 MHz e é composto por uma *tag* ou *transponder* contendo dados do veículo gravados, uma antena e um sistema associado capaz de ler e validar os dados recebidos. Sobre essa tecnologia, destacam-se os seguintes aspectos:

Existem algumas frequências possíveis para se implementar um sistema de IAV – Identificação Automática de Veículos. Dentre elas, as mais conhecidas são as de 128KHz, 14MHz, 433MHz, 915MHz, 2.4GHz, 5.8GHz e 5.9GHz. Aquelas de

frequência mais alta (5.8GHz e 5.9GHz) não permitem que sejam desenvolvidas soluções que operem sem bateria e que sejam realizadas em tamanho miniaturizado. As de frequência mais baixa (128KHz, 14MHz) não têm grande alcance (necessário para acompanhar veículos em pistas por exemplo). A frequência de 915MHz, que tem as capacidades esperadas para longo alcance, permite o desenvolvimento de soluções passivas (sem-bateria), seu tamanho é reduzido e comporta mecanismos digitais de encriptação de dados aceitos mundialmente pelos organismos internacionais.

As especificações definidas para o SINIAV foram testadas e aceitas pela comissão ISO como aquelas adequadas à segurança de RFID no âmbito global e em vários tipos de cadeias de produto e transporte (ISO18000-6C / ISO 29167) e suportadas pela organização EPC Global (*Electronic Product Code Global*). Adicionalmente, o NIST *National Institute of Standards and Technology* (em português: Instituto Nacional de Padrões e Tecnologia), organismo norte-americano de padronização e medidas, aferiu que o mecanismo de encriptação do SINIAV (AES128) é passível de verificação por todos os organismos internacionais e sem previsão de quebra (pelo menos até o ano de 2030). Acrescente-se que essas normativas ISO são abertas (sem propriedades intelectuais associadas) e geridas por uma comissão internacional da qual o Brasil é signatário e tem parte ativa na especificação.

Comparativamente, as soluções propostas pela União Europeia estão vinculadas ao aspecto de fazer adensar os veículos com comunicação veículo-veículos e veículo-infraestrutura, necessitando uma maior capacidade de processamento embarcada e de custo mais elevado. No caso Brasileiro, trata-se de identificação veicular de forma inequívoca, segura com a probabilidade de fraude, de identidade a menor possível, eficiente e de menor valor absoluto de instalação e que possa ser embarcada em qualquer tipo de veículo.

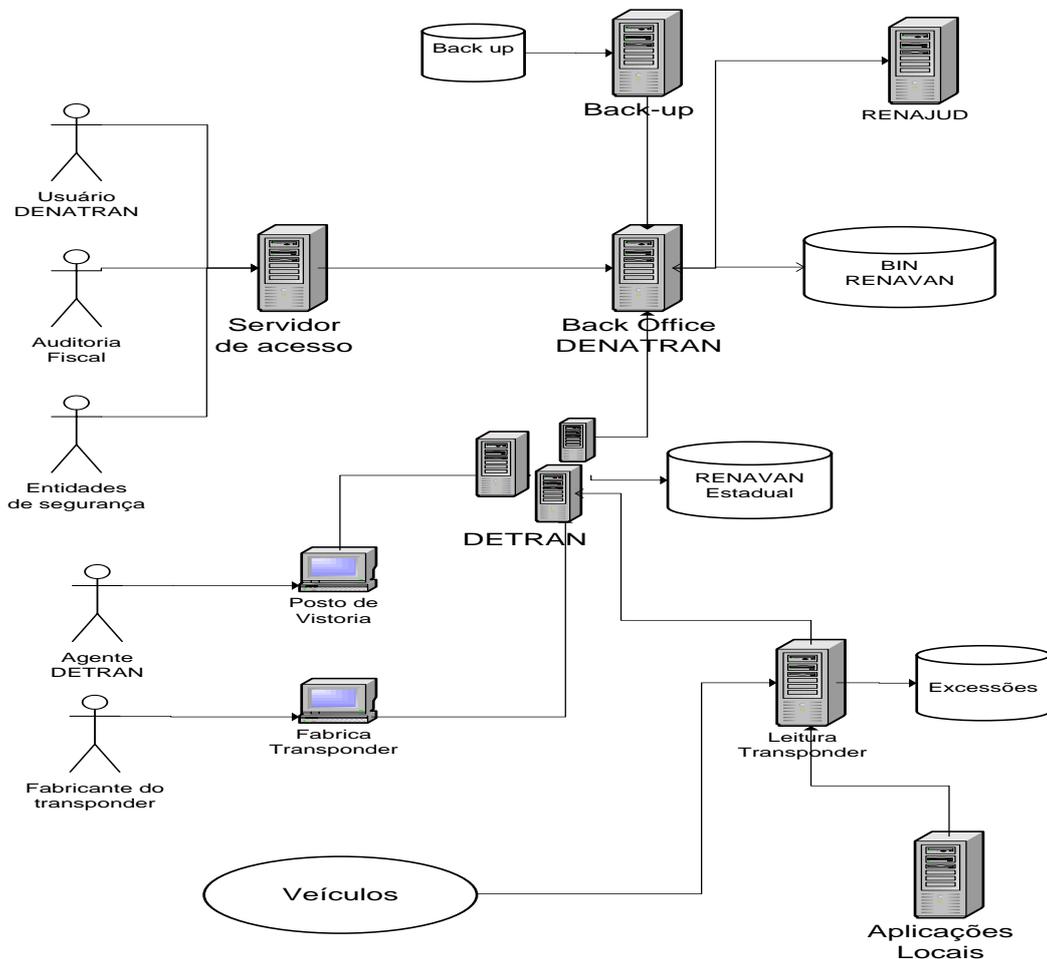
O custo dessa implementação do protocolo e funcionalidades especificadas para o sistema é estimado com valor aproximado a R\$ 10,00 (dez reais) para milhares de unidades, tendendo a R\$ 6,00 para volumes maiores a preços de 2010.

Os objetivos do SINIAV que englobam o acompanhamento do Ciclo de Vida do Veículo compreendem:

- Fiscalização Urbana e Rodoviária;
- Gestão de Trânsito;

- Recuperação de Veículos;
- Gestão de Meios de Pagamento e Seguro de Veículos;
- Gestão do Transporte de Cargas e Logística.

A visão geral de sua arquitetura está representada na Figura 4.1



**Figura 4.1:** Visão geral do SINIAV (CONTRAN, 2006)

Essa organização básica não é um projeto de sistema, tampouco um conceito de projeto. Sua finalidade é estabelecer as várias abordagens que podem ser desenvolvidas, cada uma adaptada para atender às necessidades específicas do órgão, mantendo os benefícios de uma arquitetura comum. O uso desses dados adequadamente tratados possibilitará identificar os padrões de movimentação dos veículos e conseqüentemente a matriz OD e a alocação na malha viária.

O mapa de utilização da memória da *tag* do SINIAV é mostrado no Quadro 4.1.

**Quadro 4.1** Mapa de memória da *tag* do SINIAV (CONTRAN, 2006)

Aplicação	Dado	Memória em bits
BASE/FABRICANTE	Número serial único	64
	Controle de manufatura	32
	Memória programável	928
	Total(mínimo)	1024
PLACA ELETRÔNICA	Identificação do emissor (país, estado)	64
	Número agente	32
	Data/hora aplicação	16
	Placa	88
	Número chassi	128
	RENAVAN	36
	Marca e modelo do veículo	16
	Aplicações governamentais	164
	Sub-total	544
CONTROLE DO VEÍCULO	Bloco 1 a 6 para uso pela iniciativa privada	64 cada bloco

O registro de passagem do veículo pela antena deverá gerar um registro computacional contendo informações criptografadas relativas ao número serial único, à data e hora e à identificação do equipamento que realizou a leitura. A partir da consolidação desses dados com as bases de dados do DENATRAN, será possível identificar o veículo e as demais informações correlatas de sua passagem.

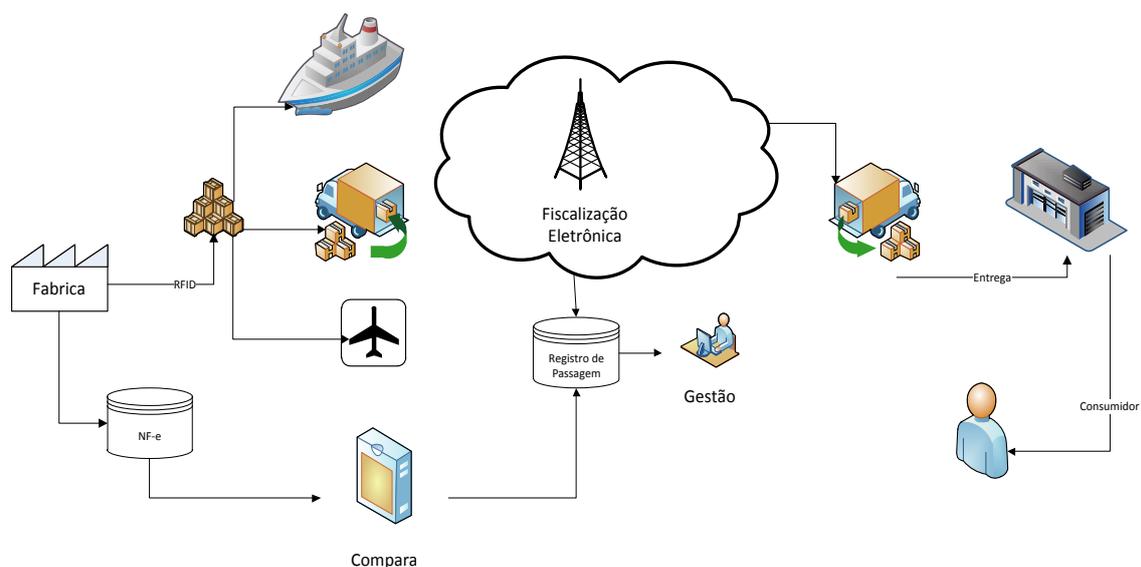
Analisando-se a estrutura de dados do mapa de memória do SINIAV e o registro de passagem coletado pelas antenas, pode-se observar que as mesmas possuem uma estrutura que permite identificar, de forma inequívoca, o veículo e sua localização e, ainda se existe a disponibilidade de memória para gravação de informações que possibilite o rastreamento do veículo e da carga.

### 4.1.3 Sistemas relacionados ao transporte de carga

- Projeto Brasil-ID

A partir da celebração de Acordo de Cooperação Técnica firmado em 31 de agosto de 2009 entre o Ministério da Ciência e Tecnologia, a Receita Federal e os Estados da União – por intermédio de suas Secretarias de Fazenda – formalizou-se o início do Sistema de Identificação, Rastreamento e Autenticação de Mercadorias (Fonte: <http://www.brasil-id.org.br>), que se baseia no emprego da tecnologia de Identificação por Radiofrequência (RFID), e outras acessórias integradas para realizar, dentro de um padrão único, a Identificação, Rastreamento e Autenticação de mercadorias em produção e circulação pelo País – no contexto de extensão da Nota Fiscal Eletrônica. O projeto é coordenado pelo Centro de Pesquisas Avançadas Wernher von Braun em parceria com o Encontro Nacional dos Administradores Tributários (ENCAT).

Segundo divulgação do Centro von Braun (2010), os Conceitos Básicos podem ser visualizados na Figura 3.2:



**Figura 4.2:** Funcionalidades Básicas do Brasil-ID (Adaptado de VON BRAUN, 2010)

O modelo apresentado na Figura 4.2 indica que os produtos em circulação no país passarão a ter uma etiqueta eletrônica que possibilitará ser rastreada por toda vida útil.

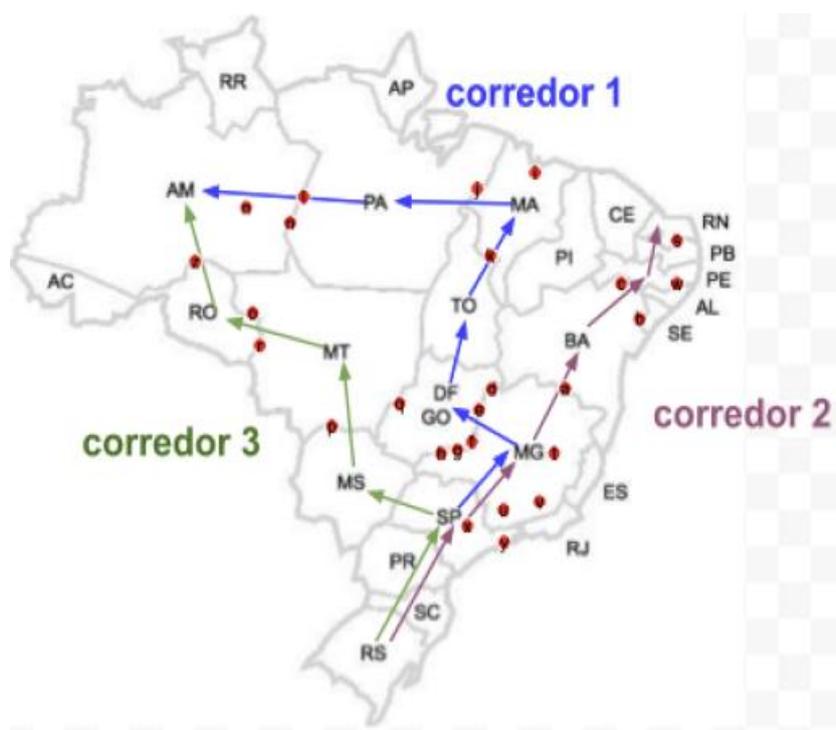
A organização do sistema apresentado mostra uma complexa interação entre os diversos atores logísticos a partir da instalação de uma infraestrutura de dados com

gestão nacional de leitura e gravação RFID. A finalidade precípua do sistema é a de controle, de forma que uma mercadoria fabricada no País é identificada eletronicamente na fábrica e monitorada até o consumidor final.

A aferição do modelo, desenvolvido e implantado em 2011, em Projeto Piloto, compõe-se de:

- 54 portais de passagem em Postos Fiscais;
- 40 Coletores de Inspeção Móveis;
- 604 Antenas RFID;
- 5000 *transponders* com tecnologia do SINIAV;
- 1 Portal de Identificação Completo;
- 80 empresas para implementação.

Nesse portfólio, os corredores logísticos analisados estão apresentados na Figura 3.3.



**Figura 4.3:** Principais Corredores do Brasil-ID, Projeto Piloto (VON BRAUM, 2010)

Os resultados obtidos pela leitura das *tags* pelas antenas posicionadas nos principais eixos rodoviários do país, de acordo com a avaliação do Centro Von Braun e Encat, foram satisfatórios.

Analisando-se o projeto Brasil-ID, pode-se concluir que sua estrutura possibilita rastrear a carga desde sua origem (fábrica ou porto de entrada) até seu destino final, passando pelos diversos modais utilizados. Esse sistema, na forma como concebido, já possui a tecnologia da informação desenvolvida e está em condições de implementação em um horizonte próximo. Trata-se, portanto, de uma tecnologia viável para os propósitos desta tese.

- Projeto CT-e

O terceiro projeto analisado nesta tese é o Conhecimento do Transporte Eletrônico CT-e. Esse documento foi criado pelo Conselho Nacional de Política Fazendária (CONFAZ) com a emissão do Ajuste do Sistema Nacional Integrado de Informações Econômico-Fiscais, SINIEF N° 09, 25 de outubro de 2007 (CONFAZ, 2007), utilizado pelos contribuintes do Imposto sobre Operações Relativas à Circulação de Mercadorias e sobre a Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação – ICMS. Esse diploma legal define que: “Considera-se Conhecimento de Transporte Eletrônico – CT-e o documento emitido e armazenado eletronicamente, de existência apenas digital, com o intuito de documentar prestações de serviço de transporte de cargas” e é aplicável ao transporte rodoviário, ferroviário, aquaviário e aéreo. Partes da estrutura dos dados gerados pelo sistema CT-e que são de interesse ao estudo para o transporte rodoviário são mostrados no Quadro 4.2

**Quadro 4.2** Estrutura dos dados do CT-e (parcial) Fonte (CT-e)

(continua)

Número	Campo	Descrição	Observações
Identificação do transporte			
13	dhMi	Data e hora da emissão	
25	modal	01- Rodoviário 02- Aéreo 03- Aquaviário 04- Ferroviário 05- Dutoviário	
27	cMunini	Código do município de início da prestação	Usar tabela IBGE 99999 para exterior
71	dProg	Data programada para entrega	

Quadro 4.2 Estrutura dos dados do CT-e (parcial) Fonte (CT-e)

(conclusão)

Dados do produto			
270	vMerc	Valor total das mercadorias	
271	prodPred	Produto predominante	
272	xOutCat	Outras características da carga	Fria, granel etc
274	cUnid	Código da unidade de medida	Kg, litro etc
276	qCarga	Quantidade	
Dados do veículo			
334	Placa	Placa do veículo	
340	toRod	Tipo de Rodado: 00 - Não aplicável 01 - Truck; 02 - Toco; 03 - Cavallo Mecânico; 04 - Van; 05 - Utilitário; 06 - Outros.	
341	tpCar	Tipo de Carroceria: 00 - Não aplicável 01 - Aberta; 02 - Fechada/Baú; 03 - Graneleira; 04 - Porta Contêiner; 05 - Sider	

A estrutura dos dados do CT-e mostra a diversidade de informações com possibilidade de identificar todos os parâmetros do transporte de carga de sua origem, destino, veículo de transporte, mercadoria transportada, entre outros.

O Conhecimento do Transporte é um sistema antigo que atualmente passa para sua versão eletrônica, segundo aspectos operacionais da Nota Fiscal Eletrônica – NF-e. Os principais estados produtores já o adotam em sua versão eletrônica.

Analisando-se a estrutura de dados do sistema, pode-se avaliar que existe a possibilidade de se identificar inúmeros aspectos relativos ao transporte de carga.

A seguir serão apresentados os diversos componentes tecnológicos a serem utilizados no desenvolvimento das diversas etapas previstas no Modelo e não retratadas por ocasião da elaboração do referencial teórico.

## 4.2 Considerações sobre a integração dos sistemas:

Analisando-se a possibilidade de integração de dados dos sistemas analisados, pode-se concluir que:

- Os dados que possibilitam correlacionar carga com o veículo estão contidos no componente do Brasil-ID que replicam os dados da NF-e e do CT-e, o que o permite considerar como elemento chave nesse processo;
- A origem e o destino da viagem são estabelecidos no CT-e;
- A data da viagem é estabelecida no CT-e;
- A característica da carga é estabelecida no CT-e;
- O início da viagem pode ser determinada pela primeira passagem do veículo pela antena em data próxima à data prevista e registrada pelo SINIAV;
- O término da viagem pode ser determinado pela última passagem próxima ao destino previsto;
- Uma viagem é considerada como uma série de leituras pelas antenas, que correspondem ao tempo em que o veículo está em movimento.

## 4.3 Modelagem de Domínio para determinação da movimentação do veículo ou da carga

A arquitetura básica dos sistemas SINIAV e Brasil-ID considera a coleta de dados dos veículos e das cargas quando de sua passagem pelas antenas. Para determinação dos padrões frequentes de movimentação, torna-se necessário reorganizar os dados por veículo a partir da “Modelagem de Domínio”, que serão utilizados no desenvolvimento das abordagens do Modelo.

A Modelagem de Domínio apresenta, de forma resumida, os elementos e suas relações na identificação dos padrões de deslocamento. Nesta tese, dois objetos de domínio se destacam como principais entidades do sistema: as antenas e as estradas.

As antenas são capazes de registrar a identificação do veículo ou da carga, o instante de passagem designado como *time stamp* e a localização da antena. Para que seja possível analisar os padrões de movimentação dos veículos, é necessário reorganizar os dados e associá-los à estrada.

A modelagem de negócio utilizada para interligar esses dois conceitos foi estabelecer que uma estrada é um grafo formado por uma sequência de portais, isto é, uma estrada é um conjunto de portais interligados. Cada par de portais interligados, ou cada aresta desse grafo, recebe o nome de Segmento. O grafo de uma estrada é um dígrafo linear.

Na Matemática e na Computação, um grafo é uma estrutura de dados composta por um conjunto de pontos – os vértices – ligados por um conjunto de retas – as arestas. Os vértices de um grafo possuem um atributo chamado “grau”. O grau de um vértice é o número de outros vértices ligados a ele. Na Teoria dos Grafos, ramo da matemática dedicado a estudar essa estrutura, se as arestas possuem orientação, isto é, se existe uma ordem para os vértices, esse grafo recebe o nome de dígrafo e as arestas passam a ser representadas por setas. Uma definição do grafo de movimento – um movimento  $G(V,E)$  – é um grafo que representa o movimento do objeto, onde  $V$  é o conjunto de localizações,  $E$  é o conjunto de transições entre locais. Um  $e(I,j)$  de borda indica que, em objetos movidos do local  $v_i$  para o local  $v_j$ , cada aresta é anotada com a história do movimento do objeto ao longo da borda, sendo cada entrada na história uma coleção de itens do formulário

Na Figura 4.4, é exemplificada a representação por Modelo de Grafo de uma malha rodoviária com alguns pontos de coleta, onde  $P_{n=7}$  são os identificadores dos Pontos de Coleta, A,B,C,D,E,F,G,H,I são as vias que ligam os pontos e  $T_1$  e  $T_2$  são os tempos Inicial e Final, respectivamente, de passagem por um trecho.





**Figura 4.5:** Exemplo de Dígrafo Linear

Uma estrada é representada por essa estrutura, onde cada vértice é um portal e cada aresta é um segmento. Um segmento possui algumas propriedades para melhor se adequar ao conceito que representa. Cada segmento tem uma designação, que corresponde ao nome do trecho da estrada compreendido entre seus portais e possui a informação da quilometragem inicial e final. Por exemplo, pode-se dizer que o segmento hipotético “Itaguaí – Mangaratiba” representa o trecho entre os quilômetros 150 e 220 da estrada “BR-101”. Para que os dados organizados por antena possam ser traduzidos em viagens realizadas por veículo, procede-se segundo o seguinte algoritmo: Determina-se o número de veículos que passam por uma estrada para poder analisá-lo a partir de diversas óticas que se calcula, entre dois portais de um segmento, o percentual de veículos que passaram pelos portais de saída e de entrada. Por exemplo, seja  $V1$  o número de veículos que passaram pelo primeiro portal do segmento  $S$  e  $V2$  o número de veículos que passaram pelo portal final desse mesmo segmento. Em condições reais, isto é, aplicando essa modelagem em uma situação em que existam portais medindo a passagem de veículos nas estradas do país, essa informação seria facilmente obtida pela comparação entre  $V1$  e  $V2$ . No caso de  $V1 > V2$ , isso significa que há uma ou mais vias de saída de veículos adjacentes à estrada nesse segmento  $S$ . Se  $V1 < V2$ , isso também significa que existe uma ou mais vias adjacentes e que nestas predomina a entrada de veículos nessa estrada.

Considerando que o objetivo desta tese é demonstrar a viabilidade do uso de diversas abordagens para integração de sistemas baseados em RFID como suporte ao planejamento de transportes, analisaram-se três diferentes tecnologias aplicáveis:

#### 4.4 Tecnologias da Informação da Abordagem Sistema de Informações

O desenvolvimento de Sistemas de Apoio à Decisão com acesso a bancos de dados e emissão de relatórios a partir de técnicas estatísticas exploratórias é a primeira tecnologia analisada e de uso mais frequente. A documentação relativa a modelagem dos softwares encontra-se disponível em

<https://skydrive.live.com/redir?resid=F27A4F06F5CFE1C9!432>

O gerenciamento do desenvolvimento do aplicativo foi realizado com as técnicas previstas no Processo RUP – *Rational Unified Process* (ou Processo Unificado Rational) (IBM, [N.d.]), que organiza o desenvolvimento de aplicativos computacionais em quatro fases, nas quais são tratadas questões sobre planejamento, levantamento de requisitos, análise, implementação, teste e implantação do aplicativo, modificado para uma pequena equipe de desenvolvimento.

**Fase de Concepção / Iniciação:** Fase do RUP que abrange as tarefas de comunicação com o cliente e o planejamento. Tendo consentimento das partes interessadas, é feita a definição do escopo do projeto e são examinados os objetivos para se decidir sobre a continuidade do desenvolvimento.

**Fase de Elaboração:** Abrange a modelagem do modelo genérico do processo que objetiva analisar de forma mais detalhada o domínio do problema. Revisando os riscos que o projeto pode sofrer, sua arquitetura começa a ter sua forma básica.

**Fase de Construção:** Desenvolve ou adquire os componentes de *software* onde o principal objetivo é a construção do sistema, com foco no desenvolvimento de componentes e outros recursos. É na fase de construção que ocorre a maior parte de codificação.

**Fase de Transição:** Abrange a entrega do *software* ao usuário e a execução de testes. O objetivo é disponibilizar o sistema, tornando-o disponível e compreendido pelo usuário final. As atividades dessa fase incluem o treinamento dos usuários finais e também a realização de testes da versão beta, visando a garantir que o mesmo possua o nível adequado de qualidade.

A passagem pelas quatro fases corresponde ao ciclo de desenvolvimento. Cada passagem produz uma geração do *software*. A menos que o produto "desapareça", ele irá se desenvolver na próxima geração, repetindo a mesma sequência de fases de iniciação, elaboração, construção e transição, mas agora com ênfase diferente nas diversas fases. Esses ciclos subsequentes são chamados de ciclos de evolução. À medida que o produto atravessa vários ciclos, são produzidas novas gerações.

