BRUNO SARNO MUGNELA

GENPOLIS:
PROTOTIPAGEM E APLICAÇÃO DE UM SIMULADOR DE TRÂNSITO
VOLTADO PARA OTIMIZAÇÃO DE SINALIZAÇÃO SEMAFÓRICA
POR MEIO DE ALGORITMOS GENÉTICOS

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de Mestre em Ciências

São Paulo
2012

BRUNO SARNO MUGNELA

GENPOLIS:
PROTOTIPAGEM E APLICAÇÃO DE UM SIMULADOR DE TRÂNSITO
VOLTADO PARA OTIMIZAÇÃO DE SINALIZAÇÃO SEMAFÓRICA
POR MEIO DE ALGORITMOS GENÉTICOS

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de Mestre em Ciências

Área de Concentração:

Sistemas Eletrônicos

Orientador:

Prof. Dr. Marcio Lobo Netto

São Paulo
2012

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, 04 de Setembro de 2012

Assinatura do autor

Assinatura do orientador

FICHA CATALOGRÁFICA

Mugnela, Bruno Sarno

Genpolis: prototipagem e aplicação de um simulador de transito voltado para otimização de sinalização semafórica por meio de algoritmos genéticos / B.S. Mugnela. -- São Paulo, 2012.

115 p.

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

1. Transporte urbano (Simulação) 2. Sinalização de tráfego 3. Otimização combinatória 4. Algoritmos genéticos I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II. t.

BRUNO SARNO MUGNELA

GENPOLIS: PROTOTIPAGEM E APLICAÇÃO DE UM SIMULADOR DE TRÂNSITO
VOLTADO PARA OTIMIZAÇÃO DE SINALIZAÇÃO SEMAFÓRICA
POR MEIO DE ALGORITMOS GENÉTICOS

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção
do título de Mestre em Ciências.

Aprovado em: 13 / 07 / 2012

Banca Examinadora:

Prof. Dr. Marcio Lobo Netto

PSI / Escola Politécnica da USP

Prof. Dr. Fuad Kassab Junior

PTC / Escola Politécnica da USP

Dr. Marcelo Camilli Landmann

LOGIT Engenharia Consultiva

DEDICATÓRIA

Dedico este trabalho à toda minha família e à minha namorada pois, mais do que qualquer coisa, é o apoio dos entes amados que possibilita, em países como Brasil, o sucesso de um programa de mestrado strictu sensu

AGRADECIMENTOS

À minha mãe, Cláudia Sarno Mugnola, e à minha namorada, Elisa Franco Feitosa, que foram as primeiras a me apoiar a continuar estudando, evoluindo e a não me satisfazer apenas com os cinco anos da graduação. Elas sempre me incentivaram a querer mais, me lembrando de que posso cada vez mais.

Ao meu irmão Marcelo, meu pai Celso, avós Marilena, Marcos, Albano e Dora e aos tios Marco Cesar e Sérgio, por sempre me darem muito amor e proporcionarem uma maravilhosa estrutura emocional e financeira, e pela compreensão e paciência por eu assumir o risco de não tomar a decisão óbvia que seria ir ao mercado logo após me graduar engenheiro eletrônico.

Ao meu Professor e, antes de tudo, veterano de Poli, Dr. Marcio Lobo Netto, por saber me passar sua experiência como ex-aluno e pesquisador da unidade, me dar oportunidades, direcionar meu trabalho para que eu evitasse cair em armadilhas e saísse dos eixos, sanar minhas dúvidas de condução do projeto, me tranquilizar, me apoiar e, sobretudo, me cobrar.

Finalmente deixo agradecimentos especiais aos especialistas e engenheiros da CET-SP, Luis, Flávio, Sérgio e Alexandre, pelas orientações dadas na concepção do modelo de simulação e pelo apoio dado durante a realização do estudo de caso.

"The most extensive computation known has been conducted over the last billion years on a planet-wide scale: it is the evolution of life. The power of this computation is illustrated by the complexity and beauty of its crowning achievement, the human brain."

