LUCIANO APARECIDO BARBOSA

**PREVISÃO DE LOCALIZAÇÃO FUTURA DE VEÍCULOS BASEADA
EM DADOS DE AVL**

São Paulo
2010

LUCIANO APARECIDO BARBOSA

**PREVISÃO DE LOCALIZAÇÃO FUTURA DE VEÍCULOS BASEADA
EM DADOS DE AVL**

Dissertação apresentada á Escola Politécnica da
Universidade de São Paulo para a obtenção do título
de Mestre em Engenharia.

Área de Concentração: Engenharia de Transportes –
Informações Espaciais

Orientador: Prof. Dr. Marcos Rodrigues

São Paulo
2010

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, de outubro de 2010.

Assinatura do autor _____

Assinatura do orientador _____

FICHA CATALOGRÁFICA

Barbosa, Luciano Aparecido

Previsão de localização futura de veículos baseada em dados

de AVL / L.A. Barbosa. -- ed.rev. -- São Paulo, 2010.

117 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Transportes.

1. Sistemas inteligentes de transportes 2. Veículos (Localização) 3. Rastreamento I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II. t.

*Dedico este trabalho aos meus queridos
pais Jamiro e Sueli, e aos meus irmãos
Leandro e Lucas.*

AGRADECIMENTOS

Meus sinceros agradecimentos...

Ao meu orientador Prof. Dr. Marcos Rodrigues pela confiança, parcimônia e apoio, que em todas as reuniões muito me ensinou, contribuindo para meu crescimento científico e intelectual.

Ao Prof. Dr. José Alberto Quintanilha e ao Prof. Dr. Carlos Eduardo Cugnasca pelas sugestões apresentadas durante a realização do exame de qualificação, que foram muito oportunas para a conclusão desta pesquisa.

Ao colega Eduardo Jun Shinowara pelas discussões e constante colaboração.

Aos colegas Mariana, Cláudia, Patrícia, Joana, Rodrigo e demais colegas do Laboratório de Geoprocessamento pela receptividade, companheirismo e amizade.

À escola Politécnica da Universidade de São Paulo, pela oportunidade de realização do curso de mestrado.

A CAPES, pela concessão de bolsa durante o período de 24 meses, para a realização deste trabalho.

À Empresa Logica América do Sul Soluções de Tecnologia Ltda., pelo apoio e incentivo para a conclusão deste trabalho.

Enfim, aos meus queridos pais Jamiro e Sueli, que sempre incentivaram seus filhos para o estudo e nunca mediram esforços para isso. Aos meus irmãos Leandro e Lucas pelo incansável apoio e afeto, e a todos os meus familiares pelo incentivo constante.

“A melhor maneira de prever o futuro é criá-lo.”

Peter Drucker

RESUMO

BARBOSA, L. A. **Previsão de localização futura de veículos baseada em dados de AVL.** 2010. 117 f. Dissertação (Mestrado) – Departamento de Engenharia de Transportes, Escola Politécnica da Universidade de São Paulo, São Paulo, 2010.

O crescente desenvolvimento de aplicações utilizadas por dispositivos móveis que fazem uso das tecnologias de posicionamento via satélite e comunicação móvel, juntamente com a popularização destes dispositivos, sejam eles celulares ou GPS's (*Global Positioning System*) automotivos reforçam ainda mais a necessidade de representação e o entendimento a cerca das entidades móveis retratadas nestes dispositivos e incentivam estudos que forneçam um significado maior do que a simples representação posicional destas entidades. Considera-se neste trabalho, que estas entidades móveis são veículos rastreados via satélite que fornecem sua posição espacial, determinada por um par de coordenadas geográficas (latitude e longitude), coletadas em intervalos de tempo regulares para sistemas AVL (*Automatic Vehicle Location*) que são responsáveis pelo monitoramento do estado destes veículos. Neste trabalho, foram desenvolvidas funções para a previsão da localização e geração de padrões dos veículos monitorados por sistemas AVL. Para tanto, as paradas efetuadas pelos veículos irão definir regiões comuns de parada ocorridas durante um intervalo de tempo passado e serão consideradas como um padrão de localização, enquanto que as trajetórias serão utilizadas para definir o padrão de movimentação. Os relacionamentos topológicos que estes padrões possuem irão definir por meio de condições espaciais/direcionais e temporais quais serão as regiões de paradas consideradas como prováveis destinos, a partir de outra região de parada, considerada como origem e, permitirão a geração de matrizes com valores de frequências relativas que consideram o número de visitas que uma região recebe a partir da outra. O que possibilita extrair valores de probabilidade condicional para cada destino provável. Portanto, a metodologia proposta e as funções desenvolvidas que foram validadas em experimentos, que utilizaram dados reais de rastreamento, fornecem uma camada inicial de inteligência aos sistemas AVL, que proporciona aos controladores destes sistemas utilizarem consultas preditivas, identificarem mais facilmente anomalias de comportamento, que possam evidenciar alguma ocorrência incomum na movimentação do veículo, além de aumentar a segurança dos veículos que possuem um dispositivo de rastreamento por meio da definição de padrões inerentes ao veículo.

Palavras-chave: Sistemas Inteligentes de Transporte. Localização Automática de Veículos. Sistema de Posicionamento Global. Sistemas de Informações Geográficas. Padrão. Previsão.

ABSTRACT

BARBOSA, L. A. **Future location prediction of vehicles based on data AVL.** 2010. 117 f. Dissertação (Mestrado) – Departamento de Engenharia de Transportes, Escola Politécnica da Universidade de São Paulo, São Paulo, 2010.

The increasing development of applications used by mobile devices that make use of the technologies of satellite positioning and mobile communications, along with the popularity of these devices, whether cell phones or GPS's (Global Positioning System) automotive further reinforce the need for representation and understanding about the mobile entities represent in these devices and encourage studies that provide a greater meaning than the simple positional representation of these entities. It is considered in this work that these entities are tracked vehicles that provide satellite spatial position, determined by a pair of coordinates (latitude and longitude), collected at regular time intervals for systems AVL (*Automatic Vehicle Location*) that are responsible for monitoring the state of these vehicles. In this work, functions have been developed to predict the location and pattern generation of vehicles monitored by AVL systems. Accordingly, these stops will define common regions of the stop occurred during a period of time past and will be considered as the pattern location, while the trajectories are used to define the pattern of movement of the vehicle. The topological relationships that have these patterns define conditions through spatial/directional and temporal, which are stops regions considered as probable destinations from another stop region, regarded as origin and allow the generation of matrices with values of frequencies on considering the number of visits that region receives from the other. What makes it possible to extract values of conditional probability for each likely destination. Therefore, the proposed methodology and functions developed that been validated in experiments using real data to tracking provide a initial layer of intelligence to the AVL system that gives drivers of these systems use predictive queries, more easily identify behavioral abnormalities that may show some unusual occurrence in moving the vehicle, in addition to increasing the safety of vehicles which have a tracking device by setting patterns relating to the vehicle.

Keywords: Intelligent Transportation Systems (ITS). Automatic Vehicle Location (AVL). Global Positioning System (GPS). Geographic Information System (GIS). Patterns. Prediction.

LISTA DE ILUSTRAÇÕES

FIGURA 1. NUVEM DE PONTOS DE RASTREAMENTO – PARADOS E EM MOVIMENTO.....	15
FIGURA 2. PONTOS DE RASTREAMENTO QUE REPRESENTAM PERÍODOS DISTINTOS ONDE O VEÍCULO ESTEVE PARADO.....	16
FIGURA 3. PARADAS OCORRIDAS EM INTERVALOS DE TEMPO DISTINTOS PARA UM ÚNICO VEÍCULO.....	17
FIGURA 4. REGIÃO DE PARADA - PARADAS OCORRIDAS EM PERÍODOS DISTINTOS PARA UM ÚNICO VEÍCULO.....	18
FIGURA 5. COMPONENTES DE UM SISTEMA AVL.....	33
FIGURA 6. EQUIPAMENTOS EMBARCADOS DE UM VEÍCULO DE PASSEIO.	34
FIGURA 7. CONSTELAÇÃO GPS.....	39
FIGURA 8. POSICIONAMENTO ABSOLUTO CINEMÁTICO.....	40
FIGURA 9. SISTEMA DE REFERÊNCIA TERRESTRE WGS84.....	41
FIGURA 10. EFEITO DO SA NOS RESULTADOS DE POSICIONAMENTO GPS COM UM RECEPTOR DE SIMPLES FREQUÊNCIA, À ESQUERDA O SA ATIVO E A DIREITA APÓS A DESATIVAÇÃO DO SA.....	42
FIGURA 11. COBERTURA NACIONAL GSM DA OPERADORA CLARO.....	45
FIGURA 12. EVOLUÇÃO 3G DOS PROTOCOLOS GSM E CDMA.....	46
FIGURA 13. EXTENSÃO GEOGRÁFICA DOS DADOS DE RASTREAMENTO.....	51
FIGURA 14. BASE CARTOGRÁFICA – ESTADO DE SÃO PAULO.	52
FIGURA 15. PROCESSO GERAL PARA A GERAÇÃO DOS PADRÕES E PREVISÃO.	58
FIGURA 16. QUANTIDADE DE REGISTROS NECESSÁRIOS PARA A DEFINIÇÃO DO ESTADO DO VEÍCULO.	
FIGURA 17. ALGORITMO 1: DEFINIÇÃO DO ESTADO DO VEÍCULO.....	62
FIGURE 18. EM (A) TRAJETO DE IDA ENTRE (O-D) E EM (B) TRAJETO DE VOLTA ENTRE (O-D).....	63
FIGURA 19. ALGORITMO 2: DEFINIÇÃO DOS GRUPOS DE PARADAS.....	64
FIGURA 20. ALGORITMO 3: GERAÇÃO DA GEOMETRIA DA PARADA E CÁLCULO DO TEMPO DE DURAÇÃO DA PARADA.	65
FIGURA 21. GEOMETRIAS GERADAS A PARTIR DO <i>BUFFER</i> DE UMA PARADA.	65

FIGURA 22. ALGORITMO 4: PREENCHE AS INFORMAÇÕES RELACIONADAS AO PADRÃO DE MOVIMENTAÇÃO.....	68
FIGURA 23. ALGORITMO 5: ARMAZENAR O ÚLTIMO MOVIMENTO REALIZADO PELO VEÍCULO;.....	69
FIGURA 24. ALGORITMO 6: AGRUPAMENTO E IDENTIFICAÇÃO DE UMA REGIÃO DE PARADA.....	72
FIGURA 25. REGIÃO DE PARADAS.....	73
FIGURA 26. INTERVALO DE ORIENTAÇÃO DOS SEGMENTOS DE UMA TRAJETÓRIA.	77
FIGURA 27. IDENTIFICAÇÃO DE UM MESMO CAMINHO.....	78
FIGURA 28. ALGORITMO 7: DEFINIÇÃO DO ESTADO DO VEÍCULO PARA A PREVISÃO.	81
FIGURA 29. ALGORITMO 8: IDENTIFICAÇÃO DA ORIGEM NO PADRÃO DE LOCALIZAÇÃO.....	83
FIGURA 30(A). CATEGORIZAÇÃO DOS POSSÍVEIS DESTINOS PARA O CÁLCULO DE PROBABILIDADE – VEÍCULO PARADO.....	88
FIGURA 30(B). CATEGORIZAÇÃO DOS POSSÍVEIS DESTINOS PARA O CÁLCULO DE PROBABILIDADE – VEÍCULO PARADO.....	88
FIGURA 30(C). CATEGORIZAÇÃO DOS POSSÍVEIS DESTINOS PARA O CÁLCULO DE PROBABILIDADE – VEÍCULO EM MOVIMENTO.	89
FIGURA 30(D). CATEGORIZAÇÃO DOS POSSÍVEIS DESTINOS PARA O CÁLCULO DE PROBABILIDADE – VEÍCULO EM MOVIMENTO.	89
FIGURA 31. ALGORITMO 9: CÁLCULO DA PROBABILIDADE CONDICIONAL.	94
FIGURA 32. ALGORITMO 10: CÁLCULO DA PROBABILIDADE DO DESTINO QUANDO A ORIGEM NÃO FOR ENCONTRADA NO PADRÃO DE LOCALIZAÇÃO.....	95
FIGURA 33. PONTOS DE RASTREAMENTO GPS PARA UM VEÍCULO, SOBREPOSTOS À BASE CARTOGRÁFICA.....	97
FIGURA 34. TRAJETÓRIAS DE UM VEÍCULO SOBREPOSTAS À BASE CARTOGRÁFICA.	98
FIGURA 35. REPRESENTAÇÃO DOS PROVÁVEIS DESTINOS DO VEÍCULO 1 DA TABELA 10.	102
FIGURA 36. REPRESENTAÇÃO DOS TRAJETOS ENTRE A ORIGEM E PROVÁVEIS DESTINOS DO VEÍCULO 1.....	103

LISTA DE TABELAS

TABELA 1 – RELACIONAMENTOS TOPOLÓGICOS ENTRE DUAS GEOMETRIAS	19
TABELA 2 – COMPARATIVO ENTRE OS SISTEMAS DE COMUNICAÇÃO	35
TABELA 3 – CARACTERÍSTICAS DOS SINAIS GPS	42
TABELA 4 – PROTOCOLOS DE TRANSMISSÃO DE DADOS	44
TABELA 5 – ESTADOS BRASILEIROS E TECNOLOGIAS DE COMUNICAÇÃO FORNECIDA PELA OPERADORA CLARO.	45
TABELA 6 – BANDAS E CANAIS UTILIZADOS PELO GSM.	47
TABELA 7 – FREQUÊNCIA DE VISITAS A UM DESTINO A PARTIR DE UMA ORIGEM (ESPACIAL/DIRECIONAL).....	86
TABELA 8 – FREQUÊNCIA DE VISITAS A UM DESTINO A PARTIR DE UMA ORIGEM (ESPACIAL/DIRECIONAL E TEMPORAL DIA DA SEMANA).....	87
TABELA 9 – FREQUÊNCIA DE VISITAS A UM DESTINO A PARTIR DE UMA ORIGEM (ESPACIAL/DIRECIONAL TEMPORAL DIA DA SEMANA E PERÍODO DO DIA).	87
TABELA 10 – PROBABILIDADE CONDICIONAL DOS DESTINOS PARA CADA VEÍCULO.....	100
TABELA 11 – PROBABILIDADE (FREQUÊNCIA RELATIVA) DOS DESTINOS PARA CADA VEÍCULO.	101
TABELA 12 – NÚMERO DE REGIÕES DE PARADA DETECTADAS PARA CADA VEÍCULO.....	104
TABELA 13 – DESVIO-PADRÃO DO NÚMERO DE PARADAS PARA CADA CLASSE DE VEÍCULO.	105

LISTA DE SIGLAS

ASP	Application Service Provider
AVL	Automatic Vehicle Location
C/A	Coarse Acquisition
CDMA	Code Division Multiple Access
DoD	Department of Defense
ESA	European Space Agency
ETSI	European Telecommunications Standards Institute
FDMA	Frequency Division Multiple Access
GLONASS	GLObal NAVigation Satellite System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPRS	General Packet Radio Service
GRS	Geodetic Reference System
GSM	Global System for Mobile Communications
ITRF	International Terrestrial Reference Frame
IERS	International Earth rotation and Reference systems Service
IRM	IERS Reference Meridian
IRP	IERS Reference Pole
NAVSTAR-GPS	NAVigation Satellite with Time And Ranging
PN	Pseudo-random
PPS	Precise Positioning Service
RF	Radiofrequência
SA	Selective Availability
SIG	Sistemas de Informações Geográficas
SGBD	Sistema Gerenciador de Banco de dados
SPS	Standard Positioning Service
SMS	Short Message Service
TDMA	Time Division Multiple Access
WGS84	World Geodetic System of 1984

LISTA DE SÍMBOLOS

λ	Latitude geodésica
φ	Longitude geodésica

SUMÁRIO

1 INTRODUÇÃO	13
1.1 OBJETIVOS	21
1.2 DESCRIÇÃO DOS CAPÍTULOS	22
2 FUNDAMENTAÇÃO.....	23
2.1 ASPECTOS CONCEITUAIS	23
2.1.1 <i>Objetos Móveis</i>	23
2.1.2 <i>Padrões de Objetos Móveis</i>	25
2.1.3 <i>Previsão de Movimentação de Objetos Móveis</i>	27
2.1.4 <i>Rastreamento de Veículos</i>	29
2.2 ASPECTOS TÉCNICOS	31
2.2.1 <i>Sistemas AVL</i>	31
2.2.1.1 Aquisição de Dados e Atuação	33
2.2.1.1.1 Os Equipamentos Embarcados.....	33
2.2.1.2 Sistemas de Comunicação.....	34
2.2.1.3 Gestão da Informação	35
2.2.2 <i>Técnicas de Posicionamento Utilizadas em AVL</i>	36
2.2.2.1 Posicionamento do Veículo utilizando o GNSS	37
2.2.2.1.1 O Sistema de Posicionamento Global - GPS	38
2.3.2 <i>Transmissão de Dados Utilizada em AVL</i>	43
2.3.2.1 Transmissão de Dados Via Telefonia Móvel	44
2.3.2.1.1 O Sistema de Comunicação GSM.....	46
2.4 CONSIDERAÇÕES FINAIS	48
3 RECURSOS UTILIZADOS NA PESQUISA	49
3.1 DADOS DO RASTREAMENTO.....	49
3.2 ÁREA DE ESTUDO	51
3.2.1 <i>Base Cartográfica</i>	52
3.3 APOIO COMPUTACIONAL.....	53
3.4 CONSIDERAÇÕES FINAIS	53

4	METODOLOGIA.....	54
4.1	ETAPA I – PREPARAÇÃO DOS DADOS	54
4.2	ETAPA II – GERAÇÃO DE PADRÕES	57
4.2.1	<i>Detecção de Trajetos.....</i>	59
4.2.2	<i>Detecção de Paradas</i>	63
4.2.3	<i>Padrões de Movimentação</i>	66
4.2.3.1	Geração dos Padrões de Movimentação	67
4.2.4	<i>Padrões de Localização</i>	69
4.2.4.1	Geração dos Padrões de Localização.....	70
4.3	ETAPA III – PREVISÃO DE DESTINOS	73
4.3.1	<i>Componente Espacial.....</i>	74
4.3.2	<i>Componente Temporal</i>	75
4.3.2.1	Dias da Semana.....	75
4.3.2.2	Períodos do Dia.....	76
4.3.3	<i>Componente Direcional</i>	76
4.3.4	<i>Processo Preditivo.....</i>	79
4.3.4.1	Estado Atual do Veículo	80
4.3.5	<i>Aspectos Estatísticos.....</i>	83
4.3.5.1	Probabilidade Condicional.....	90
4.4	CONSIDERAÇÕES FINAIS	96
5	EXPERIMENTAÇÃO	97
5.1	CONSIDERAÇÕES INICIAIS	97
5.2	EXECUÇÃO DAS FUNÇÕES DESENVOLVIDAS.....	98
5.3	RESULTADOS OBTIDOS	99
5.4	CONSIDERAÇÕES FINAIS	105
6	CONCLUSÃO E RECOMENDAÇÕES.....	106
6.1	TRABALHOS FUTUROS	107
	REFERÊNCIAS.....	109
	APÊNDICE A	112
	APÊNDICE B	113

1 INTRODUÇÃO

Consideram-se como objetos móveis para uma modelagem, no âmbito desta pesquisa, todos aqueles objetos aos quais se podem estabelecer uma relação espacial, relação esta de movimentação que na maioria das vezes está vinculada à proximidade de outros objetos e que seja variável no espaço-tempo. Estes objetos são agentes de inúmeros processos de atividades do cotidiano humano. Isto porque, assume-se que eles possam ser qualquer coisa que esteja em movimento e, esteja sendo rastreado, como por exemplo, uma pessoa, um animal, um veículo etc.

O rastreamento de veículos tem como objetivo fornecer e gerenciar sistematicamente a posição e o estado dos veículos, com vários níveis de exatidão posicional e intervalo de tempo. O intuito do rastreamento de veículos é aumentar a eficiência dos despachos de mercadorias, otimizar o uso dos veículos, reduzir o custo operacional e atender às exigências das gerenciadoras de risco, uma vez que sua utilização aumenta a segurança da carga, do veículo e do próprio motorista (RODRIGUES; CUGNASCA; QUEIROZ FILHO, 2009).

Ora, se estes objetos estabelecem relacionamentos no espaço-tempo e estes relacionamentos se referem às movimentações realizadas por estes objetos, isto permite imaginá-los como um sistema dinâmico cujo estado se deseja conhecer ao longo do tempo. Estado este que concerne a um conjunto de propriedades pertinentes ao sistema num determinado instante de tempo. Portanto, existirão diferentes estados que este sistema poderá possuir, como por exemplo: “parado” ou “em movimento” e, é na identificação destes estados que se estaria interessado num primeiro momento. Uma vez que, a partir desta identificação diferentes ações poderão ser tomadas em relação a este objeto.

No contexto desta pesquisa, estes sistemas são veículos rastreados via satélite que atualizam sua informação posicional a uma taxa temporal predefinida. Assim, a representação cartográfica desta informação posicional produzirá nuvens de pontos que representam trajetórias onde cada veículo costuma se movimentar ou regiões onde ele fica parado. Portanto, o estado que este sistema possui num determinado

instante de tempo pode ser determinado com base nos pontos de rastreamento, quando considerado, por exemplo, dois estados possíveis: “parado” e “em movimento”.

Estes sistemas compõem outro sistema responsável pelo monitoramento destes veículos. Segundo Rodrigues; Cugnasca e Queiroz Filho (2009), os sistemas que utilizam o rastreamento de veículos iniciaram sua operação por volta do final da década de 1960, em Chicago, IL Estados Unidos. Inicialmente denominado de *Automatic Vehicle Location* (AVL), foi um sistema criado para ser utilizado como ferramenta de apoio à gestão de tráfego. Porém, atualmente é utilizado principalmente como ferramenta de gerenciamento logístico e de risco. No Brasil, é empregado principalmente por motivos de segurança, com o intuito de prevenir furtos e/ou recuperar bens. O que é reforçado por Branco (2009a)

[...] as gerenciadoras de risco são as principais usuárias dos softwares, desenvolvidos pelas empresas que atuam na área de equipamentos para rastreamento e monitoramento de veículos. Os softwares para o gerenciamento de carros, caminhões e cargas, dependendo da tecnologia embarcada, são parametrizáveis e têm características que são imprescindíveis na hora da contratação do serviço.

Os sistemas AVL são sistemas que permitem o monitoramento de veículos ou entidades móveis (ex: cargas, pessoas, animais, etc.). Geralmente, estas entidades móveis são compostas por atributos como posição (Latitude e Longitude), instante de tempo da coleta da posição, e de qualquer outro atributo que será estabelecido em razão do propósito de monitoramento da entidade móvel, por exemplo, quando se considera como entidade móvel um caminhão frigorífico, um atributo que pode ser interessante monitorar é a temperatura do baú.

A movimentação de veículos monitorados por sistemas AVL pode ser representada pelas trajetórias, que podem ser consideradas como a história de mudança da posição do veículo e ainda, podem relacionar as diferentes regiões onde o veículo costuma estar parado, estabelecendo significados a estas regiões como, por exemplo: origem e destino.

Com base no que foi mencionado, imagina-se que existirão regiões de aglomeração de pontos relacionadas por meio de trajetórias e, que estas regiões de aglomeração podem ser representadas por uma área que contém todos os pontos de rastreamento que refletem uma parada realizada pelo veículo. A figura 1 ilustra a representação de uma nuvem de pontos de rastreamento onde os pontos em vermelho representam paradas efetuadas pelo veículo e os pontos na cor verde a sua movimentação.