A notação utilizada para modelar os artefatos da fase de análise e projeto é a UML (*Unified Modeling Language*). A UML ou Linguagem Unificada de Modelagem é uma linguagem gráfica para visualização, especificação, construção e documentação de artefatos de sistemas complexos de *software*. A UML proporciona uma forma-padrão

para a preparação de planos de arquitetura de projetos de sistemas, incluindo aspectos conceituais, tais como os processos de negócios e funções do sistema, além de itens concretos como as classes escritas em determinada linguagem de programação, esquemas de banco de dados e componentes de *software* reutilizáveis (Pressman, 2011).

Para atendimento aos requisitos de disponibilização de informações por todos os órgãos de transporte nacionais, é considerado que a aplicação deve seguir o paradigma RIA (*Rich Internet Application*) (Adobe, [N.d.]), que visa a juntar a usabilidade das aplicações *Desktop* com a flexibilidade de uma aplicação *Web*.

As linguagens consideradas para serem utilizadas são o Adobe Flex para a subaplicação cliente (Apresentação) e o Java para a subaplicação servidora (Controle), o acesso aos dados (Modelo) pela linguagem HQL (*Hibernate Query Language*). HQL é uma linguagem orientada a objeto de consulta, semelhante ao SQL (*Structured Query Language*), porém, em vez de operar em tabelas e colunas, trabalha com objetos persistentes e suas propriedades. Consultas HQL são traduzidas pelo *Hibernate* em consultas SQL convencionais, que por sua vez realizam ação no banco de dados. A aplicação poderá usar um misto das arquiteturas Cliente-Servidor e Três Camadas.

## 4.5 Tecnologias da Informação da Abordagem Mineração de Dados

### 4.5.1 Considerações Preliminares

O objetivo dessa abordagem é desenvolver um modelo teórico de identificação automática de padrões frequentes de movimentação de veículos baseados na integração de dados do Sistema Nacional de Identificação Automática de Veículos – SINIAV com uso de técnicas de descoberta de conhecimento em bases de dados – KDD (do Inglês *Knowledge Discovery in Databases*). Serão abordadas apenas as duas primeiras etapas do processo KDD: Pré-processamento e mineração de dados (*Datamining*).

O termo KDD foi formalizado em 1989 em referência ao amplo conceito de procurar conhecimento a partir de bases de dados. Uma das definições mais populares foi proposta em 1996 por um grupo de pesquisadores, segundo Fayyad *et al.* (1996): “KDD é um processo, de várias etapas, não trivial, interativo e iterativo, para

identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grandes conjuntos de dados”.

Na definição formal de KDD, o termo interativo indica a necessidade de atuação do homem como responsável pelo controle do processo. O homem utiliza os recursos computacionais disponíveis em função da análise e da interpretação dos fatos observados e dos resultados obtidos ao longo do processo de descoberta do conhecimento.

Segundo Han e Kamber (2001), um conceito muito difundido e equivocado sobre mineração de dados é o que define os sistemas de mineração de dados como sistemas que podem automaticamente minerar todos os conceitos valiosos que estão escondidos em um grande banco de dados sem intervenção ou direcionamento humano.

Considerando ainda a definição de KDD, um padrão deve ser interpretado como um conhecimento representado segundo as normas sintáticas de alguma linguagem formal, segundo Fayyad *et al.* (1996). Um padrão compreensível refere-se, portanto, a um padrão caracterizado pela representação do conhecimento que possa ser interpretada pelo homem. A expressão padrão válido indica que o conhecimento deve ser verdadeiro e adequado ao contexto da aplicação de KDD. Um padrão novo deve acrescentar novos conhecimentos àqueles existentes no contexto da aplicação de KDD.

Além da definição de Fayyad *et al.* (1996), diversas outras definições de KDD podem ser encontradas. Em semelhança aos autores, outras definições de KDD podem ser encontradas:

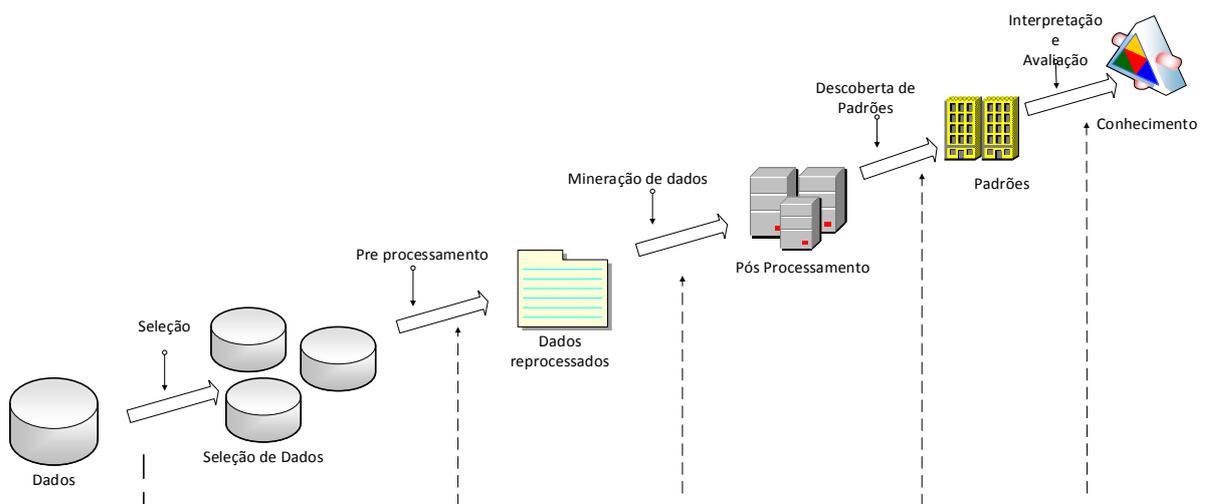
- É a busca de informações valiosas em grandes bancos de dados. É um esforço de cooperação entre homens e computadores. Os homens projetam bancos de dados, descrevem problemas e definem seus objetivos; os computadores verificam dados e procuram padrões que casem com as metas estabelecidas pelos homens (WEISS & INDURKHYA, 1998);
- É a exploração e análise de dados, por meios automáticos ou semiautomáticos, em grandes quantidades de dados, com o objetivo de descobrir regras ou padrões interessantes (BERRY & LINOFF, 1997);
- Em poucas palavras, é a análise de dados indutiva (MENA, 1999);

- É o processo de proposição de várias consultas e extração de informações úteis, padrões e tendências, frequentemente desconhecidos, a partir de grande quantidade de dados armazenada em bancos de dados (THURASINGHAM, 1999).

A importância da análise dos dados do deslocamento de veículos e mercadorias é amplamente reconhecida (GIDOFALVI e PEDERSEN, 2009; GU e XIAO-LI, 2009; HUIPING *et al.*, 2007). Dentre os modelos para rastreabilidade de objetos, destaca-se o trabalho de Lee *et al.* (2011), que apresenta novos conceitos sobre a classificação de padrões de deslocamento em rodovias.

#### 4.5.2 Etapas do Processo de KDD

O KDD é caracterizado como um processo composto por várias etapas operacionais. A Figura. 3.7 apresenta um resumo das etapas operacionais executadas em processos de KDD (FAYYAD *et al.*, 1996).



**Figura. 4.6** Etapas Operacionais do Processo de KDD

Fonte: Adaptado de Fayyad et al. (1996)

A Figura 4.6 mostra que, iniciando um processo de KDD, a primeira etapa é um agrupamento de forma organizada dos dados (seleção). A etapa da limpeza dos dados vem a seguir, através de um pré-processamento dos dados, visando a adequá-los aos algoritmos que serão utilizados. Para facilitar o uso das técnicas de mineração de dados, os dados ainda podem passar por uma transformação que os armazena adequadamente em arquivos para serem lidos pelos algoritmos. É a partir desse momento que se chega à fase de mineração de dados especificamente, que começa com a escolha das ferramentas

(algoritmos) a serem utilizadas. Essa escolha depende fundamentalmente do objetivo do processo de KDD: classificação, agrupamento, regras associativas, ou desvio. De acordo com o algoritmo utilizado, será gerado um arquivo descritivo (relatório ou gráfico) que deve ser interpretado para formular conclusões sobre o conhecimento na base de dados.

#### 4.5.2.1 Seleção de Dados

Essa atividade consiste na identificação de quais informações dentre as bases de dados existentes devem ser efetivamente consideradas durante o processo de KDD, (GOLDSMITH e PASSOS, 2005). Em geral, os dados encontram-se organizados em bases de dados transacionais que sofrem constantes atualizações ao longo do tempo.

Com esses dados e com a detecção dos padrões frequentes de deslocamento, pode-se prever a trajetória dos veículos nas vias. Essa previsão tem muitas aplicações úteis, tais como planejamento urbano e transporte, construção de estradas, projeto, manutenção, gerenciamento de tráfego, aplicação da lei e segurança nacional entre outros.

Nesse paradigma, com os recursos extraídos, cada ponto de dados é transformado em um vetor de características a partir do qual são fornecidos elementos para um classificador voltado ao treinamento e à previsão (LEE *et al.*, 2011).

Um dos requisitos mais importantes para uma efetiva previsão é descobrir características discriminativas para prever a trajetória de um objeto. Recentemente, a abordagem usando padrões frequentes (*frequent patterns*) para a classificação e previsão foi adotada com sucesso para dados relacionais (CHENG *et al.*, 2011).

Dois tratamentos são aplicáveis aos dados na análise das rotas utilizadas por objetos em movimento: uma visão espaço-temporal e uma visão puramente espacial (KORNER *et al.*, 2012).

#### 4.5.2.2 Pré-processamento

Essa etapa compreende as funções relacionadas à captação, à organização, ao tratamento e à preparação dos dados para a etapa da mineração com a correção dos dados e o ajuste de sua formatação para o uso dos algoritmos.

- Limpeza

A fidedignidade dos dados tem grande influência na eficiência dos modelos de conhecimento selecionados.

A fase de limpeza dos dados envolve uma verificação da consistência das informações, a correção de possíveis erros e o preenchimento ou a eliminação de valores desconhecidos e redundantes, além daqueles não pertencentes ao domínio (GOLDSCHMIDT e PASSOS , 2005).

- Classificação

A etapa de classificação de dados é a atividade de pré-processamento que se concentra na apresentação dos dados durante o processo de KDD. Trata-se de uma atividade criativa que deve ser realizada repetidas vezes em busca de melhores representações, segundo Goldschmidt e Passos (2005).

A escolha dos atributos a utilizar no modelo considera a sua relevância para previsão, segundo os padrões relevantes para a análise.

- Enriquecimento

Essa etapa consiste em conseguir agregar mais informações aos registros existentes, para que estes forneçam mais elementos para o processo de descoberta de conhecimento. A seguir, estão descritas algumas das operações mais usualmente utilizadas no processo de enriquecimento das bases de dados.

Nessa operação todas as iniciativas de enriquecimento que envolvem a captação de novas informações junto às fontes originais devem ser incluídas. Normalmente as equipes de pesquisa requerem a inclusão de novos atributos ou mesmo de novas tabelas nas bases de dados existentes. Diferem das operações de limpeza porque não estão restritas a preencher informações ausentes. Buscam-se, no caso do enriquecimento, agregar novas informações, o que, muitas das vezes, são inviáveis devido ao alto custo de implementação. O levantamento amostral pode considerar um único registro ou ser exaustiva. Caso os dados sejam transcritos de informações registradas em formulários, uma especial atenção é dada de maneira a facilitar a digitação e verificar a sua consistência.

Nas consultas a bases de dados externos, a atividade de enriquecimento pode ser realizada mediante a incorporação de informações fornecidas por outros sistemas. É muito comum a importação de informações advindas de outras bases de dados.

- **Construção de Atributos:**

Essa operação consiste em gerar novos atributos derivados a partir dos existentes. Como exemplo, a obtenção da idade a partir da data de nascimento e da data corrente do sistema. A importância desse tipo de operação é justificada porque novos atributos, além de expressarem relacionamentos conhecidos entre atributos existentes, podem reduzir o conjunto de dados, simplificando o processamento dos algoritmos de Mineração de Dados. A construção de atributos por meio da combinação daqueles já existentes pode incorporar ao problema informações de relacionamentos entre os dados, que sejam úteis ao processo de KDD. É conveniente enfatizar que é muito comum a substituição dos atributos existentes pelos respectivos atributos derivados.

- **Correção de prevalência:**

Essa operação é muitas vezes necessária em tarefas de classificação. Consiste em corrigir um eventual desequilíbrio na distribuição de registros com determinadas características.

### 4.5.2.3 Mineração de Dados

A mineração de dados é a fase que transforma dados fidedignos de um banco de dados a partir de algoritmos para extrair conhecimento implícito e útil.

Na fase de mineração de dados, dentro do processo de KDD, necessita-se definir a técnica e o algoritmo a ser utilizado em função da tarefa proposta. Uma vez escolhido, deve-se implementá-lo e adaptá-lo ao problema proposto para obter resultados que serão analisados na fase de interpretação e avaliação do resultado.

A mineração de dados vem sendo desenvolvida para os mais diferentes domínios, suas tarefas também vêm diversificando-se cada vez mais e permitem extrair diferentes tipos de conhecimento, sendo necessário decidir, no início do processo, qual o tipo de conhecimento que o algoritmo deve extrair.

Segundo Wu *et al.* (2008), existem dez algoritmos usados em mineração de dados dos quais, posteriormente Kumar e Verma (2012) selecionam seis principais,

como redes neurais, árvores de decisão, algoritmos genéticos, sistemas baseados em regras e programas estatísticos, que tanto isoladamente, quanto em combinação, podem ser aplicadas ao problema. Resumidamente, o processo de mineração de dados possui as seguintes abordagens:

- Probabilística: Essa família de técnicas de algoritmos usados em KDD utiliza modelos de representação gráfica para comparar diferentes representações do conhecimento. Esses modelos são baseados em probabilidades e independências de dados. Eles são úteis em aplicações que envolvam a incerteza e aplicações estruturadas de tal forma que a probabilidade pode ser atribuída a cada resultado ou com pouco conhecimento a ser descoberto. As técnicas probabilísticas podem ser utilizadas em sistemas de diagnóstico e nos sistemas de planejamento e controle;
- Estatística: A abordagem estatística utiliza a descoberta de regras e baseia-se em relações de dados. Um algoritmo de aprendizado indutivo pode selecionar automaticamente os caminhos de junção úteis e atributos para a construção de regras a partir de um banco de dados com muitas relações. Esse tipo de indução é usado para generalizar padrões nos dados e construir as regras dos padrões observados;
- Abordagem Clássica: É o mais antigo e mais utilizado de todos os métodos de KDD. A tarefa de classificação é uma função de aprendizado que mapeia dados de entrada em um número finito e pré-definido de classes. O objetivo de um algoritmo de classificação é encontrar alguma relação entre os atributos, de modo que o processo possa prever a classe de uma nova informação.

Dentre os tipos de técnicas de classificação, pode-se citar:

- (i) Abordagem Bayesiana: A Abordagem Bayesiana para KDD é um modelo gráfico que usa arcos orientados para formar um Grafo Acíclico Dirigido (DAG). Embora a abordagem Bayesiana utilize probabilidades e um meio de representação gráfica, é também considerado um tipo de classificação. Redes Bayesianas são tipicamente usadas quando a incerteza é associada com um resultado. Pode ser expresso em termos de uma probabilidade “*a priori*”;
- (ii) Descoberta de Padrões Frequentes: É outro tipo de classificação que reduz sistematicamente um grande banco de dados para alguns registros

pertinentes e informativos. Os dados não representativos são eliminados e a tarefa de descobrir padrões nos dados é simplificada;

- (iii) **Análise de Tendências:** A detecção de padrões de tendências por filtragem é a base para essa abordagem KDD. As técnicas de análise de tendências são normalmente aplicadas a bases de dados temporais. Uma boa aplicação para este tipo de KDD é a análise de tráfego em grandes redes de transporte;
- (iv) **Rede Neural Artificial (RNA):** É uma técnica que constrói um modelo matemático para um sistema neural biológico simplificado, com capacidade de aprendizado, generalização, associação e abstração. Assim como no cérebro humano, as redes neurais apresentam uma estrutura altamente paralelizada, composta por processadores simples (neurônios artificiais) conectados entre si. As RNAs tentam aprender por experiência, ou seja, diretamente dos dados, através de um processo de repetidas apresentações dos dados à rede;
- (v) **Algoritmos Genéticos:** Esses algoritmos referem-se aos sistemas evolutivos simulados, mais precisamente, ditam como, a população de um organismo deve ser gerada, avaliada e modificada;
- (vi) **Híbrida:** A abordagem híbrida para KDD combina mais de uma abordagem e é também chamada multiparadigmática. Embora a implementação possa ser mais difícil, ferramentas híbridas são capazes de combinar os pontos fortes de várias abordagens. Alguns dos métodos são utilizados para combinar técnicas de visualização, indução, redes neurais e sistemas baseados em regras para alcançar a descoberta de conhecimento desejado.
- **Pós-Processamento:** A última etapa do KDD, o Pós-Processamento, envolve a visualização, a análise e a interpretação do modelo de conhecimento gerado pela etapa de Mineração de Dados. Em geral, é nessa etapa que o especialista em KDD e o especialista no domínio da aplicação avaliam os resultados obtidos e definem novas alternativas de investigação dos dados.

A obtenção do conhecimento não é o passo final do processo de KDD. O conhecimento extraído pode ser utilizado na resolução de problemas do transporte, seja por meio de um Sistema Inteligente ou de um especialista como apoio a algum processo de tomada de decisão. Os padrões a serem descobertos pertencentes a cada transação denominam-se “padrões intratransação”.

É importante que algumas questões sejam respondidas aos usuários, segundo Liu e Hsu (1996):

- ✓ “O conhecimento extraído representa o conhecimento do especialista?”
- ✓ “De que maneira o conhecimento do especialista difere do conhecimento extraído?”
- ✓ “Em que parte o conhecimento do especialista está correto?”
- *Softwares* considerados para a mineração de dados.

Pesquisando-se na literatura os *softwares* disponíveis para mineração de dados, foram identificados os seguintes:

- ✓ Microsoft SQL Server, ferramenta originalmente de banco de dados que a cada nova versão apresenta novas funcionalidades de *Business Intelligence*. Possui diversos algoritmos e sua plataforma é extensível para integração de outros algoritmos desenvolvidos;
- ✓ IlliMine, Projeto de mineração de dados escrito em C++;
- ✓ InfoCodex, Aplicação de mineração de dados com uma base de dados linguística;
- ✓ KDB2000, uma ferramenta livre em C++ que integra acesso a bases de dados, pré-processamento, técnicas de transformação e um vasto escopo de algoritmos de mineração de dados;
- ✓ KXEN, Ferramenta de mineração de dados comercial, utiliza conceitos do Professor Vladimir Vapnik como Minimização de Risco Estruturada (*Structured Risk Minimization* ou SRM) e outros;
- ✓ KNIME, Plataforma de mineração de dados aberta que implementa o paradigma de *pipelining* de dados. Baseada no eclipse;
- ✓ LingPipe API em Java para mineração em textos distribuída com código-fonte;
- ✓ MDR Ferramenta livre em Java para detecção de interações entre atributos utilizando o método da *multifactor dimensionality reduction* (MDR);
- ✓ Orange Toolkit livre em Python para mineração de dados e aprendizado de máquina;
- ✓ Pimiento. um ambiente para mineração em textos baseado em Java;
- ✓ PolyaAnalyst, ambiente que permite a montagem de fluxos para mineração de dados e texto;
- ✓ Tanagra *Software* livre de mineração de dados e estatística;

- ✓ Weka *Software* livre em java para mineração de dados;
- ✓ *Cortex Intelligence*, Sistema de PLN para mineração de textos aplicado à Inteligência Competitiva

Para análise e seleção do *software* a ser usado nesta tese, consideraram-se os seguintes requisitos: (a) ser *software* livre; (b) ser de fácil entendimento e modelagem; e (c) existir a possibilidade de colaboração de especialistas em sua implantação. O *software* de Mineração de Dados selecionado foi o Weka (*Waikato Environment for Knowledge Analysis*), que atendeu aos requisitos estabelecidos. Esse *software* começou a ser escrito na Universidade de Waikato, Nova Zelândia, em 1993, usando a linguagem Java, sendo adquirido posteriormente por uma empresa no final de 2006. O Weka encontra-se licenciado ao abrigo da GPL – *General Public License* (FSF, S.d.), sendo, portanto, possível estudar e alterar o respectivo código fonte.

Weka é uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Os algoritmos podem ser aplicados diretamente a um conjunto de dados ou chamado a partir de seu próprio código Java. Weka contém ferramentas para pré-processamento de dados, classificação, regressão, agrupamento, regras de associação e visualização. Também é bem apropriado para o desenvolvimento de sistemas de aprendizagem de novas máquinas. Esses algoritmos e técnicas permitem ao computador um "aprendizado" (no sentido de obter novo conhecimento).

O Weka também procede à análise computacional e estatística dos dados fornecidos recorrendo a técnicas de mineração de dados, tentando, indutivamente, a partir dos padrões encontrados, gerar hipóteses de soluções.

Analisando-se a tecnologia para descoberta de conhecimento em bases de dados e o levantamento do estado da arte sobre o tema, pode-se concluir por sua aplicabilidade, notadamente para estabelecimento de padrões de deslocamento urbano. Os algoritmos disponíveis possibilitam identificar as principais variáveis integrantes do tráfego ou de um sistema de transporte e apresenta-se como uma ferramenta viável para os problemas do transporte.

#### 4.6 Modelagem ‘Representação do Conhecimento’

Um aspecto que se apresenta promissor para obtenção de novos elementos ligados ao planejamento do transporte e a logística, considerando a existência de uma

base de dados de movimentação de carga rodoviária heterogênea e confiável, é a conceituação e desenvolvimento de sistemas integrados sensíveis a contexto, usando as técnicas de modelagem da Ontologia.

O desenvolvimento de sistemas de informação sensíveis a contexto, focado em Sistemas Inteligentes de Transportes – ITS, geralmente representa um ambicioso projeto de Engenharia de Transportes. Um ponto relevante para o desenvolvimento de tais sistemas é a forma de apresentação do contexto.

As pesquisas em computação sensíveis a contexto têm relatado, em geral, a necessidade de se desenvolver infraestruturas de *software* e protótipos com o intuito de mostrar como esse novo paradigma pode beneficiar variados domínios de aplicação, como transporte, entretenimento e segurança (GU *et al.*, 2005).

Quanto à modelagem de informação de contexto, os modelos existentes, em geral, apresentam diferentes graus de expressividade, formalidade, uso de padrões, técnicas de modelagem, dimensão de informação de contexto e um domínio de aplicações.

Quanto maior a expressividade de um modelo, maior é a sua capacidade de representar a estrutura e a semântica dos conceitos manipulados por um sistema. Quanto mais formal for o modelo, maior é a capacidade de sistemas sensíveis a contexto realizarem inferências sobre as informações.

A partir do levantamento realizado sobre aspectos de modelagem de informação contextual, o trabalho elaborado por Bulcão (2006) “identifica a necessidade de modelos que apresentem avanços com respeito a sua expressividade, formalidade, uso de padrões de representação de informação, técnica de modelagem, dimensão semântica e escopo de aplicações-alvo”.

A capacidade de sistemas de se comunicar de forma transparente (ou o mais próximo disso) com outro promove a interoperabilidade entre múltiplas entidades ou empresas com fontes de conhecimento heterogêneas e torna-se um foco de investigação no domínio da gestão da cadeia de suprimentos (SCM) e do transporte em geral. Especificamente, a partilha de conhecimento em uma cadeia de suprimentos é crucial. No entanto, poucos estudos têm abordado o problema da interoperabilidade e compartilhamento de conhecimento em cadeias de suprimentos. Huang e Lin (2010) propõem uma solução para a partilha de conhecimentos com a *web* semântica que

envolve um modelo de conhecimento semiestruturado para representar o conhecimento, não só um processo de anotação baseada no agente explícito e compartilhável, mas também um formato significativo para resolver tanto problemas associados com a heterogeneidade dos documentos de conhecimento quanto um mecanismo de articulação para melhorar a eficácia da interoperabilidade entre as duas ontologias heterogêneas.

Paralelamente à identificação da necessidade de compartilhar conhecimento na cadeia de suprimento recentemente, vêm sendo direcionados esforços no sentido de aperfeiçoar os mecanismos de produção e controle com o uso de computação sensível a contexto (ZAMANI *et al.*, 2008).

A junção desses dois conceitos, computação (ou Sistemas de Informação) sensível a contexto e ontologia, apresenta-se de forma viável para conceituar um modelo que utilize dados coletados dos sistemas RFID. A seguir serão apresentados os conceitos relativos a Sistemas de Informação Sensível a Contexto e a Ontologia.

#### 4.6.1 Sistemas de Informação Sensíveis a Contexto

A computação ubíqua tem surgido nos últimos tempos como um paradigma no qual usuários interagem com ambientes instrumentados por meio de aplicações computacionais transparentes. A ideia básica é fornecer elementos ao usuário além daqueles considerados como “triviais”. As aplicações computacionais sensíveis a contexto são caracterizadas por se adaptar automaticamente às mudanças de ambiente sem a necessidade da atenção do usuário, ou ainda podem explorar características do meio ambiente, fornecendo informações ao usuário de proximidades, características relevantes, entre outras influenciadas por esse meio ambiente.

Dey (2001) define contexto como

“qualquer informação que pode ser usada para caracterizar a situação de uma entidade. Uma entidade é uma pessoa, lugar ou objeto considerado relevante para a interação entre o usuário e a aplicação incluindo o próprio usuário ou a aplicação”.

Bettini *et al.* (2010) apontam como principais técnicas para modelagem:

- Tuplas objeto-valor;
- *Markup Scheme Models*;

- Modelos Gráficos;
- Modelos Orientados a Objeto;
- Modelos Baseados em Lógica;
- Modelos Baseados em Ontologia.

Neste trabalho, o modelo proposto deriva da identificação da complexidade em se representar sistemas sensíveis a contexto.