David Rogers, no artigo "Weather Prediction Using a Genetic Memory" (1990)

RESUMO

Nas grandes cidades ao redor do mundo os congestionamentos são problemas bem conhecidos e compartilhados por todos os setores da sociedade. Não é surpreendente, então, que grande parte dos investimentos públicos caminhem na direção de reduzir paulatinamente o esforço rotineiro que a população itinerante faz para chegar ao trabalho ou retornar ao lar, melhorando sua qualidade de vida. É partilhando desse intuito que se encontrou a motivação inicial para o desenvolvimento deste trabalho. Em outro pólo encontra-se a vertente dos procedimentos evolutivos que ao longo da história da vida na Terra foram responsáveis pela emergência espontânea de uma enorme gama de soluções para os mais variados ecossistemas. A tradução disso para ambientes computacionais criou a classe dos algoritmos evolucionários, dentre os quais os Algoritmos Genéticos (AGs), que se destacaram por serem boas heurísticas de busca por conjuntos de parâmetros que resultem em ótimos globais para problemas de engenharia. A aplicação de AGs para otimizações em engenharia de tráfego possui boa base bibliográfica, mas são raras as aplicações reais em meios onde há escassez de dados e de Sistemas Inteligentes de Transporte (do inglês, *Intelligent Transportation Systems – ITS*). Neste trabalho então foi desenvolvido um novo modelo de simulação mesoscópica sobre o qual um AG é aplicado para encontrar planos semafóricos que reduzam atrasos e paradas em sub-redes congestionadas. A ferramenta é simplificada para execução rápida, usando parâmetros normalmente colhidos pela Companhia de Engenharia de Tráfego de São Paulo (CET-SP) em estudos de revisão de temporização semafórica. Ao fim do trabalho, o estudo de caso em uma sub-rede paulistana resultou em reduções da ordem de 30% no nível de atraso e paradas em relação aos valores obtidos com a simulação dos planos anteriores. Vale ressaltar que o espaço de busca foi reduzido ao sub espaço de planos aceitos pela experiência dos especialistas da CET-SP, e mesmo dentro deste escopo, o algoritmo foi bem sucedido ao descartar soluções ruins e fazer emergirem soluções ótimas coerentes.

Palavras-chave: Simulação de Trânsito. Algoritmos Genéticos. Sinalização Semafórica. Redução de Atrasos. Alinhamento à Prática (CET-SP).

ABSTRACT

In large cities throughout the world, high traffic congestion is a well known problem endured by all parts of society. Not surprisingly, a great deal of public investments is made in the direction of reducing the citizens' daily efforts for going from home to work and backwards. It is by sharing this intent that the motivation for this thesis was found. Meanwhile, in another section of human knowledge there are concepts revolving the evolutionary procedures that were responsible for the spontaneous emergence of an immeasurable quantity of solutions for adaptation to an almost as greater quantity of ecosystems. When science brought these concepts to computational environments the class of evolutionary algorithms was born, a class embodied by the Genetics Algorithms (in this text referred as AGs, for *Algoritmos Genéticos*, in portuguese) which are heuristics that stand out as good alternatives for searching parameter settings that result in global optima for engineering problems. There is a good knowledge base regarding the use of AGs for traffic engineering optimizations, but rare are the real implementations in which specialists have to deal with lack of data and the absence of Intelligent Transportation Systems (ITS). Therefore, in this thesis a new mesoscopic simulation model was developed over which an AG is applied for finding signal timing plans that reduce stops and delays in congested subnetworks. The prototyped tool is simplified for quicker execution, using data that is normally collected by the Traffic Engineering Company of São Paulo (CET-SP) in signal timing revision activities. After the development, one of the numerous city's subnetworks was adopted as the case study, for which the prototype found plans that reduced the stops and delays in approximately 30% when compared to the values measured with the simulation of the old plans. It is worth notice that the search space was reduced to the subspace that only contains solutions accepted by the experience of CET-SP's traffic signal specialists, and within this subgroup, the algorithm succeeded in discarding the bad solutions and providing means for the emergence of coherent global optima.

Keywords: Traffic simulation, Genetic Algorithms. Signal Coordination. Total Delay Reduction. Practical Approach (CET-SP).