Figura 1. Nuvem de Pontos de Rastreamento – Parados e Em Movimento.

Porém, estes pontos de rastreamento podem representar paradas ocorridas em intervalos de tempo distintos. Dessa maneira, torna-se necessário identificar quais são os pontos que definem uma parada ocorrida durante um intervalo de tempo. A figura 2 distingue os pontos de rastreamento que representam os diferentes intervalos de tempo em que o veículo esteve parado, nela cada cor representa um intervalo de tempo que o veículo esteve parado.

Figura 2. Pontos de Rastreamento que representam períodos distintos onde o veículo esteve parado.

Para tanto, todos os pontos de rastreamento que definem uma parada, ocorrida durante um intervalo de tempo, serão representados por uma área. Portanto, uma parada será definida como uma região, na qual o veículo interrompeu sua movimentação por um período de tempo e será representada por um círculo de raio igual ao erro de posicionamento considerado, com origem definida pelo posicionamento médio dos pontos de rastreamento que definem esta parada, de tal maneira que todos estes pontos estejam contidos no seu interior. A figura 3 ilustra a geometria responsável por representar cada parada.

Figura 3. Paradas ocorridas em intervalos de tempo distintos para um único veículo.

Uma vez que tenham sido definidos todos os pontos de rastreamento que compõem uma parada ocorrida durante um intervalo de tempo, é necessário avaliar e identificar quais paradas pode refletir uma região comum de parada. Considerando-se que cada uma destas paradas reflete uma visita do veículo a uma mesma região, o estabelecimento de relacionamentos topológicos entre essas paradas pode definir a área desta região comum, visitada pelo veículo em diferentes instantes de tempo.

Assim, o relacionamento topológico entre estas paradas define uma nova área, doravante denominada de região de parada que reúne todas as informações temporais das paradas que a compoem uma única região geográfica, caracterizando o local onde o veículo costuma ficar parado. A figura 4 ilustra esta região comum.

Figura 4. Região de Parada - paradas ocorridas em períodos distintos para um único veículo.

É importante ressaltar que o conceito de região de parada, adotado nesta pesquisa, se restringe ao plano bi-dimensional e, portanto, essa região será representada por círculo que contém todas as paradas que satisfazem um relacionamento topológico de proximidade e que, portanto, serão consideradas como uma mesma parada que ocorre em intervalos de tempo distintos. A tabela 1 ilustra alguns desses relacionamentos topológicos, nela é possível verificar o relacionamento (Área/Área), considerado para caracterizar uma região de parada.

Tabela 1 – Relacionamentos topológicos entre duas geometrias

A relação espacial com B	Disjuntos	Interage					
		Toca	Divide	Sobreposições			
				Dentro de	Contém	Cruza	Igual
Ponto/Ponto							
Ponto/Linha							
Linha/Ponto							
Ponto/Área							
Área/Ponto							
Linha/Linha							
Linha/Área							
Área/Linha							
Área/Área							

Fonte: Adaptada de Smallworld (2008).

Mais quais seriam os significados associados a estas regiões de parada? A periodicidade de ocorrência das paradas que compõem uma região de parada pode indicar este significado. Por exemplo, considere uma região de parada que ocorre durante todos os dias úteis da semana das 9h às 18h. Neste caso, o significado desta região parece evidente, ou seja, ela pode refletir a localização aproximada do local de trabalho do condutor do veículo.

Portanto, a semântica que cada uma das regiões possui pode ser estabelecida com base nas informações temporais de cada uma das paradas que compõem esta região de parada. Este conceito possibilita inferir que qualquer região de parada possui um significado e, portanto, a dificuldade de identificação deste significado passa a ser a correta interpretação do cotidiano do condutor do veículo, a partir da periodicidade de ocorrência das paradas que compõem uma região.

Se existe uma periodicidade de ocorrência, tanto das paradas quanto das trajetórias torna-se possível estabelecer padrões relacionados às paradas e trajetórias do veículo. Assim, estes padrões refletem os relacionamentos existentes entre as regiões de paradas e os trajetos realizados pelo veículo, ou seja, indicam por onde o veículo costuma ficar ou se movimentar.

A manutenção destes padrões num banco de dados possibilitaria a realização de previsões de localização. Porém, a correta identificação do atual estado do veículo passa a ser primordial para responder este tipo de questionamento, uma vez que dependendo do atual estado do veículo diferentes análises serão necessárias. Por exemplo, considere o fato do veículo estar parado, nesta configuração o simples relacionamento entre as regiões de paradas e as trajetórias prediz quais são os possíveis destinos deste veículo, desde que seja identificada a região de parada onde o veículo está parado, isto é, a origem. Por outro lado, quando o veículo estiver em movimento, a identificação das trajetórias armazenadas no padrão, compatíveis a trajetória atual do veículo passa a ser relevante no apontamento dos possíveis destinos, pois, serão a partir da identificação dessas trajetórias que serão determinadas as possíveis origens do veículo.

Evidente que existirão veículos, aos quais não será possível atribuir padrões confiáveis que auxiliem o processo preditivo, como por exemplo, veículos de aluguel ou táxis. Porém, a pesquisa considera que a metodologia e funções desenvolvidas serão utilizadas no monitoramento de veículos leves que pertençam a pessoas que possuem hábitos regulares e, portanto, seja possível identificar estes padrões e relacionamentos e assim, responder questionamentos relacionados aos instantes de tempo passado e futuro.

A consideração de veículos leves nesta pesquisa ocorre em função das discussões acerca da Resolução Nº 245 do Conselho Nacional de Trânsito, que se refere a um dispositivo antifurto com funções de bloqueio remoto e localização, embarcado nos veículos automotores novos, sejam eles nacionais ou importados. Estas discussões fornecem a perspectiva de que uma quantidade maior de veículos leves será rastreada e, por conseguinte surgirão novas aplicações e situações, nas quais os sistemas AVL podem ser utilizados. Enfim, as aplicações são as mais variadas,

considerando-se as diversas tecnologias disponíveis e, portanto, o surgimento de novas aplicações ficará restrito somente por conta da inspiração e imaginação da mente humana motivada pela identificação de novas necessidades.

1.1 OBJETIVOS

O objetivo deste trabalho é propor um método que possibilite a previsão de localização de veículos rastreados via satélite, propondo uma estrutura de bancos de dados para o armazenamento das informações geradas a partir dos pontos de rastreamento desses objetos móveis e validar esta estrutura proposta por meio dos resultados obtidos com o desenvolvimento de funções e a utilização de dados reais de rastreamento. Considerando-se procedimentos que identificam o estado atual do veículo e geram padrões com algum significado associado aos pontos de rastreamento, de modo que a localização futura do veículo deixe de ser apenas um par de coordenadas.

O método proposto tem como principais contribuições:

- Estender os trabalhos relacionados à extração de padrões de comportamento de veículos.
- Atribuir significado às regiões de parada efetuadas pelos veículos rastreados.
- Gerar objetos extensos (linhas e polígonos) a partir de objetos pontuais e estabelecer relacionamentos entre estes objetos.
- Predizer a localização de um veículo considerando um valor de probabilidade.

Com isso, os objetivos específicos deste trabalho são:

- Implementar funções que possibilitem a geração dos padrões de comportamento e a previsão da localização do veículo rastreado, a partir dos dados de rastreamento.
- Propor um modelo de dados para objetos móveis que possibilite o armazenamento de padrões de comportamento e previsão de localização.

- Validar o método proposto e as funções desenvolvidas pela instanciação de exemplos de consultas, utilizando dados de rastreamento reais.

1.2 DESCRIÇÃO DOS CAPÍTULOS

Este trabalho é composto por seis capítulos. No Capítulo 2 são apresentados os trabalhos que contribuíram com a fundamentação conceitual e técnica desta pesquisa. Os tópicos abordados nesse capítulo se referem a: conceitos e definições de Objetos Móveis, Rastreamento de Veículos e Sistemas AVL (*Automatic Vehicle Location*).

No Capítulo 3 são apresentados os recursos utilizados para o desenvolvimento do método proposto, como por exemplo, a amostra dos dados de rastreamento, os recursos computacionais desde configuração de máquina até *softwares* utilizados.

No Capítulo 4 é apresentada a proposta metodológica para a geração dos padrões de comportamento relacionados aos veículos rastreados e como obter a previsão de localização de um determinado veículo, considerando-se o seu estado atual e padrão de comportamento.

O Capítulo 5 descreve a experimentação realizada com os dados reais do rastreamento de veículos e apresenta os resultados obtidos.

No Capítulo 6 são apresentadas as conclusões e recomendações para pesquisas futuras.

2 FUNDAMENTAÇÃO

Este capítulo apresenta uma fundamentação conceitual e técnica sobre o rastreamento de veículos, sobretudo daqueles que utilizam como tecnologia para a aquisição da posição do veículo os sistemas de posicionamento por satélite.

2.1 ASPECTOS CONCEITUAIS

Neste item são abordados aspectos conceituais relacionados ao rastreamento de veículos, com o intuito de possibilitar uma visão mais aprofundada dos temas abordados na pesquisa, como por exemplo, objetos móveis, definição de parada e trajeto, padrões relacionados a objetos móveis, previsão de movimentação de objetos móveis e rastreamento de veículos.

2.1.1 Objetos Móveis

Pode-se considerar como objeto móvel qualquer ser ou coisa do mundo real que possa gerar movimento voluntário ou involuntário. Dessa maneira, pode-se incluir como objetos móveis desde aqueles objetos que frequentemente estejam em repouso até aqueles que apresentam um movimento contínuo (BEARD; PALANCIOGLU, 2000).

Do ponto de vista computacional, de acordo com Güting e Schneider (2005), objetos móveis são, essencialmente, geometrias dependentes do tempo, mantidas num banco de dados espaço-temporal, no qual a posição pode mudar constantemente durante um espaço de tempo considerado.

Segundo Oliveira (2003), uma das vantagens que as tecnologias de computação móvel e redes de comunicação sem fio possibilitaram foi a manutenção da informação posicional de um objeto que se movimenta. Desde que este objeto esteja equipado com dispositivos que se comuniquem com uma unidade central ou outros

objetos, as informações posicionais podem ser armazenadas num banco de dados, com o intuito de possibilitar que consultas sejam realizadas em função da localização do objeto.

Consultas relacionadas a objetos móveis podem ser divididas em consultas baseadas nas coordenadas dos objetos ou na trajetória realizada pelo objeto (JUNG; RYU, 2006). Ainda, segundo os mesmos autores, as consultas baseadas nas coordenadas do objeto consideram-no como um ponto e estão relacionadas à proximidade com algum outro objeto. Já as consultas baseadas em trajetórias consideram os movimentos realizados pelo objeto e derivam de informações relacionadas do movimento do próprio objeto, como, por exemplo, velocidade e direção. E, em ambas as consultas, é necessário uma grande quantidade de dados, uma vez que é imprescindível armazenar as mudanças espaços-temporais.

Entretanto, aos objetos móveis não se associa somente o atributo de localização e instante de tempo; outros atributos também podem estar associados, de acordo com o tipo de objeto rastreado e negócio que se destina o objeto rastreado. Estes atributos serão definidos com base no nível de informação que se deseja obter sobre o estado desse objeto.

Alguns desses atributos podem ser dinâmicos, ou seja, assumem valores de acordo com a interação que o objeto móvel possui com outros objetos, o que contextualiza o objeto rastreado num sistema dinâmico. De acordo com Rodrigues, Cugnasca e Queiroz Filho (2009), “um sistema que pode ser definido como um conjunto de entes inter-relacionados”. Um sistema dinâmico é aquele no qual eventos ocorrem e, estes eventos influenciam no estado do sistema com o passar do tempo (ACKOFF, 1971).

Portanto, o estado de um objeto móvel será atribuído em função dos valores que os seus atributos possuem num determinado instante de tempo. Este conceito permite pensar em objetos móveis como sistemas móveis, cujo estado se deseja conhecer no decorrer do tempo (RODRIGUES, 2008b).

2.1.2 Padrões de Objetos Móveis

Várias consultas podem ser realizadas considerando-se a informação posicional de um objeto móvel. De acordo com Rodrigues (2008a), na maioria das vezes, o que importa é quais informações podem ser obtidas a partir desse posicionamento como, por exemplo, se o objeto está num determinado quilômetro de uma determinada rodovia ou se o objeto está se movimentando conforme programado, ou ainda, qual a proximidade do objeto em relação a um determinado evento. Porém, o que de fato interessa são os significados associados às posições, que podem ser os mais variados, de acordo com o interesse e análise de quem recebe essa informação posicional.

Outro questionamento frequente, referente à localização de um objeto móvel, remete à sua localização futura num determinado instante de tempo (t). Dentre as diversas maneiras de se obter uma resposta a este questionamento está o histórico de movimentação do objeto ou, ainda, a obtenção, a partir desse histórico, de um padrão de movimentação, que possibilite a previsão, considerando-se o seu estado atual.

Segundo Oliveira (2003), as informações sobre a movimentação de um objeto que são armazenadas no banco de dados geram um histórico de movimentação que pode ser utilizado para identificar padrões que facilitem a previsão de movimentações deste mesmo objeto. Além disso, o histórico de movimentação possibilita responder consultas referentes a localizações ocupadas por esse objeto no passado.

A geração de padrões poderá responder questões relacionadas a localizações passadas com maior rapidez, uma vez que o histórico de movimentação do objeto já foi utilizado no processo de geração deste padrão. Portanto, o padrão fornece por si só uma informação com significado associado e reduz o volume de dados armazenados e necessários para responder as questões relacionadas ao objeto móvel.

A definição de padrões relacionados a objetos móveis indica que as mudanças temporais destes objetos tendem a possuir um único padrão regular, e este padrão pode ser traçado utilizando-se de técnicas de mineração de dados temporais (CHUNG *et al* 2002). Ainda, segundo os mesmos autores, a mudança de localização desses objetos móveis pode ocorrer com um padrão discreto ou contínuo e, assim, pode ser descrita como um ponto no tempo ou períodos de tempo e, desde que seja possível descrever estas mudanças contínuas no mundo real, a movimentação desses objetos móveis pode ser obtida a partir da localização desses pontos discretos, onde cada ponto representa o início e fim do tempo estendido.

A identificação de um padrão de movimentação para cada objeto móvel pode caracterizar particularidades inerentes a ele, definindo uma espécie de perfil ou particularidade para cada objeto. Existem diversas formas de caracterizar o padrão de movimentação de objetos móveis, como por exemplo:

- De acordo com Beard e Palancioglu (2000), a identificação e atribuição de perfis às classes de objetos podem distinguir diferentes comportamentos de movimentação que auxiliam e tornam o processo preditivo mais confiável. Ainda, segundo os mesmos autores, a identificação de padrões pode ser utilizada como modelo de pesquisa para se detectar outros padrões similares.
- Jin *et al* (2006) propõem a definição de um padrão baseado num esquema de calendário relacional, no qual os intervalos de tempo seriam especificados pelo usuário. Neste esquema de calendário relacional os atributos seriam uma granularidade do tempo, como: ano, mês, dia, etc., que representariam ciclos periódicos de ocorrência do padrão, como por exemplo, cada semana, cada mês, e assim sucessivamente. Dessa forma, o padrão estaria representado por um vetor de atributos da seguinte maneira: <2010,1,*,3>. Neste exemplo, o vetor representa um padrão que ocorreu no ano de 2010, no mês de janeiro, durante todas as semanas (representado pelo asterisco) e no terceiro dia da semana.

Wolf *et al* (2004), defendem que as informações de velocidade, localização e tempo não são suficientes para análises de comportamento de viagens quando o objeto

móvel considerado for um veículo. Especificamente, esse conjunto de dados é ausente da clara identificação dos destinos das viagens e de seus correspondentes propósitos.

De encontro a afirmação de Wolf *et al.* (2004), vem a necessidade de se buscar maneiras de identificar um padrão de movimentação para o veículo rastreado considerando a informação de origem e destino. Fundamentalmente, isto significa tornar possível atribuir um significado a esta origem ou destino como, por exemplo: município, bairro, casa, trabalho, colégio ou periodicidade de ocorrência.

Outra preocupação que surge está relacionada ao dimensionamento da amostra necessária para satisfazer a geração de um padrão. De acordo com Lee-Gosselin (2005 *apud* Chung; Shalaby, 2005), a necessidade de longos períodos de observação está aumentando, em virtude das mudanças de padrão da vida urbana. Não há mais um “dia de semana simples” versus “dia do final de semana”, mas sim variados padrões da atividade humana.

Tais afirmações corroboram com a importância que a identificação e geração de padrões de objetos móveis possuem no processo de extração das informações e processos preditivos relacionados a estes objetos, e sugere diversas maneiras de se obter e armazenar estes padrões.

2.1.3 Previsão de Movimentação de Objetos Móveis

Considerando-se que o dado espaço-temporal de um determinado objeto móvel esteja armazenado num banco de dados, torna-se possível gerar e armazenar padrões em função do histórico de movimentação desse objeto, os quais podem ser utilizados para prever o posicionamento desse objeto num instante de tempo. De acordo com Oliveira (2003), o histórico de movimentação pode melhorar a confiabilidade das previsões realizadas, e também pode caracterizar movimentos que são realizados com certa frequência, ou seja, padrões de movimentação.

Estudos como o de Nhan e Ryu (2006) enfatizam a importância da descoberta de padrões de movimentação associados a uma variável espaço-temporal para a previsão da localização de objetos móveis, baseados em regras de movimentação. Estas regras consideram uma unidade espaço-temporal definida como a mínima extensão de espaço e tempo de interesse, decompostos em cubos retangulares espaços-temporais, nos quais as trajetórias realizadas pelos objetos móveis “chocam-se” com as células espaciais num determinado instante de tempo. Esta interação entre as trajetórias e as células seria utilizada para a previsão.

Desta perspectiva, podem ser consideradas não apenas células retangulares, mas também regiões espaços-temporais com dimensões diversas e significados associados como, por exemplo, origem e destino, definidos por meio de um relacionamento topológico e temporal com as trajetórias realizadas pelos objetos móveis. Estes relacionamentos, associado com a informação temporal das regiões seria utilizado na previsão de localização da entidade móvel.

Segundo Xu e Lee (2003), se a movimentação de um objeto for um reflexo do seu padrão de movimentação total, uma parada (denominada pelo autor como nó sensor de rede) pode ser capaz de prever a direção de movimentação desse objeto, utilizando-se de um modelo de previsão direcional.

De acordo com Hellings e Aertse (1994), dois tipos de dados podem ser obtidos a partir da investigação da movimentação de veículos: a conexão dos dados de tempos de viagem e os dados de Origem-Destino (O-D). Na prática, somente dois valores são conhecidos, que seriam: o número de partidas e chegadas em relação à origem e destino; e o número total de veículos investigados que realizam viagens na rede estabelecida entre a O-D durante um período de tempo. Assim, a probabilidade de ocorrência que uma viagem recebe poderia ser representada pela frequência relativa obtida a partir do número de partidas da origem (i) em relação ao destino (j).

Considerando-se o que foi descrito nesta seção, seria possível desenvolver um modelo preditivo de localização para objetos móveis que considerasse a probabilidade associada entre uma origem e destino. Esta probabilidade seria

estimada com base no número de partidas a partir de uma origem e chegadas a um destino, considerando o histórico de movimentações do veículo investigado.

2.1.4 Rastreamento de Veículos

Segundo Rodrigues, Cugnasca e Queiroz Filho (2009), o rastreamento de veículos pode ser considerado como um sistema que monitora e gerencia a localização e o estado de um determinado veículo a cada instante de tempo, enquanto este veículo se descola sobre a superfície terrestre. Portanto, considera-se o veículo como um sistema dinâmico, pois ele possui atributos dinâmicos como, por exemplo, posição geográfica e, por consequência, diversos estados definidos com base nos significados que estes atributos possuem para o sistema. Este sistema que se integra ao sistema de rastreamento de veículo.

Ainda que o rastreamento de veículos tenha sido concebido originalmente com o intuito de auxiliar a gestão de tráfego, sua aplicação abrange outras categorias como transporte de cargas, gerenciamento de transporte coletivo urbano, gestão de viaturas policiais e hospitalares, guinchos e táxis, e recuperação de veículos roubados. Portanto, os veículos rastreados nada mais são do que entidades móveis que constantemente estão alterando seu estado em função da sua posição geográfica e instante de tempo. Além de outros atributos, quando disponíveis e de interesse para o sistema responsável pelo monitoramento como, por exemplo, uma porta aberta ou temperatura do baú. A esta constante alteração da localização, pode-se categorizar dois níveis diferentes de estado: “parado” ou “em movimento”.

O veículo terá seu estado considerado como “parado”, nesta pesquisa, quando houver uma interrupção de movimentação ocorrida durante um intervalo de tempo (independente da taxa de coleta das posições), e que seja suficiente para identificar quatro segmentos de deslocamento inferior ao erro considerado do posicionamento GPS quando analisado, no máximo, sete segmentos consecutivos.

Inicialmente, imaginou-se que quatro segmentos evidenciarão o estado do veículo. Porém, experimentos mostraram que a ocorrência de erros aleatórios no

posicionamento prejudicava esta afirmação. Assim, optou-se pela ocorrência dos quatro segmentos numa quantidade máxima de sete segmentos caracterizando um período máximo de observação (independente da taxa de coleta das posições) para uma afirmação de mudança de estado do veículo.

Segundo Alesheikh (2008), em termos geométricos, o movimento de um objeto é chamado de trajetória, ou seja, a trajetória é o resultado do rastreamento do veículo no tempo. E a representação da trajetória é suficiente para obter certas propriedades e relacionamentos do movimento do objeto, estes relacionamentos se referem a outros objetos espaciais como, por exemplo, limites municipais, estradas, etc.

Nesta pesquisa, considera-se como trajetória a sequência de pontos que representa um movimento realizado pelo veículo entre duas paradas consecutivas. O que define um caminho percorrido pelo veículo entre estas paradas num determinado período de tempo.

As paradas irão representar as possíveis origens ou destinos de um veículo num determinado período de tempo. Considerando-se como origem o ponto inicial de uma ação ou coisa, que tem continuidade no tempo e/ou no espaço e, como destino o local para onde o veículo está indo.

A representação de uma parada armazenada no padrão será considerada como uma origem quando a última parada realizada pelo veículo possuir alguma interatividade com a geometria armazenada no bando de dados do padrão. Por outro lado, uma parada será considerada como um possível destino, quando possuir um relacionamento topológico com a parada definida como origem, relacionamento este que será expresso por uma trajetória percorrida entre essas duas paradas, num determinado instante de tempo.

De acordo com Wolf *et al.* (2004), as paradas efetuadas por um veículo ocorrem em diferentes instantes de tempo e, em muitos casos, podem representar uma mesma região geográfica. As coordenadas que representam uma parada podem variar significativamente por várias razões, como por exemplo, em função do uso de diferentes vagas de estacionamento em diferentes ocasiões.

Considerando-se esta variabilidade de ocorrência da localização de uma mesma parada, torna-se necessário executar análises topológicas, inclusive de proximidade, que definam *clusters* de paradas ocorridas em diferentes ocasiões, mas que na realidade se referem a uma mesma parada. Uma vez definido um *cluster* de parada, esta aglomeração de paradas irá representar esta única parada que poderá ser considerada como uma origem ou destino, ou seja, uma região de parada.