#### 4.6.2 Ontologia – Contexto histórico

O planejamento, o projeto, a construção, a manutenção e a operação de infraestrutura de transporte requerem a aquisição e o tratamento de grandes volumes de dados. No entanto, o compartilhamento de dados e a aquisição em diferentes plataformas e aplicativos é um desafio, juntamente com questões institucionais, devido à heterogeneidade dos sistemas existentes em termos de conceitos, técnicas de codificação de dados e estruturas de armazenamento. A técnica de representação e modelagem *Unified Modeling Language* (UML) utilizada no desenvolvimento da abordagem do Sistema de Informações descrita anteriormente foi identificada como uma forma adequada para representação de dados de transporte.

O modelo de dados UML suporta o intercâmbio de dados de transporte nas dimensões espacial e temporal e possibilita gerar melhorias nesses dados pelo compartilhamento de dados. Além disso, como um modelo de dados orientado a objeto, ele aumenta tanto a eficiência de manutenção do banco de dados quanto o desempenho de aplicativos. Apesar dessas vantagens, o modelo de dados UML oferece suporte à interoperabilidade de dados somente no nível sintático e não pode resolver o problema de heterogeneidade semântica, que é inerente a diversas fontes de dados de transporte.

Sem uma padronização sobre a semântica de seus dados, as organizações e as respectivas aplicações computacionais de transporte tendem a utilizar diferentes terminologias para desenvolver seus próprios modelos de dados, o que possibilita inferir que o modelo de dados UML não pode ser apresentado como um quadro semanticamente coerente no domínio do transporte para a integração de fontes de dados distribuídos e heterogêneos. Uma abordagem possível para superar o problema da heterogeneidade semântica é por meio de ontologia (ZHANG *et al.*, 2008).

O termo “ontologia” vem do campo da Filosofia que se preocupa com o estudo do ser ou da existência. Na Filosofia, pode-se falar de uma ontologia como uma teoria sobre a natureza da existência (por exemplo, a ontologia de Aristóteles oferece categorias primitivas, tais como substância e qualidade, supondo que elas davam conta de tudo que é).

No caso da computação e da ciência da informação, ontologia é um termo técnico que denota um artefato que é projetado para uma finalidade, que é a de permitir a modelagem de conhecimento sobre algum domínio, real ou imaginário.

A expressão tem sido adotada desde o início por pesquisadores de Inteligência Artificial (IA) que reconheceram a sua aplicabilidade ao trabalho da Lógica Matemática e argumentaram que poderiam criar novas ontologias como modelos computacionais que permitiriam certos tipos de raciocínio automatizado.

### 4.6.3 Ontologia- Descrição

Ontologias são definidas como a especificação formal de um vocabulário de conceitos e axiomas relacionando-os (GUARINO, 1998).

As ontologias permitem-nos expressar as relações complexas dentro de uma área de domínio, estabelecendo relações entre conceitos que nos permitem compreender o conceito não apenas por suas propriedades, mas por sua presença em relação a outros conceitos dentro da ontologia. Dessa forma, a essência da ontologia é capturar as características naturais de realidades e relações entre realidades (GUARINO, 1998).

Considerando um domínio como uma área de estudo ou área de conhecimento específica, como transporte, medicina ou gestão financeira, uma ontologia define os termos utilizados para descrever e representar esse domínio. Nas ontologias, a semântica é concretizada por um conjunto de termos, relações entre termos e regras de inferência. Esse é o núcleo das ontologias. Os termos constituem um vocabulário controlado com definições explícitas. As relações entre termos incluem as relações entre as instâncias, entre as classes e suas instâncias e entre as classes. Enquanto as regras de inferência permitem definir conhecimento sobre um sistema que poderia estar oculto ou implícito, permitem também derivar novos conhecimentos a partir de fatos existentes (W3C, n/a).

Uma ontologia pode ser desenvolvida para vários fins, mas, de uma forma geral, ajuda as pessoas:

- A compreender melhor uma determinada área de conhecimento. No desenvolvimento de uma ontologia, as pessoas envolvidas no processo se veem diante de um desafio. Explica seu entendimento sobre o domínio em questão, o que as faz refletir e melhorar sua compreensão sobre esse domínio;
- A atingir um consenso no seu entendimento sobre uma área de conhecimento. Geralmente, em uma determinada área de conhecimento, diferentes especialistas têm entendimento diferenciado sobre os conceitos envolvidos, o que leva a problemas na comunicação. Ao se construir uma ontologia, essas diferenças são explicitadas e busca-se um consenso sobre seu significado e sua importância;
- A compreender certa área de conhecimento. Se houver uma ontologia sobre uma determinada área de conhecimento desenvolvida, uma pessoa que deseje aprender mais não precisa se reportar ao especialista, pois pode estudar a ontologia e aprender sobre o domínio em questão, absorvendo um conhecimento geral e de consenso.

Diferentes tipos de ontologias são construídos para diversas aplicações com níveis diferenciados de detalhes. As classificações das ontologias são resumidas em termos de generalidade (NOY e MCGUINNESS, 2001):

- Ontologias genéricas: são aquelas que descrevem conceitos gerais, independentes de qualquer domínio particular (Tempo, Processos);
- Ontologias de domínio: são aquelas que descrevem os conceitos de um domínio em particular (Transporte, Biologia, Física ou Direito Tributário);
- Ontologias de aplicação: são aquelas que descrevem os conceitos necessários para aplicações específicas e podem ser construídas sobre ontologias genéricas e de domínio (ontologias para auxiliar na identificação da movimentação de objetos – veículos e cargas).

Becker e Smith (1997), em um dos primeiros trabalhos para a área de transporte, definiram uma Ontologia para o planejamento do transporte militar multimodal. A

ontologia proposta aborda aspectos relativos aos serviços de transporte, às atividades, aos recursos (veículos, tripulações e recursos dos terminais) e às restrições que estabelecem como, quando, por quem e onde as atividades de transporte serão realizadas. Essa ontologia considera diferentes modais de transporte de veículos militares. O trabalho elaborado por Haller *et al.* (2008) propõe uma arquitetura baseada na *framework* “RosettaNet” com o objetivo de facilitar a comunicação entre as diversas aplicações computacionais na área logística. Outro trabalho na área de transporte baseado em ontologias de transportes públicos foi apresentado por Wang *et al.* (2005), que, com base nas informações do usuário (origem, destino e as prioridades), as pesquisas procuram definir paradas de ônibus usando um raio de pesquisa espacial. O algoritmo encontra viagens com base na origem e destino, por pares de identificação de rota de ônibus, utilizando uma matriz de relação entre a rota e estação de embarque.

A pesquisa de Niaraki e Kim (2009) apresenta uma ontologia para determinar um modelo de impedância em rodovias, usando Sistema de Posicionamento Global GPS e Sistemas Inteligentes de Transportes. A impedância é um modo de calcular o valor de custo ou de resistência que deverá passar por um *link* associado ao nó de origem e ao nó de destino. Definiram uma ontologia para um segmento de estrada com base nas preferências do usuário e critérios (ambientais) e de contexto. A partir dessa ontologia, eles definiram uma estrutura hierárquica dividida em dois ramos: um relacionado com os critérios do usuário (como informações sobre atrações turísticas e preferências) e outro com os critérios de contexto (tais como clima e segurança). Jianmu e Lei (2008) descrevem uma ontologia para sistematizar a coordenação em uma rede de comunicação baseada em RFID.

Finalmente, algumas ontologias para a área de transporte estão disponibilizadas na *Web*, de forma livre, para uso em diversas aplicações. Dentre essas aplicações, pode-se citar:

- OSM (*Ontology of Soft Mobility*). Esse projeto objetiva modelar a interação entre diversos modais na mobilidade (a pé, bicicleta, ou ambos) urbana para promover a maneira ideal de se deslocar. Considera aspectos como a duração da viagem de um tipo de usuário e que aspectos podem causar impacto a muitos potenciais usuários. O caráter sugestivo de alguns caminhos (passeios em parques) também é levado em conta (SEMCITY, 2012);

- OTN (*Ontology of Transportation Network*). O projeto *REWERSE Network of Excellence* disponibiliza a ontologia com os principais componentes dos sistemas de transporte e suas interações ( REWERSE, 2013);
- *SWEET Ontology*. Está sendo desenvolvida pela JPL (Laboratório de Propulsão a Jato) da NASA, na linguagem de ontologia OWL e disponível ao público. Detém 6.000 ontologias agrupadas em 200 conceitos separados nas mais diversas áreas da ciência. Possibilita a maioria dos desenvolvedores adicionar uma ontologia de domínio específico, usando os componentes definidos para satisfazer as necessidades do usuário final.

O desenvolvimento de aplicações sensíveis a contexto é complexo. Essas aplicações são passíveis de adaptação às alterações do contexto da informação a partir do contexto físico, computacional ou do usuário. A literatura pesquisada identifica a área de logística, em especial SCM (*supply chain management*), como promissora para o uso de sistemas sensíveis a contexto. A representação dos conceitos com o uso de Ontologia possibilita capturar e especificar de forma geral e explícita o conhecimento de um domínio por meio de terminologias consensuais, axiomas formais e restrições, abrindo caminho para maior interação entre os veículos e a infraestrutura.

## 4.7 Sistemas Simuladores ou Geradores de Dados para Transporte

Embora o planejamento de transportes deva ser embasado em dados reais, estes usualmente estão sujeitos a restrições de confidencialidade e segurança. Além disso, os projetos atuais ainda estão em fase de implantação. Em função disto, é comum a utilização de dados simulados para efetuar a análise e o desenvolvimento de aplicações. No entanto, é necessário que os dados estejam os mais próximos possível da realidade, o que requer aplicações que sejam voltadas para esse domínio e que haja dados em quantidade suficientes com qualidade.

### 4.7.1 *Softwares* Simuladores de Tráfego

A descrição confiável do tráfego nem sempre é uma tarefa trivial. Inúmeros modelos foram desenvolvidos e propostos, sem, contudo poderem ser considerados como padrões universais, posto que 60% do tráfego estão submetidos ao desejo individual de mobilidade (KRAJZEWICZ *et al.*, 2002). De uma forma geral, os

modelos podem ser classificados em quatro categorias principais, conforme grau de detalhamento: Modelo Macroscópico, Modelo Mesoscópico, Modelo Microscópico e Submicroscópico. Os modelos macroscópicos são mais utilizados quando o comportamento individual do veículo não é requerido, mas apenas a avaliação geral do fluxo de tráfego na via e são utilizados para o planejamento regional de transportes. Os modelos microscópicos consideram o comportamento individual de cada veículo. Os modelos submicroscópicos consideram partes isoladas dos veículos e o comportamento dos condutores. Cada veículo pode ser discretizado no tempo e no espaço, usando autômatas celulares ou apenas no espaço discreto. Nesta tese, optou-se pela utilização da simulação microscópica por melhor se enquadrar nos objetivos do projeto.

Na aferição do método proposto de determinação de padrões frequentes de movimentação de veículos, foi conduzida uma pesquisa para identificar *software* de simulação da movimentação de veículos em vias rodoviárias que possuíssem características adequadas ao projeto em questão. Alguns desses *softwares* foram analisados no projeto SMARTTEST (1999), em sua maior parte consideram a simulação microscópica. Outro trabalho analisado para subsidiar a escolha do *software* a ser adotado na simulação foi “*A Comparision of Microscopic Traffic Flow Simulation Systems for na Urban Area*” (MACIEJEWSKI, 2010), que compara os resultados da aplicação dos *softwares* VISSIM, SUMO e TRANSIMS.

Na simulação de eventos de transporte, existem basicamente três tipos de técnicas para a geração de dados de testes: randômico, orientado a objetivo e orientado a caminho.

A técnica randômica é a forma mais simples e, portanto, mais fácil, de gerar dados para teste. Sendo assim, pode ser usada para gerar dados de entrada em qualquer tipo de programa como inteiros, *string* e pilhas. No entanto, não tem um desempenho muito bom no sentido de cobertura dos casos de teste, pois se limita a usar a probabilidade e com isso as chances de encontrar pequenos erros semânticos são baixas.

Por exemplo, no trecho de código abaixo, ao utilizar o método randômico, as chances de gerar as entradas a e b iguais e assim escrever 1 são muito pequenas.

```
void foo(int a, int b) {  
    if (a == b) then  
        write(1);  
    else  
        write(2);  
}
```

A geração de dados orientada a objetivos é um método considerado mais expressivo do que a randômica, no sentido de que proporciona uma orientação para um determinado conjunto de caminhos. Essa característica faz com que seja possível a geração de entradas para qualquer caminho, já que depois ele vai ser direcionado a um valor específico, resultando em uma redução no risco de percorrer caminhos que podem se desviar do objetivo. Uma vez que esse método se baseia no conceito de encontrar qualquer caminho, nem sempre é possível prever a abrangência que um dado conjunto de objetivos irá ter.

Por último, o método de geração de dados de teste orientados a caminho é tido como o mais exequível quando comparados aos três. Ele não permite ao gerador que este faça a seleção em um universo de caminhos, mas que utilize apenas um. A principal diferença em relação ao método descrito anteriormente é que este prioriza o caminho pelo qual a meta vai ser alcançada. A desvantagem desse método é que ele dificulta que sejam encontrados dados de teste, porém, possibilita uma maior amplitude de cobertura em relação aos possíveis testes.

A seguir será apresentada a análise de alguns sistemas simuladores de tráfego que foram considerados para gerar as informações de entrada para os testes nesse trabalho.

#### 4.7.2 Critérios Usados para a Seleção:

Para a seleção do *software* a ser utilizado, considerou-se além das características técnicas de simulação a possibilidade de obtenção como *Software Livre*, a criação de arquivo contendo a movimentação de veículos de diversos tipos e detectores de leitura que identificassem o veículo e sua velocidade.

### 4.7.3 Softwares analisados

- TransCad

O primeiro *software* analisado foi o TransCad por ser um dos mais utilizados na área de Engenharia de Transportes. O TransCAD é um *software* de planejamento de transportes com plataforma de Sistema de Informações Geográficas (SIG), ou mais especificamente uma plataforma SIG-T (Sistema de Informação Geográficas para Transportes), desenvolvido especificamente para armazenar, visualizar, manipular e analisar dados relativos ao transporte. Trata-se de *software* proprietário de custo elevado, razão pela qual seu uso foi descartado.

- FreeSim

É um *software* livre feito na linguagem Java e têm como principais algoritmos de simulação e geração dos dados: Dijkstra's Algorithm, Bellman-Ford's Algorithm, Johnson's Algorithm e All-Pairs All-Paths Pre-Computed algorithms.

Sua principal qualidade é disponibilizar duas formas de simulação: Pelo caminho mais curto, usando "*Shortest path algorithms*", e pelo caminho mais rápido, usando "*Fastest path algorithms*".

A simulação e a configuração são por veículo. Isso torna sua configuração muito trabalhosa e complexa, pensando em um cenário em que são necessárias simulações de uma quantidade muito grande de carros, que é o caso desta tese.

O FreeSim não gera dados complementares como Tipo do Veículo, Características do Veículo, Carga e Motorista, razão pela qual não foi considerado.

- TEN – Traffic Engine

Desenvolvido na plataforma dotNet com a linguagem C#, utiliza em sua simulação o IDM (*Intelligent Driver Model*) para o cálculo de variação da velocidade dos veículos. Para inserção de novos veículos, é utilizada a Distribuição de Poisson e a Relação de Webster para cálculo de atraso por veículo.

O sistema permite parametrizar o tempo dos semáforos, a velocidade máxima das ruas, o número de carros por minuto e a distância de segurança entre os carros. Entretanto, a saída não é adequada ao projeto desta tese, pois emite apenas um relatório com o número de carros e a velocidade média deles, não mostrando a localização, o

tempo e a velocidade de cada veículo em um determinado instante. Além disso, tem a limitação de só poder ser executado em sistemas operacionais da Microsoft.

- SUMO – Simulation of Urban Mobility

“O SUMO é um simulador de tráfego urbano discreto no tempo e contínuo no espaço” (SUMO, 2012). Ele está disponível sob a licença GPL (*General Public License*) e foi desenvolvido utilizando as linguagens Python e C++.

Permite várias parametrizações do veículo como: nome, tipo, tamanho, velocidade máxima, aceleração, desaceleração, imperfeição do motorista, cor e rota. É multimodal, podendo-se modelar carros, ônibus, veículos longos e pesados como os caminhões, bicicletas entre outros.

Segundo Behrisch *et al.* (2011), o SUMO tem como características:

- Não existem colisões entre veículos;
- Pode-se modelar diferentes tipo de veículos;
- Ruas com várias faixas e mudanças de faixas;
- Cruzamentos com prioridades diferentes para os veículos;
- Saída em XML (eXtended Mark-up Language) contendo informações sobre a rede em cada instante de tempo
- Entrada em XML em que podem ser distribuídos por vários arquivos para um melhor manuseio.

O modelo utilizado no SUMO é o *Gipps model extension* descrito por Krauss (1998) e Janz (1998), capaz de exibir as principais características de tráfego como fluxo livre e congestionamentos.

O SUMO pode gerar dois tipos de saída. O primeiro tipo contém informações sobre cada veículo como: a descrição da posição, a velocidade e o nome para cada faixa de cada rua em um determinado instante de simulação. A outra saída consiste em colocar detectores em determinados pontos das pistas e esses detectores geram *logs* com as informações coletadas, tais como fluxo de veículos, velocidade média, porcentagem de ocupação da área do detector etc.

- Thomas Brinkhoff

É um *software* proprietário feito na linguagem Java e utiliza algoritmos espaço temporais (*Spatio temporal algorithms*) para simulação e geração dos dados.

Sua principal qualidade é mostrar a disposição geográfica dos objetos em uma determinada área após a movimentação, o que é importante, pois mostra a visão geral de uma área com todos seus veículos.

O autor deste trabalho manteve contato com o Prof. Thomas Brinkhoff (Alasca-EUA), que disponibilizou acesso ao *software* pelo site <http://iapg.jade-hs.de/personen/brinkhoff/generator/>. Os testes realizados demonstraram que sua configuração é muito complexa e o *software* não disponibiliza dados complementares como Tipo do Veículo, Características do Veículo, Carga e Motorista, além de não permitir a criação de arquivo contendo a movimentação dos veículos.

A velocidade fornecida é a máxima e não a velocidade na qual o suposto veículo passou no ponto de coleta dos dados.

- VISSIM

É um *software* comercial de simulação VISSIM desenvolvida para o tráfego dentro e fora das cidades e, portanto, aplica-se a diversas áreas. O VISSIM simula processos de tráfego em grandes cruzamentos e dentro de redes com mais detalhes que simuladores semelhantes. Apresenta todos os elementos de tráfego motorizados. Até mesmo pedestres atravessando a rua são levados em consideração.

Segundo Oliveira e Cybis (2008), o *software* de microssimulação VISSIM apresenta vantagens com relação a outros ferramentais quanto a (i) facilidade de elaboração das redes modeladas e (ii) flexibilidade de implementação do comportamento do tráfego relacionado a cada *link*. O comportamento de *car-following* de veículos em simulações de rodovias no VISSIM é alterado por meio de dez parâmetros comportamentais que buscam representar diferentes modos de condução do veículo: (i) condução livre; (ii) condução de aproximação; (iii) condução de seguidor e; (iv) condução de frenagem.

Embora seja um dos sistemas melhores avaliados pela comunidade de transportes, trata-se de *software* proprietário que o autor desta tese não teve acesso e, portanto, foi desconsiderado.

- Arena

O Arena foi desenvolvido nas linguagens SIMAN e VBA e é proprietário e considerado um dos principais *softwares* de simulação de uma forma geral.

Para simulação e geração dos dados, utiliza um Modelo lógico-matemático que representa a dinâmica do sistema em estudo. Esse modelo normalmente incorpora valores para tempos, distâncias e recursos disponíveis. Ao modelo são anexados dados sobre o sistema. Dessa forma, a simulação se diferencia, pois não são utilizados valores médios para os parâmetros no modelo e sim distribuições estatísticas geradas a partir de uma coleção de dados sobre o parâmetro a ser inserido.

Sua configuração é simples e não é necessário escrever nenhuma linha de código na configuração, pois todo o processo de criação do modelo de simulação é gráfico e visual, de maneira integrada.

É um simulador genérico, onde dados referentes ao sistema são incluídos, diferenciando a simulação. Existem várias versões do ARENA, porém não existe uma versão específica para transporte, o que dificulta o uso do mesmo nesta tese, pois será necessária a criação de um *template* para transportes, tornando seu uso muito complexo. O *software* não disponibiliza dados complementares como Tipo do Veículo, Características do Veículo, Carga e Motorista.

Com base no levantamento dos *softwares*, foi possível observar que o SUMO mostrou-se mais adequado, devido às suas principais características, que não foram encontradas nos outros sistemas analisados, mas atenderam às necessidades deste trabalho. São elas: (1) permissão da modelagem de diferentes tipos de veículos, (2) possibilidade da criação de seu próprio tipo de veículo, (3) disponibilidade de um alto número de parametrizações, (4) retorno da saída em um formato compatível com outras ferramentas (no caso, XML) e (5) geração de dados discretos por se tratarem de valores finitos. Vale explicitar que a importância do uso de dados discretos no presente projeto se deve ao fato de que os dados obtidos pelas antenas de leitura do RFID também o são, ou seja, sabe-se que um determinado veículo passou por dois pontos, mas não se tem o conhecimento do que aconteceu entre os mesmos. Além dessas propriedades essenciais

ao trabalho, existem outras interessantes, como o fato de o SUMO possuir seu código aberto e estar sobre a licença GPL, ser multiplataforma, ser direcionado a pesquisas acadêmicas e possuir uma vasta documentação.

## Capítulo 5 - Método Proposto

Nos capítulos anteriores foram apresentados e identificados aspectos relativos ao planejamento do transporte dentre os quais se destacam:

- O modelo das Quatro Etapas é o mais empregado para o planejamento dos transportes no Brasil e no Exterior;
- A construção da Matriz OD é o item que demanda o maior esforço. Em resumo, custos elevados para pesquisa de campo, periodicidade elevada de levantamento de dados, coleta de informações insuficientes, entre outros;
- A literatura pesquisada aponta aspectos críticos dos modelos e enfatiza a necessidade de serem adotados processos que possibilitem a coleta e tratamento de dados confiáveis da movimentação de carga com padrões mais refinados e correlacionados a geografia local.

Diante das limitações descritas na revisão bibliográfica, o Método Proposto deverá atender aos seguintes requisitos básicos:

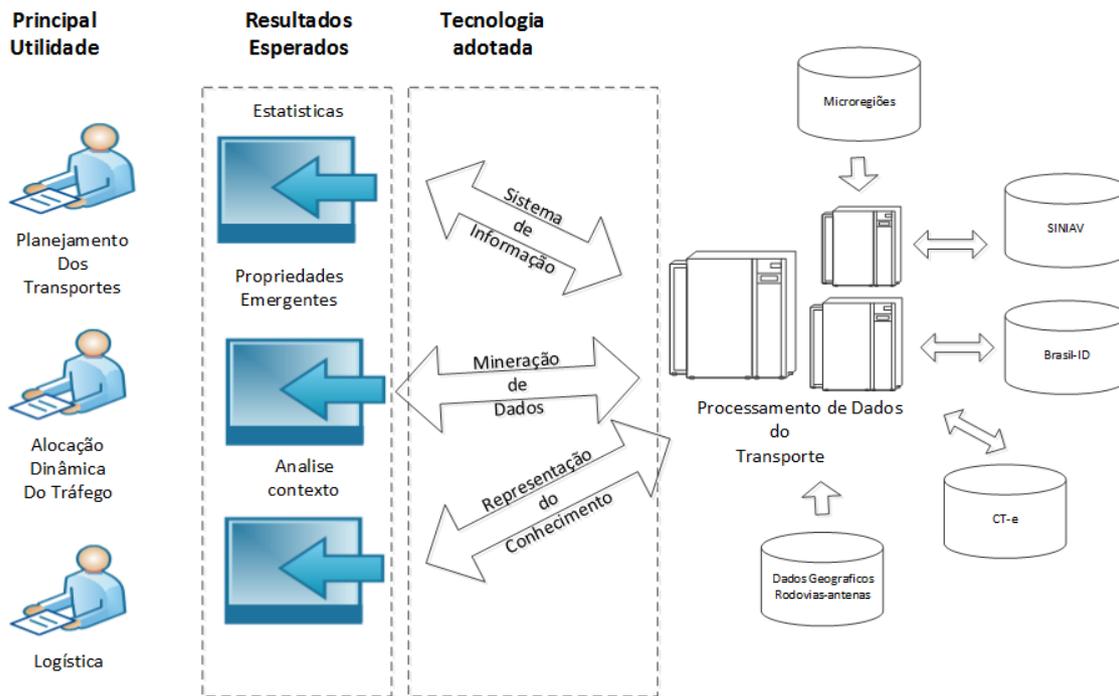
- Aumento da base de dados em quantidade, qualidade e variedade sem aumento de custos a partir de sistemas com:
  - Abrangência em todo território nacional;
  - Identificação automática de veículos;
  - Identificação automática da carga transportada;
  - Identificação automática da Origem e Destino.
  - Atualização permanente.

A seguir será apresentado o método resultante da pesquisa realizada para coleta e tratamento de dados do transporte.

### 5.1 Descrição do Método

O Método proposto objetiva coletar e tratar os dados de transporte para atendimento ao planejamento do transporte rodoviário de carga e para o planejamento do tráfego urbano, segundo os paradigmas dos Sistemas Inteligentes de Transportes – ITS, atendendo a todas as etapas do Modelo das Quatro Etapas.