LISTA DE ABREVIATURAS E SIGLAS

AG	Algoritmo genético
CET-SP	Companhia de Engenharia de Tráfego de São Paulo
ITS	Intelligent Transportation Systems
EPUSP	Escola Politécnica da Universidade de São Paulo

GLOSSÁRIO

Linha de retenção: listra branca pintada antes da faixa de pedestres, ou simplesmente o lugar onde os veículos param quando o farol está vermelho, para o cruzamento de pedestres ou para dar a preferência a outros veículos

Intersecção/Nó: cruzamento

Segmento: pedaço de rua que compreende uma face de um quarteirão, ou em outras palavras, a seção de rua que vai de um cruzamento qualquer até o mais próximo (em caso de mão dupla, há um segmento para cada mão)

Segmento à Montante: como aquilo que se encontra acima no que concerne o fluxo de um rio, por exemplo, é o segmento localizado antes do ponto de observação ou referência. Usa-se o termo aqui para referenciar qualquer segmento que passa veículos para um segmento observado através de um cruzamento

Segmento à Jusante: como aquilo que se encontra abaixo no que concerne o fluxo de um rio, por exemplo, é o segmento localizado após o ponto de observação ou referência. Usa-se o termo aqui para referenciar qualquer segmento que recebe veículos de um segmento observado, ou o semáforo por onde estes passam

Segmento Concorrente: segmento montante a uma intersecção que apresenta movimento conflitante com algum movimento do segmento observado

Escoamento: saída de quantidades de veículos pela linha de retenção de um segmento

Pelotão: agrupamento de veículos que estejam muito próximos

Demanda: Contingente de veículos que passa por uma rede de tráfego

Oferta: Conjunto formado pela infraestrutura de transporte disponível em uma rede urbana

Plano Semafórico: Perfil de temporização que determina o ciclo de chaveamento de um semáforo entre sinal vermelho e sinal verde

Passo de simulação: ciclo em que todos os cálculos referentes a um único momento de simulação são feitos

Cromossomo: Código que parametriza uma solução candidata em um AG

Fitness: Função objetivo/índice de mérito que avalia soluções candidatas em um AG

Otimização: Procedimento que busca boas conformações paramétricas ou topológicas para problemas de engenharia para executar melhorias através de minimizações ou maximizações

SUMÁRIO

1. INTRODUÇÃO	15
1.1 Motivação.....	17
1.2 Estrutura da Dissertação.....	19
2 SIMULADORES DE TRÂNSITO	21
2.1 Conceitos Gerais.....	21
2.2 Tipos de Simuladores.....	23
2.2.1 Simuladores Macroscópicos.....	26
2.2.2 Simuladores Microscópicos	27
2.2.3 Simuladores Mesoscópicos	29
2.3 Coletas Manuais de Dados para Simulações.....	31
3 MÉTODOS DE OTIMIZAÇÃO.....	34
3.1 Tipos de Heurísticas.....	34
3.1.1 Algoritmos Estocásticos.....	34
3.1.2 Algoritmos Físicos	34
3.1.3 Algoritmos de Enxame (<i>Swarm Algorithms</i>)	35
3.2 Algoritmos Evolucionários	36
3.2.1 Procedimentos Evolutivos	36
3.2.2 Introdução ao Algoritmo Genético	37
4 AGs APLICADOS À ENGENHARIA DE TRÁFEGO	43
5 O SIMULADOR GENPOLIS.....	48
5.1 Porque Fazer um Simulador Dedicado	48
5.2 Resumo Descritivo do Modelo	49
5.3 Arquitetura da Ferramenta	50
5.3.1 City Engine e OpenStreetMaps	51
5.3.2 Construtor	53
5.3.3 Simulador	54
5.3.4 Algoritmo Genético	55
5.3.5 Fluxograma da Ferramenta	55

5.4	Detalhes do Funcionamento (Modelo de Simulação).....	56
5.4.1	Movimentação dos Veículos Pelo Segmento	56
5.4.2	Servidor de Filas.....	57
5.4.3	Cálculo Explícito de Filas e de Bloqueio de Segmento.....	59
5.4.4	Manipulação de Demanda.....	66
5.4.5	Programação Semafórica	69
5.4.6	Algoritmo Genético	73
5.4.7	Testes Iniciais de Funcionamento	77
5.4.8	Teste de Simulação Microscópica	78
6	ESTUDO DE CASO	81
6.1	Rede Escolhida	81
6.2	Coleta de Dados	83
6.3	Calibragem.....	87
6.4	Otimizações	88
6.5	Resultados	91
7	CONCLUSÕES	93
7.1	Futuros Desenvolvimentos.....	94
	REFERÊNCIAS.....	96
	REFERÊNCIAS COMPLEMENTARES.....	102
	Apêndice A: Controle de velocidade de um modelo mesoscópico	104
	Apêndice B: Dados coletados e parâmetros do estudo de caso	106
	Apêndice C:	112