2.2 ASPECTOS TÉCNICOS

No item anterior foram vistos vários conceitos relativos a objetos móveis, sobretudo no que se refere ao rastreamento de veículos. Também foram descritos os desafios e trabalhos propostos para geração de padrões e previsão de localização dos objetos móveis. Neste item, é descrito a estruturação que um sistema AVL possui, bem como alguns conceitos técnicos relacionados a este tipo de sistema.

Os assuntos foram organizados para representar a importância do rastreamento de veículos num sistema AVL, bem como os componentes que um sistema desse tipo possui, às técnicas utilizadas para a aquisição da posição do veículo e a maneira de transmissão de dados entre o objeto rastreado e a central de monitoramento e vice-versa.

2.2.1 Sistemas AVL

Como mencionado anteriormente, o atributo temporal é responsável por informar a data e hora de aquisição da posição de um veículo monitorado por um sistema AVL. Este atributo é o único de valor esperado e conhecido, pois, pode ser pré-estabelecido na concepção do sistema.

Portanto, o atributo temporal define a taxa de coleta dos dados, que pode ocorrer a cada 15 ou 3 minutos, ou ainda, minuto a minuto, estabelecido em função do propósito do sistema. Caso o valor deste atributo não esteja dentro do intervalo pré-

estabelecido o sistema pode estar sofrendo algum tipo de interferência ou anomalia. Isso poderá ocorrer em virtude do veículo trafegar em regiões onde não é possível enviar as informações coletadas ou no caso de não ser possível determinar a posição do veículo.

Estas regiões serão chamadas de “regiões de sombra do sistema” e serão consideradas como interferências, uma vez que impossibilitam a comunicação entre a central de monitoramento e o dispositivo móvel. As anomalias estarão relacionadas ao mau funcionamento dos dispositivos que compõem o sistema e merecem, portanto, uma atenção especial para efeitos de manutenção destes dispositivos.

É importante ressaltar que, se o usuário optar por tecnologias de posicionamento como a GPS, e as informações do rastreamento forem enviados via satélite, dificilmente surgirão estas regiões de sombra, mas esta escolha pode encarecer demasiadamente o custo operacional do sistema e torná-lo inviável dependendo do seu propósito estabelecido. Além disso, esta escolha não garantirá totalmente a exclusão dessas regiões de sombra, o que será detalhado nas seções: 2.2.2 (Técnicas de Posicionamento Utilizadas em AVL) e 2.3.2 (Transmissão de Dados Utilizada em AVL).

Os sistemas de Localização Automática de Veículos são compostos basicamente por quatro componentes e podem ser visualizados na figura 5:

- 1 Aquisição de Dados e Atuação – dispositivos embarcados no veículo;
- 2 Sistema de Comunicação – são as tecnologias de transmissão e recepção dos dados entre uma central de operação e o veículo;
- 3 Gestão da Informação – são os processos que gerenciam as informações obtidas a partir dos atributos;
- 4 Aplicações – são as finalidades a que se destina o monitoramento das entidades móveis, como: aplicações comerciais, segurança, logística etc.

Figura 5. Componentes de um Sistema AVL. (Fonte: Adaptado de Rodrigues; Cugnasca e Queiroz Filho 2009, p. 17).

2.2.1.1 Aquisição de Dados e Atuação

Este componente do sistema AVL compreende desde as diversas tecnologias existentes para a aquisição da posição geográfica do veículo, até os equipamentos embarcados no veículo, os quais são responsáveis pelas informações que traduzem o estado do veículo. As tecnologias disponíveis para a obtenção da localização do veículo serão descritas, oportunamente, na seção 2.2.2 (Técnicas de Posicionamento Utilizadas em AVL).

2.2.1.1.1 Os Equipamentos Embarcados

Os equipamentos embarcados, como mencionado previamente, são aqueles dispositivos responsáveis pela tradução do estado do veículo. De acordo com Rodrigues, Cugnasca e Queiroz Filho (2009), os equipamentos embarcados são responsáveis pelas mais diversas funções, como por exemplo, o armazenamento e o processamento de dados, a comunicação e detecção de estado do veículo, a

atuação da central de monitoramento sobre o veículo e a interação com o motorista. Numa configuração simples, um veículo estaria equipado com:

1. **Módulo de processamento** – responsável pelo processamento e armazenamento dos dados a bordo do veículo e pela integração com os sensores, atuadores e outros periféricos como os leitores de código de barra;
2. **Módulo de interação** – responsável pela troca de informações entre o tripulante ou motorista e a central de monitoramento, além de gerenciar algumas funções do veículo;
3. **Sensores e Atuadores** – são os dispositivos responsáveis pelo fornecimento dos dados que indicam o estado do veículo. Enquanto os atuadores são os dispositivos que irão permitir a intervenção da central de monitoramento sobre o veículo.

A figura 6 ilustra os equipamentos embarcados que um veículo de passeio pode possuir:

Figura 6. Equipamentos Embarcados de um veículo leve.

2.2.1.2 Sistemas de Comunicação

Similarmente às diversas tecnologias existentes para a obtenção da localização geográfica dos veículos, existem várias tecnologias disponíveis que podem ser utilizadas pelos sistemas de comunicação. Este sistema será responsável pela comunicação entre o veículo rastreado e a central de operação responsável pelo monitoramento do mesmo. A comunicação pode ser tanto por texto (transmissão de

dados), quanto por voz, e pode ocorrer em ambos os sentidos (veículo → central, ou central → veículo).

As tecnologias de comunicação mais utilizadas pelos sistemas AVL são: a comunicação via rádio, a comunicação via celular ou via satélite. A opção por uma dessas tecnologias de comunicação depende do propósito de implantação do sistema AVL, uma vez que cada uma possui suas vantagens e desvantagens, a tabela 2 ilustra as vantagens e desvantagens dos sistemas de comunicação.

Tabela 2 – Comparativo entre os Sistemas de Comunicação

Sistemas de Comunicação	Vantagens	Desvantagens	Utilização recomendada
Satélite Geoestacionário	Disponibilidade de sinal	Preço elevado	Viagens longas; Localidades desprovidas de cobertura celular; Monitoramento de objetos de alto valor econômico.
Satélite de Órbita Baixa	Equipamento e comunicação mais baratos do que o geoestacionário; Cobertura mundial.	Sinal nem sempre disponível	Viagens longas, Monitoramento de objetos de menor valor ou risco; Menor taxa de frequência de posições/dia.
Celular	Equipamento simples e barato; Os protocolos CDMA, GSM e GPRS têm custos de comunicação menores.	Sinal restrito às áreas de cobertura (em contínua expansão)	Localidades urbanas e rodovias com sinal de celular
Rádio	Comunicação extremamente barata; Sinal sempre presente na área de cobertura.	Pequeno raio de cobertura; Custos altos para manter e ampliar cobertura.	Localização em regiões conhecidas e com cobertura; Localidades urbanas.

Fonte: Rodrigues; Cugnasca e Queiroz Filho (2009, p. 33)

2.2.1.3 Gestão da Informação

Todo o gerenciamento das informações dos dados de localização e estado do veículo é realizado pelos Sistemas de Gestão da Informação. A gestão da informação pode ser centralizada ou descentralizada e requer equipamentos de comunicação, armazenamento e processamento dos dados recebidos.

A gestão centralizada pressupõe a existência de uma central de controle. É nesta central que se dá a integração das informações levantadas do veículo, a partir dos dados de rastreamento com recursos de geoprocessamento. (RODRIGUES, 2008a).

Optar pela gestão centralizada implica numa abordagem na qual uma central de controle será a unidade responsável pelo gerenciamento de toda a rede de infraestrutura referente aos modos de comunicação que serão utilizados (rádio, celular, satélite etc.), integração desses modos, processamentos dos dados, servidores de dados e mapas, bancos de dados, interfaces de visualização, execução de *software* específico para cada tipo de gestão (ex. controle de frotas, sistemas de entregas urbanas, gerenciamento de resgates, etc.) e, finalmente, divulgação dos dados (ex. sob consulta ou relatórios periódicos).

No caso dos sistemas AVL que optam pela gestão descentralizada, Rodrigues (2008a) afirma que:

[...] utilizam a Web e estão relacionadas ao modelo Application Service Provider (ASP). Esses serviços oferecem soluções específicas para uma grande variedade de operações a custos baixos, uma vez que o usuário deixa de implantar um sistema próprio de gestão. Com isso, há grande redução de custo, pois se economiza com hardware, software, compra e manutenção da base de dados, treinamento e gestão de pessoal, desenvolvimento de aplicativos e, sobretudo, o custo de atualização tecnológica. Este último onera pesadamente uma empresa, mas quando compartilhado por inúmeras empresas de um serviço ASP, torna-se baixo.

2.2.2 Técnicas de Posicionamento Utilizadas em AVL

Em virtude dos avanços tecnológicos na área de posicionamento, diferentes técnicas de posicionamento podem ser utilizadas em AVL, para distintas circunstâncias. A técnica de posicionamento utilizada pode ser definida de acordo com a natureza da operação do veículo, a oferta de tecnologia para a região onde o veículo trafega, bem como a natureza do sistema de gestão da informação.

De acordo com Rodrigues, Cugnasca e Queiroz Filho (2009), as principais técnicas de posicionamento baseiam-se nos sistemas de navegação global por satélites, na

trilateração das antenas de frequências de rádio, na rede de telefonia móvel e em postos de sinalização (*signpost*). Ainda, segundo os mesmos autores, a opção por uma dessas técnicas de posicionamento é determinada pelo tipo de atuação que se terá sobre este veículo e que pode variar conforme:

- O tipo de carga;
- O trajeto a ser percorrido;
- As características da via de transporte;
- O tipo de região percorrida pelo veículo; e
- O risco previsto.

O presente trabalho se restringirá a descrever a técnica de posicionamento que utiliza os Sistemas de Navegação Global por Satélites, pois foi esta a técnica empregada para a obtenção dos dados utilizados nesta pesquisa.

2.2.2.1 Posicionamento do Veículo utilizando o GNSS

Uma das técnicas utilizadas para a determinação das posições dos veículos em sistemas AVL é o posicionamento por meio de satélites artificiais. A opção por este tipo de técnica se deve à abrangência fornecida por estes sistemas de satélites artificiais e também pela acurácia¹ do posicionamento proporcionada pelos mesmos. A sigla GNSS é o acrônimo de *Global Navigation Satellite System* e, segundo Seeber (2003), é a nomenclatura geral dada a todos os Sistemas de Navegação Global por Satélite.

Podem-se citar pelo menos três sistemas desse tipo: o GPS, o GLONASS e o GALILEO. Dentre eles, o mais utilizado em sistemas AVL é o GPS e, será referente a este sistema que, o presente trabalho se aterá na seção seguinte. Por hora, basta saber que o GLONASS (*GLObal NAVigation Satellite System*) é operado pela Federação Russa e está disponível para a comunidade civil, tendo sido concebido na década de 70 pela extinta União Soviética para fins militares, enquanto que o GALILEO é um programa de navegação por satélite europeu, projetado pela União

¹ “Acurácia é a proximidade entre o valor obtido experimentalmente e o valor verdadeiro na medição de uma grandeza física.” (HOUAISS, 2001).

Europeia em cooperação com a Agência Espacial Europeia (ESA), totalmente controlado por autoridades civis.

2.2.2.1.1 O Sistema de Posicionamento Global - GPS

Segundo Monico (2000, p. 19): “[...] posicionar um objeto nada mais é do que atribuir-lhe coordenadas [...]”. Pensando nisto, na década de 1970 uma solução para este problema foi apresentada com a proposta do GPS (*Global Positioning System*), desenvolvido pelo Departamento de Defesa dos Estados Unidos da América – DoD (*Department of Defense*). Inicialmente denominado como NAVSTAR-GPS (*NAVigation Satellite with Time And Ranging*) e com o objetivo de auxiliar a navegação das forças armadas americanas, o sistema GPS surgiu da fusão de outros dois sistemas (*Timation* e o *System 621B*), sob responsabilidade da Marinha e da Força Aérea norte-americanas, respectivamente.

Em virtude da acurácia e disponibilidade de uso (sob quaisquer condições climáticas e de abrangência global) o sistema GPS logo passou a ser utilizado com grande frequência pela comunidade civil, principalmente a partir de dois de maio de 2000, após a eliminação do código SA (*Selective Availability*), que era responsável pela degradação do sinal do sistema e, conseqüentemente, da precisão fornecida pelo mesmo. Essa medida possibilitou aos usuários civis do sistema um posicionamento com uma acurácia dez vezes superior do que se tinha até então (MONICO, 2008).

O método de posicionamento por satélite utilizado para determinar a posição dos veículos nos sistemas AVL é o Posicionamento Absoluto Cinemático. De acordo com Wells *et al.* (1986) “[...] o posicionamento cinemático é utilizado para aplicações em que se deseja determinar a trajetória do objeto móvel [...]”. Ainda segundo os mesmos autores, nestes tipos de aplicações o objeto móvel é o receptor GPS que pode ser transportado por um veículo que percorre uma estrada, um barco, ou uma aeronave.

A concepção do sistema GPS permite ao usuário, a qualquer instante, rastrear no mínimo quatro satélites em qualquer parte da superfície terrestre, ou próxima a ela.

Essa disponibilidade de satélites permite ao usuário um posicionamento em tempo real (MONICO, 2000).

A disponibilidade mínima de satélites oferecida pelo GPS se deve ao fato do segmento espacial ser constituído de 24 satélites distribuídos em seis planos orbitais igualmente espaçados, com quatro satélites por plano, a uma altitude aproximada de 20200 km. Devido a inclinação dos planos orbitais (55°) em relação a linha do Equador, o sistema GPS disponibiliza no mínimo quatro satélites e garante que a posição do satélite se repita diariamente, 4 minutos antes em relação ao dia anterior (MONICO, 2008). A figura 7 ilustra a constelação GPS.

Figura 7. Constelação GPS. (Fonte: http://infosecurity.us/images/gps_constellation.png)

Ainda, segundo o mesmo autor, o princípio básico de navegação por GPS consiste na medida de distâncias entre o usuário (receptor) e os quatros satélites rastreados, na qual, conhecendo-se as coordenadas dos satélites num sistema de referência apropriado, é possível calcular as coordenadas da antena do receptor do usuário no mesmo sistema de referência dos satélites e, assim, determinar a posição da antena em tempo real.

[...] O objeto a ser posicionado pode estar em repouso ou em movimento, o que gera um complemento à classificação com respeito ao referencial adotado. No primeiro caso, trata-se do posicionamento estático, enquanto o segundo diz respeito ao

posicionamento cinemático. No que concerne ao posicionamento utilizando o GNSS, independente do estado do objeto, ele pode ser realizado pelos métodos absoluto e relativo. (MONICO, 2008, p. 279).

De acordo com Monico (2008), pode-se ter, portanto, o posicionamento absoluto estático ou o posicionamento absoluto cinemático. No posicionamento absoluto, também conhecido como posicionamento por ponto, quando são utilizadas efemérides² transmitidas, a posição obtida se refere ao um Sistema de Referência Terrestre utilizado pelo sistema em uso, como, por exemplo, o WGS84 utilizado pelo sistema GPS. A figura 8 ilustra o princípio deste tipo de posicionamento e, considera o posicionamento absoluto cinemático.

Figura 8. Posicionamento Absoluto Cinemático. (Fonte: adaptada de Wells et al. 1986).

Considerando-se a figura 8, pode-se dizer que a distância r_1 (em vermelho) é a distância entre o satélite **S1** e o Geocentro (Centro de Massa da Terra), previamente conhecida e fornecida pelas efemérides transmitidas pelo satélite, ou seja, para cada satélite (**S_n**) há uma distância **r_n**. A distância entre os satélites (**S_n**) e o receptor (**P_i**), formada pelo vetor **S_nP_i** é a distância medida pelo sistema, e o vetor **R_i** é a distância procurada. Portanto, no Posicionamento Absoluto Cinemático, o vetor **R_i** é continuamente determinado. Quando o receptor (antena GPS) está em repouso, a

²"efemérides – tábua astronômica que registra, em intervalos de tempo regulares, a posição relativa de um astro." (Houaiss, 2001).

determinação pode conduzir sempre para um mesmo resultado, produzindo uma redundância de resultados, a menos é claro do erro fornecido pelo sistema.

Como já mencionado, os sistemas AVL possuem como atributos principais a posição do veículo e o instante de tempo da determinação. Esta posição é geralmente expressa em coordenadas geográficas referenciadas a um Sistema de Referência Terrestre; no caso do GPS é o WGS84(G1150), compatível ao ITRF2000. A figura 9 ilustra a definição do sistema de referência WGS84.

Figura 9. Sistema de Referência Terrestre WGS84. Fonte: Seeber (2003).

Segundo Seeber (2003), com a desativação do SA, a acurácia disponível para os usuários do SPS (*Standard Positioning Service*), os quais são usuários civis, se tornou semelhante àquela obtida pelos usuários militares que utilizam o PPS (*Precise Positioning Service*).

O valor da acurácia global média de posicionamento é definida como um erro menor ou igual a 13 metros, considerando-se as componentes X e Y (erro máximo horizontal), e menor ou igual a 22 metros para a componente vertical (Z) com 95% de confiança (DoD, 2001³ apud SEEBER, 2003, p. 230). Ainda, segundo o mesmo

³ DoD (2001): **Global Positioning System Standard Positioning Service Performance Standard**. Department of Defense, Document, October 2001.

autor, experimentos mostram que, na prática, a acurácia alcançável é muito mais alta. A figura 10 mostra o efeito da desativação do SA e ilustra o que o autor sugere.

Figura 10. Efeito do SA nos resultados de posicionamento GPS com um receptor de simples frequência, à esquerda o SA ativo e a direita após a desativação do SA. Fonte: Adaptada de Seeber (2003).

Os satélites GPS transmitem duas ondas portadoras, denominadas de L1 e L2, respectivamente. Estas ondas são geradas a partir de uma frequência fundamental (10,23 MHz) e são multiplicadas por 154 e 120, respectivamente (MONICO, 2000). A tabela 3 demonstra os valores da frequência e comprimento de onda (λ) para ambas as ondas portadoras.

Tabela 3 – Características dos sinais GPS

Onda Portadora	Frequência	Comprimento de Onda
L1	1575,42 MHz	19 cm
L2	1227,60 MHz	24 cm

Fonte: (Monico, 2000, p. 26).

As entidades móveis rastreadas com a tecnologia GPS utilizam o código C/A (*Coarse Acquisition*) que é modulado somente sobre a onda portadora L1. De acordo com Monico (2000), esse código possui comprimento de onda de aproximadamente 300m, transmitido a uma razão de 1,023MHz, e gerado a partir do produto de duas sequencias PN (*Pseudo-random*), denominadas G1 e G2.

Segundo Leick⁴ (1995 *apud* MONICO, 2000, p. 27) o código C/A tem como característica básica a baixa correlação entre seus membros, o que possibilita a rápida distinção dos sinais recebidos, simultaneamente, de vários satélites. Dessa maneira, os receptores GPS que costumam ser usados em rastreamento de veículos são destinados a levantamentos expeditos. Assim, estes receptores podem ser classificados como receptores de navegação ou receptores para SIG (*Sistemas de Informações Geográficas*), quando classificados segundo sua aplicação, ou receptores de código, quando a classificação considerar o tipo de dados proporcionados pelo receptor (MONICO, 2000).

Como mencionado na seção 2.2.1 (Sistemas AVL), a opção por este tipo de tecnologia para a determinação da posição do veículo não garante totalmente a eliminação das chamadas regiões de sombra, uma vez que, em razão do comprimento de onda do sinal GPS, o veículo pode estar em lugares onde as estruturas de metal e concreto não permitem a propagação do comprimento de onda do sinal GPS, criando, assim, uma espécie de barreira para este sinal. São exemplos deste tipo de estrutura: túneis, garagens subterrâneas etc. Alguns sistemas AVL adotam alternativas para contornar este inconveniente do sistema GPS, mas estas alternativas devem ser consideradas satisfatórias com certo critério. Um exemplo dessas alternativas utilizadas é a atribuição da última posição do veículo obtida pelo GPS, antes da perda de sinal, para os demais intervalos de tempo onde não é possível determinar a posição atual do veículo.

2.3.2 Transmissão de Dados Utilizada em AVL

Como mencionado em seções anteriores, existem três técnicas diferentes de comunicação, utilizadas pelos sistemas AVL (via rádio, via celular e via satélite). Nesta seção, no entanto, somente será detalhada aquela utilizada para a obtenção dos dados empregados nesta pesquisa, ou seja, a que faz uso da telefonia móvel.

⁴ LEICK, A. *GPS Satellite Surveying*. New York: John Wiley & Sons, 1995. 560p.

2.3.2.1 Transmissão de Dados Via Telefonia Móvel

Com a expansão da cobertura dos sistemas de telefonia móvel, o uso dessa alternativa de comunicação passou a ser considerado com mais frequência por outros sistemas, como o AVL. Além da cobertura de sinal cada vez maior, os custos de comunicação oferecidos por esta tecnologia e os avanços tecnológicos na transmissão de dados se tornaram atrativos mais que suficientes para os sistemas que optam por este tipo de comunicação. Outro fator importante na expansão desta tecnologia de comunicação, em se tratando de AVL, foi que sua cobertura atingiu as rodovias mais utilizadas do país.

Esta tecnologia de comunicação não só permite a comunicação via voz, como também por mensagens de texto (SMS) e protocolos de troca de dados (por ex. CDMA e GSM). A escolha por um protocolo de troca de dados dependerá, primordialmente, da aplicação do sistema AVL e da área de cobertura do sistema de telefonia móvel, além dos seus recursos disponíveis, a tabela 4 ilustra os recursos disponíveis para o protocolo GSM.

Tabela 4 – Protocolos de transmissão de dados

PROTOCOLO	RECURSOS
CDMA	1xRTT – Comunicação de dados via IP (144Kbs)
	SMS – Via mensagens de texto
	CDMA – Voz e comunicação discada de dados (9600 bps)
	AMPS – Voz e comunicação discada de dados (1200 bps)
GSM	GPRS – Comunicação de dados via IP (115Kbs)
	SMS – Via mensagens de texto
	GSM – Voz e comunicação discada de dados (9600 bps)

Fonte: Rodrigues, Cugnasca e Queiroz Filho. (2009, p. 30).

A cobertura oferecida por estes protocolos de transmissão de dados no Brasil pode ser ilustrada na figura 11. Na qual, de acordo com a UCEL, é apresentada a cobertura fornecida por uma das duas operadoras que possuem cobertura nacional. Em seguida, a tabela 5 destaca o tipo de tecnologia disponibilizada para os estados brasileiros destacadas na figura 11.

Figura 11. Cobertura Nacional GSM da operadora Claro. Fonte: ucel⁵ (2010).

Tabela 5 – Estados brasileiros e tecnologias de comunicação fornecida pela operadora Claro.

Estados Brasileiros	Tecnologias
SP	CDMA/GSM
RJ e ES	CDMA/GSM
BA e SE	GSM
PR e SC	GSM
RS	CDMA/GSM
MT, MS, GO e DF	CDMA/GSM
MA, PI, BA, CE, RN, PB, PE, AL e SE	CDMA/GSM
AM, AP, AC, PA, RO, RR e TO	GSM
MG	GSM

Fonte: adaptada de ucel (2008).

Atualmente, as operadoras esforçam-se em melhorar a velocidade de transmissão dos dados, a fim de proporcionar aos seus clientes serviços de conteúdo multimídia mais elaborados, a evolução dessas redes sem fio digitais de alta velocidade é apresentada na figura 12.