Os principais componentes do método são apresentados na Figura 5.1:



**Figura 5.1:** Diagrama Geral

O diagrama mostrado na Figura 5.1 integra os dados obtidos em tempo real com componentes da geometria das vias e divisões em microrregiões, com os modelos de planejamento e encapsula os algoritmos de modelagem do tráfego de forma dinâmica, retroalimentando as etapas de planejamento. Sua estrutura mostra um Modelo de Sistema de Planejamento de Transporte baseado na arquitetura computacional cliente-servidor, integrando dados dos sistemas SINIAV, Brasil-ID e CT-e, além de dados geográficos das rodovias e microrregiões estabelecidas pelo IBGE.

O Método considera três modelos ou abordagens distintas a seguir detalhadas:

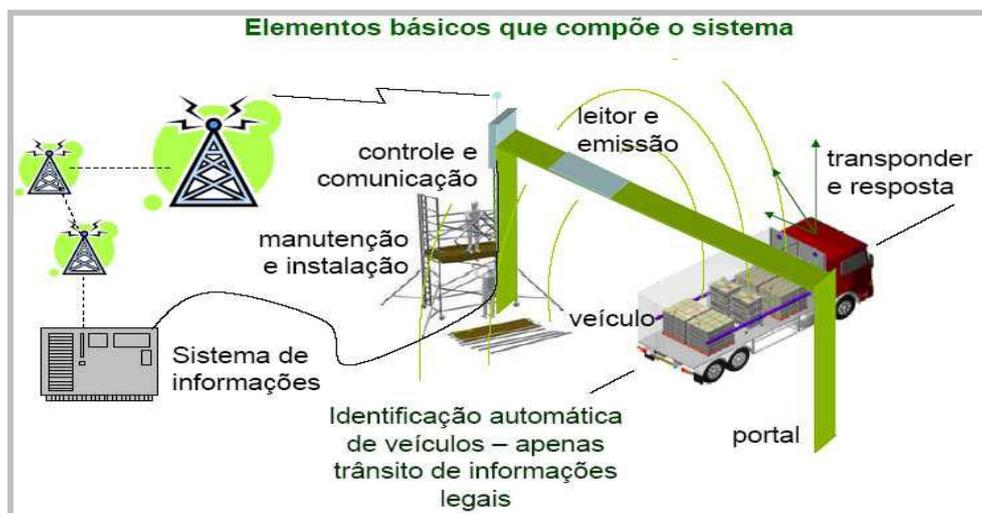
- A primeira abordagem visa ao estabelecimento da Matriz OD para carga e sua alocação na malha rodoviária. Seu desenvolvimento é baseado em um Sistema de Informações (também designado como Sistema de Apoio à Decisão) com o tratamento de dados “on line” (OLAP do inglês *on line analytical process*) e tecnologia WEB para a emissão de relatórios, segundo métodos estatísticos descritivos. A fonte de dados são os dos sistemas de controle fiscal: Brasil-ID, CT-e e de Identificação Automática de Veículos SINIAV. Os principais relatórios consolidam dados estatísticos de origem, destino, tipo de carga, tipo de veículo de transporte e rota utilizada;

- Uma segunda abordagem, voltada ao planejamento de trafego urbano, utiliza *Softwares Livres* para a descoberta de conhecimento em base de dados (do Inglês *KDD Knowledge Discovery in Databases*). Este trabalho considera as duas primeiras etapas previstas nessa metodologia: pré-processamento e mineração de dados (*datamining*) para identificação de propriedades emergentes da movimentação de veículos. Os resultados são comparados com trabalhos científicos similares publicados recentemente;
- A terceira e última abordagem apresenta uma visão logística que considera o desenvolvimento de Modelo Conceitual para representação do conhecimento segundo o paradigma da ‘computação sensível a contexto’ e para o qual serão utilizados os conceitos da Ontologia.

Em todos os casos, os resultados do processamento de cada modelo podem ser expressos em gráficos ou em tabelas definidas pelo usuário.

## 5.2 Considerações Preliminares

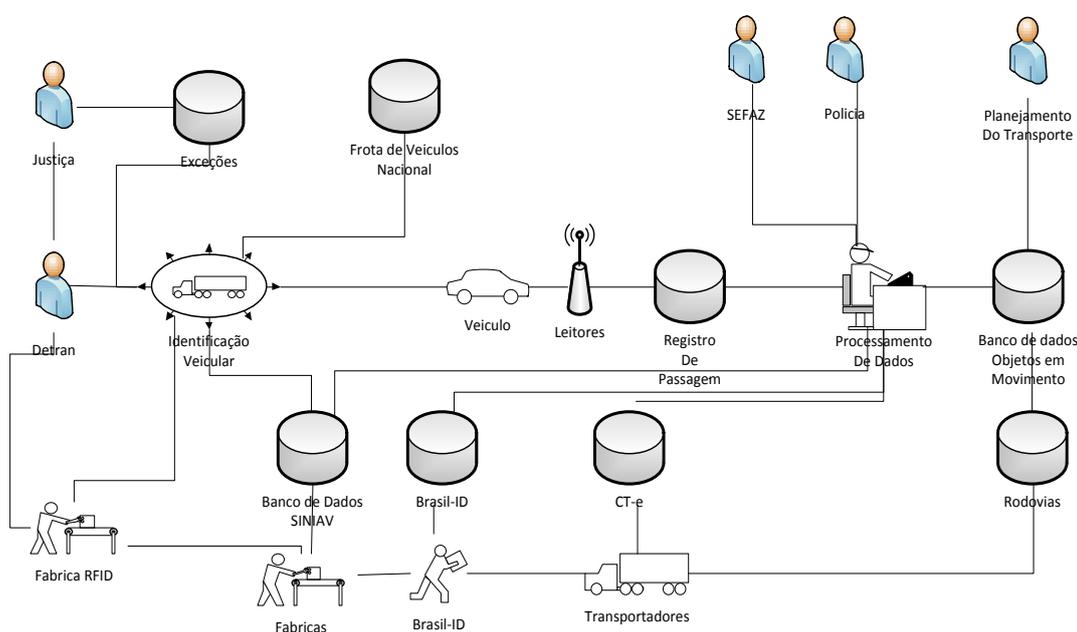
Nos modelos considera-se que todos os veículos – de passageiros ou de carga – contenham etiqueta RFID segundo o projeto SINIAV e que a carga seja igualmente etiquetada segundo o padrão do Projeto Brasil-ID. Assim, quando um veículo passar pela antena de leitura, tanto os dados do veículo (número do chassi, placa do veículo, código RENAVAM e outros) quanto os dados da carga serão armazenados em uma base e poderão ser correlacionados com os dados do CT-e e posteriormente tratados (Figura 5.2).



**Figura 5.2:** Projeto SINIAV – Monitoramento dos Veículos

A Figura 5.2 ilustra a proposta do monitoramento de veículos do SINIAV, no qual ao longo do trajeto o veículo passará por antenas (portais) conectadas a leitores RFID. O veículo contém uma *tag* e um receptor de sinais que recebe e armazena os dados transmitidos pelos leitores. Quando o veículo passa por um leitor, a *tag* recebe um pulso gerado pelo sistema e então responde a esse pulso com as informações armazenadas em sua memória para, em seguida, processar esses dados.

Para o melhor entendimento do processo de aquisição de dados, apresenta-se na Figura 5.3 a concepção lógica do processo de coleta de dados e suas principais dimensões.



**Figura 5.3:** Concepção Lógica para aquisição de dados

A sequência de etapas de aquisição de dados pode ser assim resumida:

- Os veículos são identificados pelos departamentos de trânsito dos estados por uma placa eletrônica com tecnologia RFID segundo os critérios estabelecidos para o SINIAV. Durante esse processo, o DENATRAN é consultado e libera a identificação, incluindo em seu cadastro o veículo com suas características. A regulamentação do SINIAV considera que todos os veículos da frota nacional sejam identificados utilizando-se essa tecnologia;
- Para o transporte de carga, o transportador emite o CT-e (para o serviço de transporte) e a NF-e (para a mercadoria). Segundo a regulamentação prevista pelo projeto Brasil-ID, a carga também será identificada por componentes RFID

e a DANFE igualmente conterá um ‘*chip*’ RFID identificando os documentos fiscais;

- Estão previstas instalações de antenas nas cidades para controle dos veículos segundo o projeto SINIAV e nas estradas em complementação às barreiras fiscais dos estados. Outras aplicações, como por exemplo *free flow* para pagamento eletrônico de pedágio, poderão, igualmente, serem desenvolvidas mediante acordo com os órgãos de trânsito.

## 5.3 Abordagem – Sistema de Informações

### 5.3.1 Considerações Preliminares

O objetivo desse sistema é demonstrar a adequabilidade do uso das informações dos sistemas nacionais para calcular os padrões de movimentação de cargas e veículos pelas rodovias.

Para tratar essa abordagem foram desenvolvidos dois *softwares* distintos e complementares descritos em

<https://skydrive.live.com/redir?resid=F27A4F06F5CFE1C9!432>:

O primeiro *software* tratou de identificar a passagem dos veículos pelas antenas em um determinado período com a determinação da antena inicial, antena final e rota utilizada no deslocamento.

Considerou-se a hipótese de antenas móveis a partir das quais o *software* deveria permitir uma nova localização com facilidade de alteração para o usuário. Nesse caso, foram usados os componentes de *software* disponibilizados pelo Google Maps para identificação da rodovia e para a localização da antena. Para validação do modelo, optou-se por criar outro aplicativo que gere uma base de dados para fluxo de veículo em rodovias à semelhança dos dados disponibilizados pelo DNIT (2001) com as informações sobre cargas, veículos, dia, horário, entre outras. Assim como, a simulação da leitura dos dados contidos nas *tags* das cargas pelos portais implantados nas estradas.

Para simular o processo de conversão do veículo na via, foi desenvolvido outro componente denominado *Fluxo Alternativo*, composto por um portal especial que pode ser “Entrada”, “Saída” e “Entrada/Saída”. Isso sinaliza ao gerador que a partir daquele portal existe uma via auxiliar que pode interferir na quantidade de veículos que passam pelo portal final de um segmento. Fluxos alternativos não existiriam em casos reais, eles

servem simplesmente como artifício de programação para que o aplicativo possa simular o tráfego com um grau maior de perfeição e em conformidade com a realidade.

A partir dessas informações, por meio de relatórios e gráficos gerados pelo sistema, foram elaboradas análises, permitindo aos Analistas de Transportes identificar possíveis irregularidades, estradas mais utilizadas, principais origens e destinos das cargas.

O segundo aplicativo objetivou ampliar o escopo do projeto no âmbito nacional, considerando as práticas usuais para o planejamento do transporte, em que as Zonas de Tráfego foram tratadas com a orientação do IBGE e DNIT, respectivamente.

A identificação da Origem-Destino é baseada na associação das antenas a Zonas de Tráfego de Carga, ou pela sigla em Inglês FAZ (*Freight Analysis Zones*), que são previamente determinadas por pesquisa estatística. As antenas são posicionadas no sistema em visualização gráfica ou por suas coordenadas geográficas, segundo sua localização prevista. A movimentação é complementada por tratamento informatizado dos dados dos postos de fiscalização a partir da leitura eletrônica por identificação por radiofrequência.

A seguir serão apresentados os principais critérios e técnicas utilizadas no desenvolvimento do Sistema.

### 5.3.2 Estrutura dos Dados

Uma avaliação preliminar indicou que os padrões frequentes de movimentação de cargas e veículos podem ser calculados com o uso dos recursos de Sistemas Nacionais de Identificação por Radiofrequência e/ou Controle Fiscal.

Os projetos considerados foram:

- CT-e: fornece Código de Endereçamento Postal (CEP) da Origem e do Destino, Código da Mercadoria, NF-e, Identificação do Veículo. Possibilita identificar a matriz OD, caso as zonas de tráfego estejam associadas ao CEP.;
- Brasil-ID: fornece o código da mercadoria e a NF-e a partir de sua passagem pelos postos de fiscalização. Possibilita rastrear a mercadoria pela rota utilizada;
- SINIAV: fornece a identificação do veículo e o registro de passagem pelos postos de fiscalização. A integração dos dados da *tag* do veículo

com os registros do DENATRAN possibilitam obter as características do veículo.

### 5.3.3 Localização das Antenas

Mirchandani *et al.* (2009) apresentam pesquisa relatando estudo relativo à localização ótima de antenas para antenas de coleta RFID para identificação automática de veículos. O estudo objetiva a responde duas questões: onde localizar as antenas para maximizar a monitoração dos veículos e qual a quantidade mínima de antenas para minimizar a variância da previsão do tempo de viagem. O modelo formulado de otimização indicou tratar-se de um problema NP-complexo. Na modelagem do problema, usou-se o sistema da IBM CPLEX com heurísticas aplicadas a malhas hipotéticas e comparadas com aquelas da cidade de Harris County, Texas.

Nesta tese, consideraram-se aspectos práticos relativos à instalação das antenas, visto que, o sistema RFID foca primordialmente a fiscalização do pagamento de tributos, sejam eles relativos ao veículo (IPVA) ou para a carga (IPI/ICMS).

Na medição da movimentação de carga pelas rodovias brasileiras estão previstas a instalação de pórticos com antenas para os sistemas SINIAV e Brasil-ID em complementação às barreiras fiscais localizadas nas principais rodovias nos limites dos Estados. Entretanto, o sistema possibilita a localização das antenas segundo critério definido pelo usuário.

### 5.3.4 Linguagens Utilizadas no Desenvolvimento do Modelo

- Visual Basic.Net

Visual Basic.NET é uma linguagem de programação orientada a objetos. Criada pela Microsoft, é uma linguagem que se destaca pela facilidade de se desenvolverem sistemas juntos com a IDE Visual Studio.NET e o .NET Framework SDK;

- JavaScript

O JavaScript é uma linguagem de programação interpretada que foi originalmente desenvolvida para ser parte dos navegadores *web*. Dessa maneira, os *scripts* podem ser executados no lado do cliente sem precisar ser executado pelo servidor. Com o JavaScript é possível controlar o navegador *web*, realizar comunicação assíncrona e alterar o conteúdo dos dados exibidos no *browser*;

- SQL

*Structured Query Language*, ou, em português, linguagem de consulta estruturada, é uma linguagem declarativa que serve para manipular bancos de dados relacionais. Com ela é possível fazer pesquisas, manipular dados e alterar a estrutura de tabelas e colunas de um banco de dados. A álgebra relacional inspirou muitas das características do SQL;

- Visual Studio.Net 2008

O Visual Studio.Net 2008 é uma plataforma IDE (*Integrated Development Environment*), ou seja, um ambiente integrado para desenvolvimento de *software*, criada para facilitar e dinamizar a construção de aplicações. Utiliza interações com o usuário do tipo arrastar-e-soltar do mouse, onde código é gerado em segundo plano, evitando atrapalhar o desenvolvimento dos programas, apenas gerando trechos de códigos óbvios. Possui forte orientação ao desenvolvimento baseado na orientação a objeto e sua *framework* possui inúmeros requisitos para atender às necessidades dos desenvolvedores;

- MySql 5.5.16

MySql é um SGBD (Sistema Gerenciador de Banco de Dados) de código aberto, sendo um dos mais populares do mundo, com mais de 10 milhões de instalações individuais. Utiliza o SQL como linguagem padrão e seus pontos positivos são o alto desempenho, escalabilidade e alto grau de compatibilidade;

- IIS 7.5

IIS é um servidor de aplicações para *web* seguro e fácil de gerenciar para desenvolver e hospedar aplicativos e serviços *web*. É uma tecnologia incluída no Windows Sistema Operacional Microsoft que implementa as tecnologias. Responsável por gerenciar páginas ASP.NET;

- Astah Community 6.6.4

Ferramenta de modelagem UML. Dentre os recursos da ferramenta estão:

- a) Suporte a UML 2.1;
- b) Diagramas de Classe, Caso de Uso, Sequência, Atividade, Comunicação, Máquina de Estado, Componentes, Implantação, Estrutura de Composição, Objetos e Pacotes;

- c) Ajustes de alinhamento e tamanho dos diagramas;
- d) Impressão dos diagramas;
- e) Exportação das imagens dos diagramas;

- MySQL Workbench

MySQL *Workbench* é um editor visual de banco de dados que dá suporte a conexões com bancos MySQL, de modo a oferecer uma interface gráfica para a edição da estrutura do banco de dados e execução de consultas SQL;

- ReportViewer 2008

O *ReportViewer* 2008 é um editor visual para a confecção de relatórios. Ele se conecta a aplicação por meio de uma conexão realizada no Visual Studio.NET ou recupera os dados por um objeto de dados ADO.NET. Com essa ferramenta, é possível montar graficamente a estrutura de um relatório pelo conceito de arrastar componentes com o mouse. É capaz de exportar relatórios em diversos formatos, mas o principal deles é o PDF (*Portable Document Format*).

### 5.3.5 Arquitetura do Sistema

Pode-se descrever a arquitetura do projeto como um misto de dois padrões arquiteturais bastante conhecidos: O padrão Cliente-Servidor e o Padrão de três camadas (Apresentação, Controle e Modelo).

- Aspecto Cliente Servidor

Na criação de aplicações ASP.NET, o visual studio.NET fornece suporte a todas as etapas necessárias para o desenvolvimento de uma aplicação Cliente – Servidor. O *Front-End* do sistema é exibido em *Web Forms* ASPX.

O *Back-End* é, neste caso, a subaplicação VB.NET rodando em um servidor de aplicação, em uma máquina diferente daquele do usuário. Essa separação é feita para centralizar o acesso ao banco de dados, de modo que todos os usuários do sistema acessem sempre a mesma base de dados, centralizada pela parte VB.NET da aplicação.

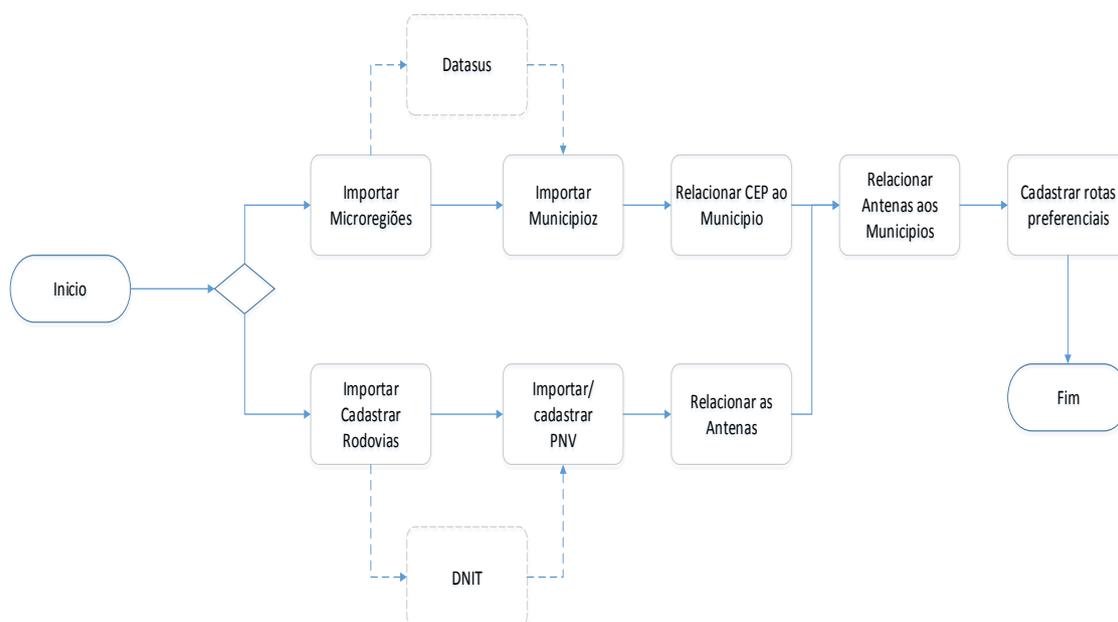
Nessa arquitetura, as subaplicações *Web Forms* e os demais projetos VB.NET são sistemas de informação, a princípio isolados um do outro, que se comunicam pelo protocolo http e TCP/IP para formar um único sistema distribuído.

### 5.3.6 Modelo de Requisitos

O modelo considera a existência de duas etapas de processamento para a determinação da matriz Origem-Destino-Carga-Modo (ODCM).

As Figuras 5.4. e 5.5 ilustram os passos necessários:

Na primeira, na fase de configuração do sistema, obtêm-se os dados necessários para dar suporte para determinação da rota utilizada, das análises, das extrações do CT-e SINIAV e da geração da Matriz ODCM.



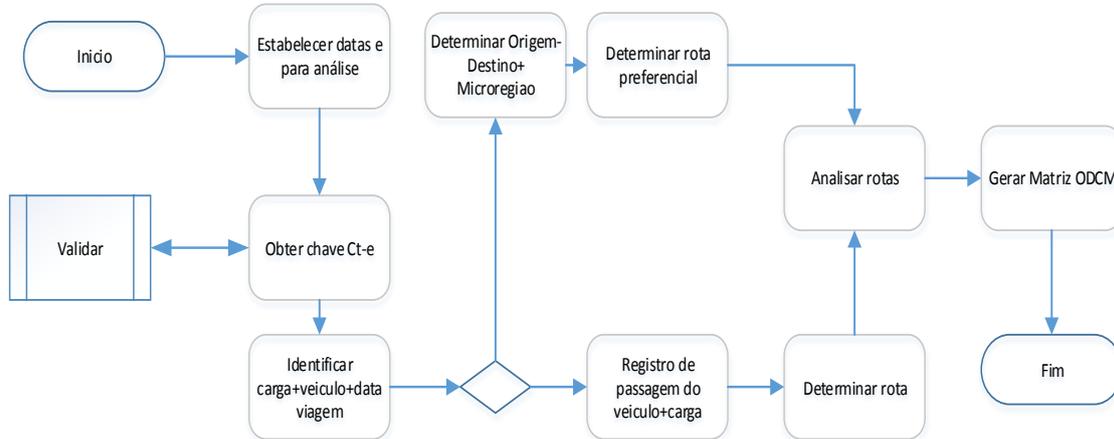
**Figura 5.4:** Passos para Configuração do Sistema

Da Figura 5.4 pode-se constatar as seguintes características:

- Os dados relativos às Zonas de Tráfego são associados às microrregiões do IBGE. A pesquisa realizada identificou que o sistema do DATASUS que tem como base os dados do IBGE possibilita a importação de dados atualizados. Cada microrregião tem relacionado um ou mais municípios e a ele o prefixo de 5 dígitos do CEP, que possibilita identificar o município e a microrregião;
- Os dados da rede rodoviária são importados do PNV (Plano Nacional de Viação) do DNIT (2013);
- As antenas são dispostas nas rodovias, sendo que duas antenas em sequencia formam um segmento e o conjunto desses segmentos forma uma rota. Cada antena se relaciona com as demais próximas a ela;

- As rotas são cadastradas e ponderadas quanto a sua escolha pelos motoristas. Cada rota é prevista para existir entre duas microrregiões.

Na fase de geração da matriz ODCM, as diversas informações são agrupadas, conforme mostrado na Figura 5.5.



**Figura 5.5:** Geração da Matriz ODCM

Os seguintes aspectos podem ser destacados desta etapa:

- O processo tem início com a identificação da viagem pela chave do CT-e existente na Secretaria da Fazenda dos Estados (SEFAZ) de origem. A partir dessa chave, pode-se obter as informações do transporte, tais como veículo, carga endereço da origem e do destino, data da viagem, entre outros;
- Segue-se à determinação das microrregiões de origem e destino e à rota preferencial. Com base nessa rota preferencial, são estabelecidas as antenas que poderão registrar a passagem do veículo/carga.

### 5.3.7 Principais Atores do Sistema

Os principais atores do sistema que foram identificados durante a fase de análise são:

- ✓ Administrador: Responsável por todo o sistema, desde manutenção de usuários até a geração de relatórios;
- ✓ Analista de Tráfego: Analisa os dados relacionados às cargas e aos veículos nas estradas e à infraestrutura utilizada para esses fins;
- ✓ Analista de Planejamento e Infraestrutura: Responsável pela análise dos transportes de cargas e das rotas utilizadas para gerar a Matriz ODCM.