LISTA DE FIGURAS

Figura 1: Gráfico de nível de congestionamento medido por porcentagem de vias monitoradas com lentidão, retirado em 03/2011, do portal da CET-SP (http://cetsp1.cetsp.com.br/monitransmapa/agora/graficolimite.asp)	22
Figura 2: Panorama mundial da simulação de tráfego	25
Figura 3: Curva de evolução em função do tempo quando há escolha de parceiros para reprodução (azul) e quando não há essa escolha (verde)	37
Figura 4: Processo de recombinação (<i>crossover</i>) com mutação em um AG	40
Figura 5: Fases de um Algoritmo Genético. Busca e Descobrimento / Exploração (esquerda), Clusterização (centro), Mineração ou Garimpo / Ajuste (direita)	41
Figura 6: Curva de adaptação em sistemas com um agente face à curva de adaptação em sistemas multiagente	42
Figura 7: Conceito inicial de rede viária para o simulador	51
Figura 8: Captura de tela do recurso de exportação de redes viárias do site www.openstreetmap.org	52
Figura 9: Captura de tela do software City Engine para planejamento urbano, num momento em que uma rede extraída pelo openstreetmap é editada para ser simulada. A rede em questão é a região da Avenida Paulista, em São Paulo	53
Figura 10: Captura de tela apresentando a rede apresentada na Figura9, após edição no software City Engine e compilada no construtor do GenPolis	54
Figura 11: Captura de tela de simulação de teste. Valendo ressaltar que esta rede se assemelha somente geometricamente à região da Av.Paulista pois aqui os parâmetros utilizados para as vias foram todos padronizados, sendo que existem semáforos em todos os cruzamentos para se estressar o modelo	55
Figura 12: Fluxograma do GenPolis	56
Figura 13: Em um segmento/link, veículos em movimento (verdes)	57
Figura 14: Métodos de roteamento por i:porcentagens de conversão no segmento montante (esquerda) e ii:fluxo total dos segmentos à jusante (direita)	58
Figura 15: Exemplo de cálculo do fim da fila realizado em um passo de simulação	60
Figura 16: Exemplo de propagação de espaços com bloqueio e liberação de segmento	62
Figura 17: Detalhe legendado do ambiente de simulação	65
Figura 18: Perfis contínuo (vermelho) e discreto (azul) de inserção de demanda	66
Figura 19: Perfil de Fluxo em ambos sentidos de uma linha de retenção e suas respectivas cartas de tempos. Acima a linha azul se refere à carta de tempos A abaixo e a linha laranja se refere à carta de tempos B abaixo. As linhas contínuas se referem ao comportamento real enquanto as linhas pontilhadas se referem ao modelo com as devidas simplificações	71

Figura 20: Exemplo de cromossomo usado pelo GenPolis na codificação de um plano de temporização semaforica para o estudo de caso	74
Figura 21: Exemplos de corte duplo circular para cruzamento de cromossomos, com uma mutação no terceiro caso (bit em vermelho)	75
Figura 22: Perfis de inserção de veículos em veículos equivalentes por segundo (acima e perfis de comprimento somado de filas em metros para cada perfil de demanda (abaixo/em vermelho).....	77
Figura 23: Captura de tela da simulação com modelo microscópico testado	80
Figura 24: Imagem de Satélite da Região Simulada	82
Figura 25: Mapa Correspondente da Região Simulada	82
Figura 26: Curvas de evolução do melhor fitness de cada geração em função do número de gerações. Foram realizadas seis otimizações (cores distintas) para cada nível de demanda (de 100 a 130% do esperado), cada qual com uma semente de pseudo-aleatoriedade distinta.....	89
Figura 27: Curvas de evolução do melhor fitness de cada geração em função do número de gerações (esquerda); Tabela com os <i>fitness</i> do plano escolhido e do plano aplicado (centro); Planos anterior e escolhido para implantação (direita).....	91