⁵ <http://www.ucel.com.br/mapa.html>

Figura 12. Evolução 3G dos protocolos GSM e CDMA. Fonte: adapta de Apple⁶ (2008).

É importante ressaltar que essa cobertura anunciada pelas operadoras não se aplica a toda extensão geográfica dos Estados brasileiros supracitados e que, portanto, alguns municípios não contam com nenhum tipo de tecnologia de comunicação. O que vem a corroborar o que foi dito sobre a existência de regiões de sombra para um sistema AVL, na quais ou o sistema de posicionamento é obstruído ou o sistema de comunicação não é abrangente.

2.3.2.1.1 O Sistema de Comunicação GSM

Segundo Rodrigues, Cugnasca e Queiroz Filho (2009), o GSM é um padrão de telefonia celular definido principalmente pela ETSI (*European Telecommunications Standards Institute*) que especifica toda a infraestrutura de telefonia celular e não apenas a interface de rádio, que funciona dividindo o espectro de frequência em canais (FDMA), seguida pela divisão do tempo de transmissão em *time-slots* atribuídos para cada conversão (TDMA). A tabela 6 ilustra as bandas utilizadas pelo GSM e número de canais que cada frequência GSM possui. De acordo com os mesmos autores, o primeiro e o último canal da banda de operação não são

⁶ www.apple.com/br/quicktime/technologies/3gpp/

utilizados e servem como banda de guarda para minimizar o efeito de possíveis interferências causadas por outros sistemas.

Ainda, segundo os mesmos autores, os canais de controle do sistema GSM não necessitam estar alocados e definidos *a priori* e podem assumir suas funções atribuídas no momento de demanda, possibilitando que os canais permaneçam na maioria dos casos, disponíveis para o tráfego, o que contribui para a disponibilidade e número de usuários no sistema.

Tabela 6 – Bandas e Canais utilizados pelo GSM.

Frequências	Bandas Transmissão do celular (MHz)		Bandas Transmissão da ERB (MHz)		Número de Canais
GSM-800	880	915	925	960	124
GSM-1800	1710	1785	1805	1880	373
GSM-1900	1850	1910	1930	1990	298

Fonte: adaptada de Rodrigues , Cugnasca e Queiroz Filho (2009, p. 58).

A utilização da tecnologia GSM permite, na maioria dos casos, configurar a identificação dos usuários. Esta configuração é realizada por meio de circuitos integrados *smart-card* denominados SIM (*Subscriber Identity Modules*) que contêm um número com a identificação unívoca do assinante, independente do equipamento empregado para acessar a rede de telefonia móvel (RODRIGUES; CUGNASCA; QUEIROZ FILHO, 2009).

Alguns dos serviços adicionais oferecidos pelo sistema GSM é o envio de mensagens de texto. Entretanto, como essa tecnologia foi concebida para aplicações de voz, as aplicações em comunicações de dados não são satisfatórias. Uma medida utilizada para contornar essa limitação foi desenvolver um mecanismo adicional de acesso exclusivo para dados, denominado GPRS (*General Packet Radio Service*) que permite aos *time-slots*, não utilizados de um determinado canal de frequências, serem momentaneamente empregados para o envio de um pacote de dados.

2.4 CONSIDERAÇÕES FINAIS

Como observado, este capítulo teve por objetivo fundamentar aspectos técnicos e conceituais que são abordados nessa pesquisa. A contextualização desses aspectos irá subsidiar os capítulos seguintes e esclarecer eventuais dúvidas relacionadas a termos técnicos e conceituais.

3 RECURSOS UTILIZADOS NA PESQUISA

Neste capítulo são abordados assuntos referentes aos recursos utilizados na pesquisa. Será exposto o tipo de dados de rastreamento, a área de estudo considerada como base de informação cartográfica e os recursos computacionais utilizados.

Na primeira seção detalha os arquivos do rastreamento dos veículos, quanto à sua aquisição, período de observação e número de veículos rastreados. A seção seguinte descreve a região geográfica de estudo e o detalhamento das informações cartográficas disponíveis para essa área. Por fim, na última seção, serão detalhados os recursos computacionais utilizados, como a configuração da máquina e o sistema operacional.

3.1 DADOS DO RASTREAMENTO

Beard e Palancioglu (2000) consideram que os dados de rastreamento se referem a qualquer tipo de medida que adquira a localização de um ou mais objetos do mundo real, desde imagens de satélite, fotografias aéreas, vídeos, observações GPS ou observações de telemetria.

Os dados utilizados nesta pesquisa foram cedidos por empresas de rastreamento, mediante acordo de garantia de sigilo. Portanto, a numeração de identificação dos veículos foi alterada e os resultados provenientes da geração de padrões e previsão de destinos não estão especificados com os verdadeiros nomes das localidades em que os veículos da amostra estiveram durante o período observado.

As observações do rastreamento dos veículos foram obtidas utilizando a tecnologia de posicionamento GPS e transmissão de dados GSM/GPRS. A taxa de coleta dessas informações foi de três minutos e a amostra foi constituída por 13 veículos, observados por um período ininterrupto superior a 30 dias.

Segundo Chung e Shalaby (2005), quando se utiliza o GPS para levantamentos de viagem, existem algumas limitações que afetam a qualidade dos dados coletados. O primeiro problema é a “partida” a quente ou frio. Geralmente, os receptores GPS necessitam de um tempo de aquisição para obter sua posição atual depois de serem ligados. O período consumido (durante o qual não se tem dados GPS disponíveis) varia entre 15 segundos e 5 minutos dependendo do estado do receptor GPS. A segunda limitação está relacionada ao sinal do GPS que sofre obstrução quando o receptor está próximo a cânions urbanos (edificações altas em ambas as faces da via).

De acordo com os mesmos autores estes cânions não apenas podem obstruir o sinal dos satélites GPS, como também podem refleti-los de forma irregular. Dessa maneira, os receptores GPS quando próximos a edifícios altos podem não produzir dados confiáveis de localização. A terceira limitação é o teto dos veículos ou estacionamentos cobertos que também podem bloquear o sinal.

Os dados de rastreamento utilizados nesta pesquisa possuem um comportamento que evidencia a ocorrência de situações de perda do sinal GPS. Este comportamento foi contornado pela empresa responsável pelo rastreamento e configurado no dispositivo rastreador. O trabalho de contorno consiste em informar a última posição obtida pelo receptor GPS antes da perda de sinal, ou seja, se o período de perda de sinal GPS for superior à taxa de coleta/envio da informação, a base de dados do rastreamento irá conter sequencias de pontos de mesma localização em instantes de tempo distintos. Esta medida é uma maneira de garantir uma informação mais próxima da realidade atual do veículo antes da perda do sinal GPS.

Além disso, a base de dados de rastreamento está sujeita a outros tipos de anomalias comuns em bancos de dados como, por exemplo, registros duplicados. Nestes casos, além das coordenadas duplicadas, o instante de tempo da aquisição dessas informações posicionais também estará duplicado. A amostra de dados referente aos 13 veículos considerados, durante o período de observação, com taxa de coleta de três minutos, gerou um volume de registros no banco de dados de aproximadamente 240.000 registros.

3.2 ÁREA DE ESTUDO

A região geográfica considerada nesta pesquisa abrange todos os municípios do Estado de São Paulo. O dimensionamento da área de estudo a ser considerada, levou em conta somente a extensão do Estado de São Paulo devido ao fato da distribuição dos dados de rastreamento estar restrita ao limite deste Estado, como pode ser verificado na figura 13.

Figura 13. Extensão Geográfica dos dados de rastreamento.

Apesar da grande maioria da concentração dos dados de rastreamento estar vinculada à região metropolitana do Município de São Paulo, optou-se por considerar toda a extensão territorial do Estado para avaliar o desempenho das funções desenvolvidas na pesquisa, tanto no processo preditivo quanto no processo de geração dos padrões daqueles veículos que eventualmente realizam grandes viagens. O objetivo principal era avaliar o comportamento dos algoritmos quando expostos às amostras obtidas em longos trajetos e sujeitos à falta de comunicação, ou seja, a não transmissão dos dados de rastreamento por um ou vários períodos de tempo ao longo do trajeto.

3.2.1 Base Cartográfica

A base cartográfica foi composta pelos limites municipais, bem como pelos limites dos respectivos bairros quando esta informação era disponível, ou seja, nem todos os Municípios do Estado de São Paulo possuíam os limites dos seus bairros. A figura 14 ilustra a base cartográfica utilizada na pesquisa.

Figura 14. Base Cartográfica – Estado de São Paulo.

O sistema de coordenadas considerado para a base cartográfica foi o mesmo utilizado pelo sistema GPS, isto é, o WGS84. Optou-se por esse sistema de coordenadas em virtude dos dados de rastreamento utilizados na pesquisa também estarem neste sistema de coordenadas. Portanto, como estes dados eram obtidos a partir do posicionamento GPS, para evitar a necessidade de conversões de coordenadas decidiu-se utilizar uma base cartográfica que estivesse no mesmo sistema de coordenadas.

A base cartográfica possui somente a informação referente aos nomes dos municípios (quando disponíveis), o código do Estado, o código do próprio Município e a geometria do tipo área responsável por representar o limite municipal. No caso

da base de dados dos bairros as informações são semelhantes, contêm somente os nomes dos bairros, o código do Município, o código do Estado e a geometria do tipo área responsável por representar a extensão do bairro. Esta base cartográfica foi armazenada no banco de dados, disposta em duas tabelas denominadas CITIES_SP e NEIGHBORHOODS_SP, e as informações presentes em cada uma delas pode ser vinculada utilizando o código do município comum às duas tabelas.

3.3 APOIO COMPUTACIONAL

Esta seção tem por finalidade descrever a configuração de máquina utilizada para os experimentos e desenvolvimento das funções desta pesquisa. O desenvolvimento foi realizado sobre o ambiente operacional Windows® e a configuração da máquina utilizada foi:

- Sistema Operacional – Microsoft Windows XP *Home Edition* Versão 2002 *Service Pack 3*;
- Processador – Intel® Core™ T5500 (2M *Cache*, 1.66GHz, 997 MHz);
- Memória RAM – 1GB;
- Disco Rígido (*HD*) – 80GB.
- Aplicativos:
 - PostGIS – versão 1.3.6
 - PostgreSQL – versão 8.2
 - ArcMap – versão 9.1
 - gvSIG – versão 1.9;

3.4 CONSIDERAÇÕES FINAIS

O objetivo principal deste capítulo foi descrever os recursos que foram utilizados nesta pesquisa, tanto no que tange às características e origem dos dados, sejam eles oriundos do rastreamento dos veículos ou aqueles que compõem a base cartográfica, quanto à configuração de máquina utilizada para executar as funções desenvolvidas.

4 METODOLOGIA

Este capítulo apresenta a metodologia proposta para a geração dos padrões relacionados aos veículos rastreados, e o procedimento executado para obtenção da previsão dos prováveis destinos de um determinado veículo. É importante ressaltar que a metodologia proposta considera um sistema no qual os dados do histórico de rastreamento do veículo foram ou são armazenados continuamente por um período mínimo de 30 dias, e a previsão é fornecida considerando-se valores de probabilidade obtidos pelas funções desenvolvidas.

Além disso, os algoritmos desenvolvidos nesta pesquisa consideram apenas como objetos móveis aqueles objetos que possam ser representados por geometrias pontuais (ex: veículos, pessoas etc.) e cuja localização foi obtida utilizando o posicionamento via satélite. Toda a modelagem desenvolvida considera veículos leves (do tipo passeio).

4.1 ETAPA I – PREPARAÇÃO DOS DADOS

O objetivo principal desta seção é descrever os procedimentos realizados sobre a amostra de dados de rastreamento, descrevendo o formato original, os procedimentos executados que permitiram sua utilização e, discutir a necessidade de cada procedimento.

Os arquivos de dados cedidos para a pesquisa estavam no formato ***.csv** (arquivo com valores separados por vírgulas) e, antes da utilização destes dados, foi realizado um tratamento prévio para a inserção desses dados num banco de dados. O SGBD (Sistema Gerenciador de Bancos de Dados) utilizado na pesquisa foi o PostgreSQL na versão 8.2, acrescido da sua extensão espacial PostGIS na versão 1.3.6.

O PostGIS possui um executável chamado **shp2pgsql.exe**, o qual é utilizado para fazer a inserção de dados espaciais no banco de dados a partir de um arquivo

shapefile. Este executável interpreta os dados do arquivo shapefile e gera um arquivo no formato ***.sql** com a estrutura da tabela que será criada no banco de dados, juntamente com os dados que serão inseridos nesta tabela.

Quando a opção de inserção dos dados for a partir desse executável, os dados do arquivo ***.csv** devem ser previamente geocodificados em algum aplicativo GIS (por exemplo: ArcGIS – ArcMap ou gvSIG) e posteriormente exportados para o formato shapefile para que possam ser utilizados neste executável.

De acordo com ESRI (1998) um arquivo shapefile é composto por um arquivo principal no formato ***.shp**, um arquivo de indexação (***.shx**) e uma tabela dBASE (***.dbf**). O arquivo principal corresponde ao arquivo de acesso direto que descreve, em forma de lista, o tipo de variável, tamanho do registro (ex: 4bit) e os vértices que cada registro possui. No arquivo de índice cada registro contém um apontamento para o registro correspondente no arquivo principal. A tabela dBASE contém os atributos que cada registro possui, e o relacionamento um-para-um entre a geometria e atributos baseia-se no número do registro. Os registros dos atributos na tabela dBASE devem estar na mesma ordem que os registros estão no arquivo principal. Portanto, um arquivo shapefile será composto de no mínimo três arquivos como, por exemplo:

- Arquivo Principal: bairros.shp
- Arquivo de Índice: bairros.shx
- Tabela dBASE: bairros.dbf

Uma vez que tenham sido gerados os arquivos shapefile, eles devem ser armazenados no mesmo diretório em que se encontra o executável, disponível em: **...\\PostgreSQL\\8.2\\bin**. Depois, utilizando a interface de linha de comando, disponível na instalação do SGBD, é necessário executar os procedimentos descritos no APÊNDICE B.

Portanto, a etapa de preparação dos dados consistiu da:

1. Geocodificação dos dados contidos no arquivo ***.csv**;
2. Exportação dos dados geocodificados para o formato shapefile e,

3. Utilização dos executáveis **shp2pgsql.exe** e **psql.exe** para a inserção dos dados no banco de dados.

A etapa de inserção dos dados também considera a necessidade de validação dos formatos considerados no momento de inserção dos dados no banco, por exemplo, a informação referente ao instante de tempo da aquisição das coordenadas deve estar num formato de tempo e, de acordo com a figura 1 do APÊNDICE B, esta informação está sendo inserida com uma string de 254 posições. Sendo assim, é necessário fazer uma conversão desta informação, de tal forma que o dado passe a ser do tipo *“timestamp without time zone”*.

É importante ressaltar que este procedimento de carga dos dados garante que eles quando inseridos no banco de dados possuam o tipo da geometria associada a cada registro, como pode ser verificado na coluna ***the_geom*** do arquivo de inserção, ilustrado pela figura 1 do APÊNDICE B. Porém, o desenvolvimento dos algoritmos desta pesquisa desconsidera a necessidade desse campo geométrico, pois essa geometria é gerada em tempo de execução pelos algoritmos, a partir das coordenadas presentes nos campos da tabela AVL_ALL denominados “latitude” e “longitude”, descritos no APÊNDICE A.

Esta desconsideração do campo geométrico vinculado ao registro foi feita em razão da forma que os dados são transmitidos pelo rastreador à central de monitoramento e são armazenados no banco de dados de rastreamento, ou seja, o banco de dados de rastreamento não possui uma geometria vinculada a cada registro. Geralmente, possui apenas um par de coordenadas associado a um instante de tempo.

O objetivo desta consideração foi não alterar o ciclo executado por um sistema AVL, que venha a considerar a implantação das rotinas desenvolvidas nesta pesquisa. Portanto, uma vez que as coordenadas do rastreamento de um veículo esteja armazenada numa tabela do PostGIS não é necessária essa etapa de preparação dos dados.

4.2 ETAPA II – GERAÇÃO DE PADRÕES

O processo de Geração dos Padrões tem como objetivo principal fornecer subsídios que auxiliem no processo de previsão da movimentação de um determinado veículo, fundamentados nos dados de rastreamento. Este processo deverá fornecer informações suficientes para responder questões como:

- Para onde um determinado veículo está indo?
- Quais são os locais que um determinado veículo costuma estar em determinados dias da semana ou período do dia?
- Qual é a localidade que o veículo costuma estar na maior parte do tempo?
- O que caracteriza uma parada ou um trajeto realizado pelo veículo?
- Quais informações adicionais seriam interessantes na definição de uma parada ou trajeto?

O processo geral, no qual são gerados os Padrões e a Previsão, está representado na figura 15 e pode ser interpretado da seguinte maneira:

1. O banco de dados irá armazenar durante um período de tempo consecutivo todas as informações do rastreamento dos veículos;
2. Uma função desenvolvida no SGBD será executada periodicamente para que sejam gerados ou atualizados os padrões (aqui divididos em Padrões de Movimentação e Padrões de Localização);
3. Os padrões gerados serão armazenados em tabelas distintas criadas no próprio banco de dados;
4. As informações armazenadas nas tabelas dos padrões serão utilizadas por outra função responsável pela previsão do provável Destino de um determinado veículo.

Figura 15. Processo Geral para a Geração dos Padrões e Previsão.

O processo de geração de padrões é executado periodicamente, de maneira que as informações armazenadas nas tabelas de Padrão de Movimentação e Padrão de Localização sejam atualizadas constantemente. Já o processo de previsão será executado sob demanda, ou seja, quando o usuário desejar obter a informação de **“Para onde o veículo está indo ou irá?”**, desde que o veículo consultado possua um Padrão de Movimentação e Localização para responder a este questionamento.

As fontes de informação utilizadas no processo de geração de padrões foram: uma base cartográfica, e os dados de rastreamento obtidos com o posicionamento GPS que, na concepção de um sistema AVL, são armazenados diretamente numa tabela do banco de dados. No caso desta pesquisa esta tabela foi denominada AVL_ALL (APÊNDICE A), caracterizada com os seguintes atributos: VEHICLE_ID, LATITUDE, LONGITUDE, DATE_TIME, responsáveis por armazenar a informação referente ao identificador do veículo, o valor das coordenadas da latitude e longitude e o instante de tempo da aquisição destas coordenadas, respectivamente. Assim, o processo de geração de padrões inicia-se a partir da distinção do estado de cada ponto de rastreamento, ou seja, da detecção dos pontos que definem um trajeto e daqueles que compõem uma parada.

4.2.1 Detecção de Trajetos

Este procedimento é o principal na geração dos padrões, por que é responsável por identificar o estado de cada ponto de rastreamento do veículo. Com base no resultado desse procedimento serão definidos os pontos de rastreamento que compõem as paradas realizadas pelos veículos e, a partir destas paradas que serão gerados os Padrões de Localização.

A importância da descoberta dos pontos de rastreamento que definem uma parada para o Padrão de Movimentação também é reforçada por Cunha (2008), que faz a diferenciação dos pontos em movimento e dos pontos parados, utilizando distâncias relativas entre o ponto de rastreamento e uma determinada rodovia de estudo. Nesta diferenciação, é avaliado se a distância entre um ponto de rastreamento e a rodovia é superior a um valor previamente estipulado e, caso isso aconteça, este ponto será considerado com uma saída de percurso (uma parada, saída de percurso etc.).

A premissa considerada nesta pesquisa é que todos os registros de rastreamento que não são classificados como em movimento, estão parados. A utilização dessa premissa se deve ao fato de ser mais confiável afirmar quais são os pontos de rastreamento que definem um movimento, do que quais estão parados quando se está avaliando os registros de rastreamento de um objeto móvel. Isto porque, a movimentação do objeto é um evento mais característico, pois, na maioria dos casos, os deslocamentos entre os pontos de rastreamento são maiores que o erro de posicionamento considerado.

A afirmação do estado “em movimento” de um determinado veículo considera a comparação entre o valor de deslocamento calculado entre pontos consecutivos e um valor pré-estabelecido. Além de uma quantidade mínima de pontos sucessivos com esse mesmo comportamento, que, dependendo da taxa de coleta das posições do veículo, determinarão o período mínimo de observação para tal afirmação.

Nesta pesquisa, o valor pré-estabelecido de deslocamento entre pontos considerado como limite, foi de dezesseis metros. Este valor foi considerado como a pior acurácia

na determinação de uma posição do veículo e, foi obtido considerando-se a soma aproximada de dois valores:

1. O primeiro valor refere-se ao erro obtido a partir da resultante utilizando-se as coordenadas X e Y do exemplo apresentado por Monico (2008, p. 299), no qual a valor da acurácia posicional apresentou o valor de 15.45 m, considerando-se as coordenadas X e Y.
2. O segundo valor é proveniente do erro causado pelo multicaminho do sinal GPS (0.6 m), destacado por Rodrigues, Cugnasca e Queiroz Filho (2009, p. 28).

É importante ressaltar que outras literaturas apresentam um valor menor e, conseqüentemente, uma melhor da acurácia posicional para esta técnica de posicionamento. Mas o objetivo, neste trabalho, foi estipular o valor na pior das hipóteses. Portanto, nesta pesquisa o valor considerado foi de 16 m, ou seja, se entre pontos de rastreamento a distância calculada for inferior a 16 metros, considera-se que entre estes dois instantes o veículo não se movimentou. Como mencionado previamente, o período de observação mínimo para avaliação de uma mudança de estado do veículo será estipulado em função da taxa de coleta das posições do veículo. No caso desta pesquisa a taxa de coleta foi de três minutos.

Para esta pesquisa, definiu-se que a afirmação do estado do veículo deve considerar o maior número de ocorrências de deslocamentos entre pontos de rastreamento numa amostra de, no máximo, oito registros de rastreamento consecutivos, ou seja, se nesta amostra prevalecer a ocorrência de deslocamentos maiores que 16 metros assume-se que o veículo está em movimento. A figura 16 ilustra as possibilidades das quantidades de registros (pontos de rastreamento) necessários para a definição do estado do veículo.

Figura 16. Quantidade de registros necessários para a afirmação do estado do veículo.

Observe que os três primeiros casos podem sofrer variações em relação a figura 16. Para efeitos de ilustração, essas variações foram desconsideradas. Entretanto, são consideradas na análise executada pelo algoritmo responsável por identificar o estado atual do veículo.

Ressalta-se ainda que o algoritmo considera a contagem de deslocamentos que prevalecem e foram obtidas na sequência anterior de pontos. Por exemplo, no Caso (1) da figura 16 se o nono registro apresentar um deslocamento inferior a 16 metros a contagem de deslocamentos inferiores será iniciada novamente em 1 (um), caso ele apresente um deslocamento superior à contagem continua. Este comportamento visa eliminar possíveis aberrações nos dados causadas por fontes de erros adversas. O Algoritmo 1, representado na figura 17, ilustra o processo de definição dos registros que são classificados como “em movimento”.