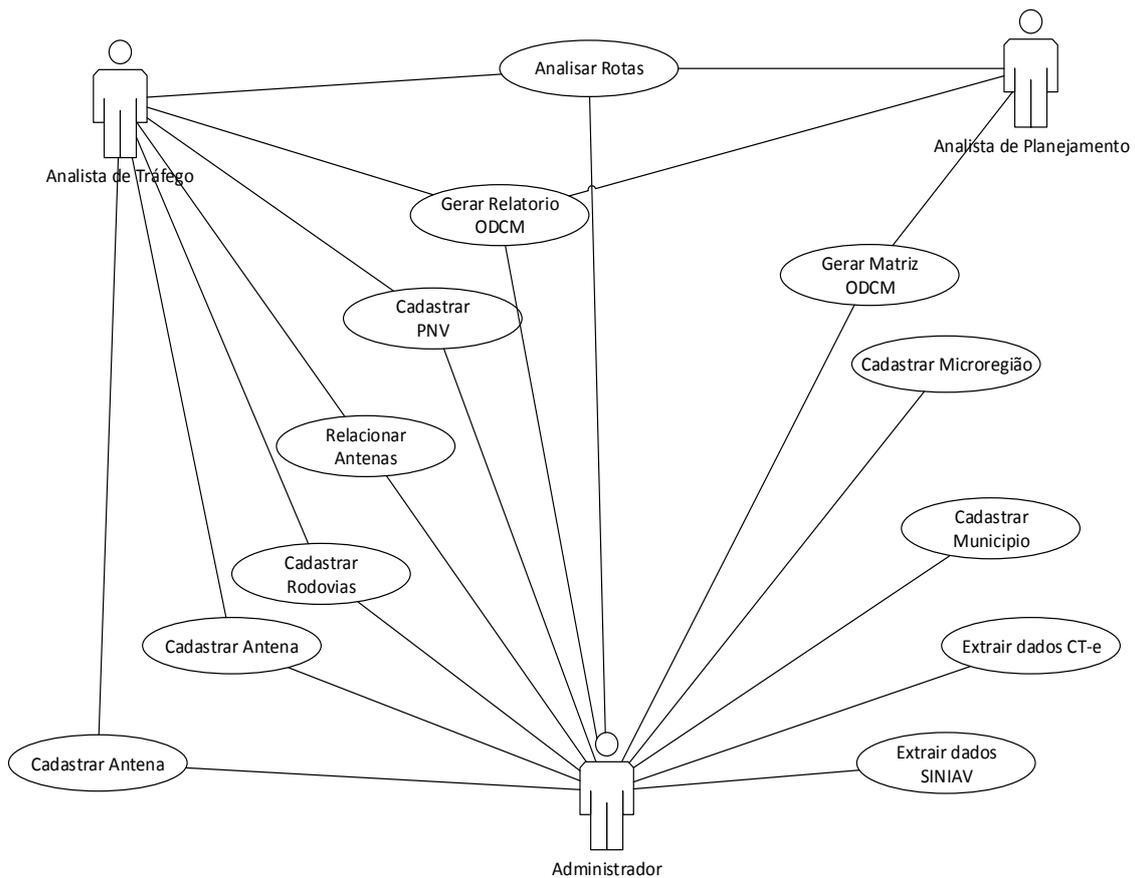
As principais responsabilidades desses atores encontram-se no quando 5.1

**Quadro 5.1:** Atores e responsabilidades

Responsabilidade	ATOR		
	Administrador	Analista de Tráfego	Analista de Planejamento e Infraestrutura
Cadastrar Município	SIM	NÃO	NÃO
Cadastrar Microrregião		SIM	
Cadastrar Rodovia			
Cadastrar PNV			
Cadastrar Antena			
Relacionar Antenas			
Cadastrar Rotas		NÃO	SIM
Extrair Dados CT-e		NÃO	
Analisar Rotas		SIM	
Gerar Matriz ODRCM		NÃO	
Gerar Relatório		SIM	SIM
Extrair Dados SINIAV		NÃO	
Gerar Relatório ODRCM			SIM

### 5.3.8 Casos de Uso

O diagrama apresentado Figura 5.6 mostra o conjunto de casos de uso com os atores do sistema, bem como seus relacionamentos (BOOCH *et al.*,2005), ou seja, as interações homem-máquina que serão disponibilizadas. Cada caso de uso representa uma unidade discreta de interação e é simbolizado por uma elipse.



**Figura 5.6:** Casos de uso do sistema

### 5.3.9 Modelagem dos Dados

- Modelo Entidade Relacionamento

O Modelo de Entidade e Relacionamento (MER) é uma representação da realidade e pode ser representado por entidades, relacionamentos e atributos. Existem muitas notações para diagrama de entidades e relacionamentos. A notação original foi proposta por Chen (1990) e é composta de entidades (retângulos), relacionamentos (losangos), atributos (círculos) e linhas de conexão (linhas) que indicam a cardinalidade, no sentido matemático, de uma entidade em um relacionamento que pode ser 1:1, 1:N ou N:N.

A Modelagem dos Dados e seus relacionamentos são apresentados no Modelo de Entidade-Relacionamento (MER) da Figura 5.7



Esse diagrama representa as estruturas de dados envolvidas na coleta e do tratamento de dados de transporte de carga da forma mais próxima ao mundo real dos negócios

## 5.3.10 Validação do Modelo

### 5.3.10.1 Parte 1 – considerando apenas o deslocamento na via

O modelo lógico desse *software* encontra-se descrito em <https://skydrive.live.com/redir?resid=F27A4F06F5CFE1C9!432>, do qual se destacam os seguintes relatórios:

- Tela Principal – Usuário Gerador de Tráfego

A tela apresentada na Figura 5.8 objetiva simular o tráfego para análise do *software*. Os dados coletados das antenas são reorganizados por veículo e por data de passagem.

**Datas**

Data Inicial   Data Final

Horário Inicial 00 : 00 Horário Final 00 : 00

**Estrada**

Estrada

Segmento Inicial

Segmento Final

**Entradas**

Nome	%

**Saídas**

Nome	%

**Veículos**

Caminhões  Outros Veículos  Total: 0

**Velocidades**

Entre  e  Km/h

**Tipos de Cargas**

- Material de Construção
- Alimentos
- Bebidas
- Produtos Químicos e troquímicos

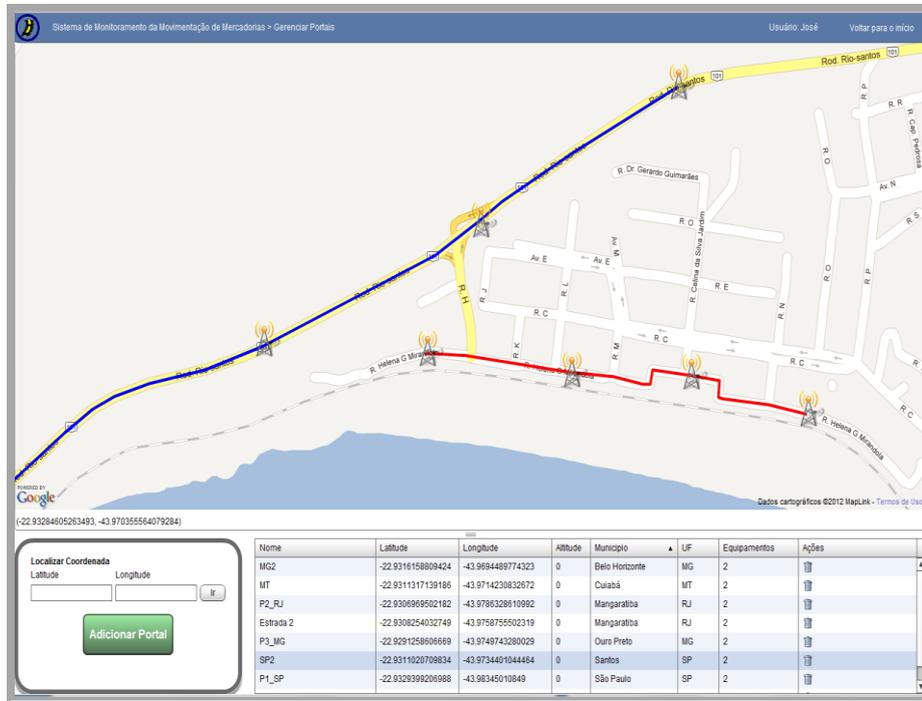
Selecionar Todos

**Gerar**

Figura 5.8: Tela Principal – Perfil Gerador de Tráfego

- Tela de Gerenciamento de Portais

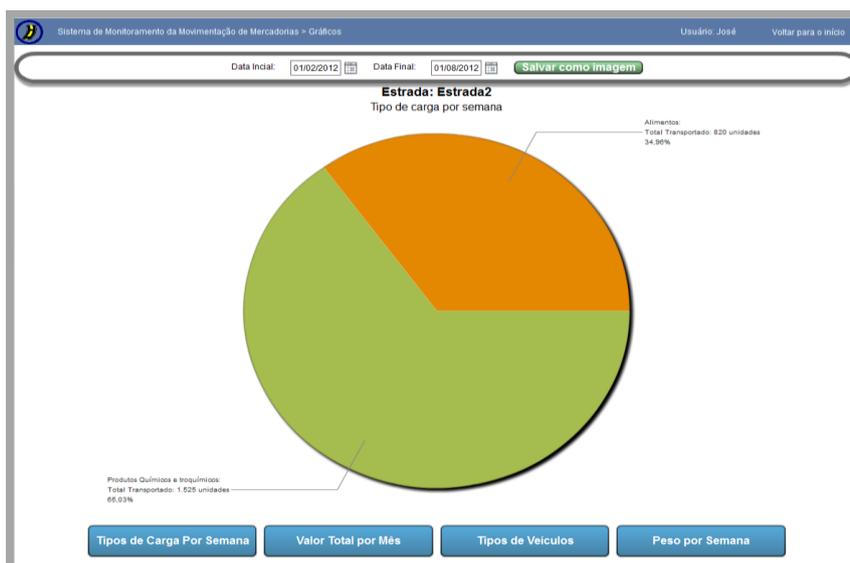
A tela mostrada na Figura 5.9 possibilita ao analista posicionar as antenas de acordo com sua real posição no campo. Para tanto, são utilizadas coordenadas geodésicas disponibilizadas pelo Google Maps.



**Figura 5.9:** Tela de gerenciamento de portais

- Gráfico Tipo de Cargas por Semana

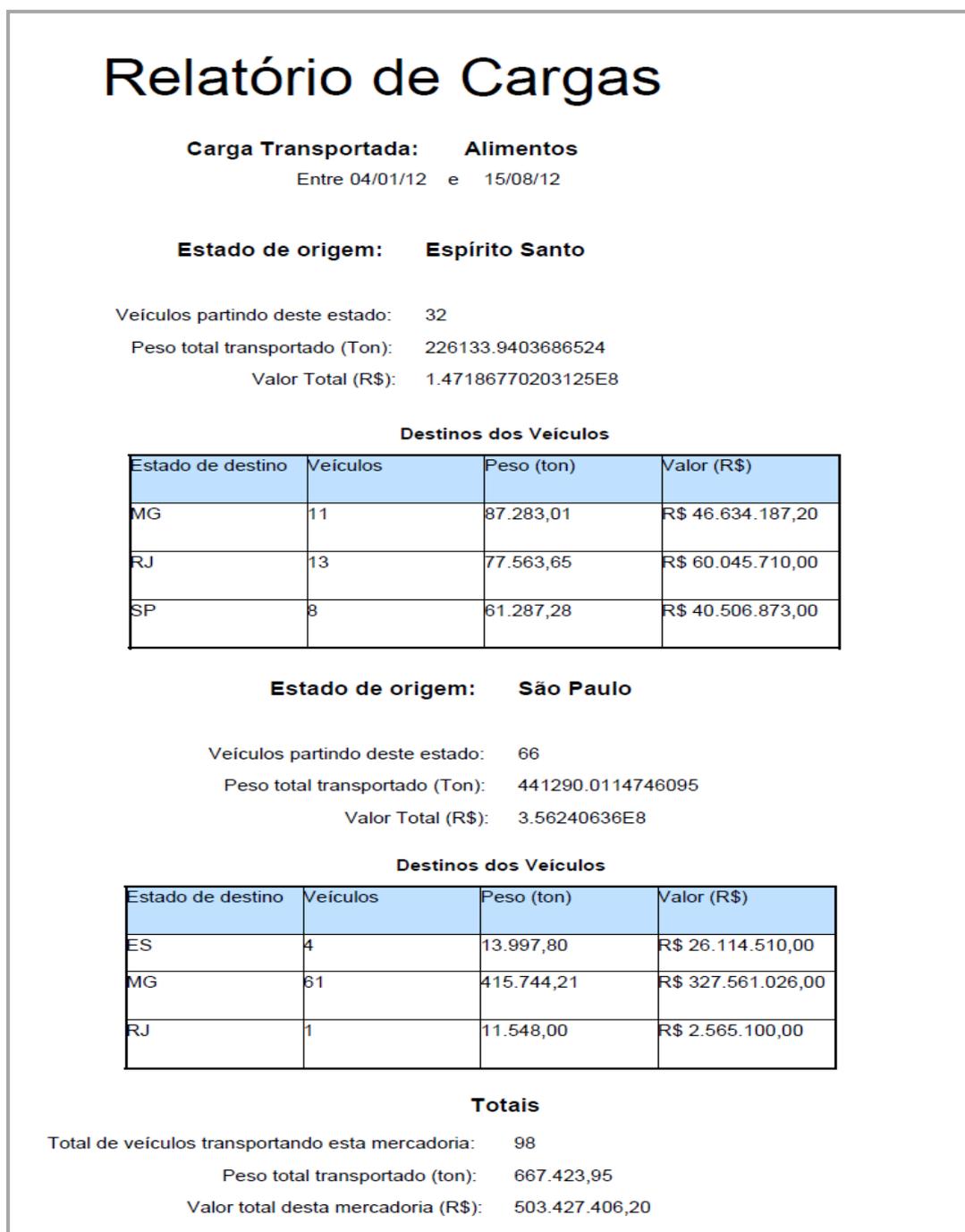
O relatório mostrado na Figura 5.10 permite ao analista visualizar as estatísticas de uso da rodovia por tipo de carga.



**Figura 5.10:** Gráfico Tipo de Cargas por Semana

- Relatório de Cargas

O relatório mostrado na Figura 5.11 permite identificar as Origens e Destinos das cargas selecionadas em um determinado período.



**Figura 5.11:** Relatório de Cargas

### 5.3.10.2 Parte 2 – considerando as origens e os destinos do CT-e

Essa parte considera a integração das Origens e dos Destino da carga estipuladas no CT-e e o rastreamento do deslocamento por meio da passagem dos veículos pelas antenas distribuídas nas rodovias.

O modelo lógico desse *software* encontra-se descrito no anexo, do qual se destacam os seguintes relatórios. Esse *software* integra as funcionalidades do *software* anterior e avança para os requisitos necessários à identificação da Matriz OD de carga e aos padrões frequentes de movimentação rodoviária.

- Tela Principal – Usuário Administrador



Figura 5.12: Tela Principal – Perfil Administrador

Essa tela de abertura do sistema (Figura 5.12) possibilita ao administrador realizar as principais tarefas do *software*.

- Tela de Geração do Relatório ODRCM

The screenshot shows the 'Gerar Relatório ODRCM' interface. At the top, the user is identified as 'José da Silva dos Santos' with a 'Sair do Sistema' link. The main form is titled 'Gerar Relatório' and is divided into several sections: 'Dados de Pesquisa' with date and location filters; 'Demais Dados' with dropdowns for route, cargo, and mode; 'Opções' with checkboxes for displaying specific data; and 'Formato' with radio buttons for 'Sintético' and 'Analtico'. A vertical sidebar on the left is labeled 'Gerar Relatório ODRCM'. On the right, a menu lists various administrative tasks like 'Cadastrar Município', 'Cadastrar Microrregião', 'Cadastrar Rodovia', 'Cadastrar PNV', 'Cadastrar Antena', 'Relacionar Antena', 'Cadastrar Rotas', 'Extrair Dados CT-e', 'Analisar Rotas', 'Gerar Matriz ODRCM', and 'Extrair Dados SINIAV'. At the bottom of the form are 'Visualizar' and 'Sair' buttons.

**Figura 5.13:** Tela de Geração do Relatório ODRCM

Essa tela (Figura 5.13) permite ao analista obter os relatórios da matriz OD, tipo de carga, principais rotas e veículo utilizado.

O Modelo tem por objetivo principal obter a matriz OD da carga, informações sobre a carga transportada, a rota que foi utilizada para realizar o transporte e o modal do veículo que realizou o transporte. O sistema permite ao usuário cadastrar as informações necessárias e obter as demais informações dos órgãos responsáveis por tais informações, sendo que, o SINIAV monitora os movimentos dos veículos e o CT-e controla o transporte de cargas. Essas informações tornam-se, assim, mais acessíveis e centralizadas, visto que o usuário pode acessar o sistema pelo *browser* na *Web*.

Com base nessas informações, foram elaboradas análises sobre os dados, permitindo ao Analista de Planejamento identificar não só a origem e o destino de uma determinada carga, como também a rota utilizada e o modal que realizou esse

transporte, quais as estradas e rotas mais utilizadas, quais os modais que transportam determinado tipo de carga, entre outros, pelo Relatório Matriz ODCM, atendendo, desta forma, às lacunas apontadas no relatório *Freight Demand Modeling and Data Improvement* (SHRP2/TRB, 2013).

## 5.4 Abordagem – Mineração de Dados

### 5.4.1 Objetivo da Abordagem

O objetivo dessa abordagem é proporcionar a descoberta de propriedades emergentes a partir da identificação dos padrões frequentes de movimentação com o uso de técnicas de mineração de dados (*datamining*). Neste contexto, é possível formular novas perguntas a respeito de correlações entre a rota e o tempo gasto para o deslocamento por tipo de mercadorias, do veículo, a quantidade de veículos por rodovia e outros dados necessários ao planejamento eficiente do transporte.

Justifica-se o uso dessa técnica a partir do trabalho “*A Review of KDD-Datamining Framework and Its Application in Logistics and Transportation*” (Desa & Wibowo, 2011), que aponta que o uso de mineração de dados na área de transporte como uma tecnologia viável e de amplo emprego.

De modo a corroborar a ideia defendida nesta tese, optou-se por dividir o processo em etapas que melhor o exemplifiquem. Primeiramente, foi realizada a obtenção dos dados de testes, dados estes que foram obtidos pelo sistema gerador de dados SUMO. Em segundo lugar, esses dados necessitaram de um tratamento, conforme segunda etapa de pré-processamento do KDD, para que finalmente fossem usados como base de dados para a etapa de detecção de padrões frequentes e previsão de movimentação. . Para essa etapa foi selecionado o software Weka (Waikato Environment for Knowledge Analysis).

Essa abordagem objetiva demonstrar a aplicabilidade do uso de técnicas de descoberta de conhecimento em bases de dados para “Alocação Dinâmica do Tráfego” conforme exposto na seção 2.2.5.2 dessa Tese. O primeiro passo para delimitação dos estudos é estabelecer as questões que podem ser respondidas pelo sistema.

Diante das inúmeras possibilidades de perguntas, optou-se em tratar do planejamento do tráfego urbano, considerando as questões:

- É possível identificar qual o próximo destino de um veículo com base no seu histórico de movimentação?
- Quais informações que podem ser coletadas pelo sistema e quais devem ser utilizadas?
- Qual o nível de agregação a ser adotado?
- Qual o melhor algoritmo que proporcionará um resultado satisfatório?
- Os dados selecionados para essa etapa foram simulados por *software*, não ocorrendo levantamento de dados de campo. Todos os dados gerados foram selecionados.

A modelagem proposta para essa etapa considera a simulação de uma rede viária e a instalação de antenas fixas capazes de coletar informações espaço-temporais relativas à passagem dos veículos: o ID (identificador do veículo) dos veículos, a velocidade e o tempo em que ele entrou e saiu da área da antena, o seu tipo e o tamanho.

O modelo a seguir apresentado atende às etapas previstas por Fayad (1996), descritas no Capítulo 4.

✓ Seleção de Dados

Para responder a essas questões, foram identificadas as seguintes grandezas:

- Tempo: Hora em que o veículo passa na via, pois existe padrão de deslocamento nas horas de pico, em que a maioria dos motoristas está indo ou retornando do trabalho/escola. Um exemplo de padrão desse atributo é de uma pessoa que tem um carro particular e utiliza-o para trabalho e faculdade. Em um determinado tempo, ele vai do ponto X para o Y (casa-trabalho) e em outro momento o seu veículo se movimenta do ponto Y para o Z (trabalho-faculdade);

- ID\_Veículo: O motivo da escolha desse atributo se assemelha ao anterior. Uma pessoa que utiliza seu veículo para sua rotina de estudo e trabalho tende a percorrer o mesmo percurso e, portanto, o identificador do veículo se torna importante para a previsão do movimento;

- Velocidade: Atributo escolhido que menos contribui com padrões, porém, tem relevância em casos específicos. Imaginando que o trajeto normal de alguns veículos em uma via sem engarrafamento (carros na velocidade média da via) é do ponto A para o B e, em dias de engarrafamento (carros em baixa velocidade), alguns motoristas conhecem um caminho alternativo e escolhem o mesmo para sair do engarrafamento da via que

leva ao ponto B, percorrendo o trajeto A => C. Nesse cenário, os algoritmos através da velocidade identificam um padrão de deslocamento, porém esses casos não são frequentes, pois não são todas as vias que possuem um caminho alternativo;

- Tipo do Veículo: Pode-se encontrar padrão de movimentação de acordo com o tipo de veículo. Os ônibus possuem itinerários fixos e, na maioria das vezes, exceto quando estão vindo ou indo para garagem, repetem o mesmo trajeto. Já os veículos de táxi não apresentam padrões na movimentação;

- Antena Via Destino: Esse é o atributo que será previsto pelo algoritmo por meio das saídas conhecidas. Cada valor desse atributo representa uma antena de coleta de dados instalada na via.

#### ✓ Pré-processamento

Esse modelo necessita organizar os dados simulados pelo SUMO para que possam ser analisados pela ferramenta de mineração de dados que no KDD chama-se “fase de pré-processamento”. Foi desenvolvido um programa em VB.NET que fizesse a leitura de todos os arquivos XML que foram gerados por cada uma das antenas e organizasse-os em um único arquivo XML com as seguintes informações: (1) Antena anterior, (2) Tempo, (3) Velocidade, (4) Tipo do veículo e (5) Antena de destino. Sendo assim, é possível obter-se a informação do tempo que o veículo levou de uma antena a outra, além do caminho percorrido. Notam-se duas etapas da fase de pré-processamento: a seleção dos dados, que consiste na escolha dos dados que melhor atendem ao processo, e a etapa de enriquecimento, quando pelo algoritmo se criam novas informações, partindo-se dos dados existentes.

O programa tem como entrada os arquivos gerados pelas antenas do SUMO a partir de arquivos onde “id” é o identificador da antena, “time” é o tempo do veículo e “state” é o estado do veículo, sendo “enter” o momento em que ele entrou na área da antena. “Stay” se permanece parado por um tempo e “leave” é o momento em que ele deixa a área coberta pela antena. “vehID” é o identificador do veículo, “speed” a velocidade na qual o veículo se encontra naquele estado, “length” o comprimento do veículo e “type” o tipo do veículo.

Em um primeiro momento, o programa lê cada arquivo gerado pelas antenas e percorre cada item “instantOut” do arquivo gerado pelo SUMO. Para cada item, ele verifica o estado do veículo “enter” e registra os atributos “AntenaAtual”, “Tempo”,

“TipoVeiculo” e “Velocidade”, deixando o atributo “AntenaAnterior” para ser preenchido em um próximo passo. Uma lista com os identificadores dos veículos também é montada.

✓ Enriquecimento

No modelo proposto foi possível, pela simulação com diversas agregações de dados, melhorar seu desempenho por meio de mudanças na organização dos dados. O atributo do tipo numérico “velocidade” não apresenta vantagem para o modelo, pois não é relevante a informação da velocidade exata que o veículo passou pelo ponto de coleta e sim a informação se o veículo está em uma via engarrafada ou não. Partindo desse princípio, pode-se “agrupar” esse atributo. Assim, foi criado filtro para “agrupar” os registros em faixas de velocidade, representando as condições de trânsito, considerando que: até 18,45 km/h de velocidade – ENGARRAFADO; de 18,45 até 41,31 km/h – LENTO; maior que 41,31km/h – NORMAL.

✓ Mineração de dados (*datamining*)

O objetivo dessa etapa é identificar a próxima antena a ser percorrida. Com as ferramentas de Mineração de dados é possível prever uma saída a partir de algoritmos de classificação. Dentre os algoritmos disponíveis, selecionou-se o “Vizinho mais Próximo” (*Nearest Neighbour*), também conhecido como filtragem colaborativa ou de aprendizagem baseada em instância, que é uma técnica de mineração de dados útil, pois permite a utilização de suas instâncias de dados do passado, com valores de saída conhecidos, para prever um valor de saída desconhecido de uma instância de novos dados.

Para previsão com os algoritmos de mineração de dados, é necessária a criação de um arquivo de configuração descrevendo os atributos. Com os dados do passado, escolher quais os atributos serão utilizados no modelo. Para isso, considerou-se a relevância de cada atributo para previsão, levando em consideração o cenário proposto para prever a movimentação de veículos.

Descoberta de Padrões

Os padrões de movimentação atendem à questão sobre o próximo movimento a ser realizado pelo veículo.

Para tanto, foram realizadas simulações considerando:

- Todos os veiculo sem agregação;
- Todos os veículos com agregação;
- Veículos de carga com agregação;
- Todos os veículos excluindo táxi e com agregação.

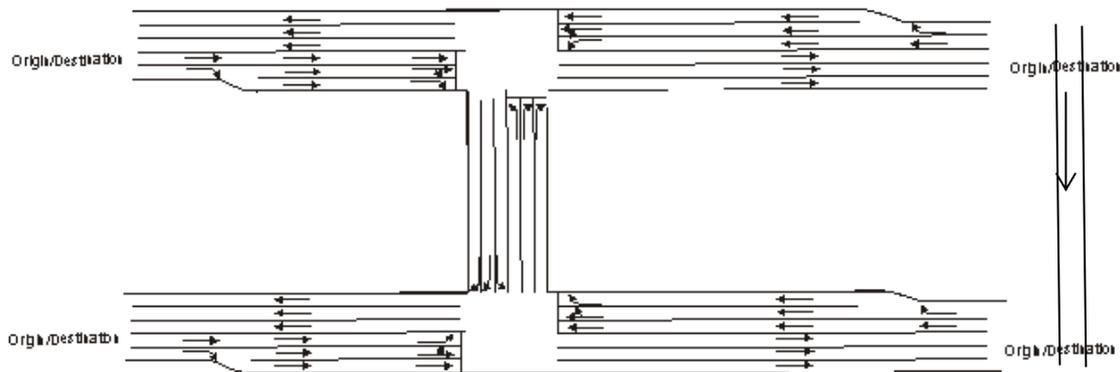
A interpretação dos resultados foi realizada mediante a comparação com estudos similares anteriormente publicados.

Outro aspecto considerado foi que, durante a realização da pesquisa, o autor desta tese participou na condição de Observador Acadêmico do protótipo de funcionamento do SINIAV durante o ano 2011. A melhor taxa de acerto de leitura foi de 70%, para os veículos equipados com a *tag*. Essa medida também foi designada por outros autores, como MPR (*market penetration rate*).