LISTA DE TABELAS

Tabela 1: Plano antigo da manhã (esq.) e plano atual otimizado (dir.)	90
Tabela 2: Desempenhos de planos otimizados e do anterior para pico da manhã ...	90

1. INTRODUÇÃO

Este trabalho tem por objeto o estudo de mecanismos de controle de tráfego urbano e a proposição de uma alternativa heurística para melhorar de modo generalista a temporização de sub-redes de semáforos que, de forma coordenada, melhorem o fluxo de tráfego nas respectivas sub malhas viárias de grandes centros urbanos. O objetivo é criar mecanismos de apoio que levem à diminuição dos atrasos de percurso causados principalmente por congestionamentos, reduzindo assim prejuízos tanto materiais quanto psicológicos e a emissão de poluentes. O trabalho se apoia num simulador de tráfego, desenvolvido com o propósito de permitir uma análise da interação dos planos de sinalização e seu impacto no desempenho do tráfego. Vale citar já aqui que esta proposta deve ser vista como complementar a outras ações que tratem da realização de um planejamento urbano cada vez mais eficiente.

O interesse em uma ação operacional como a coordenação da operação de semáforos em rede decorre da dimensão e complexidade que o problema assume em grandes áreas urbanas e do fato de ser bastante estudado. No que concerne à temporização de uma rede sincronizada de semáforos, existem três parâmetros principais a serem configurados: tempo de ciclo, tempos de verdes, e as defasagens que determinam o instante de início de funcionamento entre os ciclos de cada semáforo (em relação a um semáforo de referência).

Para planos de temporização fixa, quando se leva em conta apenas o comportamento de uma única intersecção, a determinação dos dois primeiros parâmetros envolve um processo relativamente simples. Entretanto, mesmo esse procedimento faz uso de dados cuja atribuição de valores é baseada largamente na experiência dos órgãos operadores de tráfego com o comportamento temporal do trânsito, adquirida ao longo de anos realizando ciclos de aplicação e observação (VILANOVA, 2005a).

Quando a questão envolve o funcionamento de vários semáforos em uma rede sincronizada essa tarefa fica mais difícil, pois a interação é essencial e uma importante parte do trabalho reside em se determinar o correto conjunto de defasagens entre os diferentes semáforos. Pelo fato de controlarem o fluxo de diversos movimentos concorrentes, realizar um ajuste baseado apenas nos tempos

de percurso de cada via de forma isolada não é suficiente, pois cada um de seus trechos se comportará distintamente devido à influência de outras vias com as quais tiver conexão, tanto daquelas que a alimentam, quanto das outras para as quais vão veículos que por ela passaram. Naturalmente vale lembrar a validade da reciprocidade, ou seja, trata-se de um problema que só pode ser resolvido do ponto de vista sistêmico (e não individualizado). Além disso, no caso de malhas viárias a interação entre seus elementos não permite uma descrição matemática formal favorecedora do uso de métodos analíticos e ao mesmo tempo envolve explosões combinatórias que criam um espaço de buscas vasto que não permite uma busca exaustiva pela solução ótima global (LEBDEH, 1997).

Em função da característica mencionada, os métodos clássicos de coordenação semafórica em rede (como o do TRANSYT-7F, em Hale (2008)) não utilizam algoritmos de otimização convencionais e não representam a interação de forma analítica, usando ao invés disto simulação de tráfego. Nessas ferramentas, métodos de busca de ótimos locais foram inicialmente aplicados de forma a ajustar a sinalização a partir de configurações baseadas nas observações, no conhecimento e na insubstituível experiência dos engenheiros de tráfego. Um exemplo importante desses métodos complementares é o hill-climbing (HC) (BROWNLEE, 2011; CANTARELLA et al., 2004) que possui em sua formulação o compromisso de encontrar o ótimo local pertencente à região do plano de busca que compreende a solução inicial dada como entrada. Mais recentemente, meta-heurísticas de procura global foram propostas e vêm sendo usadas com sucesso para otimização de sinalização e uso das vias (revisão tratada mais à frente) e outros problemas de engenharia (não observados neste trabalho) que envolvam explosões combinatórias. Para uma descrição formal de muitos deles, recomenda-se Brownlee (2011).