Algoritmo 1: Define o estado do veículo (parado ou em movimento)

```

1.  OPEN cursor1;
2.  OPEN cursor2;
3.  n = 1;
4.  n2 = 1;
5.  dgrupo = 0;
6.  dgrupo2 = 0;
7.  FETCH cursor2 INTO vv_gid, vv_latitude, vv_longitude;
8.  FETCH cursor1 INTO a_gid, a_latitude, a_longitude, a_datahora;
9.  LOOP
10.     FETCH cursor2 INTO v_gid, v_latitude, v_longitude, v_datahora;
11.     IF NOT FOUND THEN
12.         EXIT;
13.     END IF;
14.     IF a_latitude <> v_latitude AND a_longitude <> v_longitude THEN -- Longitudes e Latitudes Iguais entre os Pontos
15.         distmetros = ROUND(DISTANCE_SPHEROID(SETSRID(('POINT(' || a_longitude || ' ' || a_latitude ||
        '))::GEOMETRY, 4326), SETSRID(('POINT(' || v_longitude || ' ' || v_latitude || '))::GEOMETRY, 4326),
        'SPHEROID["WGS 84",6378137,298.257223563]')::NUMERIC, 3);
16.         IF distmetros > 16 THEN -- O veículo está Em Movimento
17.             status = 'MOVIMENTANDO';
18.             n = (n + 1);
19.             IF n = 1 THEN
20.                 datahora1 = a_datahora;
21.             ELSIF n >= 4 THEN
22.                 n2 = 0;
23.                 IF n = 4 THEN
24.                     dgrupo = (dgrupo + 1);
25.                     UPDATE avl_calc SET point_movi = dgrupo WHERE date_time >= datahora1 AND
                    date_time <= v_datahora AND point_stops ISNULL;
26.                     datahora1 = v_datahora;
27.                 ELSE
28.                     UPDATE avl_calc SET point_movi = dgrupo WHERE date_time >= datahora1 AND
                    date_time <= v_datahora AND point_stops ISNULL;
29.                     datahora1 = v_datahora;
30.                 END IF;
31.             END IF;
32.         ELSE -- O veículo está Parado
33.             status = 'PARADO';
34.             n2 = (n2 + 1);
35.             IF n2 >= 4 THEN
36.                 n = 0;
37.             END IF;
38.         END IF;
39.         UPDATE avl_calc SET dist_meters = distmetros, the_geom = SETSRID(('POINT(' || v_longitude || ' ' ||
        v_latitude || '))::GEOMETRY, 4326) WHERE date_time = v_datahora;
40.     ELSE -- Longitudes e Latitudes Iguais entre os Pontos = O veículo está Parado
41.         status = 'PARADO';
42.         n2 = (n2 + 1);
43.         IF n2 >= 4 THEN
44.             n = 0;
45.         END IF;
46.         UPDATE avl_calc SET dist_meters = 0, the_geom = SETSRID(('POINT(' || v_longitude || ' ' || v_latitude ||
        '))::GEOMETRY, 4326) WHERE date_time = v_datahora;
47.     END IF;
48.     a_latitude = v_latitude;
49.     a_longitude = v_longitude;
50.     a_datahora = v_datahora;
51. END LOOP;
52. CLOSE cursor2;
53. CLOSE cursor1;

```

Figura 17. Algoritmo 1: Definição do Estado do Veículo.

Uma sequência de registros classificados como “em movimento” recebe um identificador de grupo. É a partir desse identificador de grupo que serão geradas as geometrias (lineares) que irão representar os trajetos percorridos pelo veículo entre

uma parada e outra - origem e destino, respectivamente. Essas geometrias serão geradas ordenando-se cronologicamente as geometrias pontuais dos registros que foram classificados como “em movimento”, considerando-se também a geometria que representa os dois centróides provenientes das geometrias do tipo área que representam a origem e o destino. Na figura 18-(a) a linha realçada em laranja ilustra o trajeto gerado para a movimentação do veículo entre a origem (O) e o destino (D) e, na figura 18-(b) há uma inversão entre a origem e o destino e a linha em azul representa o trajeto de retorno entre O-D.

Figure 18. Em (a) trajeto de ida entre (O-D) e em (b) trajeto de volta entre (O-D)

4.2.2 Detecção de Paradas

O processo que define as paradas realizadas pelos veículos rastreados é dependente e decorrente do processo que define quais pontos de rastreamento são classificados como “em movimento”. Portanto, o algoritmo que define as paradas

efetuadas pelos veículos é bem mais simples. Na verdade, esse algoritmo apenas percorre os registros dos pontos de rastreamento cronologicamente e atribui àqueles que não foram classificados como “em movimento”, identificadores de grupos, ou seja, todos os pontos de rastreamento que estão entre dois grupos de pontos classificados como “em movimento” recebem o mesmo valor do identificador de grupo. A figura 19 ilustra o Algoritmo 2 que é responsável pela definição dos pontos de rastreamento que são considerados como “parados”, ou seja, as paradas efetuadas pelo veículo.

Algoritmo 2: *Define os Grupos de Paradas Efetuadas pelo veículo*

```

1.  n = MAX(point_movi) FROM avl_viewcalc;
2.  FOR i IN 1..n LOOP
3.      datahora1 = MIN(date_time) FROM avl_viewcalc WHERE point_movi = i;
4.      UPDATE avl_calc SET point_stops = i WHERE date_time < datahora1 AND point_movi ISNULL AND point_stops
      ISNULL;
5.  END LOOP;

```

Figura 19. Algoritmo 2: Definição dos Grupos de Paradas.

Uma vez definidos os grupos de pontos de rastreamento que sinalizam uma parada efetuada pelo veículo, um terceiro algoritmo define a geometria que representa a área da parada realizada, bem como o intervalo de tempo que o veículo permaneceu nesse local. A definição dessa geometria é de suma importância, pois ela será utilizada tanto no processo de geração de Padrões de Localização quanto no processo de geração do Padrão de Movimentação.

Como o erro máximo considerado no posicionamento GPS foi de 16 metros, a geometria que representa uma parada é definida por um *buffer* de 16 metros a partir do ponto médio, calculado com base nos pontos de rastreamento que possuem o mesmo identificador de grupo de parada. Este *buffer* passa a ser, portanto, uma área de entorno ao ponto médio, representado por um círculo com origem neste ponto e raio igual a 16 metros. O algoritmo responsável pela geração destas áreas está representado na figura 20. O resultado gráfico do Algoritmo 3 pode ser verificado na figura 21, na qual é representada a geometria gerada pelo *buffer* a partir do ponto central de uma determinada parada.

Algoritmo 3: *Geração da Geometria que representará a Parada e cálculo do Tempo de Permanência na Parada*

```

1.  n = MAX(point_stops) FROM avl_viewcalc;
2.  FOR i IN 1..n LOOP
3.    datahora1 = MIN(date_time) FROM avl_viewcalc WHERE point_stops = i;
4.    datahora2 = MAX(date_time) FROM avl_viewcalc WHERE point_stops = i;
5.    deltatarada = (datahora2 - datahora1);
6.    i2 = 0;
7.    n2 = 0;
8.    FOR v_the_geom IN (SELECT the_geom FROM avl_viewcalc WHERE point_stops = i) LOOP
9.      i2 = (i2 + 1);
10.     IF i2 >= 2 THEN
11.       distmetros = ROUND(DISTANCE_SPHEROID(v_the_geom, a_the_geom, 'SPHEROID["WGS
12.       84",6378137,298.257223563]')::NUMERIC, 3);
13.       IF distmetros < 16 THEN
14.         n2 = (n2 + 1);
15.         m_the_geom = (m_the_geom + v_the_geom)/n2;
16.       END IF;
17.     ELSE
18.       n2 = (n2 + 1);
19.       m_the_geom = v_the_geom;
20.     END IF;
21.   END LOOP;
22.   bufferparadas = SETSRID(BUFFER(m_the_geom, 0.000144)::GEOMETRY, 4326);
23.   UPDATE avl_calc SET deltatarada = deltatarada, the_geom_stop = bufferparadas WHERE date_time_path =
   datahora2;
24. END LOOP

```

Figura 20. Algoritmo 3: Geração da geometria da parada e cálculo do tempo de duração da parada.

Figura 21. Geometrias geradas a partir do *buffer* de uma parada.

Uma particularidade encontrada nos dados de rastreamento quando o veículo está parado, e que ocorre em virtude do veículo estar estacionado em estacionamentos cobertos ou subterrâneos, é que durante o seu período de permanência nesses tipos

de estacionamentos, suas coordenadas geográficas (Latitude e Longitude) não são obtidas pelo receptor GPS. Quando isso ocorre, a localização do veículo, transmitida para a central de monitoramento se refere ao último posicionamento obtido pelo receptor GPS, antes da perda de sinal. Entretanto, a informação referente ao instante de tempo tem contínua transmissão, ou seja, o banco de dados armazena registros de localização repetidos em instantes de tempo diferentes.

Segundo Volpato *et al.* (2008), edifícios, morros, cobertura densa de uma floresta e interferências eletrônicas podem bloquear a recepção do sinal GPS, o que causa erros de posicionamento ou possivelmente nenhuma leitura do atual posicionamento do receptor. Assim, os receptores GPS geralmente não funcionam dentro de casas, túneis, estacionamentos cobertos, debaixo d'água, etc..

Caso ocorra perda na comunicação entre o dispositivo embarcado e a central de monitoramento, os dispositivos geralmente são capazes de armazenar temporariamente as informações de rastreamento e, assim que for reestabelecida a comunicação, essas informações serão transmitidas à central de monitoramento (RODRIGUES; CUGNASCA; QUEIROZ FILHO, 2009).

Em razão destas circunstâncias, pode-se dizer que a condição mais favorável para o posicionamento do veículo seria quando ele estivesse em movimento ou parado num estacionamento aberto, uma vez que, na maioria das vezes, esta condição pode garantir um número maior de satélites disponíveis para o posicionamento do veículo, salvo os momentos em que o veículo estiver transitando em túneis ou entre edifícios. Sendo assim, são considerados, para representação das paradas efetuadas pelo veículo, os pontos de rastreamento que estão entre dois trajetos subsequentes, por exemplo: *trajeto 1; parada 1; trajeto 2; parada 2; ...; trajeto n; parada n; trajeto n+1*.

4.2.3 Padrões de Movimentação

Cada trajeto realizado entre uma parada e outra será considerados como um registro de Padrão de Movimentação, isto é, de onde pra onde o veículo se

movimentou num determinado instante de tempo. Este padrão será representado por uma geometria linear gerada a partir dos registros classificados como “em movimento”, ordenados cronologicamente, incluindo-se um ponto inicial e final, referentes ao último ponto parado antes do início da movimentação e primeiro ponto parado após o término da movimentação, respectivamente.

As informações referentes ao Padrão de Movimentação são armazenadas em uma tabela denominada AVL_PATTERNS_MOV. Os atributos que compõem esta tabela são responsáveis por armazenar informações referentes a identificação do veículo, a geometria linear que representa o trajeto realizado, a geometria pontual que representa o início deste trajeto, a geometria pontual do final deste trajeto, os instantes de tempo inicial e final deste trajeto, o ângulo de orientação de cada segmento que compõem esse trajeto, um identificador para o trajeto e um identificador para cada segmento que compõem o trajeto, respectivamente. Esse conjunto de informações caracteriza um padrão de movimentação.

A partir destes atributos, podem ser realizados questionamentos referentes ao próximo destino do veículo considerando sua atual informação espaço-temporal. Portanto, com base nessa geometria linear, são extraídas as informações necessárias para realizar o relacionamento espacial entre as geometrias das regiões de paradas, armazenadas na tabela do Padrão de Localização. Estes atributos permitem filtragens temporais no questionamento como, por exemplo, considerando-se que o veículo esteja numa determinada parada, num determinado dia da semana e período do dia, quais seriam os seus prováveis Destinos?

4.2.3.1 Geração dos Padrões de Movimentação

Como mencionado anteriormente, todo processo de geração de um Padrão de Movimentação se inicia a partir dos pontos de rastreamento classificados como “em movimento” e da geometria proveniente desses pontos. Portanto, pode-se considerar como início desse processo as rotinas do Algoritmo 1. O algoritmo 4, representado na figura 22, é responsável por preencher as informações na tabela AVL_PATTERNS_MOV.

Algoritmo 4: Extrai e Insere as Informações relacionadas ao Padrão de Movimentação

-- Preparação dos dados que serão utilizados na geração do Padrão de Movimentação

```

1. FOR i IN 1..(SELECT MAX(point_movi) FROM avl_viewcalc) LOOP
2.   datahora1 = MAX(date_time) FROM avl_viewcalc WHERE point_stops = i;
3.   datahora2 = MIN(date_time) FROM avl_viewcalc WHERE point_stops = (i + 1);
4.   pto_ini = CENTROID(the_geom_stop) FROM avl_viewcalc WHERE point_stops = i AND the_geom_stop NOTNULL;
5.   pto_fim = CENTROID(the_geom_stop) FROM avl_viewcalc WHERE point_stops = (i + 1) AND the_geom_stop NOTNULL;
6.   UPDATE avl_calc SET the_geom_path = pto_ini, date_time_path = datahora1, trajeto = i WHERE date_time = datahora1;
7.   UPDATE avl_calc SET the_geom_path = pto_fim, date_time_path = datahora2, path = i WHERE date_time = datahora2;
8.   FOR datahora3, pto_central IN (SELECT date_time, the_geom FROM avl_viewcalc WHERE point_movi = i) LOOP
9.     UPDATE avl_calc SET the_geom_path = pto_central, date_time_path = datahora3, path = i WHERE date_time = datahora3;
10.  END LOOP;
11. END LOOP;

```

-- Construir a geometria do trajeto executado pelo veículo

```

12. FOR i IN 1..(SELECT MAX(path) FROM avl_viewcalc) LOOP
13.   geometria = (MAKELINE(SETSRID(foo.the_geom_path, 4326))) FROM (SELECT the_geom_path, path FROM avl_viewcalc WHERE path = i ORDER BY date_time_path, path) AS foo GROUP BY foo.path;
14.   a_datahora = MIN(date_time) FROM avl_viewcalc WHERE path = i;
15.   v_datahora = MAX(date_time) FROM avl_viewcalc WHERE path = i;
16.   UPDATE avl_calc SET the_geom_path_line = geometria WHERE date_time = v_datahora;
17. END LOOP;

```

-- Inserir dados na tabela de Padrões de Movimentação

```

18. FOR i IN 1..(SELECT MAX(path) FROM avl_viewcalc WHERE path NOTNULL) LOOP
19.   v_gid = gid FROM avl_viewcalc WHERE path = i AND the_geom_path_line NOTNULL;
20.   IF v_gid NOTNULL THEN
21.     geometria = the_geom_path_line FROM avl_viewcalc WHERE path = i AND the_geom_path_line NOTNULL;
22.     pto_ini = STARTPOINT(the_geom_path_line) FROM avl_viewcalc WHERE path = i AND the_geom_path_line NOTNULL;
23.     pto_fim = ENDPOINT(the_geom_path_line) FROM avl_viewcalc WHERE path = i AND the_geom_path_line NOTNULL;
24.     a_datahora = MIN(date_time_path) FROM avl_viewcalc WHERE path = i AND the_geom_path NOTNULL;
25.     v_datahora = MAX(date_time_path) FROM avl_viewcalc WHERE path = i AND the_geom_path NOTNULL;
26.     INSERT INTO avl_patterns_mov ( vehicle_id, the_geom_path, the_geom_beg, the_geom_end, date_time_beg, date_time_end, day_of_week) VALUES ( codveic_min, geometria, pto_ini, pto_fim, a_datahora);
27.   END IF;
28. END LOOP;

```

-- Calcula e preenche a orientação dos segmentos que compõem a trajetória do veículo

```

29. FOR i IN SELECT DISTINCT vehicle_id FROM avl_patterns_mov LOOP
30.   FOR geom_first, v_gid IN SELECT the_geom_beg, gid FROM avl_patterns_mov WHERE vehicle_id = i LOOP
31.     FOR geom_first_centro IN SELECT CENTROID(t1.the_geom) FROM avl_patterns_loc t1 WHERE WITHIN(geom_first, t1.the_geom) AND t1.vehicle_id = i LOOP
32.       num_ptos = NUMPOINTS(the_geom_path) FROM avl_patterns_mov WHERE gid = v_gid;
33.       n_pto = 2;
34.       n_pto2 = 1;
35.       WHILE n_pto < num_ptos LOOP
36.         geom_next = POINTN(the_geom_path, n_pto) FROM avl_patterns_mov WHERE gid = v_gid;
37.         IF geom_next <> geom_first_centro THEN
38.           azim = ROUND(DEGREES(AZIMUTH(geom_first_centro, geom_next)::NUMERIC)::NUMERIC, 6);
39.           IF n_pto = 2 THEN
40.             UPDATE avl_patterns_mov SET orient_seg_path = azim, the_geom_path_point = geom_next, path_pt_id = v_gid, path_pt = 1 WHERE gid = v_gid;
41.           ELSE
42.             n_pto2 = (n_pto2 + 1);
43.             INSERT INTO avl_patterns_mov (orient_seg_path, the_geom_path_point, path_pt_id, path_pt) VALUES (azim, geom_next, v_gid, n_pto2);
44.           END IF;
45.           geom_first_centro = geom_next;
46.         END IF;
47.         n_pto = (n_pto + 1);
48.       END LOOP;
49.     END LOOP;
50.   END LOOP;
51. END LOOP;

```

Figura 22. Algoritmo 4: Preenche as informações relacionadas ao Padrão de Movimentação.

Após o preenchimento das informações relacionadas ao Padrão de Movimentação, outro algoritmo é responsável por armazenar as informações referentes ao último movimento realizado pelo veículo, e que foi considerado na geração dos padrões. O objetivo desse processo é assegurar que, na próxima vez que a função de geração de padrão for executada, ela continue a partir do último movimento detectado na execução anterior. A figura 23 ilustra este algoritmo.

Algoritmo 5: Armazenar o último movimento realizado pelo veículo

--Popular a tabela de padrões dos últimos movimentos

```

1. i = MAX(point_stops) FROM avl_viewcalc;
2. n = DISTINCT vehicle_id FROM avl_viewpatterns_last_mov WHERE vehicle_id = codveic_min;
3. IF n ISNULL AND i NOTNULL THEN
4.   INSERT INTO avl_patterns_last_mov (the_geom, date_time, vehicle_id, group_stops) SELECT DISTINCT
     the_geom_path, date_time, vehicle_id, group_stops FROM avl_viewcalc WHERE group_stops = i AND
     the_geom_path NOTNULL;
5. ELSIF n NOTNULL AND i NOTNULL THEN
6.   a_datahora = date_time FROM avl_viewcalc WHERE group_stops = i AND the_geom_path NOTNULL;
7.   geometria = the_geom_path FROM avl_viewcalc WHERE group_stops = i AND the_geom_path
     NOTNULL;
8.   UPDATE avl_patterns_last_mov SET group_stops = i, the_geom = geometria, date_time = a_datahora
     WHERE vehicle_id = n;
9. END IF;

```

Figura 23. Algoritmo 5: Armazenar o último movimento realizado pelo veículo;

4.2.4 Padrões de Localização

Consideram-se como um Padrão de Localização, nesta pesquisa, todas as informações possíveis que estejam vinculadas a uma determinada região de parada como, por exemplo, o número de vezes que ela ocorre, quais dias da semana e períodos do dia são mais frequentes, além de informações extraídas a partir de uma base cartográfica que forneçam um significado associado como, por exemplo, o nome do bairro e município desta parada.

Para tanto, uma tabela denominada AVL_PATTERNS_LOC armazena as informações referentes a cada região de parada efetuada pelo veículo. As informações armazenadas nesta tabela visam evitar uma grande quantidade de registros na tabela, bem como que as operações topológicas que definem uma região de parada sejam executadas na função responsável pelo cálculo de previsão. O principal objetivo desta abordagem é garantir que a previsão de movimentação do

veículo seja executada num intervalo de tempo menor, uma vez que esta função é requisitada sob demanda do usuário e a função que gera os padrões de localização é executada periodicamente.

Os atributos armazenados nessa tabela fornecem informações espaço-temporal que são utilizadas em conjunto com as informações armazenadas na tabela AVL_PATTERNS_MOV, para prever a movimentação do veículo. Em razão dos vários atributos da tabela AVL_PATTERNS_LOC armazenada, não cabe aqui uma descrição detalhada dos mesmos. Sendo assim, maiores detalhes podem ser observados no APÊNDICE A, que descreve o diagrama do banco de dados utilizado nesta pesquisa. Por hora, é suficiente saber que esta tabela é capaz de armazenar a informação espacial de onde uma ou mais paradas ocorrem, por meio da região de paradas e, todas as informações temporais provenientes de cada parada que compõem esta região.

4.2.4.1 Geração dos Padrões de Localização

A geração de um Padrão de Localização inicia-se a partir dos grupos de pontos de rastreamento que foram identificadas sequencialmente como “parados”. Cada um desses grupos é representado por um círculo de raio igual a 16 metros, que será utilizado na análise espacial que define uma região de parada, a qual reúne todas as informações temporais pertinente a cada um destes círculos. A análise espacial considera a proximidade entre as geometrias das paradas (os círculos).

Como o valor da face de quadra, que é uma variável no cálculo da área de um imóvel, é uma característica física e peculiar de cada imóvel, devido à variação que ocorre bairro a bairro, loteamento a loteamento, onde as características do terreno, especialmente a área, são padronizadas de formas distintas, em função do zoneamento, características sócio-econômicas. É importante considerar os valores de face de quadra, obtidos numa região de estudo, que define uma amostra que melhor se adequa a massa de imóveis avaliados (CTAGEO, 2010).

Nesse contexto, o valor de proximidade considerado nesta pesquisa foi de 50 metros. Este valor foi obtido a partir do valor médio calculado com base numa amostra de tamanhos de faces de estacionamentos contidos numa fração da área de estudo considerada nesta pesquisa. O objetivo dessa análise de proximidade é identificar as paradas que ocorrem no mesmo local, mas que, devido a dimensão do estacionamento e/ou ofertas de vagas, estejam distantes em relação a parada que ocorreu nesse mesmo estacionamento no passado. O algoritmo responsável pela identificação desse agrupamento de paradas está representado na figura 24.