## 5.4.2 Validação do Modelo

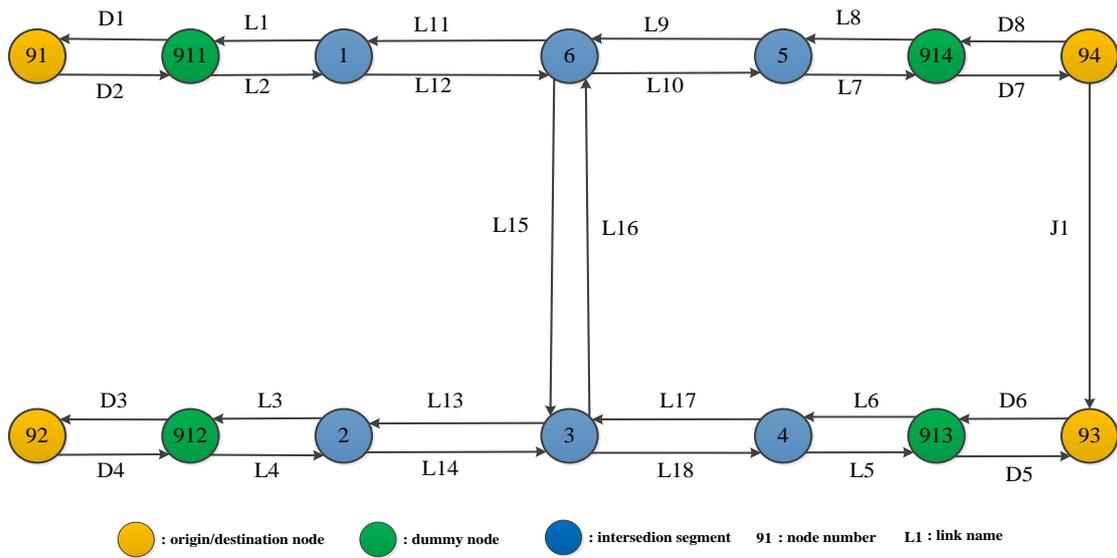
### 5.4.2.1 Configurações das ruas

Para a simulação, foi utilizado como base um modelo já existente no SUMO, o *quickstart*, modificado para atender aos requisitos desse modelo (Figura 5.14).



**Figura 5.14:** *Layout* da rede usado como base para simulação.

Foi realizada uma alteração no modelo existente para introduzir mais uma estrada à direita da rede com a finalidade de simular a existência de um caminho alternativo para os veículos que saem do nó 91 para o nó 92. Essa estrada foi configurada para ter apenas uma pista e uma única direção, do nó 94 ao 93, como mostra a figura 5.15.



**Figura 5.15:** Novo *layout* da rede de simulação.

Os seguimentos D2, D4, D6 e D8 (vistos na Figura 5.2) possuem duas pistas, enquanto os demais possuem três, com exceção da estrada nova que tem apenas uma pista.

Todas as vias foram configuradas com velocidade máxima de 13,889 m/s, com exceção à nova via, onde foi estipulada a velocidade máxima como sendo 10 m/s.

#### 5.4.2.2 Configurações dos veículos

Os veículos de carga foram definidos com base nas principais cargas transportadas no Rio de Janeiro e indicadas no PDTU (SECTRANS, 2002).

**Quadro 5.2:** Principais cargas transportadas

Produto	Fluxo Anual (mil ton/ano)	% de participação
Carga Geral	13.212.483	37,58
Material de construção	4.860.518	13,83
Alimentos	3.960.403	11,27
Bebidas	3.006.958	9,55
Produtos químicos	1.931.877	5,50

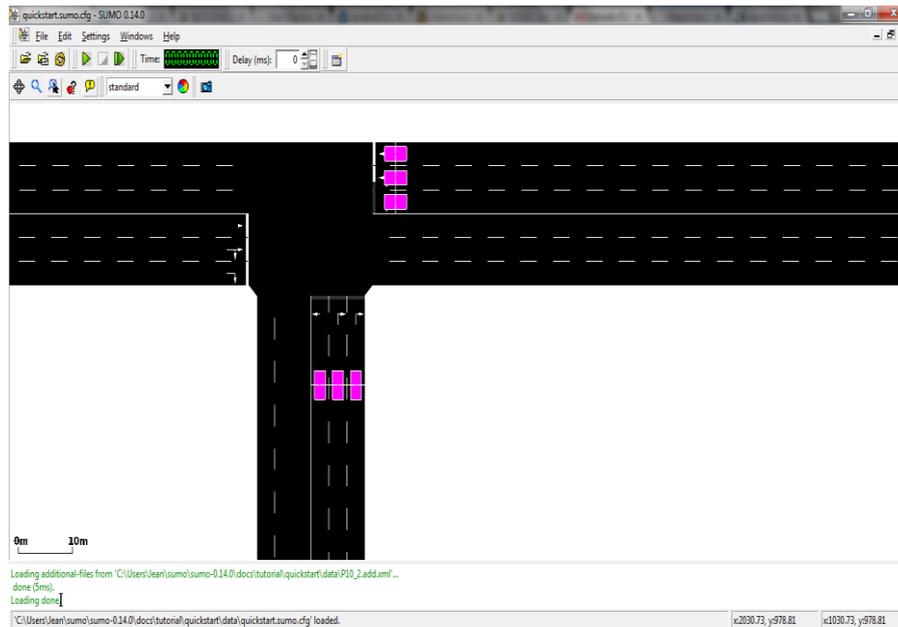
Fonte: PDTU

Os tipos de veículos que entraram na simulação foram: veículos de passeio, táxi, ônibus e caminhão. Considerou-se que Caminhao\_1, Caminhao\_2, Caminhao\_3 e Caminhao\_4 são respectivamente: material de construção, alimentos, bebidas e produtos químicos e petroquímicos. A carga geral foi considerada como sendo transportada pelos demais tipos de caminhão.

### 5.4.2.3 Simulação

Para a simulação, foram definidas quinze rotas segundo as possibilidades para a configuração do modelo. Os dados foram gerados para sete dias da semana e para cada um desses dias foram definidos os fluxos de veículos. Os veículos foram configurados segundo sua rota em duas categorias: fluxos fixos, que todos os dias realizam o mesmo caminho, como é o caso dos ônibus, e fluxos aleatório, como os táxis.

As antenas de leitura RFID também foram simuladas no SUMO. Vinte e nove delas foram espalhadas pela rede viária em dez pontos. Como o SUMO permite apenas que se coloque uma antena por pista, cada ponto escolhido recebeu duas ou três antenas, dependendo do número de pistas que a via possui (Figura 5.16). Nela podem-se ver dois pontos cobertos pelas antenas. Em cada um desses dois pontos existem três antenas representadas pelos retângulos em rosa, pois as vias onde esses pontos estão localizados possuem três pistas.



**Figura 5.16:** Simulação das antenas de RFID

Ao passar por essas antenas, é guardado o ID (identificador do veículo) dos veículos, a velocidade e o tempo em que ele entrou e saiu da área da antena, o seu tipo e tamanho. Ao final da simulação, cada antena gera um arquivo XML contendo essas informações. Esse arquivo é a base de dados que, depois de ser tratada, servirá como entrada para etapa da Mineração de Dados.



**Figura 5.17:** Simulação em andamento.

A Figura 5.17 mostra uma simulação em andamento, onde em amarelo estão os veículos que podem ser identificados pelo seu tamanho: ônibus, caminhões e carros. Na parte inferior da figura, pode-se notar algumas informações de colisões entre os

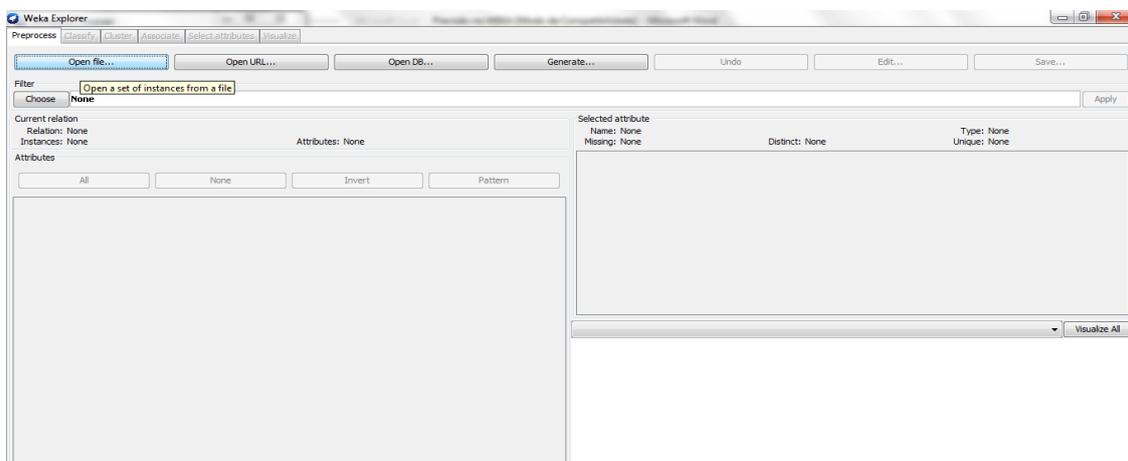
veículos, que geralmente acontecem em horário de pico em que a frequência de veículos das ruas é superior a sua capacidade.

Finalmente, com todos os elementos preenchidos, a lista de elementos é percorrida para criar o arquivo XML de saída que será usado na mineração dos dados.

Para validação do Modelo, foram criados três cenários, a seguir apresentados:

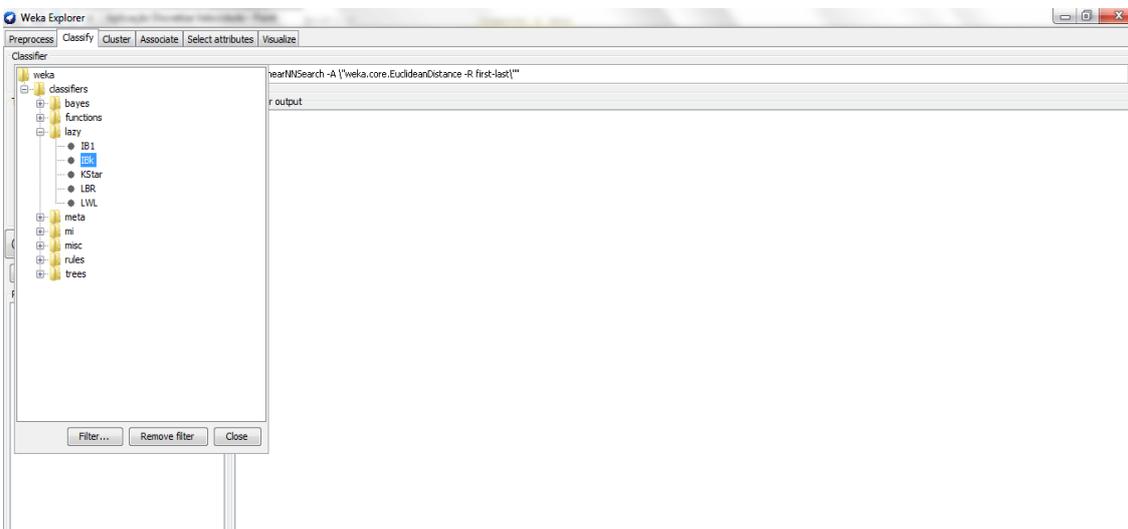
- Cenário 1 – Previsão utilizando todos os tipos de veículos

A Figura 5.18 mostra a tela inicial da versão *explorer* do WEKA em que o arquivo ARRF é escolhido.



**Figura 5.18:** Weka Explorer

A tela representada na Figura 5.19 disponibiliza a lista de algoritmos disponíveis. O algoritmo selecionado para previsão é o IBK, algoritmo do “Vizinho mais Próximo”.



**Figura 5.19:** Tela de escolha do algoritmo

Após a execução, o WEKA disponibiliza uma saída para análise da precisão do algoritmo escolhido que é mostrado na Figura 5.20.

Os conceitos básicos relativos aos campos que tratam da precisão dos resultados do WEKA são :

- **Correctly Classified Instances:** Quantidade de instancias classificadas corretamente em valores absolutos e relativos;
- **Incorrectly Classified Instances:** Quantidade de instancias classificadas incorretamente em valores absolutos e relativos;
- **Kappa Statistic:** Índice que compara o valor encontrado nas observações com aquele que se pode esperar do acaso. É o valor calculado dos resultados encontrados nas observações e relatado como um decimal (0 a 1). Quanto menor o valor de Kappa, menor a confiança de observação, o valor 1 implica a correlação perfeita;
- **Mean Absolute Error:** média da diferença entre os valores atuais e os preditos em todos os casos, é a média do erro da predição.
- **Root Mean Squared Error:** reduz o quadrado do erro relativo na mesma dimensão da quantidade sendo predita incluindo raiz quadrada. Assim como a raiz quadrada do erro significativo (root mean-squared error), este exagera nos casos em que o erro da predição foi significativamente maior do que o erro significativo;
- **Relative Absolute Error:** É o erro total absoluto. Em todas as mensurações de erro, valores mais baixos significam maior precisão do modelo, com o valor próximo de zero temos o modelo estatisticamente perfeito;
- **Total Number of instances:** Numero total de instancias;

No segmento Detailed Accuracy By Class tem-se:

- **TP Rate - True Positives:** são os valores classificados verdadeiramente positivos;
- **FP Rate - False Positives:** são os falsos positivos, são os dados classificados erroneamente como positivos pelo classificador;

- Precision: É o valor da predição positiva (número de casos positivos por total de casos cobertos), muito influenciada pela especificidade e pouco pela sensibilidade. Sensibilidade é o número de casos positivos que são verdadeiramente positivos e especificidade é o número de casos negativos que são verdadeiramente negativos;
- Recall : É o valor da cobertura de casos muito influenciada pela sensibilidade e pouco pela especificidade. É calculada por número de casos cobertos pelo número total de casos aplicáveis;
- F-Measure: Usada para medir o desempenho, pois combina valores de cobertura e precisão de uma regra numa única fórmula [ $2 * \text{Prec} * \text{Rec} / (\text{Prec} + \text{Rec})$ ].

As principais grandezas analisadas para identificar os padrões foram:

- Percentual de acerto de previsão (*Correctly Classified Instances*);
- Erro relativo absoluto (*Relative Absolute Error*).

A seguir serão apresentados os relatórios dos estudos elaborados:

Cenário 1 – Considerando todos os veículos

IB1 instance-based classifier  
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Stratified cross-validation ===  
=== Summary ===

Correctly Classified Instances	742	84.6066 %
Incorrectly Classified Instances	135	15.3934 %
Kappa statistic	0.7583	
Mean absolute error	0.053	
Root mean squared error	0.2257	
Relative absolute error	25.0107 %	
Root relative squared error	69.4 %	
Total Number of Instances	877	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.857	0.007	0.833	0.857	0.845	0.932	P2_0
	0.78	0.006	0.865	0.78	0.821	0.884	P2_1
	0	0	0	0	0	?	P2_2
	0.954	0.013	0.971	0.954	0.962	0.974	P5_0
	0.861	0.132	0.867	0.861	0.864	0.863	P5_1
	0.42	0.073	0.37	0.42	0.393	0.678	P5_2
Weighted Avg.	0.846	0.078	0.853	0.846	0.849	0.885	

=== Confusion Matrix ===

	a	b	c	d	e	f	<-- classified as
30	5	0	0	0	0	0	a = P2_0
6	32	0	0	2	1	0	b = P2_1
0	0	0	0	0	0	0	c = P2_2
0	0	0	268	9	4	0	d = P5_0
0	0	0	8	378	53	0	e = P5_1
0	0	0	0	47	34	0	f = P5_2

**Figura 5.20:** Relatório – Todos os Veículos

Nesse relatório, as informações analisadas são:

- Correctly Classified Instances: Quantidade e percentual de acerto da previsão.

Nessa execução, o percentual de acerto foi de 84,6066%.

- Incorrectly Classified Instances: Quantidade e percentual de erro da previsão.

Nessa execução, o percentual de erro foi de 15,3934%.

- Relative Error-Erro relativo: 25,0107%;

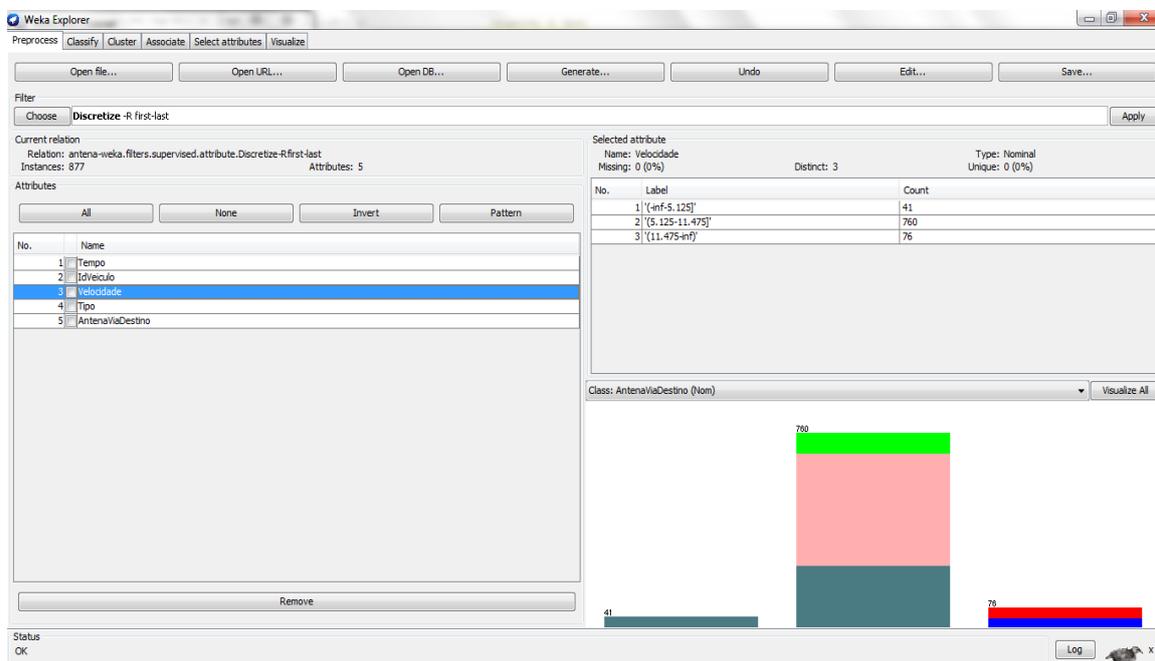
- Confusion Matrix: Os acertos são representados na diagonal principal da matrix e os erros nos outros elementos.

Por exemplo: Na primeira linha e na segunda coluna está marcado de vermelho o número “5”. Isso representa que o algoritmo errou 5 vezes, prevendo que o movimento era para a=P2\_0, mas o correto seria b=P2\_1.

Este relatório foi considerado insatisfatório face aos elevados índices de erros.

Na próxima etapa agrupou-se os veículos segundo sua velocidade.

A tela seguinte (Figura 5.21) mostra as agregações realizadas para o atributo velocidade.



**Figura 5.21:** Utilização do filtro Discretize

Executando o algoritmo novamente após o agrupamento, obtém-se uma melhora considerável de aproximadamente 5% (aumentando o percentual de acertos de 84% para 89% e diminuindo o erro relativo de 25% para 20%), conforme mostrado na Figura 5.22, comprovando a importância da fase de pré-processamento na Mineração de Dados.

```

01:34:30 - lazy.IBk

IB1 instance-based classifier
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      781      89.0536 %
Incorrectly Classified Instances    96       10.9464 %
Kappa statistic                     0.8191
Mean absolute error                 0.0439
Root mean squared error            0.1495
Relative absolute error             10.7296 %
Root relative squared error        45.9603 %
Total Number of Instances          877

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
          1      0.011    0.795     1      0.886     0.994    P2_0
          0.78    0      1      0.78    0.877     0.995    P2_1
          0      0      0      0      0      ?      P2_2
          1      0      1      1      1      1      P5_0
          0.973   0.171    0.851    0.973    0.908     0.965    P5_1
          0.074   0.015    0.333    0.074    0.121     0.895    P5_2
Weighted Avg.  0.891    0.088    0.855    0.891    0.862     0.972

=== Confusion Matrix ===

  a  b  c  d  e  f  <-- classified as
35  0  0  0  0  0 |  a = P2_0
 9  32  0  0  0  0 |  b = P2_1
 0  0  0  0  0  0 |  c = P2_2
 0  0  0 281  0  0 |  d = P5_0
 0  0  0  0 427 12 |  e = P5_1
 0  0  0  0 75  6 |  f = P5_2

```

Figura 5.22: Relatório com agrupamento por faixas de velocidade

Embora os índices tivessem melhorado com o agrupamento, o que demonstra o acerto deste filtro os valores de erros permanecem elevados e o experimento é considerado insatisfatório.

- Cenário 2 – Previsão utilizando o tipo caminhão

Procedendo-se novo processamento mantendo o agrupamento por velocidade e considerando apenas a categoria caminhão, os seguintes resultados foram obtidos (Figura 5.23):

```

Classifier output
IB1 instance-based classifier
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      1492      89.4484 %
Incorrectly Classified Instances    176      10.5516 %
Kappa statistic                    0.8763
Mean absolute error                 0.022
Root mean squared error             0.1449
Relative absolute error             12.8659 %
Root relative squared error        49.5734 %
Total Number of Instances          1668

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.943    0.024    0.93      0.943    0.936      0.96      P4_0
      0.812    0.02     0.789    0.812    0.8        0.896     P4_1
      0         0         0         0         0          ?         P4_2
      1         0         1         1         1          1         P5_0
      0.774    0.005    0.9       0.774    0.832     0.823     P7_1
      0.828    0.02     0.848    0.828    0.838     0.915     P7_2
      0.967    0.018    0.914    0.967    0.94       0.966     P8_0
      0.829    0.018    0.856    0.829    0.842     0.908     P8_1
      0.865    0.011    0.726    0.865    0.789     0.923     P8_2
      0.874    0.006    0.942    0.874    0.907     0.925     P9_0
Weighted Avg.  0.894    0.016    0.896    0.894    0.894     0.935

```

**Figura 5.23:** Relatório com dados da categoria caminhão

O índice de acertos nessa simulação também foi próximo de 89%, indicando regularidade do modelo. Houve também uma redução considerável no erro relativo (de 20% para 12%). Isso se justifica pela exclusão dos registros de veículos do tipo táxi, que são veículos que não possuem padrão frequente de deslocamento devido à mudança de trajeto a cada nova “corrida”.

- Cenário 3 – Previsão sem os registros do tipo táxi

O último cenário simulado tem todos os tipos de veículos, exceto os táxis que não possuem origem e destino definidos. Esses veículos não apresentam padrões de deslocamento, dificultando a previsão feita pelo algoritmo. Com a exclusão desses

registros, o algoritmo de classificação obteve o melhor resultado, com 96,88% de acertos e 6,21% de erro relativo. A Figura 4.24 mostra os resultados encontrados.

```

Classifier output

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      466          96.8815 %
Incorrectly Classified Instances    15           3.1185 %
Kappa statistic                    0.9511
Mean absolute error                 0.0158
Root mean squared error            0.1113
Relative absolute error             6.2103 %
Root relative squared error        31.2026 %
Total Number of Instances          481

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.961    0.02    0.978     0.961   0.969     0.964    P4_0
                0.923    0.024  0.857     0.923   0.889     0.953    P4_1
                0        0       0         0       0         0.954    P4_2
                1        0       1         1       1         1        P5_0
                1        0       1         1       1         1        P5_1
Weighted Avg.   0.969    0.013  0.968     0.969   0.968     0.976

=== Confusion Matrix ===

  a  b  c  d  e  <-- classified as
222  9  0  0  0 |  a = P4_0
  5 60  0  0  0 |  b = P4_1
  0  1  0  0  0 |  c = P4_2
  0  0  0 162  0 |  d = P5_0
  0  0  0  0  22 |  e = P5_1

```

Figura 5.24: Relatório excluindo os registros do tipo táxi

#### 5.4.2.4 Considerações Finais

Na previsão utilizando padrões frequentes de deslocamento feita no Weka, foi possível obter a previsão da malha viária e também a matriz OD.

Os pesquisadores Guozhen *et al.* (2011), no artigo “*Dynamic OD Estimation Using Automatic Vehicle Location Information*”, propuseram um modelo matemático determinístico designado “*OD estimation with automatic vehicle location*” ou ODAVL para estimar a Matriz Dinâmica de OD. Os resultados obtidos serão utilizados para comparar com aqueles obtidos nesta pesquisa.

A base de ambos os estudos são dados gerados automaticamente pelos próprios veículos, porém, nem todos os veículos estarão equipados com dispositivos que possam gerar esses dados, como GPS e RFID. Para resolver esse aspecto, alguns pesquisadores primeiro estimam a taxa de penetração no mercado de veículos com esses dispositivos e utilizam os dados com a taxa de penetração no mercado (percentual de veículos equipados) para estimar a Matriz OD.

A taxa de penetração no mercado MPR pode ser conceituada, no presente estudo, como sendo a relação entre o total de veículos equipados com a *Tag RFID* que estiveram na área de cobertura da antena e o total de *Tags* de veículos lidos corretamente.

Para apurar o desempenho do modelo ODAVL, foi desenvolvido um método para calcular o erro absoluto (RE – *Relative Error*) relativo em relação à taxa de penetração no mercado (MPR – *Market Penetration Rate*) desses equipamentos, dando origem ao Quadro 5.3.

**Quadro 5.3:** Erro relativo por MPR do modelo ODAVL

MPR	10%	20%	30%	40%	50%	60%	70%	80%	90%
RE	12,2	10,96	9,83	7,63	6,46	5,41	4,74	4,16	2,62

No modelo ODAVL, veículos que não possuem origem ou destino regular em um determinado tempo não são considerados pelo modelo, pois não é possível determinar a matriz OD. Os táxis, por exemplo, são veículos excluídos do modelo ODAVL.

No Quadro 5.3, observa-se que o erro relativo dos dois modelos se aproxima. É importante salientar que a comparação é do ODAVL com 70% de MPR, considerando que o percentual médio de veículos que são captados pelo protótipo da antena do projeto SINIAV é de 70%, conforme constatou o autor desta tese, que participou de alguns testes desse protótipo na qualidade de observador. Outro critério de comparação é que todos os veículos sem origem e destino definidos são excluídos do ODAVL, mas não do Modelo Proposto, pois não é possível identificar todos os veículos que se enquadram nesse critério.

**Quadro 5.4:** Comparação dos modelos ODAVL e Modelo Proposto

Modelo	Critério de filtragem	Erro Relativo (%)
ODAVL	Sem veículos com rota predeterminada	4,74%
Modelo Proposto	Apenas os táxis não foram considerados	6,21%

Dessa forma, pode-se concluir, pela adequabilidade para previsão da movimentação de veículos usando-se dados de sistemas RFID, o algoritmo do vizinho mais próximo, agrupando-se dados de velocidade de veículo com a exclusão dos táxis.

## 5.5 Abordagem Representação do Conhecimento

### 5.5.1 Considerações Iniciais

O objetivo da abordagem apresentada é capturar e especificar, de forma geral e formal o conhecimento para o domínio da Engenharia de Transportes aplicável à determinação de padrões frequentes de movimentação da carga rodoviária a partir da identificação da carga e da localização dos veículos por sistemas RFID, considerando aspectos relativos ao contexto dessa movimentação.

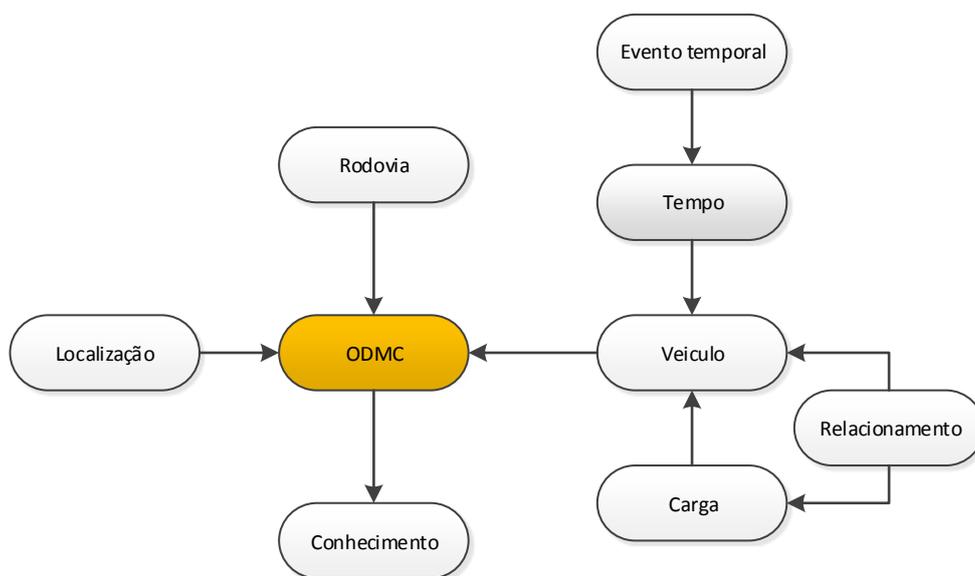
O processo de concepção do modelo pode ser generalizado como aquele proposto por Noy e Mcguinness (2001), no qual ontologias podem ser construídas desde o início, ou pelo reuso de outras ontologias. As ontologias que compõem o modelo foram inspiradas no modelo SeCom (BULCÃO NETO, 2006), que possibilita a concepção de sistemas sensíveis a contexto.

Prevalece, em geral, o reuso de definições de metadados e de ontologias existentes na *Web*. Com vistas a facilitar a sua extensibilidade, o modelo é composto de um conjunto modular de ontologias inter-relacionadas baseadas nas dimensões semânticas de identidade do veículo, da carga, geometria da rodovia, localização e características do equipamento de leitura e tempo de deslocamento.

A disposição das ontologias que compõem o modelo segue uma abordagem em duas camadas: a camada superior de ontologias, Figura 5.25, que representa o contexto do modelo, e a camada inferior de ontologias, que poderá ser construída por um projetista de uma aplicação sensível a contexto em trabalho futuro. Nesse caso, o modelo pode ser reusado e/ou mesmo estendido com o conhecimento que é particular dessa aplicação.

### 5.5.2 Diagrama de Contexto

A versão preliminar do diagrama de contexto para o sistema proposto é apresentado na Figura 5.25, que apresenta as principais entidades envolvidas e seu relacionamento. Cada uma das ontologias que compõe o modelo proposto é apresentada a seguir:



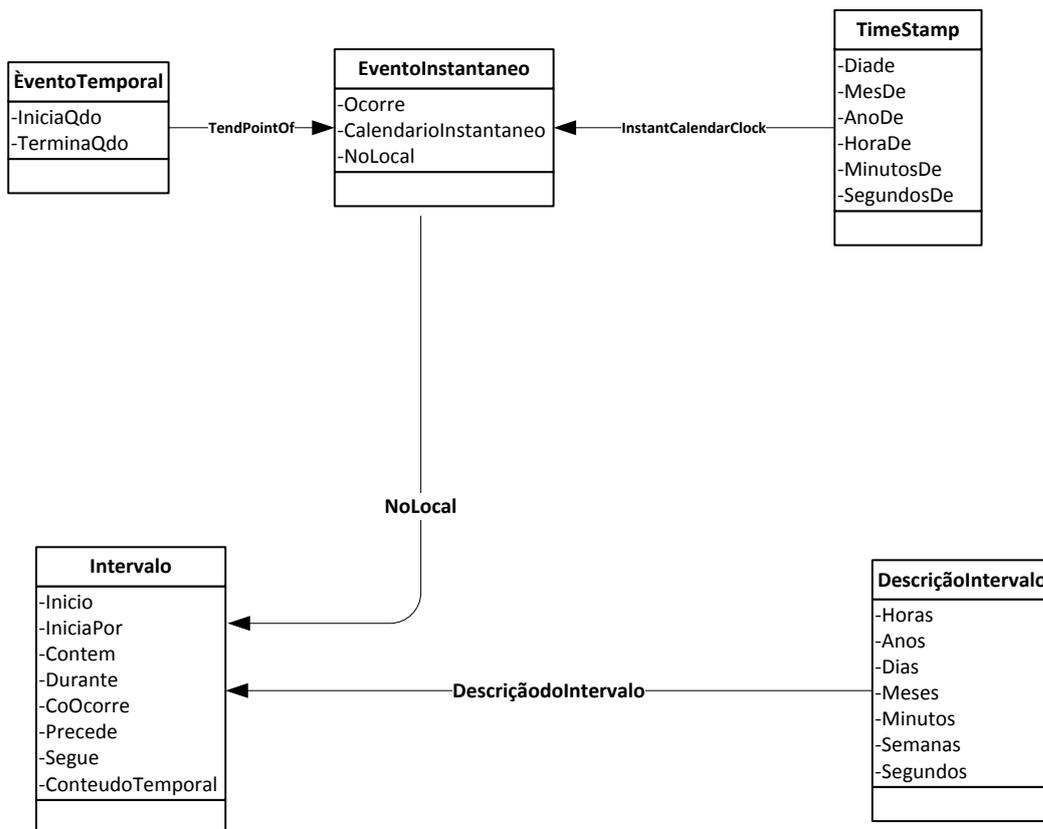
**Figura 5.25:** Diagrama de Contexto

### 5.5.3 Ontologia Tempo

A ontologia representa um tipo de informação que envolve conhecimento de senso comum e não necessita de especialista para aquisição de conhecimento e apresenta as seguintes diretrizes de projeto:

- Utilização de intervalos de tempo como primitivas do modelo, pois, na prática, em um sistema baseado em dados RFID, trabalha-se mais com informação temporal discreta;
- Representação de relações temporais entre entidades que possuam um componente temporal, como as relações de passado, passado imediato, presente, futuro imediato e futuro;
- Múltiplas granulosidades de representação de informação temporal, como horas, dias da semana, meses e anos;
- Associação de informação temporal a conceitos do modelo, o que permite não apenas inferências sobre os inter-relacionamentos temporais entre entidades de um ambiente de computação sensível a contexto, mas também o registro histórico de padrões de movimentação.

A Figura 5.26 ilustra a ontologia Tempo e suas principais classes, atributos e relações.



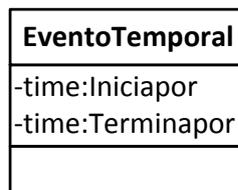
**Figura 5.26:** Ilustração da Ontologia Tempo

- Descrição semântica

A principal classe da ontologia apresentada na Figura 5.26 é a Intervalo, que descreve um evento que contenha uma extensão temporal, quer seja uma extensão por instante de tempo (classe EventoInstantaneo), quer seja uma extensão por evento temporal (classe EventoTemporal). É possível representar qualquer veículo ou carga, a partir do modelo, que possua uma extensão temporal, quer seja um instante de tempo, quer seja um intervalo de tempo, bem como as possíveis relações temporais que podem existir entre esses elementos.

### 5.5.4 Ontologia EventoTemporal

A Figura 5.27 descreve a ontologia EventoTemporal. Esta ontologia tem o papel de complementação da modelagem de informação temporal realizada pela ontologia Tempo, descrita no item anterior.



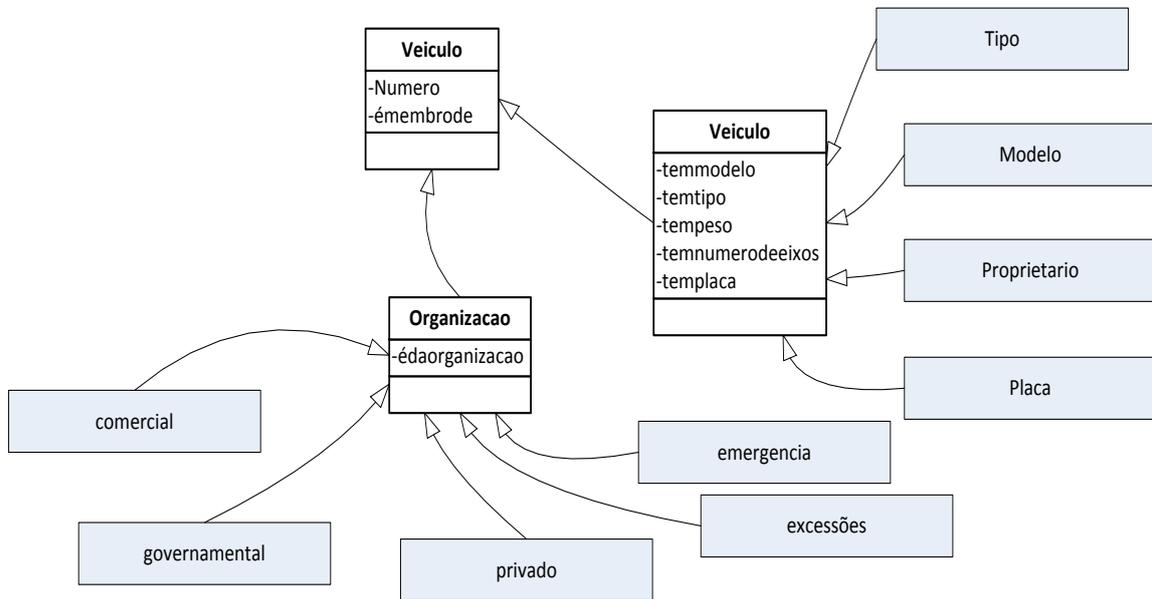
**Figura 5.27:** Ontologia EventoTemporal

- Descrição Semântica

A ontologia EventoTemporal é uma extensão da ontologia Tempo que modela eventos com extensão temporal, onde o tempo representa o espaço de nomes associados à ontologia Tempo e é composta pelas dimensões Iniciapor e Terminapor (Figura 5.27), que representam os instantes de início de término do evento.

### 5.5.5 Ontologia Veículo

A Figura 5.28 descreve a ontologia Veículo, suas principais classes, atributos e relações.



**Figura 5.28:** Ontologia Veículo

- Descrição Semântica

A Ontologia Veículo representa as informações do veículo armazenadas nos órgãos oficiais ligados aos Ministérios das Cidades (DENATRAN) e do Transporte (ANTT).

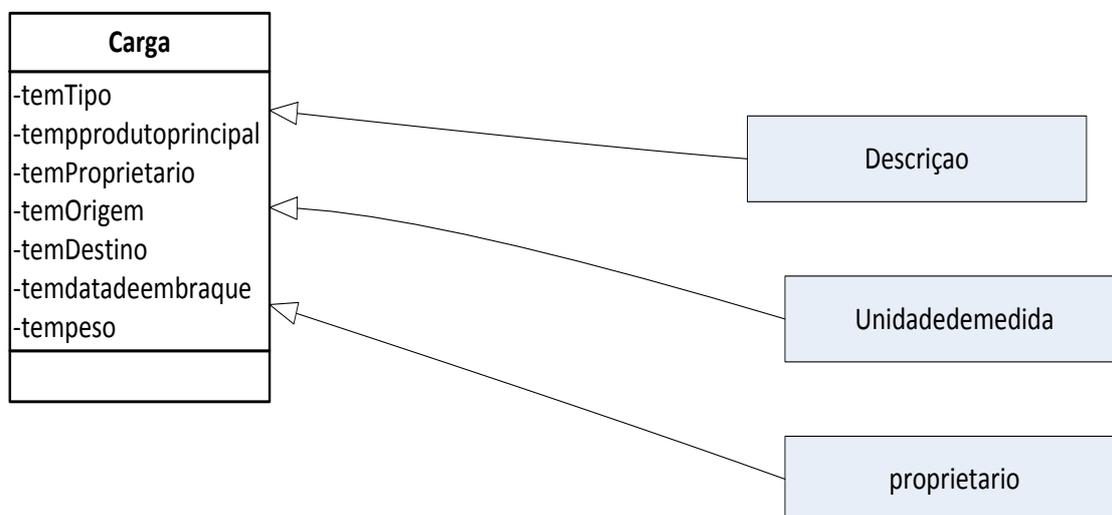
### 5.5.6 Ontologia Localização

A Figura 5.29 representa a localização das antenas onde são realizadas as leituras das passagens dos veículos.



### 5.5.7 Ontologia Carga

A Figura 5.30 representa a ontologia carga, suas classes, atributos e relações.



**Figura 5.30:** Ontologia Carga

A Ontologia Carga representa as informações obtidas pelo acesso aos Bancos de Dados do CT-e e modelam as principais características da carga transportada.

### 5.5.8 Ontologia Rodovia

A ontologia rodovia modela as redes rodoviárias a partir dos segmentos das estradas. As redes rodoviárias deste trabalho podem ser representadas pelos arcos de redes conectados por nós, onde os arcos correspondem aos trechos rodoviários delimitados pelos nós relacionados às intersecções com outras vias.

### 5.5.9 Ontologia ODCM

A ontologia ODCM modela o conjunto de entidades que podem realizar ações em um ambiente de computação sensível a contexto. Essa ontologia representa atores relacionados ao domínio do transporte de carga e é composta de conceitos que descrevem o perfil de atores sob a visão da movimentação de carga, cuja passagem por postos de fiscalização pode gerar informações das *tags* RFID instaladas nas mercadorias e nos veículos.

A ampliação do foco de sistemas identifica padrões frequentes de movimentação rodoviária de carga para uma ontologia de atores que integrem, além destes, os aspectos

logísticos independente de domínio. As informações que descrevem o perfil de atores foram modularizadas de forma a representar as ontologias de apoio, conforme ilustra a Figura 5.1. Dessa forma, a ontologia ODCM modela apenas classes, atributos e relações básicas com respeito à descrição de atores.

## 5.6 Trabalhos apresentados em Congressos Internacionais

Os conceitos relativos ao modelo Representação do Conhecimento com o uso de ontologia foram apresentados em três congressos internacionais, a saber:

- Context-Aware Computing Applicable to Brazilian Nationwide Automatic Vehicle Identification (AVI), 18Th World Congress on Intelligent Transportation Systems, 2011, Orlando EUA;
- Freight Truck Assignment from RFID Flow Analysis, 19Th ITS World Congress, 2012, Viena, Austria;
- Study on RFID Data based Freight Transportation Planning, 13Th WCTR World Congress on Transportation Research, 2013, Rio de Janeiro, Brasil.

## 5.7 Considerações Finais

O modelo ontológico apresentado considera aspectos da movimentação de carga segundo as alterações no ambiente que a circunda. O desenvolvimento de aplicações sensíveis a contexto e sua representação segundo os paradigmas da Ontologia começam a ter destaque na área de Engenharia de Transportes, uma vez que é centrado nos objetivos do usuário e não nos dados.

A abordagem apresentada foi capaz de capturar e especificar, de forma geral e formal o conhecimento para o domínio da Engenharia de Transportes aplicável à determinação de padrões de movimentação da carga rodoviária a partir de elementos sensível a contexto com o uso da Tag prevista para o sistema SINIAV. considerando aspectos relativos ao contexto dessa movimentação. Sua maior aplicabilidade se encontra para os processos logísticos onde o entendimento de aspectos relativos a historia da movimentação poderão contribuir para a melhoria do sistema.

## Capítulo 6 – Conclusões e Recomendações

Os estudos de condução dessa Tese envolveram a análise de sistemas nacionais que pudessem fornecer elementos para o planejamento do transporte sem acréscimo de custos, com alta confiabilidade e diversidade de informações. A seguir serão apresentadas as principais conclusões obtidas:

### 6.1 Conclusões do trabalho

O objetivo desta tese consistiu em desenvolver um método para identificar os padrões frequentes de movimentação de cargas e veículos a partir de sistemas nacionais de identificação por radiofrequência segundo os paradigmas dos Sistemas Inteligentes de Transportes – ITS. de Transportes – ITS.

Da revisão bibliográfica relativa aos processos de planejamento dos transportes pode-se concluir que:

- O principal modelo empregado para o Planejamento dos Transportes é o das Quatro Etapas;
- O levantamento de dados para a construção da Matriz OD emprega percentual significativo de recursos ficando limitado quanto a quantidade de dados, precisão dos dados levantados e variedade de dados;
- Para a construção da Matriz OD de carga, os dados levantados por fontes secundárias, junto aos operadores logísticos entre outros, apresentam estrutura de dados distintos envolvendo significativo esforço para a compatibilização com os estudos elaborados;
- Dentre as críticas aos modelos, dentre outros, pode-se citar a ausência de análises relativas a :
  - Políticas Públicas nas definições dos modelos;
  - Desempenho dos modelos;
  - Corredores de comércio internacionais;
  - Acidentes, roubos de cargas, tamanhos de caminhões e excesso de peso.

O método proposto nessa Tese demonstrou que a integração dos dados dos sistemas de informação nacionais identificadas possibilitam:

- Que as deficiências no levantamento e tratamentos de dados de transporte de carga apontados pelo relatório Freight Demand Modeling and Data Improvement (TRB, 2013), entre outros, para a construção do Planejamento do Transporte segundo o Modelo das Quatro Etapas podem ser parcialmente supridas pelo Método Proposto;
- A possibilidade de esses dados serem reusados para a construção da Matriz OD e Alocação de Viagens necessárias ao planejamento dos transportes sem acréscimo de custo na implantação de infraestrutura.
- Que a coleta e validação permanente de dados recomenda seu uso para os propósitos formulados nesta tese, reduzindo-se os custos e os esforços no levantamento e tratamento de dados.
- Que o método sustenta elementos de três pilares distintos dentro do domínio da Engenharia de Transportes, a saber: Planejamento do transporte rodoviário de carga; Movimentação urbana e Logística.

Os modelos propostos nesta tese possuem um aspecto qualitativo. Sua originalidade deriva das limitações e das potencialidades de outros trabalhos, já que contempla o levantamento de dados necessários ao Planejamento de Transportes que não são usualmente realizados. Além disso, sua originalidade se destaca pela capacidade de obter informações a baixo custo e compreender o deslocamento de cargas em trechos rodoviários rurais.

Esta tese apresenta um aprimoramento, adaptado do Modelo PNLT, contudo tem características diferenciadas que se ajustam às necessidades deste estudo, visto a possibilidade de utilização de dados de transporte de carga obtidos com fins de fiscalização. Esta proposta contempla um diferencial para o tratamento de dados desagregados para veículos e carga.

Por meio deste método, pretende-se construir uma base de conhecimento que possibilite definir a necessidade e as estratégias para estabelecer uma infraestrutura rodoviária que favoreça a redução de custos logísticos. Como resultado, foi possível

identificar, nesta etapa da pesquisa e por simulação computacional, a viabilidade do procedimento proposto.

## 6.2 Recomendações para trabalhos futuros

Os estudos demonstraram a possibilidade de serem desenvolvidos estudos relativos ao planejamento de transporte que contemple novos aspectos, para tanto são recomendados trabalhos voltados ao:

- Aprimoramento das Ontologias para a *computação sensível a contexto* considerando a movimentação dos veículos de carga;
- Implementação de uma Política Pública que contemple os reais aspectos da movimentação de carga e a segurança da carga e do motorista;
- Desenvolvimento de modelos que considerem informações decorrentes dos índices de acidentes e de roubos de cargas;
- Recomenda-se, também, que os estudos para reuso dos dados relativos aos projetos Brasil-ID e SINIAV não fiquem contidos dentro do enfoque para controle fiscal e tributário, mas que sejam disponibilizados para o planejamento do transporte e para a logística empresarial.

## BIBLIOGRAFIA

ADOBE. Rich Internet Applications. **ADOBE**, [N.d.]. Disponível em: <[http://www.adobe.com/resources/business/rich\\_internet\\_apps/](http://www.adobe.com/resources/business/rich_internet_apps/)>. Acesso em: 21 jun. 2013.

AGRAWAL, R.; IMIELINSKI, T.; SWAMI, A. A Mining association rules between set of items in large databases. In: Proceedings of the ACM SIGMOD INTERNACIONAL CONFERENCE ON MANAGEMENT OF DATA, 1993. New York, 1993. P. 207-216.

ALIBABAI, H.; MAHMASSANI, H. S. Dynamic origin-destination demand estimation using turning movement counts. **Transportation Research Record: Journal of the Transportation Research Board**, Washington, DC, v. 2085, p. 39-48, 2008.

AGENCIA NACIONAL DE TRANSPORTES TERRESTRES-ANTT. **Relatório Anual 2010**. Brasília, 2011. Disponível em: <<http://www.antt.gov.br/relatorioanual/RelatorioAnualAntt2010.pdf>>. Acesso em: 05 out. 2011.

\_\_\_\_\_. **Registro Nacional de Transportadores Rodoviários de Cargas - RNTRC em números**, 2012. Disponível em: <[http://appweb2.antt.gov.br/rntrc\\_numeros/rntrc\\_TransportadorFrotaVeiculo.asp](http://appweb2.antt.gov.br/rntrc_numeros/rntrc_TransportadorFrotaVeiculo.asp)>. Acesso em: 12 jun. 2012.

BECKER, M.; SMITH, S. F. **An Ontology for Multi-Modal Transportation Planning and Scheduling**. Robotics Institute Carnegie Mellon University. Pittsburgh, PA 15213, p. 79. 1997.

BELL, M. G. H. The Estimation of an origin-destination matrix from traffic counts. **Transportation Science**, v. 1, p. 198-217, 1983.

BERRY, M. J. A.; LINOFF, G. **Data Mining Techniques for Marketing, Sales, and Customer Support**. New York: [s.n.], 1997.

BETTINI, C. et al. A survey of context modelling and reasoning techniques. **Pervasive and Mobile Computing**, v. 6, n. 2, p. 161-180, 2010.

BIERLAIRE, M.; CRITTIN, F. An Efficient Algorithm for Real-Time Estimation and Prediction of Dynamic OD Tables. **Operations Research**. v. 52, n.1, p. 116-127, January/February 2004.

BOLLA, R.; DAVOLI, F. Road Traffic Estimation from location Tracking Data in the Mobile Cellular Network. IN: IEEE WIRELESS COMMUNICATIONS AND NETWORK CONFERENCE, 2000. v. 3, p. 1107-1112.

BOOCH, G.; RUMBAUGH, G.; JACOBSON, I. **The unified modeling language user guide**. [S.l.]: Redwood:Addison Wesley, 2005.

MINISTERIO DA CIENCIA E TECNOLOGIA. **Sistema de Identificação, Rastreamento e Autenticação de Mercadorias**. Centro de Pesquisas Avançadas Wernher von Braun, 2009.

BULCÃO NETO, R. D. F. **Um processo de software e um modelo ontológico para apoio ao desenvolvimento de aplicações sensíveis a contexto**. 2006. Tese (Doutorado)-Instituto de Matemática e Ciência da Computação, Universidade de São Paulo, São Carlos, 2006.

CALABRESE, F. et al. Estimating Origin-Destination Flows Using Mobile Phone Location Data. **IEEE Pervasive Computing**, v. 10, n. 4, p. 36-44, nov. 2011. 36-44.

CAMBRIDGE SYSTEMATICS INC. A guidebook for forecasting freight transportation demand NCHRP report 388. Washington, DC. 1997.

CAREY, M.; RELVELLI, R. Constrained estimation of direct demand functions and trip matrices. **Transportation Science**, v. 3, p. 143-152, 1986.

CASCETTA, E. Estimation of Trip matrices traffic counts and survey data: a generalized least squares estimator. **Transportation Research, Part B.**, v. 16, n. 4-5, p. 289-299, 1984.

CASCETTA, E.; INAUDI, D.; MARQUIS, G. Dynamic Estimators of origin/destination matrices using traffic count. **Transportation Science**, v. 27, n. 4, p. 363-373, 1993.