Algoritmo 6: Agrupamento e Identificação de uma Única Parada

```

1.  i = 1;
2.  i2 = 0;
3.  LOOP
4.    n = COUNT(gid)::INT FROM avl_viewcalc WHERE the_geom_stop NOTNULL AND group_stops ISNULL;
5.    a_gid = gid FROM avl_viewcalc WHERE the_geom_stop NOTNULL AND group_stops ISNULL ORDER BY
      date_time ASC LIMIT 1;
6.    v_datahora = MAX(date_time) FROM avl_viewcalc c WHERE c.gid IN (SELECT DISTINCT c2.gid FROM
      avl_viewcalc c1, avl_viewcalc c2 WHERE (OVERLAPS(c1.the_geom_stop, c2.the_geom_stop) AND (c2.gid <>
      a_gid) AND (c1.gid = a_gid)) OR ((DISTANCE(c1.the_geom_stop, c2.the_geom_stop) < 0.000449) AND (c2.gid <>
      a_gid) AND (c1.gid = a_gid))) OR gid = a_gid;
7.    m_the_geom = CENTROID(GEOMUNION(the_geom_stop)) FROM avl_viewcalc WHERE gid IN (SELECT
      DISTINCT c2.gid FROM avl_viewcalc c1, avl_viewcalc c2 WHERE (OVERLAPS(c1.the_geom_stop,
      c2.the_geom_stop) AND (c2.gid <> a_gid) AND (c1.gid = a_gid)) OR ((DISTANCE(c1.the_geom_stop,
      c2.the_geom_stop) < 0.000449) AND (c2.gid <> a_gid) AND (c1.gid = a_gid))) OR gid = a_gid;
8.    distmetros = MAX(DISTANCE(c.the_geom_stop, m_the_geom))::NUMERIC FROM avl_viewcalc c WHERE c.gid IN
      (SELECT DISTINCT c2.gid FROM avl_viewcalc c1, avl_viewcalc c2 WHERE (OVERLAPS(c1.the_geom_stop,
      c2.the_geom_stop) AND (c2.gid <> a_gid) AND (c1.gid = a_gid)) OR ((DISTANCE(c1.the_geom_stop,
      c2.the_geom_stop) < 0.000449) AND (c2.gid <> a_gid) AND (c1.gid = a_gid))) OR gid = a_gid;
9.    UPDATE avl_calc SET group_stops = i WHERE gid IN (SELECT DISTINCT c2.gid FROM avl_viewcalc c1,
      avl_viewcalc c2 WHERE (OVERLAPS(c1.the_geom_stop, c2.the_geom_stop) AND (c2.gid <> a_gid) AND (c1.gid
      = a_gid)) OR ((DISTANCE(c1.the_geom_stop, c2.the_geom_stop) < 0.000449) AND (c2.gid <> a_gid) AND (c1.gid
      = a_gid))) OR gid = a_gid;
10.   i3 = COUNT(the_geom_stop)::INT FROM avl_viewcalc WHERE gid IN (SELECT DISTINCT c2.gid FROM
      avl_viewcalc c1, avl_viewcalc c2 WHERE (OVERLAPS(c1.the_geom_stop, c2.the_geom_stop) AND (c2.gid <>
      a_gid) AND (c1.gid = a_gid)) OR ((DISTANCE(c1.the_geom_stop, c2.the_geom_stop) < 0.000449) AND (c2.gid <>
      a_gid) AND (c1.gid = a_gid))) OR gid = a_gid;
11.   IF i3 > 1 THEN
12.     LOOP
13.       bufferparadas = SETSRID(BUFFER(m_the_geom, (distmetros + (2*0.000144))):GEOMETRY, 4326);
14.       i2 = COUNT(the_geom_stop)::INT FROM avl_viewcalc WHERE gid IN (SELECT DISTINCT c1.gid
      FROM avl_viewcalc c1 WHERE (OVERLAPS(c1.the_geom_stop, bufferparadas) AND
      c1.group_stops ISNULL) OR ((DISTANCE(c1.the_geom_stop, bufferparadas) < 0.000449) AND
      c1.group_stops ISNULL);
15.       IF i2 <> 0 THEN
16.         m_the_geom = CENTROID(GEOMUNION(the_geom_stop)) FROM avl_viewcalc WHERE gid
      IN (SELECT DISTINCT c1.gid FROM avl_viewcalc c1 WHERE
      (OVERLAPS(c1.the_geom_stop, bufferparadas) OR ((DISTANCE(c1.the_geom_stop,
      bufferparadas) < 0.000449)));
17.         distmetros = MAX(DISTANCE(c.the_geom_stop, m_the_geom))::NUMERIC FROM
      avl_viewcalc c WHERE gid IN (SELECT DISTINCT c1.gid FROM avl_viewcalc c1 WHERE
      (OVERLAPS(c1.the_geom_stop, bufferparadas) OR ((DISTANCE(c1.the_geom_stop,
      bufferparadas) < 0.000449)));
18.         UPDATE avl_calc SET group_stops = i WHERE gid IN (SELECT DISTINCT c1.gid FROM
      avl_viewcalc c1 WHERE (OVERLAPS(c1.the_geom_stop, bufferparadas) OR
      ((DISTANCE(c1.the_geom_stop, bufferparadas) < 0.000449)));
19.         v_datahora = MAX(date_time) FROM avl_viewcalc c WHERE c.gid IN (SELECT DISTINCT
      c1.gid FROM avl_viewcalc c1 WHERE (OVERLAPS(c1.the_geom_stop, bufferparadas) OR
      ((DISTANCE(c1.the_geom_stop, bufferparadas) < 0.000449)));
20.       ELSE
21.         UPDATE avl_calc SET the_geom_stop = bufferparadas WHERE date_time = v_datahora;
22.         i = i + 1;
23.         EXIT;
24.       END IF;
25.     END LOOP;
26.   ELSE
27.     UPDATE avl_calc SET the_geom_stop_cluster = the_geom_stop WHERE gid = a_gid;
28.     i = i + 1;
29.   END IF;--
30.   EXIT WHEN n = 0;
31. END LOOP;

```

Figura 24. Algoritmo 6: Agrupamento e identificação de uma região de parada.

Todas as paradas, que satisfazem a análise de proximidade, são consideradas na geração de um novo centróide, calculado a partir da localização média dos pontos médios das paradas que satisfazem a análise de proximidade. Este novo centróide

representa o centro da geometria que exprimi uma região de parada, essa geometria será representada por um *buffer* circular com origem neste centróide e raio igual à maior distância calculada entre este centróide e o centro da geometria representada pela parada mais distante, acrescido de 16 m, que é a distância referente ao raio do círculo desta parada. Desta forma, a nova geometria irá conter todas as paradas que foram encontradas na análise de proximidade, como ilustra a figura 25, na qual é representada esta nova geometria, denominada de região de paradas.

Figura 25. Região de paradas.

Portanto, a geometria que representa uma região de parada, possui informações pertinentes a cada uma das paradas, realizadas em momentos diferentes, que foram utilizadas para compor esta região. Cada uma destas paradas fornece informações como dia da semana, período do dia e, a partir destas informações será possível fazer inferências como o quão frequente é esta parada e qual o seu significado (ex: residência, trabalho, local de estudo etc.). Além, de diminuir a quantidade de registros na tabela de Padrão de Localização.

4.3 ETAPA III – PREVISÃO DE DESTINOS

Segundo Oliveira (2003), um objeto móvel possui, associado a si, uma trajetória correspondente a um determinado intervalo de tempo, e as consultas relacionadas a um objeto móvel são respondidas com base nas relações espaciais entre as

geometrias que definem a trajetória e as outras geometrias que compõem o mapa, por onde os objetos se movimentam.

Estas relações espaciais são doravante denominadas de componente espacial, temporal e direcional e, fazem parte do processo preditivo da localização do veículo investigado. Todas as três componentes serão consideradas no cálculo da probabilidade de um provável destino para o veículo considerado.

O detalhamento a que se referem cada componente será realizado nas seções seguintes. As seções seguintes descrevem o processo de previsão dos prováveis destinos e o cálculo de probabilidade considerado neste processo preditivo.

4.3.1 Componente Espacial

Como mencionado anteriormente, a componente espacial irá representar o relacionamento espacial entre uma região de parada (considerada como a origem), e as demais regiões, consideradas como possíveis destinos a partir desta região. Este relacionamento é obtido a partir das informações armazenadas nos Padrões de Movimentação e Localização.

A utilização das geometrias armazenadas em ambos os padrões permite estabelecer relacionamentos espaciais entre elas, o que possibilita identificar quais foram os destinos a partir de uma determinada região de parada (origem). Essa informação é obtida relacionando-se topologicamente uma região de parada, considerada como a origem, com as demais regiões armazenadas no Padrão de Localização, utilizando-se das trajetórias do veículo, que são armazenadas no Padrão de Movimentação, ou seja, se duas regiões estiverem conectadas por meio de uma trajetória tem-se o relacionamento espacial Origem-Destino.

A quantidade de trajetórias que se iniciam a partir da região, considerada como origem, será a frequência que o veículo foi a este possível destino a partir desta origem, ou seja, o número de saídas para este destino a partir desta origem. Portanto, a componente espacial indica qual é a origem, quais são os possíveis

destinos e quantas vezes que cada região foi destino daquela considerada como origem. Essa informação de frequência de ocorrências de cada destino é suficiente para indicar o destino mais provável a partir de uma origem, considerando-se apenas o relacionamento espacial, ou seja, a componente espacial.

4.3.2 Componente Temporal

A componente temporal nada mais é que uma particularização da componente espacial, ou seja, ela informará quais são os possíveis destinos a partir de uma região de parada, considerada como origem, utilizando o mesmo relacionamento espacial descrito na componente espacial, porém considerando um período de tempo. Este período de tempo poderá ser expresso em dias da semana, períodos do dia (ex: manhã, tarde e noite) ou até mesmo da combinação de ambos (ex: segunda-feira de manhã).

A informação obtida com a componente temporal pode indicar quais são os possíveis destinos de um veículo num contexto de espaço e tempo. As seções seguintes (Dias da Semana e Períodos dos Dias) detalham a divisão de tempo considerada nesta pesquisa para obter a componente temporal.

4.3.2.1 Dias da Semana

Em virtude do rastreamento de veículos ser contínuo ao longo do tempo, todos os dias da semana foram considerados nessa pesquisa. Assim, a componente temporal pode informar em quais dias da semana uma região de parada ocorre com mais frequência. Essa informação é armazenada e pode ser obtida no Padrão de Localização de um veículo. Para tanto, foram criados atributos booleanos para cada dia da semana, e quando um padrão é gerado ou atualizado, essas informações também são geradas ou atualizadas.

Portanto, um Padrão de Localização deve possuir a informação espacial, representada pela geometria da região de parada e, quais dias da semana ele

ocorre. Com base nessas informações, a componente temporal poderá ser gerada, considerando um determinado período de tempo. Isso implica em mais uma condição no questionamento realizado pela componente espacial e essa condição retorna quais são os possíveis destinos a partir de uma origem num determinado período de tempo, por exemplo, dia da semana.

4.3.2.2 Períodos do Dia

Com intuito de refinar ainda mais a qualidade da informação dos possíveis destinos a partir de uma origem, os dias da semana foram seccionados em quatro períodos do dia. Estes quatro períodos foram denominados de madrugada, manhã, tarde e noite e foram divididos da seguinte maneira:

- Madrugada – referente aos dados rastreados entre 00h00min e 05h59min;
- Manhã – referente aos dados rastreados entre 06h00min e 11h59min;
- Tarde – referente aos dados rastreados entre 12h00min e 17h59min;
- Noite – referente aos dados rastreados entre 18h00min e 23h59min.

Da mesma maneira que as informações de ocorrência dos dias da semana são armazenadas no Padrão de Localização de um veículo, as informações dos períodos também são armazenadas ou atualizadas. Isso permite que os possíveis destinos sejam obtidos a partir de um refinamento temporal ainda maior, o que permite responder questionamentos como: *“Quais são os possíveis destinos de um veículo considerando sua origem, dia da semana e período do dia?”*.

4.3.3 Componente Direcional

A componente direcional foi desenvolvida com o intuito de refinar as respostas para os possíveis destinos quando o estado atual do veículo sob investigação for “em movimento”. Para tanto, foi desenvolvida uma rotina que obtém a orientação dos segmentos que compõem uma trajetória realizada pelo veículo e esta informação é

armazenada, juntamente com a informação da trajetória, no Padrão de Movimentação. A rotina responsável por extrair e armazenar essa informação está representada no último bloco do algoritmo 4, representado na figura 22.

Com base nesta informação direcional dos segmentos das trajetórias realizadas pelo veículo, pode-se refinar ainda mais a resposta dos possíveis destinos, considerando-se somente aqueles trajetos do padrão, nos quais os segmentos correspondentes estejam no mesmo intervalo de orientação dos segmentos que compõem a trajetória atual do veículo. O intervalo de orientação considerado foi de $\pm 22.5^\circ$ em relação ao valor de orientação do segmento correspondente armazenado no Padrão de Movimentação, como ilustra a figura 26. Nesta figura é representado o intervalo de orientação considerado, em relação a segmento armazenado no banco.

Figura 26. Intervalo de orientação dos segmentos de uma trajetória.

Considerando-se que uma trajetória pode ser definida como o caminho percorrido por um corpo ou partícula em movimento, e que este caminho é composto por segmentos consecutivos, no qual o ponto terminal de um segmento coincide com o ponto inicial do segmento sucessor, é possível identificar se a trajetória que o veículo está executando no presente é parte de alguma trajetória armazenada no Padrão de Movimentação deste veículo. O que significa identificar que o veículo está percorrendo um mesmo caminho realizado anteriormente.

Para tanto, é necessário verificar se a orientação de cada segmento que define a atual trajetória está no mesmo intervalo de orientação do segmento correspondente,

armazenado no padrão, além de verificar se pontos inicial e final que definem o segmento atual também estão dentro da tolerância do erro de posicionamento considerado (16 m) em relação aos pontos inicial e final do segmento correspondente, armazenado no padrão. Caso algum segmento, armazenado no padrão, satisfaça estas condições será avaliado o próximo segmento do trajeto atual até que todos os segmentos do trajeto atual sejam avaliados ou algum segmento não satisfaça as condições.

A figura 27 ilustra a identificação de um mesmo caminho considerando duas trajetórias em instantes diferentes. O resultado desta avaliação será uma lista de identificadores dos trajetos armazenados no Padrão de Movimentação do veículo, que será utilizada para obter os possíveis destinos considerando-se sua atual localização e trajetória.

Figura 27. Identificação de um mesmo caminho.

4.3.4 Processo Preditivo

O processo de previsão da localização de um determinado veículo é composto das seguintes etapas:

1. Identificação do estado atual do veículo;
2. Geração das Matrizes de Frequências O-D:
 - a. Caso o veículo esteja parado:
 - i. Com base na componente espacial;
 - ii. Com base na componente espacial e a primeira variação da componente temporal (dia da semana);
 - iii. Com base na componente espacial e a segunda variação da componente temporal (dia da semana e período do dia);
 - b. Caso o veículo esteja em movimento:
 - i. Com base na componente direcional;
 - ii. Com base na componente direcional e a primeira variação da componente temporal (dia da semana);
 - iii. Com base na componente direcional e a segunda variação da componente temporal (dia da semana e período do dia);
3. Identificação da Origem no Padrão de Localização;
4. Cálculo das probabilidades;
 - a. Caso seja conhecida a Origem:
 - i. Cálculo da Probabilidade da Origem;
 - ii. Cálculo da Probabilidade Condicional dos prováveis Destinos;
 - b. Caso não seja conhecida a Origem:
 - i. Cálculo da probabilidade para todos os prováveis Destinos, considerando a frequência relativa;
 - c. Geração da mensagem de retorno para o usuário.

As subseções seguintes têm por objetivo detalhar cada uma das etapas do processo preditivo.

4.3.4.1 Estado Atual do Veículo

O processo de previsão da localização inicia-se a partir da identificação do estado atual do veículo. Independente do estado atual do veículo, o objetivo é determinar a última parada executada por ele, ou seja, a origem. O motivo de detectar a origem do veículo no padrão de localização vem de encontro à possibilidade de calcular a probabilidade condicional de um provável destino.

Segundo Bussab e Morettin (2006), se $P(A)$ for a probabilidade *a priori* e a informação adicional de que B ocorreu permitir o cálculo da probabilidade *a posteriori* $P(A|B)$, ou seja, a probabilidade condicional de A dado B. A probabilidade condicional $P(A|B)$ será maior que $P(A)$ e, portanto, a informação de que B ocorreu aumenta a chance de A ocorrer. Assim, se $P(A)$ for considerada como a probabilidade do destino e $P(B)$ a probabilidade da origem, verifica-se a importância da determinação da origem no padrão.

O algoritmo 7 representado na figura 28 ilustra esse processo de definição do estado atual do veículo. Este processo é semelhante ao processo anteriormente apresentado no Algoritmo 1. Porém, neste Algoritmo somente são considerados os registros de rastreamento mais recentes e a análise é realizada na ordem cronológica inversa.

Algoritmo 7: Definição do Estado do Veículo para a Previsão

```

1.  OPEN cursor1;
2.  OPEN cursor2;
3.  n = 1;
4.  n2 = 1;
5.  dgrupo = 0;
6.  dgrupo2 = 0;
7.  FETCH cursor2 INTO vv_gid, vv_the_geom;
8.  FETCH cursor1 INTO a_gid, a_the_geom, a_datahora;
9.  LOOP
10. FETCH cursor2 INTO v_gid, v_the_geom, v_datahora;
11. IF NOT FOUND OR dgrupo = 2 THEN
12.     EXIT;
13. END IF;
14. IF NOT EQUALS(a_the_geom, v_the_geom) THEN
15.     distmetros=ROUND(DISTANCE_SPHEROID(SETSRID((a_the_geom)::GEOMETRY, 4326),
        SETSRID((v_the_geom)::GEOMETRY, 4326), 'SPHEROID["WGS 84",6378137,298.257223563]')::NUMERIC,
        3);
16.     IF distmetros > 16 THEN --Movimentando
17.         status = 'MOVIMENTANDO';
18.         n = (n + 1);
19.         IF n = 1 THEN
20.             datahora1 = a_datahora;
21.         ELSIF n >= 4 THEN
22.             n2 = 0;
23.             IF n = 4 THEN
24.                 dgrupo = (dgrupo + 1);
25.                 UPDATE avl_auxprediction SET point_movi = dgrupo WHERE date_time <= datahora1 AND
                    date_time >= v_datahora AND group_mov ISNULL;
26.                 datahora1 = v_datahora;
27.             ELSE
28.                 UPDATE avl_auxprediction SET point_movi = dgrupo WHERE date_time >= v_datahora AND
                    group_mov ISNULL AND dist_metros > 16;
29.                 datahora1 = v_datahora;
30.             END IF;
31.         END IF;
32.     ELSE --Parado
33.         status = 'PARADO';
34.         n2 = (n2 + 1);
35.         IF n2 >= 4 THEN
36.             n = 0;
37.         END IF;
38.     END IF;
39.     UPDATE avl_auxprediction SET dist_meters = distmetros WHERE date_time = v_datahora;
40. ELSE --Parado geometrias iguais
41.     status = 'PARADO';
42.     n2 = (n2 + 1);
43.     IF n2 >= 4 THEN
44.         n = 0;
45.     END IF;
46.     UPDATE avl_auxprediction SET dist_meters = 0 WHERE date_time = v_datahora;
47. END IF;
48.     a_the_geom = v_the_geom;
49.     a_datahora = v_datahora;
50. END LOOP;
51. CLOSE cursor2;
52. CLOSE cursor1;

```

Figura 28. Algoritmo 7: Definição do Estado do Veículo para a Previsão.

Observe que na linha 11 do Algoritmo 7 a condição de saída do *Loop* é que o grupo referente a identificação de trajetos seja igual a 2 ou que todos os registros tenham sido avaliados. Isso porque o processo de geração de padrões finaliza em registros referentes a uma parada. Portanto, se os últimos registros avaliados na geração do

padrão estiveram em movimento, os registros mais recentes terminarão num conjunto de dados referentes ao início de um trajeto (considerando a ordem cronológica inversa).

Assim, o limite máximo para identificação de uma parada passa a ser o segundo trajeto encontrado na ordem cronológica inversa. Isso garante que entre estes dois trajetos ocorreu uma parada, que é a origem. É a partir da nuvem de pontos de rastreamento que definem a última parada que será possível determinar a origem do veículo, e verificar se ela existe no Padrão de Localização do mesmo. A figura 29 ilustra o algoritmo utilizado para identificar a origem no Padrão de Localização.

Algoritmo 8: Identificação da Origem no Padrão de Localização

```

1. a_datahora = MAX(date_time) FROM avl_auxprediction WHERE dist_meters NOTNULL;
2. estado_atual_veic = point_movi FROM avl_auxprediction WHERE date_time = a_datahora;
3. IF estado_atual_veic ISNULL THEN
4.     status = 'PARADO';
5.     datahora1 = MAX(date_time) FROM avl_auxprediction WHERE point_movi = 1;
6.     n = COUNT(gid)::INT FROM avl_auxprediction WHERE date_time > datahora1;
7.     IF n > 4 THEN
8.         geom_origem = CENTROID(GEOMUNION(the_geom)) FROM avl_auxprediction WHERE date_time > datahora1
            AND dist_meters < 16;
9.         v_datahora = MAX(date_time) FROM avl_auxprediction;
10.        dia_sem = dia FROM avl_auxprediction WHERE date_time = v_datahora;
11.        dia_sem_ult_mov = EXTRACT(dow FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            v_datahora;
12.        per_dia = DISTINCT period_of_day FROM avl_auxprediction WHERE date_time = v_datahora;
13.        per_dia_ult_mov = (EXTRACT(HOUR FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            v_datahora;
14.    ELSE
15.        geom_origem = the_geom FROM avl_auxprediction WHERE date_time = datahora1;
16.        dia_sem = day_of_week FROM avl_auxprediction WHERE date_time = datahora1;
17.        dia_sem_ult_mov = EXTRACT(dow FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            datahora1;
18.        per_dia = DISTINCT period_of_day FROM avl_auxprediction WHERE date_time = datahora1;
19.        per_dia_ult_mov = (EXTRACT(HOUR FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            datahora1;
20.    END IF;
21. ELSE
22.     status = 'MOVIMENTANDO';
23.     datahora1 = MIN(date_time) FROM avl_auxprediction WHERE point_movi = 1;
24.     datahora2 = MAX(date_time) FROM avl_auxprediction WHERE point_movi = 2;
25.     geom_origem = CENTROID(GEOMUNION(the_geom)) FROM avl_auxprediction WHERE date_time < datahora1 AND
            date_time > datahora2 AND dist_meters < 16;
26.     v_datahora = MAX(date_time) FROM avl_auxprediction WHERE date_time < datahora1 AND date_time > datahora2
            AND dist_meters < 16;
27.     dia_sem = day_of_week FROM avl_auxprediction WHERE date_time = v_datahora;
28.     dia_sem_ult_mov = EXTRACT(dow FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            v_datahora;
29.     per_dia = DISTINCT period_of_day FROM avl_auxprediction WHERE date_time = v_datahora;
30.     per_dia_ult_mov = (EXTRACT(HOUR FROM MAX(date_time))::INT FROM avl_auxprediction WHERE date_time =
            datahora1;
-- Agora precisa obter o primeiro ponto em movimento do atual trajeto do veículo
31.     geom_movim = the_geom FROM avl_auxprediction WHERE date_time = datahora1;
32. END IF;
-- Agora precisa obter a Origem no Padrão de Localização
33. origem_padrao_paradas = DISTINCT c2.the_geom FROM avl_patterns_loc c2 WHERE WITHIN(geom_origem,
    c2.the_geom) AND c2.vehicle_id = i;
34. IF origem_padrao_paradas ISNULL THEN
35.     origem_padrao_paradas = DISTINCT c2.the_geom FROM avl_patterns_loc c2 WHERE
        INTERSECTS(SETSRID(BUFFER(geom_origem,0.000144)::GEOMETRY, 4326), c2.the_geom) AND c2.vehicle_id
        = i;
36.     IF origem_padrao_paradas ISNULL THEN
37.         origem_padrao_paradas = DISTINCT c2.the_geom FROM avl_patterns_loc c2 WHERE
            (DISTANCE(SETSRID(BUFFER(geom_origem,0.000144)::GEOMETRY, 4326), c2.the_geom) <
            0.000449) AND c2.vehicle_id = i;
38.     END IF;
39. END IF;

```

Figura 29. Algoritmo 8: Identificação da Origem no Padrão de Localização.

4.3.5 Aspectos Estatísticos

O cálculo da probabilidade de cada possível destino do veículo poderá ser realizado de duas maneiras:

1. A primeira, como mencionado na seção anterior, por meio da probabilidade condicional, desde que a origem tenha sido encontrada no Padrão de Localização do veículo;
2. A segunda, na qual a origem não foi detectada no Padrão de Localização, será informado o destino mais provável a partir da frequência relativa obtida pelo número de visitas a cada possível destino.