CAYFORD, R.; JOHNSON, T. Operational Parameters Affecting Use of nanonymous Cellphone Tracking for Generating Traffic Information. Proc. Transportation Research Board Ann. Meeting., 2003.

CHANG, G. L.; WU, J. Recursive estimation time-varyng OD flows from traffic counts in freeway corridor. **Transportation Research**, v. 28B, n. 2, p. 141-160, 1994.

CHEN, P. **Modelagem de Dados: a Abordagem Entidade Relacionamento para Projeto Lógico**. São Paulo: McGraw-Hill, 1990.

CHENG, H. et al. Discriminative Frequent Pattern Analysis for Effective Classification, 2011. Disponível em:  
<[http://www1.se.cuhk.edu.hk/~hcheng/paper/icde07\\_hcheng.pdf](http://www1.se.cuhk.edu.hk/~hcheng/paper/icde07_hcheng.pdf)>. Acesso em: 21 jun. 2013.

CHOW, J. Y.; YANG, C. H.; REGAN, A. C. State-of-Art of freight forecast modeling: lessons learned and the road ahead. **Transportation**, v. 37, p. 1011-1030, 2010.

CNT-CONFEDERAÇÃO NACIONAL DO TRANSPORTE. **Plano CNT Logística - 2011**. Disponível em:  
<<http://www.cnt.org.br/Imagens%20CNT/PDFs%20CNT/Plano%20CNT%20de%20Log%C3%ADstica/PlanoCNTdeLog2011.pdf>>. Acesso em: 06 mar. 2002.

\_\_\_\_\_. CNT- CONFEDERAÇÃO NACIONAL DO TRANSPORTE. **Pesquisa CNT de Rodovias 2012**, 2012. Disponível em:  
<[http://pesquisarodovias.cnt.org.br/Relatorios/2012/RelatorioGeral2012\\_BaixaResolucao.pdf](http://pesquisarodovias.cnt.org.br/Relatorios/2012/RelatorioGeral2012_BaixaResolucao.pdf)>. Acesso em: 05 jul. 2013.

CONFAZ-CONSELHO NACIONAL DE POLITICA FAZENDÁRIA. **Conhecimento do Transporte Eletrônico**, 2007. Disponível em:  
<[http://www.fazenda.gov.br/confaz/confaz/ajustes/2007/AJ\\_009\\_07.htm](http://www.fazenda.gov.br/confaz/confaz/ajustes/2007/AJ_009_07.htm)>. Acesso em: 31 jan. 2012.

CONSTANTINOS, A.; BEN-AKIVA, M.; KOUTSOPOULOS, H. Incorporating automated vehicle identification data into origin-destination estimation. **Transportation Research Record**, Washington, DC, v. 1882, p. 37-44, 2004.

CONSTANTINOS, A.; RAMACHANDRAN, B.; HARIS, K. N. A Synthesis of emerging data collection technologies and their impact on traffic management applications. **European Transportation Research Rev.**, 2011. 139-148.

CRAINIC, T. G.; LAPORTE, G. Planning models for freight transportation. **European Journal for Operation Research**, v. 97, n. 3, p. 409-488, 1997.

CREMER, M.; KELLER, H. Dynamic identification of flows from traffic counts at complex intersections. In: INTERNATIONAL SYMPOSIUM TRANSPORTATION AND TRAFFIC THEORY, 8. Toronto, 1981. p. 121-142.

DENATRAN. Manual de Procedimentos para o tratamento de Polos Geradores de Tráfego. DENATRAN. Brasília, DF, 2001.

DESA, M. I.; WIBOWO, A. A review of KDD-data mining framework and its application in logistics and transportation. Networked Computing and Advanced Information Management (NCM), In: INTERNACIONAL CONFERENCE ON, 7., 2011 p. 175 - 180.

DEY, A. K. Understanding and Using Context. **Personal and Ubiquitous Computing**, London UK, v. 5, n. 1, p. 4-7, 2001.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTE-DNIT. **Estatísticas**. Departamento Nacional de Infraestrutura de Transportes, 2001. Disponível em: <<http://www.dnit.gov.br/rodovias/operacoes-rodoviaras/postos-de-monitoramento/estatisticas>>. Acesso em: 17 abr. 2012.

\_\_\_\_DNIT. Manual de Estudos de Tráfego - IPR 723, [S.l.], 2006.

\_\_\_\_DNIT. **Plano Nacional de Contagem de Trânsito**. Disponível em: <[http://www1.dnit.gov.br/rodovias/contagem/plano\\_contagem.htm#Como%20%C3%A9%20a%20Contagem%20Volum%C3%A9trica?](http://www1.dnit.gov.br/rodovias/contagem/plano_contagem.htm#Como%20%C3%A9%20a%20Contagem%20Volum%C3%A9trica?)>. Acesso em: 26 out. 2011.

\_\_\_\_DNIT **Plano-nacional-de-viacao**, 2013. Disponível em: <[http://www.dnit.gov.br/plano-nacional-de-viacao/snv-2012/04\\_SNV2013\\_Internet.pdf](http://www.dnit.gov.br/plano-nacional-de-viacao/snv-2012/04_SNV2013_Internet.pdf)>. Acesso em: 19 jun. 2013.

DU, J. Using spatial analysis to estimate link travel times on local roads. In: TRANSPORTATION RESEARCH BOARD ANNUAL MEETING, 85T. [S.l.]: [s.n.]. 2006.

ERLANDER, S.; NGUYEN, S.; STEWART, N. On calibration of the combined distribution/assignment model. **Transportation Research**, **13 B**, 1979.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, v. 17, n. 3, Fall 1996.

FDOT. Development of Revised Methodology for Collecting Origin-Destination Data. Florida Department of Transportation. Tampa, FL. 2006.

FHWA. **A Summary of Vehicle Detection and Surveillance Technologies use in Intelligent Transportation Systems**. Federal Highway Administration. [S.l.]. 2007. Disponível em  
<<http://www.fhwa.dot.gov/policyinformation/pubs/vdstits2007/04.cfm#Toc171306369>>  
Acesso em: 05/08/2013

\_\_\_\_\_. FHWA. Freight Management and Operations. **Federal Highway Administration**, 2010. Disponível em: <<http://ops.fhwa.dot.gov/freight/index.cfm>>. Acesso em: 25 jan. 2012.

\_\_\_\_\_. FHWA. The Freight Analysis Framework version 3 - **A Description of FAF3 Regional database and How It Is Constructed**, 16 June 2011. Disponível em: <<http://www.ops.fhwa.dot.gov/freight/>>. Acesso em: 01 mar. 2012.

FIORIELLO; FERMI, F.; BIELANSKA, D. The ASTRA model for strategic assessment of transport policies. **System Dynamics Review**, v. 26, n. 3, p. 283-290, jul./set. 2010.

FISK, C. S. On combining maximum entropy trip matrix estimation with user optimal assignment. **Transportation Research part B**, v. 22, p. 66-79, 1988.

FISK, C. S.; BOYCE, D. E. A note on trip matrix estimation from link traffic count data. **Transportation Research 17B**, 1983.

FRIESZ, T. L.; HOLGUÍN-VERAS, J. Dynamic game-theoretic models of urban freight: formulation and solution approach. In: \_\_\_\_\_ **Methods and Models in Transportation and Telecommunications, Advances in Spatial Science**. Berlin: Springer, 2005.

FSF. Free Software Foundation, S.d. Disponível em: <[www.fsf.org](http://www.fsf.org)>. Acesso em: 01 ago. 2013.

GIDOFALVI, ; PEDERSEN, T. B. Mining Long, Sharable Patterns in Trajectories. **GeoInformatica**, v. 13, n. 1, p. 27-55, mar. 2009.

GOLDSMITH, R.; PASSOS, E. L. **Data Mining - um Guia Prático**. [S.l.]: Campus, 2005.

GONZALEZ, H.; HAN, J.; LI, X. Flowcube: constructing RFID flowcubes for multi-dimensional analysis of commodity flows. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF VERY LARGE DATA BASES,32. [S.l.]: [s.n.]. 2006. p. 834-845. Disponível em <<http://dl.acm.org/citation.cfm?id=1164199>>. Acesso em: 05/08/2013

SECTRANS - SECRETARIA ESTADUAL DE TRANSPORTES. **Plano Diretor de Transporte Urbano**. Secretaria Estadual de Transportes. Rio de Janeiro. 2005.

GRUBER, T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. **International Journal Human-Computer Studies**, v. 43; n. 5-6, p. 907-928, 1993.

GU, C.-K.; XIAO-LI, D. Efficient mining of local frequent periodic patterns in time series database. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERTICS, 2009 [S.l.]: [s.n.]. 2009.

GU, Q.; PUNG, H. K.; ZHANG, D. Q. A service-oriented middleware for building. **Journal of Network and Computer Applications**, 2005, v. 28, n. 1, p. 1-18, 2005.

GUARINO, N. Guarino, N. (1998) Formal Ontology and Information. In: PROCEEDINGS OF FORMAL ONTOLOGY INFORMATION SYSTEMS, 1998, Amsterdam, 1998. 3-15. Disponível em: <[www.loa-cnr.it/Publications.html](http://www.loa-cnr.it/Publications.html)>. Acesso em: 02 fev. 2012.

GUOZHEN, T.; LINDONG, L.; YAODONG, W. Dynamic OD Estimation Using Automatic Vehicle Location Information. In: INFORMATION TECHNOLOGY AND ARTIFICIAL INTELLIGENCE CONFERENCE (ITAIC), 6, 2011; IEEE Joint International. [S.l.]: [s.n.]. 2011. p. 352-355.

GUR, Y. Determining an Origin-Destination trip table based on observed volumes. ORSA-In: TIMS ANNUAL MEETING. New York: [s.n.]. 1978.

GUR, Y. et al. Estimation of an origin-destination trip table based on observed link volumes and turning movements. Department of Transportation. Washington, DC. 1980.

HALLER, ; GONTARCZYK, J.; KOTINURMI, P. Towards a complete SCM ontology: the case of ontologising RosettaNet. In: ACM SYMPOSIUM ON APPLIED COMPUTING. Fortaleza, : ACM. 2008. p. 1467-1473.

HAODE, L.; JIAN, S.; ZHENGWANG, Z. Study on Process Reengineering of Transportation Planning. In: INTERNACIONAL CONFERENCE ON INTELLIGENT COMPUTATION TECHNOLOGY. [S.l.]: [s.n.]. 2008. p. 494-497.

HELAL, S. Programming pervasive spaces. **IEEE Pervasive Computing**. [S.l.]: [s.n.], v. 4, n. 1, p. 84-87, jan./mar. 2005.

HENSHER, D. A. **Handbook of transport modelling**. [S.l.]: Emerald Group Publishing, v. 1, 2001.

HENSHER, D. A.; BUTTON, K. J. Handbook of Transport Modelling. In: \_\_\_\_\_ **Handbooks in Transport**. Amsterdam: Pergamon, v. 1, 2000.

HENSHER, D.; FIGLIOZZI, M. A. Behavioural insights into the modeling of freight transportation and distribution systems. **Transportation Research B**, v. 41, n. 9, p. 921-923, 2007.

HESSE, M.; RODRIGUES, J. P. The transport geography of logistics and freight distribution. **Journal of Transportation**, v. 2, n. 3, p. 171-184, 2004.

HOLGUIN-VERAS, J.; THORSON, E.; OSBAY, K. Preliminary results of experimental economics application to urban goods modeling research, v. 1873, p. 9-16, 2004.

HUANG, C.-C.; LIN, S.-H. Sharing knowledge in a supply chain using the semantic web. **Expert Systems with applications**, v. 37, n. 4, p. 3145-3161, abril 2010.

HUIPING , C.; NIKOS , M.; CHEUNG, D. W. Discovery of Periodic Patterns in Spatiotemporal Sequences. **IEEE Transactions on Knowledge and Data Engineering**, Piscataway, NJ, USA , v. 19, n. 4, abril 2007.

IBGE. **Regioes de Influencia das Cidades**. IBGE. Rio de Janeiro, p. 201. 2007.

\_\_\_\_\_ Pesquisa Mensal de Emprego. **Indicadores**, 2011. Disponível em: <[www.ibge.gov.br](http://www.ibge.gov.br)>. Acesso em: 05 out. 2011.

IBM. IBM Rational Unified Process (RUP). **RUP: Best practices for design, implementation and effective project management**, [N.d.]. Disponível em: <<http://www-01.ibm.com/software/rational/rup/>>. Acesso em: 19 jun. 2013.

IRF. **The Vienna Manifestoon ITS**. International Road Federation - IRF. Vienna, p. 35. 2012.

ITSINTERNATIONAL. Airborne traffic monitoring - the future? **Its International**, [S.l.], n. feb 2013.

JIANMU, Y.; LEI, C. Ontology Based Coordination in RFID Network. In: **WIRELESS COMMUNICATIONS, NETWORKING AND MOBILE COMPUTING International Conference on**, 4, 2008. [S.l.]: [s.n.]. 2008. p. 1-4.

JONG, G. D.; HUGH, G.; WALKER, H. National and International Freight Transport Models: Overview and Ideas for Future Development. **Transportation Review**, v. 1, n. 24, p. 103-124, 2004.

KORNER, C.; MAY, M.; WROBEL, S. Spatiotemporal Modeling and Analysis - Introduction. **Kusntl Intell**, p. 2015-221, 2012. 215-221.

KUMAR, R.; VERNA, R. D. KDD Techniques a Survey. **International Journal of Electronics and Computer Science Engineering**, v.1, n. 4, p. 2042, 2012

KUTZ, M. **Handbook of transportation engineering**. 1. ed. [S.l.]: McGraw-Hill Prof Med/Tech, v. 1, 2011.

L.G., W. **Estimating time-dependent trip matrices from traffic counts**. In: PROCEEDING OF THE NINTH INTERNATIONAL SYMPOSIUM ON TRANSPORTATION AND TRAFFIC THEORY. [S.l.]: VUN Science Press. 1984. p. 397-411.

LEE, J.-G. et al. Mining Discriminative Patterns fo Classifyng Trajectories on Road Networks. **IEEE Transactions on Knowledge ans Data Engineering**, v. 23, n. 5, p. 713-726, maio 2011..

LEUNG, H.; EL FAOUZI, N.-E.; KURIAN, A. Intelligent Transportation System (ITS). **Information Fusion**, v. 12, n. 1, p. 2-3, 2011.

LIEDTKE, G.; SCHEPPERLE, H. Segmentation of the transportation market with regard to activicty-based freight transport modelling. **International Journal of Logistics**, p. 199-218, 2004.

LIU, Y. et al. Mining Frequent Trajectory Patterns Activity Monitoring Using Radio Frequency Tags. In: PROCEEDINGS OF THE FIFTH IEEE INTERNATIONAL CONFERENCE ON PERVASIVE COMPUTING AND COMMUNICATIONS,7. [S.l.]: [s.n.]. 2007. p. 37-46.

MCNALLY, M. G. **The Four Step Model**. Institute of Transporttion Studies - Univerity of California. Irvine, CA, p. 19. 2007.

MEIXELL, M. J.; GARGEYA, V. B. Global Supply Chain Design: a literature review and critique. **Transportation Research Part E. Logistics and Transportation Review**, v. 41, n. 6, p. 531-550, 2005.

MENA, J. **Data Mining Your Website**. [S.l.]: [s.n.], 1999.

MICHAEL BEHRISCH, L. B. J. E. A. D. K. SUMO - Simulation of Urban MObility: An Overview. In: THE THIRD INTERNATIONAL CONFERENCE ON ADVANCES IN SYSTEM SIMULATION. Barcelona, Spain: [s.n.]. 2011. p. 63-68.

MIRCHANDANI, P. B.; GENTILI, M.; HE, Y. Location of vehicle identification sensors to monitor travel-time performance. **IET Intelligent Transport Systems**, v. 3, n. 3, p. 289-303, 20 feb 2009..

MINISTERIO DOS TRANSPORTES. DESENVOLVIMENTO DE ESTUDOS PARA INTEGRAÇÃO DA BASE DE DADOS GEORREFERENCIADA DO PNLT AO SISTEMA DE INFORMAÇÕES GEOGRÁFICAS EM TRANSPORTE – SIG-T DO MINISTÉRIO DOS TRANSPORTES. Ministerio dos Transportes. Brasília, DF, p. 117. 2010.

\_\_\_\_\_'\_\_\_\_ - PNLT. Plano Nacional de Logística e Transportes. Brasília-DF 2007.

NATIONAL COOPERATIVE RESEARCH PROGRAM-REPORT 606. **Forecasting Statewide Freight Toolkits**. Washington, DC. 2008.

NGUYEN, S. **Estimating an OD matrix from network data: a network equilibrium approach**. [S.l.]: Centre de Recherche sur les Transports, 1977.

NIELSEN, O. A. A new method for estimating trip matrices from traffic counts. **Institute of Roads, Traffic and Town Planning**, The Technical University of Denmark, 1993.

NOY, F.; MCGUINNESS, D. L. **Ontology Development 101: Guide to Creating Your First Ontology**. Stanford Knowledge Systems Laboratory. [S.l.]. 2001.

OLIVEIRA, M. L.; CYBIS, H. B. B. Revisão da experiência de calibração do software VISSIM aplicado a um estudo de caso de autoestrada brasileira. In: SEMANA DE ENGENHARIA DE PRODUÇÃO SUL AMERICANA, 8. Bento Gonçalves - RGS: [s.n.]. 2008. p. 1-12.

ORTÚZAR, J. D. D.; WILLUMSEN, L. G. Modelling Transport. 4a. ed. U.K.: John Wiley & Sons Ltd, 2011.

PENDYALA, R. M.; SHANKAR, V. N.; MCCULLOUGH, R. G. Freight travel demand modelling synthesis of research and development of a framework.

**Transportation Research Record: Journal of the transportation Research Board**, v. 1725,p. 9-16 Washington DC, 2000.

PRESSMAN, R. S. **Engenharia de Software: ma abordagem Profissional**. 7a. ed. Porto Alegre: Mc GrawHill Bookman, 2011.

QUEIROGA, C. A.; BULLOCK, D. Travel time studies with global positioning and geographic information systems: an integrated methodology. **Transportation Research Part C Emerging Technologies**, v. 6, n.1-2, p. 101-127, 1998. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0968090X98000102> > Acesso em 10 jun 2012

RAKHA, H.; PARAMAHAMSAM, H.; VAN AERDE, M. **Comparasion of Static maximum likehod Origin-Destination formulations**. In: 16th INTERNATIONAL SYMPOSIUM ON TRANSPORTATION AND TRAFFIC THEORY, 16. [S.l.]: Hani S Mahmassani. 2005. p. 693-715.

REWERSE. urban\_ontologies. **SemCity**, 2013. Disponível em: <[http://semcity.unige.ch/semcity/doku.php/urban\\_ontologies](http://semcity.unige.ch/semcity/doku.php/urban_ontologies)>. Acesso em: 23 jan. 2013.

RITTO, A. C.; DA LUZ, M. M. **Tecnologias para Gestão do Transito de Veiculos-Solução Seagull para o SINIAV**. 1. ed. Rio de Janeiro: [s.n.], v. 1, 2012.

ROORDA, M. J. et al. A conceptual framework for agent-based modelling of logistics services. **Transportation research Part E**, v. 46, n. 1, p. 18-31, 2010.

SADEGHI-NIARAKI, A.; KIM, K. Ontology based personalized route planning system using a multi-criteria decision making approach. **Journal of Expert Systems with Applications**, USA, v. 36, n. 2, parte 1, p. 2250-2259, 2009. ISSN 10.1016/j.eswa.2007.12.053.

SECTRAN. PDTU-Carga. **Plano Diretor de Transporte Urbano da Região Metropolitana do Rio de Janeiro - PDTU/RMRJ**, Rio de Janeiro, 2002. Disponível em: <<http://www.central.rj.gov.br/pdtu/indexini.php>>. Acesso em: 27 fev. 2012.

SEMCITY. Urban Ontologies. **TU0801: Semantic Enrichment of 3D City Models for Sustainable Urban Development**, 2012. Disponível em:

<[http://semcity.unige.ch/semcity/doku.php/urban\\_ontologies](http://semcity.unige.ch/semcity/doku.php/urban_ontologies)>. Acesso em: 23 jan. 2013.

SHRP2/TRB. **Freight Demand Modeling and Data Improvement (Draft)**. Strategic Highway Research Program - TRB. [S.l.], p. 121. 2013.

SOUZA, D. R. D.; D'AGOSTO, D. A. Modelo de quatro etapas aplicado ao planejamento. **Journal of Transport Literature**, 7, n. 2, p. 207-234, Abr. 2013.

STREAMS. Strategic Transport Research for European Member States, 2000. Disponível em: <[ftp://ftp.cordis.europa.eu/pub/transport/docs/summaries/strategic\\_streams\\_report.pdf](ftp://ftp.cordis.europa.eu/pub/transport/docs/summaries/strategic_streams_report.pdf)>. Acesso em: 17 abr. 2012.

STREITZ, N.; NIXON, P. **The disappearing computer**. Communications of the ACM. [S.l.]: [s.n.]. 2005. p. 32-35.

SUMO. SUMO Simulation of Urban MObility, 2012. Disponível em: <<http://sumo.sourceforge.net/>>. Acesso em: 17 abr. 2012.

SUSSMAN, J. M. **Perspectives on intelligent transportation systems**. New York, USA: Springer Science+Business Media Inc., 2005.

TAVASSZY, L. A. **Freight Modeling - An overview of international experiences**. TRB- Conference on Freight Demand Modelling: Tools for Public Sector Decision Making. Washington DC: [s.n.]. 2006.

THURASINGHAM. **Bhavani. Data Mining**. [S.l.]: [s.n.], 1999.

TRB. **Highway Capacity Manual 2000**. NCHRP, FHWA, TCRP, and TRB. [S.l.], p. 1134. 2000.

\_\_\_\_\_. **Freight Demand Modeling: tools for public sector decision-making**. Transportation Resesarch Board, National Research Council. [S.l.]: [s.n.]. 2008. p. 40.

\_\_\_\_\_ **Highway Capacity Manual - HCM**. Transportation Research Board. [S.l.]. 2010.

\_\_\_\_\_ Transportation Knowledge Networks: A Management Strategy for the 21st Century. **TRB- Transportation Research Board**, 23 fev. 2012. Disponível em: <<http://www.trb.org/Main/Blurbs/156626.aspx>>. Acesso em: 12 jun. 2012.

**TRKC. Freight Transport Thematic Research**. Transport Research Knowledge Research. Disponível em <http://www.trb.org/FreightTransportation/Blurbs/161737.aspx> [S.l.]. 2010 Acesso em 22 ago 2012.

VAN ZUYLEN, H. J.; WILLUMSEN, L. G. Consistent Link Flow Estimation from Counts. **Transportation Research**, n. 148, p. 281 -293, 1982.

Von Braun. **Centro de Pesquisas Avançadas** Wernher von Braun. Campinas SP. 2010 Disponível em <http://www.vonbraunlabs.com.br/> Acesso em 22 ago 2012.

W3C. Web Ontology Language- Use Cases Requirements. **W3C Recommendation**, 2012. Disponível em: <<http://www.w3.org/TR/webont-req/>>. Acesso em: 10 jan. 2013.

\_\_\_\_\_ Vocabularies. **Semantic Web**, n/a. Disponível em: <<http://www.w3.org/standards/semanticweb/ontology>>. Acesso em: 22 jan. 2013.

WANG, J.; DING, ; JIANG,. An Ontology-based Public Transport Query System. In: PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE ON SEMANTICS, KNOWLEDGE AND GRID. Washington, DC, USA: IEEE Computer Society. 2005. p. 62

WARDROP, J. G. Some theoretical aspects of road traffic research. In: ICE PROCEEDINGS: ENGINEERING DIVISIONS. [S.l.]: [s.n.]. 1952. p. 325-362.

WEISS, S. M.; INDURKHYA, N. **Predictive Data Mining: A Practical Guide**. San Francisco: [s.n.], 1998.

WEKA. Weka 3: **Data Mining Software in Java**. University of Waikato. Disponível em: <<http://www.cs.waikato.ac.nz/ml/Weka/index.html>>. Acesso em: 25 abr. 2012.

WILLUMSEN, L. **Estimation of an OD matrix from traffic counts: a review.** Institute for Transport Studies, working paper 99, 1978a.

WILLUMSEN, L. G. **Simplified transport models based on traffic counts.** Transportation, v. 10, p. 257-278, 1981.

WU, X. et al. **Top 10 algorithm in data mining.** Knowledge and Information Systems, v. 14, n. 1, p. 1-37, 2008.

XIAOLIN, L. Infrastructure for Intelligent Transportation Planning CSCW System based on WEB-GIS Technology. In: WORLD CONGRESS ON INTELLIGENT CONTROL AND AUTOMATION, 5 [S.l.]: [s.n.]. 2004. p. 15-19.

ZAMANI, Z. et al. Context-Aware Payment for supply chains: software architecture and formal verification. In: WORLD CONGRESS ON ENGINEERING. London-UK: [s.n.]. 2008.

ZHANG , L.; WEI, T. Research on Flexible Architecture of Transportation System Based on GIS-T. In: WASE INTERNATIONAL CONFERENCE ON INFORMATION ENGINEERING, 2009. ICIE 09. [S.l.]: [s.n.]. 2009. p. 224-227.

ZHANG, C. et al. Transformation of Transportation Data Models from Unified Modeling Language to Web Ontology Language. **Transportation Research Record: Journal of the Transportation Research Board**, n. 2064, p. 81-89, dez. 2008.

ZHENGYU DUAN, L. L. S. W. **MobilePulse:** Dynamic Profiling of Land Use Pattern and OD Matrix Estimation from 10 Million Individual Cell Phone Records in Shanghai. Geoinformatics, 2011 In: INTERNATIONAL CONFERENCE ON ZHENGYU, 19. 2011. p. 1.