Assim, para calcular a probabilidade dos prováveis destinos são geradas três matrizes. As linhas de cada matriz representarão as possíveis origens e as colunas os possíveis destinos a partir de cada origem. A geração destas matrizes considera o estado atual do veículo:

- a primeira matriz representa o relacionamento espacial/direcional entre as regiões de parada consideradas como origem e as regiões consideradas como destino.
 - Caso o veículo esteja parado, todas as regiões de parada do veículo armazenadas no seu padrão de localização serão simuladas como possíveis origens, desde que possuam trajetos que iniciam a partir dela. É importante ressaltar que pode existir uma região de parada no padrão que somente pode ser considerada como destino, isto porque, o processo de geração de padrão finaliza sempre em uma parada e, portanto, esta parada pode ainda não ter sido uma origem.
 - Caso o veículo esteja em movimento, o relacionamento entre as regiões de parada é exposto por meio das trajetórias armazenadas no Padrão de Movimentação do veículo, desde que estas trajetórias sejam iguais a atual trajetória do veículo. Dessa forma, cada trajetória definirá uma possível origem e destino, caso o destino ainda não exista na matriz, caso exista o destino na matriz, a célula na matriz que for correspondente a origem e destino recebe a contagem de visita a partir desta origem;
- a segunda matriz será uma especificação temporal da matriz anterior, ou seja, será constituída das regiões de parada que possuem trajetos que iniciam-se

no mesmo dia da semana do dia atual do veículo, ou seja, as possíveis origens e destinos para aquele dia da semana.

- Por fim, a terceira será uma especificação temporal da segunda matriz considerando-se o período do dia correspondente ao atual período do dia do veículo.

A partir deste relacionamento espacial existente entre os Padrões de Localização e Movimentação pode-se atribuir valores as células da matriz. Este valor é obtido considerando-se o número de visitas a um destino a partir de uma origem. Portanto, cada célula representa o número de visitas que um determinado destino (coluna da matriz) recebeu em relação à origem (linha correspondente na matriz), o somatório de cada linha ou coluna reflete qual origem e destino são mais frequentados pelo veículo, respectivamente. São exemplos dessas matrizes as tabelas 7, 8 e 9.

Tabela 7 – Frequência de visitas a um Destino a partir de uma Origem (espacial/direcional).

Origens	Destinos																		TOTAL
	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈	D ₉	D ₁₀	D ₁₁	D ₁₂	D ₁₃	D ₁₄	D ₁₅	D ₁₆	D ₁₇	D ₁₈	
O ₁	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
O ₂	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₃	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₄	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₅	0	0	0	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9
O ₆	11	3	4	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	23
O ₇	2	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	5
O ₈	0	0	1	14	0	0	0	0	0	0	0	1	1	1	1	1	0	0	20
O ₉	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₁₀	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₁₁	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₁₂	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
O ₁₃	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2
O ₁₄	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
O ₁₅	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
O ₁₆	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
O ₁₇	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
O ₁₈	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
TOTAL	19	5	9	24	1	1	1	1	2	1	1	1	1	2	2	1	1	1	74

Tabela 8 – Frequência de visitas a um Destino a partir de uma Origem (espacial/direcional e temporal dia da semana).

Origens	Destinos			TOTAL
	D ₁	D ₂	D ₃	
O ₁	1	0	0	1
O ₂	0	4	1	5
O ₃	5	0	0	5
TOTAL	6	4	1	11

Tabela 9 – Frequência de visitas a um Destino a partir de uma Origem (espacial/direcional temporal dia da semana e período do dia).

Origens	Destinos	
	D ₁	TOTAL
O ₁	1	1
TOTAL	1	1

As tabelas 7, 8 e 9 evidenciam a importância que a filtragem temporal possui na afirmação dos possíveis destinos. É importante ressaltar que essa filtragem pode eliminar da matriz de frequência a origem detectada, o que não é interessante, pois a evidência da origem na amostra considerada possibilita o cálculo da probabilidade condicional.

As figuras 30(a) e 30(b) ilustram as três maneiras de se obter a frequência relativa que será utilizada no cálculo da probabilidade de um destino quando o veículo está parado, considerando-se as três componentes (espacial, temporal/dia e temporal/dia/período do dia) que definem as matrizes. Além de ilustrar uma ideia da variação de frequência que os destinos podem assumir dependendo do valor considerado na componente temporal.

Já as figuras 30(c) e 30(d) representam as três maneiras de se obter a frequência relativa quando o veículo está em movimento. Nelas são exibidas as quatro componentes (espacial, direcional, temporal/dia e temporal/dia/período do dia) com intuito de enfatizar que a filtragem no número de visitas que um destino sofre é maior quando considerada a componente direcional ao invés da componente espacial.

Nas figuras 30(a), 30(b), 30(c) e 30(d), as cores dos retângulos representam a diferenciação entre os locais de parada, enquanto que cada retângulo representa uma visita. Assim, a quantidade de retângulos de uma mesma cor representa o número de visitas a uma parada, a variável **N** representa o número total de visitas para cada componente, enquanto que **P** representa a probabilidade para cada parada. Portanto, quando se considera somente a componente espacial, a pergunta é: “**Quais são os possíveis Destinos do veículo ‘V’ considerando-se a Origem ‘O’?**”.

CÁLCULO DA FREQUENCIA RELATIVA DO DESTINO CONSIDERANDO:		
ESPACIAL	ESPACIAL e 1° DIA	ESPACIAL, 1° DIA e 1° PERIODO
N = 12	N = 9	N = 5

Figura 30(a). Categorização dos Possíveis Destinos para o Cálculo de Probabilidade – Veículo Parado.

CÁLCULO DA FREQUENCIA RELATIVA DO DESTINO CONSIDERANDO:		
ESPACIAL	ESPACIAL e 1° DIA	ESPACIAL, 1° DIA e 2° PERIODO
N = 12	N = 9	N = 2

Figura 30(b). Categorização dos Possíveis Destinos para o Cálculo de Probabilidade – Veículo Parado.

CÁLCULO DA FREQUENCIA RELATIVA DO DESTINO CONSIDERANDO:			
ESPACIAL	ESPACIAL e DIRECIONAL	ESPACIAL e DIRECIONAL, 1º DIA	ESPACIAL e DIRECIONAL, 1º DIA e 1º PERIODO
N = 12	N = 7	N = 5	N = 2
$P_1 = 5/12$ $P_2 = 3/12$ $P_3 = 2/12$ $P_4 = 1/12$ $P_5 = 1/12$	$P_1 = 5/7$ $P_3 = 2/7$	$P_1 = 4/5$ $P_2 = 1/5$	$P_1 = 2/2$

Figura 30(c). Categorização dos Possíveis Destinos para o Cálculo de Probabilidade – Veículo em Movimento.

CÁLCULO DA FREQUENCIA RELATIVA DO DESTINO CONSIDERANDO:			
ESPACIAL	ESPACIAL e DIRECIONAL	ESPACIAL e DIRECIONAL, 1º DIA	ESPACIAL e DIRECIONAL, 1º DIA e 1º PERIODO
N = 12	N = 7	N = 4	N = nulo
$P_1 = 5/12$ $P_2 = 3/12$ $P_3 = 2/12$ $P_4 = 1/12$ $P_5 = 1/12$	$P_1 = 5/7$ $P_3 = 2/7$	$P_1 = 4/4$	

Figura 30(d). Categorização dos Possíveis Destinos para o Cálculo de Probabilidade – Veículo em Movimento.

As componentes relacionadas à temporalidade utilizam como critério de definição para os possíveis destinos as componentes, espacial e temporal, o que caracteriza uma análise espaço-temporal. Para especificar essa componente temporal considera-se a data referente ao início do trajeto realizado entre a origem e o destino.

A informação temporal diminui o número de destinos possíveis, pois elimina casos nos quais o veículo se deslocou da origem para o destino em dias da semana ou períodos diferentes do em questão. Nesta etapa, as perguntas são: **“Quais são os possíveis Destinos do veículo ‘V’ considerando-se a Origem ‘O’ no dia da semana ‘D’?”** ou **“Quais são os possíveis Destinos do veículo ‘V’ considerando-se a Origem ‘O’ no dia da semana ‘D’ e o período ‘P’ do dia?”**.

Por fim, a componente direcional somente será considerada quando o veículo estiver em movimento, ela considera quais são os trajetos armazenados no padrão de movimentação que correspondem ao atual trajeto do veículo. Por meio do relacionamento espacial que estes trajetos possuem com as regiões de parada é determinada quais são as possíveis origens e destinos.

Portanto, quando se utiliza a componente direcional a pergunta passa a ser: **“Qual é a Origem e o Destino do veículo ‘V’ considerando-se que ele trafega pelo caminho ‘C’?”**. Além das suas variações temporais: **“Qual é a Origem e o Destino do veículo ‘V’ considerando-se que ele trafega pelo caminho ‘C’ no dia da semana ‘D’?”** ou **“Qual é a Origem e o Destino do veículo ‘V’ considerando-se que ele trafega pelo caminho ‘C’ no dia da semana ‘D’ e o período ‘P’ do dia?”**.

4.3.5.1 Probabilidade Condicional

De acordo com Bussab e Morettin (2006), pode-se obter a Probabilidade Condicional de um determinado evento A ocorrer, desde que tenha ocorrido outro evento B e que a probabilidade do evento B seja maior que zero. Portanto, conhecendo-se a origem pode-se calcular a probabilidade da origem e, conseqüentemente calcular a

Probabilidade Condicional dos possíveis destinos a partir desta origem, esta probabilidade será denotada por $P(D_n|O_i)$ e calculada pela equação 1.

$$P(D_n|O_i) = \frac{P(D_n \cap O_i)}{P(O_i)} \quad (1)$$

O cálculo da probabilidade que uma determinada região de parada, considerada como origem, é realizado a partir do somatório do número de visitas das regiões consideradas como destinos a partir dessa região, dividido pelo somatório do número de visitas de todas as regiões consideradas como destino. Portanto, a região de parada que possuir a os destinos mais frequentados será aquela que apresentará a maior frequência relativa e, conseqüentemente, a maior probabilidade de ser considerada como Origem.

A equação 2 é utilizada para obter o valor de probabilidade $P(O_n)$ de uma região de parada ser considerada como origem, onde (a_{nj}) representa o número de visitas ao um destino a partir dessa região considerada como origem e $\sum_{i=1}^x \sum_{j=1}^y a_{ij}$ representa o somatório de todos os valores de frequência de visita obtidos para cada possível destino da matriz.

$$P(O_n) = \frac{\sum_{j=1}^y a_{nj}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}} \quad (2)$$

Análogamente, o cálculo executado para obter a probabilidade do destino utiliza o mesmo conceito e pode ser obtido pela equação 3.

$$P(D_n) = \frac{\sum_{i=1}^x a_{in}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}} \quad (3)$$

A Probabilidade Condicional de um possível destino pode ser calculada consultando-se os valores armazenados na matriz de frequências das visitas aos destinos, desde que a origem esteja presente na matriz de frequências. Por exemplo, considerando-

se a matriz da Tabela 8, e que a origem do veículo foi O_2 notam-se dois possíveis destinos: D_2 e D_3 . Assim o cálculo da probabilidade condicional de D_2 seria:

$$P(D_2|O_2) = \frac{\frac{a_{22}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}}}{\frac{\sum_{j=1}^y a_{2j}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}}} = \frac{a_{22}}{\sum_{j=1}^y a_{2j}} = \frac{4}{5} = 0.8$$

Analogamente, a probabilidade condicional de D_3 seria:

$$P(D_3|O_2) = \frac{\frac{a_{23}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}}}{\frac{\sum_{j=1}^y a_{2j}}{\sum_{i=1}^x \sum_{j=1}^y a_{ij}}} = \frac{a_{23}}{\sum_{j=1}^y a_{2j}} = \frac{1}{5} = 0.2$$

Após o cálculo das probabilidades são geradas as mensagens ao usuário informando a localização atual do veículo (município e bairro) e os prováveis destinos, detalhando o valor de probabilidade calculado para o destino e sua identificação (município e bairro). Por exemplo: **“O veículo V está parado no bairro B1 do município M1 e tem 0.5 de probabilidade de ir para o bairro B2 do município M1, 0.25 de probabilidade de ir para o bairro B3 do município M1, 0.25 de probabilidade de ir para o bairro B4 do município M1”** ou **“O veículo V saiu do bairro B1 do município M1 e tem 0.5 de probabilidade de estar indo para o bairro B2 do município M1, 0.25 de probabilidade de estar indo para o bairro B3 do município M2, 0.25 de probabilidade de estar indo para o bairro B1 do município M3”**.

O objetivo no cálculo da probabilidade é fornecer ao usuário o valor de probabilidade que possua o maior grau de refinamento, ou seja, considerando a relação O-D espacial/direcional, o dia da semana e o período do dia, caso esse filtro não retorne nenhum provável destino, o questionamento dos prováveis destinos irá considerar um nível inferior de refinamento até que somente reste a relação O-D espacial/direcional.

Caso não seja possível obter nenhum provável destino, significará que o veículo não possui padrão suficiente para análise de previsão e uma mensagem de alerta indicará que o veículo possui um comportamento atípico. Para uma avaliação dos resultados obtidos, todos os valores de probabilidade encontrados em todos os níveis de refinamento foram exibidos nas mensagens de previsão.

É importante ressaltar que, como são utilizados os três níveis de refinamento para o cálculo das probabilidades dos destinos, para se obter a Probabilidade Condicional de um determinado destino, utiliza-se a probabilidade obtida para a região de parada considerada como origem no mesmo nível de refinamento. Por exemplo, se for utilizada a probabilidade do destino considerando o nível de refinamento espaço-temporal (período do dia), utiliza-se a probabilidade de origem correspondente a esse nível de refinamento, ou seja, espaço-temporal (período do dia). A figura 31 ilustra o algoritmo responsável pelo cálculo da probabilidade condicional e a figura 32 está representando o algoritmo responsável pelo cálculo do destino mais provável quando a origem é desconhecida.

Algoritmo 9: Cálculo da Probabilidade Condicional de um Destino.

```

1.  -- Calcular a Probabilidade Condicional ESPACIAL
2.  gid_origem = gid FROM avl_padroes_paradas WHERE the_geom = origem_padrao_paradas AND cod_veic = i;
3.  ntotal = SUM(total)::INT FROM avl_distrib_part_cheg_espac;
4.  norigem = total FROM avl_distrib_part_cheg_espac WHERE origem = gid_origem;
5.  prob_origem = norigem::NUMERIC/ntotal::NUMERIC;
6.  IF prob_origem <> 0 AND ntotal <> 0 THEN
7.      FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname = 'avl_distrib_part_cheg_espac'
AND NOT a.attisdropped LOOP
8.          n4 = pegavalorcoluna(column_name::TEXT, 'avl_distrib_part_cheg_espac',
gid_origem);
9.          IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
10.             prob_inter = n4::NUMERIC/ntotal::NUMERIC;
11.             prob_cond = TRUNC(ROUND((prob_inter/prob_origem), 2)*100, 0);
12.         END IF;
13.     END LOOP;
14. END IF;
15. -- Calcular a Probabilidade Condicional TEMPORAL DIA
16. gid_origem = gid FROM avl_padroes_paradas WHERE the_geom = origem_padrao_paradas AND cod_veic = i;
17. ntotal = SUM(total)::INT FROM avl_distrib_part_cheg_espac_dia;
18. norigem = total FROM avl_distrib_part_cheg_espac_dia WHERE origem = gid_origem;
19. prob_origem = norigem::NUMERIC/ntotal::NUMERIC;
20. IF prob_origem <> 0 AND ntotal <> 0 THEN
21.     FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_dia' AND NOT a.attisdropped LOOP
22.         n4 = pegavalorcoluna(column_name::TEXT, 'avl_distrib_part_cheg_espac_dia',
gid_origem);
23.         IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
24.             prob_inter = n4::NUMERIC/ntotal::NUMERIC;
25.             prob_cond = TRUNC(ROUND((prob_inter/prob_origem), 2)*100, 0);
26.         END IF;
27.     END LOOP;
28. END IF;
29. -- Calcular a Probabilidade Condicional TEMPORAL PERIODO DO DIA
30. gid_origem = gid FROM avl_padroes_paradas WHERE the_geom = origem_padrao_paradas AND cod_veic = i;
31. ntotal = SUM(total)::INT FROM avl_distrib_part_cheg_espac_per_dia;
32. norigem = total FROM avl_distrib_part_cheg_espac_per_dia WHERE origem = gid_origem;
33. prob_origem = norigem::NUMERIC/ntotal::NUMERIC;
34. IF prob_origem <> 0 AND ntotal <> 0 THEN
35.     FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_per_dia' AND NOT a.attisdropped LOOP
36.         n4 = pegavalorcoluna(column_name::TEXT, 'avl_distrib_part_cheg_espac_per_dia', gid_origem);
37.         IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
38.             prob_inter = n4::NUMERIC/ntotal::NUMERIC;
39.             prob_cond = TRUNC(ROUND((prob_inter/prob_origem), 2)*100, 0);
40.         END IF;
41.     END LOOP;
42. END IF;

```

Figura 31. Algoritmo 9: Cálculo da Probabilidade Condicional.

Algoritmo 10: Cálculo da Probabilidade de um Destino.

```

1.  -- Calcular a Probabilidade ESPACIAL
2.  ntotal = 0;
3.  FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname = 'avl_distrib_part_cheg_espac'
AND NOT a.attisdropped LOOP
4.      n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac');
5.      IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
6.          ntotal = ntotal + n4;
7.      END IF;
8.  END LOOP;
9.  IF ntotal > 0 THEN
10.     FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname = 'avl_distrib_part_cheg_espac'
AND NOT a.attisdropped LOOP
11.         n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac');
12.         IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
13.             prob_freq = TRUNC(ROUND((n4::NUMERIC/ntotal::NUMERIC), 2)*100, 0);
14.         END IF;
15.     END LOOP;
16. END IF;
17. -- Calcular a Probabilidade TEMPORAL DIA
18. n = COUNT(origem)::INT FROM avl_distrib_part_cheg_espac_dia;
19. IF n > 0 THEN
20.     ntotal = 0;
21.     FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_dia' AND NOT a.attisdropped LOOP --LISTA OS NOMES DE COLUNA DA TABELA
22.         n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac_dia');
23.         IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
24.             ntotal = ntotal + n4;
25.         END IF;
26.     END LOOP;
27.     IF ntotal > 0 THEN
28.         FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_dia' AND NOT a.attisdropped LOOP
29.             n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac_dia');
30.             IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
31.                 prob_freq = TRUNC(ROUND((n4::NUMERIC/ntotal::NUMERIC), 2)*100, 0);
32.             END IF;
33.         END LOOP;
34.     END IF;
35. END IF;
36. -- Calcular a Probabilidade TEMPORAL PERIODO DIA
37. n = COUNT(origem)::INT FROM avl_distrib_part_cheg_espac_per_dia;
38. IF n > 0 THEN
39.     ntotal = 0;
40.     FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_per_dia' AND NOT a.attisdropped LOOP
41.         n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac_per_dia');
42.         IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
43.             ntotal = ntotal + n4;
44.         END IF;
45.     END LOOP;
46.     IF ntotal > 0 THEN
47.         FOR column_name IN SELECT a.attname FROM pg_catalog.pg_attribute a INNER JOIN
pg_stat_user_tables c ON a.attrelid = c.relid WHERE a.attnum > 0 AND c.relname =
'avl_distrib_part_cheg_espac_per_dia' AND NOT a.attisdropped LOOP
48.             n4 = SOMADESTINO(column_name::TEXT, 'avl_distrib_part_cheg_espac_per_dia');
49.             IF column_name <> 'origem' AND column_name <> 'total' AND n4 > 0 THEN
50.                 prob_freq = TRUNC(ROUND((n4::NUMERIC/ntotal::NUMERIC), 2)*100, 0);
51.             END IF;
52.         END LOOP;
53.     END IF;
54. END IF;

```

Figura 32. Algoritmo 10: Cálculo da Probabilidade do Destino quando a Origem não for encontrada no Padrão de Localização.

4.4 CONSIDERAÇÕES FINAIS

O objetivo principal deste capítulo foi descrever a metodologia utilizada para previsão de movimentações dos veículos. Esta metodologia inclui desde a preparação dos dados, passando pelos processos de geração de padrão e previsão, até o detalhamento das rotinas expressas na forma dos algoritmos desenvolvidos, empregadas para extrair resultados.

Espera-se que o capítulo possibilite uma visão detalhada e esquemática das rotinas empregadas e como foi proposta a solução para previsão da movimentação de um veículo, considerando-se apenas as informações de localização e instante de tempo do rastreamento. No capítulo seguinte são apresentados os resultados dos experimentos realizados considerando-se esta metodologia.

5 EXPERIMENTAÇÃO

Com o intuito de validar os métodos utilizados nesta pesquisa para a previsão da localização de um determinado veículo, considerando-se o processo de geração de Padrões de Localização e Movimentação, foram realizados experimentos que utilizaram uma amostra de dados reais de rastreamento. Os resultados desses experimentos são apresentados nas seções seguintes.

5.1 CONSIDERAÇÕES INICIAIS

Nos experimentos realizados, utilizou-se uma amostra com dados reais referente ao rastreamento de treze veículos. A figura 33 ilustra a distribuição dos pontos de rastreamento de um desses veículos sobreposto à base cartográfica. O objetivo principal dessa figura é mostrar que a distribuição espacial dos pontos sobre a base cartográfica demonstra certos padrões visuais como: corredores ou caminhos e aglomerações de pontos, que claramente representam trajetos comumente percorridos e locais frequentemente visitados.

Figura 33. Pontos de rastreamento GPS para um veículo, sobrepostos à base cartográfica.

A figura 34 ilustra a representação dos trajetos realizados por um veículo (Padrões de Movimentação) sobrepostos à Base Cartográfica. Observa-se que na maioria dos casos a linha que define um trajeto realizado não coincide com os limites de quadras. Isso ocorre não somente pela acurácia da informação posicional, mas principalmente devido à taxa de coleta das posições do veículo, da velocidade desenvolvida no trecho de via e principalmente, pela pouca frequência com que o trajeto ocorre.

Figura 34. Trajetórias de um veículo sobrepostas à base cartográfica.

5.2 EXECUÇÃO DAS FUNÇÕES DESENVOLVIDAS

Como mencionado anteriormente, foram desenvolvidas duas funções: a primeira responsável pela geração dos Padrões de Movimentação e Localização dos

veículos, e a segunda responsável pela previsão da localização de um determinado veículo.

O propósito é que a função responsável pela geração dos padrões seja executada periodicamente para manter as informações do padrão atualizadas. Portanto, o primeiro passo do experimento foi executar a função responsável pela geração de padrões, denominada **avl_patterns()**. Para tanto, utilizou-se a interface do SGBD (pgAdim III) para executar o seguinte comando em SQL: **SELECT avl_patterns();**.

A função responsável pela geração dos padrões obteve, para os treze veículos, cerca de trezentos registros de padrões de Localização e cerca de dez mil registros de Padrões de Movimentação. Uma vez gerados os padrões para os treze veículos, qualquer um deles poderá ser questionado quanto à sua localização futura a partir dos seus dados de rastreamento atuais.

A previsão de localização de um veículo é feita utilizando-se da segunda função desenvolvida e denominada **avl_prediction()**. A execução dessa função consiste no seguinte comando SQL: **SELECT avl_prediction(VEHICLE_ID);**, na qual a variável **VEHICLE_ID** representa o identificador do veículo no banco de dados. Este procedimento foi realizado para cada um dos treze veículos da amostra e os resultados obtidos são discutidos na seção seguinte.

5.3 RESULTADOS OBTIDOS

Após executar a função de previsão para cada veículo chegou-se aos seguintes resultados expressos nas tabelas 10 e 11. Por questões de privacidade os identificadores dos veículos foram alterados para números sequenciais, bem como os resultados dos nomes obtidos para os prováveis destinos foram ocultados, pois, a utilização dos nomes reais dos bairros e municípios poderiam facilitar na identificação do proprietário do veículo.

Os significados das siglas utilizadas para representar as componentes nas tabelas 10 e 11 são:

- E – Espacial;
- ETD – Espacial e Temporal para Dia da semana;
- ETPD – Espacial e Temporal para Período do Dia;
- ED – Espacial e Direcional;
- EDTD – Espacial – Direcional e Temporal para Dia da semana;
- EDTPD - Espacial – Direcional e Temporal para Período do Dia.

A tabela 10 apresenta os valores de probabilidade condicional obtidos para cada provável destino de cada um dos sete veículos, nos quais foi possível determinar a origem no seu padrão de localização. Na tabela 11 são apresentadas as probabilidades dos prováveis destinos para os veículos, nos quais não foi possível determinar a origem.

Tabela 10 – Probabilidade Condicional dos destinos para cada veículo.

VEHICLE_ID	DESTINOS	Probabilidade Condicional (%)					
		PARADO			EM MOVIMENTO		
		E	ETD	ETPD	ED	EDTD	EDTPD
1	DESTINO 1, 3, 4 e 6	6	-	-	-	-	-
	DESTINO 2	65	100	-	100	-	-
	DESTINO 5	12	-	-	-	-	-
2	DESTINO 1	100	-	-	-	-	-
3	DESTINO 1	33	100	-	-	-	-
	DESTINO 2, 3, 4, 5, 6, e 7	11	-	-	-	-	-
4	DESTINO 1 e 3	25	-	-	-	-	-
	DESTINO 2	50	-	-	-	-	-
9	DESTINO 1	56	33	-	-	-	-
	DESTINO 2, 4 e 5	6	-	-	-	-	-
	DESTINO 3	28	67	-	-	-	-
11	DESTINO 1	48	80	100	-	-	-
	DESTINO 2	13	-	-	-	-	-
	DESTINO 3	17	20	-	-	-	-
	DESTINO 4, 5, 6, 7 e 8	4	-	-	-	-	-
12	DESTINO 1, 3 e 6	5	-	-	-	-	-
	DESTINO 2 e 7	9	-	-	-	-	-
	DESTINO 4	55	100	100	-	-	-
	DESTINO 5	14	-	-	-	-	-

Tabela 11 – Probabilidade (Frequência Relativa) dos destinos para cada veículo.

VEHICLE_ID	DESTINOS	Probabilidade (%)					
		PARADO			EM MOVIMENTO		
		E	ETD	ETPD	ED	EDTD	EDTPD
5	DESTINO 1 - 9	11	-	-	-	-	-
6	DESTINO 1	28	38	20	-	-	-
	DESTINO 2, 3, 5, 8, 9, 11 - 13, 16 - 18	3	-	-	-	-	-
	DESTINO 4	9	13	20	-	-	-
	DESTINO 6	9	25	20	-	-	-
	DESTINO 7	3	13	20	-	-	-
	DESTINO 10, 15	6	-	-	-	-	-
	DESTINO 14	3	13	20	-	-	-
7	DESTINO 1, 5 - 7, 9 - 12 e 15	3	-	-	-	-	-
	DESTINO 2	37	33	67	-	-	-
	DESTINO 3	17	17	-	-	-	-
	DESTINO 4	6	-	-	-	-	-
	DESTINO 8	3	17	-	-	-	-
	DESTINO 13	6	17	-	-	-	-
	DESTINO 14	6	17	33	-	-	-
8	DESTINO 1	12	17	7	-	-	-
	DESTINO 2, 23 e 25	2	-	-	-	-	-
	DESTINO 3	4	6	-	-	-	-
	DESTINO 4, 11 e 18	4	-	-	-	-	-
	DESTINO 5	11	28	36	-	-	-
	DESTINO 6	24	22	21	-	-	-
	DESTINO 7, 8, 10, 12, 13, 15, 16, 20 - 22, 24, 26 - 36	1	-	-	-	-	-
	DESTINO 9 e 19	4	6	7	-	-	-
	DESTINO 14, 17 e 35	1	6	7	-	-	-
10	DESTINO 1, 3, 6 - 11, 15 - 18, 20 - 22, 25, 26, 28, 30 - 35, 37 - 54, 56 - 62	1	-	-	-	-	-
	DESTINO 2	22	29	38	-	-	-
	DESTINO 4	7	4	-	-	-	-
	DESTINO 5	16	17	-	-	-	-
	DESTINO 12	2	-	-	-	-	-
	DESTINO 13	2	8	-	-	-	-
	DESTINO 14	9	4	13	-	-	-
	DESTINO 19	2	4	-	-	-	-
	DESTINO 23	3	8	25	-	-	-
	DESTINO 24 e 36	1	4	-	-	-	-
	DESTINO 27	4	4	-	-	-	-
	DESTINO 29	1	8	13	-	-	-
	DESTINO 55	1	4	13	-	-	-
13	DESTINO 1 e 3	17	-	-	-	-	-
	DESTINO 2, 4 - 6 e 8	8	-	-	-	-	-
	DESTINO 7	25	-	-	-	-	-

A figura 35 ilustra a distribuição dos prováveis destinos para o veículo 1, representado na tabela 10. Para cada destino ilustrado em círculos vermelhos existe uma Probabilidade Condicional associada e, neste caso a origem, ilustrada pelo círculo verde, também é considerada como um provável destino, o que significa que no caso deste veículo existe a possibilidade dele sair e retornar para o mesmo local.

Figura 35. Representação dos prováveis destinos do veículo 1 da tabela 10.

Na figura 36 podem-se observar os trajetos executados entre a origem e os prováveis destinos do veículo 1, nela são representados os trajetos executados por esse veículo entre a origem, determinada na função de previsão de localização, e os destinos possíveis a partir dessa origem.

Figura 36. Representação dos trajetos entre a origem e prováveis destinos do veículo 1.

Com base nos resultados exibidos na Tabela 11, na qual em seis dos 13 veículos avaliados não foi possível determinar a origem no padrão de localização, considerando-se sua atual localização. Uma nova indagação surgiu: ***“A variabilidade de ocorrência das regiões de parada é maior, no caso destes seis veículos, do que no grupo de veículos onde foi possível determinar a Origem?”***.

Com base nesta indagação uma análise sobre a amostra de dados de rastreamento dos 13 veículos foi realizada. Essa análise teve por objetivo identificar a existência de algum relacionamento entre o número de paradas detectadas e a não definição

de um padrão suficiente para responder à localização futura do veículo, com base num valor de probabilidade condicional, ou seja, verificar se a não determinação da origem no padrão poderia estar associada ao fato do veículo possuir uma grande quantidade de registros no Padrão de Localização (variabilidade de paradas). Sendo assim, a Tabela 12 informa o número de regiões de parada detectadas para cada veículo da amostra.

Tabela 12 – Número de regiões de parada detectadas para cada veículo.

ID do Veículo	Quantidade de Regiões de Parada	Origem Detectada?
1	12	Sim
2	49	Sim
3	20	Sim
4	18	Sim
5	09	Não
6	18	Não
7	15	Não
8	36	Não
9	15	Sim
10	62	Não
11	18	Sim
12	23	Sim
13	08	Não

Considerando-se dois grupos: o primeiro representando os veículos aos quais foi possível detectar a origem no padrão e o segundo representando àqueles que não foram detectadas a origem, foi calculado um valor médio das regiões de parada encontradas para cada grupo. O valor médio de regiões para o primeiro grupo de veículos foi igual a 22.143, ou seja, 22 regiões, enquanto que, para o segundo grupo de veículos o valor foi de 24.667, aproximadamente 25 regiões de parada. Com base nestes valores médios e nos valores da Tabela 12, uma medida de variabilidade foi aplicada como ilustra a Tabela 13.

Tabela 13 – Desvio-Padrão do número de paradas para cada classe de veículo.

Classes (Origem Detectada?)	n	$\sum_{i=1}^n x_i - \bar{x} $	$\sum_{i=1}^n (x_i - \bar{x})^2$	dm(X)	var(X)	dp(X)
Sim	07	55.715	915.429	07.959	130.776	11.436
Não	06	97.334	2183.333	16.222	363.889	19.076

Considerando-se a afirmação de Bussab e Morettin (2002), a respeito de que a variância é uma medida de dimensão igual ao quadrado da dimensão dos dados, e que comumente é utilizado o desvio-padrão (dp(X)) para indicar a medida de dispersão do erro cometido ao tentar substituir cada observação pela média do conjunto de dados. Pode-se dizer, então que, de acordo com o desvio médio (dm(X)) e variância (var(X)) da tabela 13, o grupo de veículos que possuem a origem detectada é mais homogêneo que o grupo de veículos que não possui origem detectada no padrão de localização. Pois, a variabilidade na quantidade de regiões de parada para o segundo grupo é quase três vezes maior que a variabilidade encontrada no primeiro grupo. O que confirma a indagação a respeito da variabilidade do grupo onde não foi possível determinar a origem no padrão de localização.

5.4 CONSIDERAÇÕES FINAIS

Os resultados dos experimentos mostram que é possível determinar a localização futura de um veículo, considerando-se apenas seu histórico de movimentação, desde que seja gerado um padrão de movimentação e localização a partir dos dados de rastreamento. Porém, os experimentos ressaltam que a identificação da origem do veículo no padrão de localização, que possibilita o cálculo da probabilidade condicional dos prováveis destinos do veículo, está relacionada à homogeneidade da amostra dos dados de rastreamento.

6 CONCLUSÃO E RECOMENDAÇÕES

Os métodos desenvolvidos nesta pesquisa predizem a localização do veículo, atribuindo-se valores de probabilidade, levando em consideração a sua posição atual e, o padrão gerado a partir do histórico das movimentações realizadas em instantes de tempo anteriores, que correspondem o período de tempo atual do veículo.

Este trabalho teve como contribuição principal considerar o histórico das movimentações por meio da geração de padrões e da utilização destes padrões no cálculo da probabilidade condicional de cada provável destino do veículo investigado, o que possibilita ao usuário uma interpretação e contextualização dos prováveis destinos para uma melhor tomada de decisão.

Além de propor uma metodologia com base nas funções desenvolvidas que apoia nos processos de monitoramento dos sistemas AVL (voltado para mercado de segurança), esta metodologia desenvolveu instrumentos que possibilitam ao sistema, com o passar do tempo, um maior conhecimento do Padrão de Movimentação e Localização de cada veículo rastreado, desde que os dados de rastreamento sejam armazenados no banco de dados e os padrões atualizados periodicamente.

É importante ressaltar que a metodologia empregada não está isenta de ressalvas, a primeira está relacionada ao fato de que existem veículos ou indivíduos que são muito prováveis e outros não, e a segunda de que todo resultado obtido é probabilístico. Porém, a primeira ressalva se aplica a todas as análises de comportamento, pois, o comportamento de cada veículo é uma qualidade intrínseca de cada condutor e por isso não pode ser generalizado, ao passo que a segunda ressalva fundamenta o resultado obtido porque considera valores de probabilidade.

A utilização de dados reais do rastreamento de veículos, associada ao desconhecimento da propriedade dos dados, reforça ainda mais as funções desenvolvidas e, sugere que esta implementação, considerando outras fontes de

informação (ex: ignição do motor, apontamento de locais frequentados pelo próprio condutor do veículo, etc.), podem fundamentar e tornar mais confiável o processo de previsão.

Considerando o desenvolvimento de novas tecnologias de rastreamento, que cada vez mais se tornam parte da rotina humana, este estudo amplia as possibilidades de utilização deste tipo de tecnologia. Sua aplicabilidade ao ambiente urbano e o fato de que o objeto rastreado pode ou não ser um veículo, podendo ser também um cartão de crédito, celular, crachá de escritório ou ainda uma combinação de todas essas fontes de informação, o que só aumenta o valor das informações obtidas a partir do rastreamento e torna o processo de extração de padrões mais confiável e, conseqüentemente, o processo de previsão da localização do ente rastreado no sistema.

6.1 TRABALHOS FUTUROS

Como continuação desse trabalho, sugere-se um estudo que especifique um intervalo de tempo confiável para a obtenção de uma amostra de dados de rastreamento, que seja suficiente para garantir um Padrão de Movimentação e Localização homogêneo para todos os veículos rastreados nesta amostra, independente da variabilidade de movimentações e paradas que o veículo executa.

Além disso, sugerem-se também outras formas de se considerar a componente direcional na análise de previsão, algo que não esteja somente relacionado às mudanças de direção no trajeto executado, mas que também considere o melhor caminho entre os pontos subsequentes do trajeto, a fim de se identificar maiores detalhes relacionados às mudanças de direção. Isto é importante uma vez que, da maneira que foi considerada nesta pesquisa, as mudanças de direção podem passar despercebidas, em virtude da taxa de coleta dos pontos de rastreamento. Assim, a determinação do melhor caminho entre pontos subsequentes pode evidenciar as mudanças de direção para a análise considerada na componente direcional.

Como um subproduto espera-se que a previsão da localização auxilie nos processos de monitoramento dos veículos rastreados, tornando o sistema cada vez mais independente de intervenções humanas e que isto aumente a segurança dos clientes que possuem seus veículos rastreados, principalmente em deslocamentos rotineiros, nos quais se torna mais fácil identificar um padrão.

Por fim, espera-se que a previsão da movimentação dos veículos possa ser utilizada por outros sistemas, para apoiar estudos de oferta de expansão da rede viária urbana ou mudanças de trajetos, considerando-se a informação de que num determinado horário de um dia da semana, exista uma tendência de deslocamentos de vários veículos para uma mesma região de destino da cidade. E, ainda, que as informações oriundas do processo de previsão forneçam um volume de informação mais confiável para as gerenciadoras de risco, de modo a reduzir o risco especificamente pelo padrão de cada cliente, discriminando e particularizando cada vez mais o valor do bem segurado de acordo com o perfil de movimentação de cada cliente.

REFERÊNCIAS

ACKOFF, R. L. Towards a system of systems concepts. **Management Science**, v. 17, n. 11, July 1971, p. 661-671.

ALESHEIKH, A.A. **Design and implementation of a moving object database for truck information systems**. Disponível em:
http://www.gisdevelopment.net/proceedings/mapmiddleeast/2008/mme08_32.pdf.
 Acesso em: 4 de março de 2010.

BEARD, K. ; PALANCIOGLU, H. M. Estimating positions and paths of moving objects. In: INTERNATIONAL WORKSHOP ON TEMPORAL REPRESENTATION AND REASONING, 7. (TIME'00), 2000. **Proceedings**. p. 155.

BOLFARINE, H. ; BUSSAB, W. O. **Elementos de amostragem**. São Paulo: Edgard Blücher, 2005.

BRANCO, A. Softwares para gerenciamento e monitoramento. **InfoGPS**, n. 10, 2009a.

BRANCO, A. Tecnologia embarcada para o futuro. **InfoGPS**, n. 11, 2009b.

BUSSAB, W. O. ; MORETTIN, P. A. **Estatística básica**. 5.ed. São Paulo: Editora Saraiva, 2006.

CHUNG, J. D.; PAEK, O. H.; LEE, J. W.; RYU, K. H.: **Temporal patterns mining of moving objects for location-based service**. Berlin: Springer, 2002. (Lecture Notes in Computer Science, v. 2453, p. 331-340).

CHUNG, E.-H. ; SHALABY, A.: A trip reconstruction tool for GPS-based personal travel surveys. **Transportation Planning and Technology**, v. 28, n.5, p. 381 – 401, 2005.

CTAGEO (CTAGEO ENGENHARIA E GEOPROCESSAMENTO): Fundamentos de base cartográfica e geoprocessamento aplicados à avaliação imobiliária. **Utilidades - Artigos**. Disponível em: <http://www.ctageo.com.br/insercoes/ESCOPO.pdf>.
 Acesso em: Janeiro de 2010.

CUNHA, N. J. **Metodologia de geração dinâmica de padrões de viagens rodoviárias para monitoramentos inteligentes de veículos de carga em sistemas AVL**. 2008. 138p. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2008.

ESRI (Environmental Systems Research Institute) **ESRI shapefile technical description**. Disponível em:
<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf> >. Acesso em: 15 de novembro de 2009.

GÜTING, R. H. ; SCHNEIDER, M.: **Moving objects databases**. San Francisco: Morgan Kaufmann, 2005.

HELLINGA, B. ; AERDE, M. V. A.: Statistical analysis of the reliability of using RGS vehicle probes as estimators of dynamic O-D departure routes. **Journal of Intelligent Transportation Systems**, v. 2, n. 1, p. 21 – 44, 1994.

HOUAISS, I. A. **Dicionário eletrônico Houaiss da língua portuguesa**. São Paulo: Editora Objetiva, 2001.

JIN, L.; LEE, Y.; SEO, S.; RYU, K. H. **Discovery of temporal frequent patterns using TFP-Tree**. Berlin:Spring, 2006. (Lecture Notes in Computer Science, v. 4016, p. 349-361).

JUNG, Y. J.; RYU, K. H. **The vehicle tracking system for analyzing transportation vehicle information**. Berlin: Spring-Verlag, 2006. (APWeb Workshops 2006, p. 1012 -1020).

MONICO, J. F. G. **Posicionamento pelo NAVSTAR-GPS: descrição, fundamentos e aplicações**. São Paulo: Editora UNESP, 2000.

MONICO, J. F. G. **Posicionamento pelo GNSS: descrição, fundamentos e aplicações**. 2.ed. São Paulo: Editora UNESP, 2008.

NHAN, V. T. H. ; RYU, K. H. **Future location prediction of moving objects based on movement rules**. Berlin: Spring, 2006. (Lecture Notes in Computer Science, v. 344, p. 875-881).

OLIVEIRA, S. R. **Modelo e implementação de objetos móveis em banco de dados objeto-relacional**. 2003. 113p. Dissertação (Mestrado) – Centro de Ciências e Tecnologia da Universidade Federal de Campina Grande, Campina Grande, 2003.

RODRIGUES, M. **AVL: logística e segurança integradas**. Disponível em: < http://www.geocargo.com.br/midia_show.asp?codigo=6>. Acesso em: 11 de março de 2008a.

RODRIGUES, M. Objetos móveis. **InfoGPS**, n. 2, 2008b.

RODRIGUES, M.; CUGNASCA, C. E. ; QUEIROZ FILHO, A. P. **Rastreamento de veículos**. São Paulo: Oficina de Textos, 2009.

SEEBER, G. **Satellite geodesy: foundations, methods, and applications**. 2.ed. Berlin: Walter de Gruyter, 2003.

SIMONCIC, M. A bayesian network model of two-car accidents. **Journal of Transportation and Statistics**, v. 7, n. 2/3, 2004. Disponível em: < http://www.bts.gov/publications/journal_of_transportation_and_statistics/volume_07_number_23/html/paper_02/index.html>. Acesso em: 26 de Outubro de 2009.

SMALWORLD (Core Spatial Technology). **Version 4.1.1 Main Documentation.**, 2008.

UCEL. Disponível em: < www.ucel.com.br/operadoras/vivo.asp >. Acesso em: 20 de Novembro de 2008.

VOLPATO, M. M. L.; VIEIRA, T. G. C.; ALVES, H. M. R. ; SOUZA, V. C. O.: GPS de navegação: dicas ao usuário. **Circular Técnica EPAMIG**, n. 45. Disponível em: < http://www.epamig.br/index.php?option=com_search&Itemid=99999999&searchword=circular+t%E9cnica&searchphrase=any&ordering=newest&limit=30&limitstart=30 >. Acesso em: 22 de Novembro de 2009.

WELLS, D.; BECK N.; DELIKARAOGLU D.; KLEUSBERG A.; KRAKIWSKY E. J.; LACHAPPELLE G.; LANGLEY R. B.; NAKIBOGLU M.; SCHWARZ K. P.; TRANQUILLA J. M. ; VANÍCEK P. **Guide to GPS positioning**. New Brunswick, Canadá: Canadian GPS Associates, 1986.

WOLF, J., SCHONFELDER, S., SAMAGA, U., OLIVEIRA, M. ; AXHAUSEN, K. W. Eighty week of GPS traces: approaches to enriching trip information. **Transportation Research Record**, n. 1870 , 2004.

XU, Y.; LEE, W. C. **On localized prediction for power efficient object tracking in sensor networks**. New York: IEEE, 2003.

Yi, B. **Um modelo de dados para objetos móveis**. 2004. 86p. Dissertação (Mestrado) – Instituto de Computação da Universidade Estadual de Campinas, Campinas, 2004.

APÊNDICE A

Diagrama do banco de dados da pesquisa:

APÊNDICE B

Procedimentos de carga dos dados de rastreamento no banco de dados:

Procedimento 1 – linha de comando de criação do arquivo de carga *.sql:

```
$ shp2pgsql -c -D arquivos_shapefile nome_da_tabela > nome_do_arquivo.sql
```

As variáveis dessa linha de comando se referem:

- **arquivos_shapefile** = nome do arquivo ESRI® shapefile (ex: bairros.shp);
- **nome_da_tabela** = nome da tabela que será criada no Banco de dados para armazenar os dados referentes ao arquivo ESRI® shapefile (ex: tabela_bairros);
- **nome_do_arquivo.sql** = nome do arquivo SQL que será criado com a estrutura de importação dos dados e criação da tabela no Banco de dados (ex: carga_bairros.sql).

Essa linha de comando gera um arquivo no formato ***.sql** com a estrutura de inserção de dados para o banco de dados e criação da tabela que foi indicada na linha de comando. A figura 1 ilustra a estrutura de inserção de dados criada num arquivo ***.sql**.

The screenshot shows a PostgreSQL command window with the following content:

```
File Edit View Query Favourites Help
mestrado_postgis on localhost:5...
BEGIN;
CREATE TABLE "avl_all" (gid serial PRIMARY KEY,
"cod_veic" integer,
"latitude" numeric,
"longitude" numeric,
"dataabr" varchar(254));
SELECT AddGeometryColumn('','avl_all','the_geom','-1','POINT',2);
COPY "avl_all" ("cod_veic","latitude","longitude","dataabr",the_geom) FROM stdin;
124507 -23.54720 -46.49370 2008-08-28 00:02:00 010100000008104C58F313F47C0075F984C158C37C0
124507 -23.54750 -46.49310 2008-08-28 00:05:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:08:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:11:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:14:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:17:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:23:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:26:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
124507 -23.54750 -46.49310 2008-08-28 00:29:00 010100000002CD49AE61D3F47C05C8FC2F5288C37C0
```

Figura 1. Arquivo de Inserção dos Dados de Rastreamento (Pontos GPS).

Com base na figura 1, pode-se dizer, por exemplo, que os valores que as variáveis, na linha de comando assumiram, foram:

```
$ shp2pgsql -c -D dados_rastreamento.shp avl_all > carga_rastreamento.sql
```

Procedimento 2 – linha de comando de inserção dos dados *.shp:

Uma vez que tenha sido criado o arquivo ***.sql**, a inserção dos dados pode ser realizada executando a seguinte linha de comando:

```
$ psql -d nome_do_banco -U usuario_do_banco -f nome_do_arquivo.sql
```

As variáveis dessa linha de comando se referem:

- **nome_do_banco** = nome do Banco de dados criado utilizando a extensão espacial PostGIS;
- **usuario_do_banco** = usuário de acesso ao Banco de dados onde serão importados os dados;
- **nome_do_arquivo.sql** = nome do arquivo SQL que foi criado com a estrutura de importação dos dados e criação da tabela no Banco de dados .