Portanto, para atacar tal problema, foi proposta uma nova ferramenta de estudo de trânsito denominada GenPolis. Tal proposta incorpora tarefas de otimização de sinalização semafórica através da heurística evolutiva de algoritmos genéticos (MITCHELL, 1995), referenciados pela sigla AGs. Esse método é acoplado a um simulador de tráfego simplificado e rápido que usa, para estabelecer o comportamento do trânsito, apenas os parâmetros que são contemplados na realização de estudos para revisão de planos semafóricos na cidade São Paulo (EJZENBERG, 2005; VILANOVA, 2005a) cuja malha viária contém a sub-rede que

foi o estudo de caso retratado ao final deste trabalho. Devido a isso, a ferramenta possui as devidas simplificações para a realização em tempo hábil das numerosas (cite-se milhares de) simulações requeridas. Assim acredita-se, mas não se garante, ser possível abreviar o trabalho exaustivo que reside na tentativa de encontrar boas soluções para realizar melhoras no fluxo de tráfego. De qualquer maneira os resultados alcançados demonstram que houve de fato melhoria nos tempos de percurso e redução dos congestionamentos. Por consequência, neste trabalho pretende-se apresentar um panorama experimental acerca da aplicabilidade de AGs para a realização de melhorias no trânsito de cidades como São Paulo, visto que os contínuos aprimoramentos da infra-estrutura em malhas viárias não conseguem acompanhar o crescimento da demanda imposta pela frota circulante.

1.1 Motivação

A magnitude dos problemas cotidianos da operação do trânsito em grandes cidades é um tema recorrente na imprensa diária. Em dados retirados de Giovanelli, (2011) temos que só na cidade de São Paulo a frota oficial é de aproximadamente 7 milhões de veículos, sendo que, destes, 3,8 milhões circulam todos os dias pela capital. Lidando com essa quantidade de veículos, a malha viária da cidade, com cerca de 17.000 quilômetros de extensão, registra oficialmente uma velocidade média de circulação de 32 quilômetros por hora no pico da manhã (07h00 às 10h00) e 18 quilômetros por hora no pico vespertino (17h00 às 20h00), com trechos críticos em que essa leitura pode chegar a 8 quilômetros por hora. Toda essa lentidão se deve a essa elevada quantidade de veículos, que na sua grande maioria circula com apenas um ocupante, e leva à consolidação de um prejuízo da ordem de 30 bilhões de reais anuais em produtividade perdida e gastos materiais, principalmente, com o combustível desperdiçado.

Para solução desses problemas e redução dessas perdas, as políticas governamentais e os órgãos de monitoramento e controle agem tanto no sentido de aumentar a infra-estrutura viária e melhorar o seu funcionamento quanto no intuito de controlar a frota de veículos que usa essa estrutura. Em Lopes (2003) são descritos inúmeros tipos de iniciativas relevantes.

No que se refere às melhorias infra-estruturais temos as iniciativas mais evidentes que seriam as obras e reformas voltadas tanto para o transporte privado

quanto para o transporte público, feitas primeiramente no intuito de expandir a capacidade da malha viária e metroviária existente. Enquanto isso, outras ações complementares tratam de viabilizar e otimizar o uso das vias, e envolvem basicamente regulamentações (como a criação de faixas exclusivas, definição dos sentidos de circulação das vias, da quantidade e disponibilidade horária das vagas de estacionamento nas ruas) e atividades de controle (que se dão na forma de sistemas de monitoração e de atuação como as que envolvam sistemas de ajuste de sincronização semafórica, uso de painéis de mensagem variável e câmeras de fiscalização).

Ainda segundo Lopes (2003), foram levantados pontos importantes relativos às principais limitações dessas medidas:

- Impacto reduzido no longo prazo (aumento da oferta atrai demanda de forma recorrente, efeito que faz as vias expandidas voltarem a operar de forma congestionada após algum tempo da implantação);
- Efeito de barreira causado por limitações impostas à circulação, principalmente durante a construção de grandes vias expressas.

Além destes há o problema crucial das deficiências no direcionamento e no planejamento dos investimentos em infraestrutura de transportes, comumente confundida com escassez de recursos financeiros.

Outro campo da gestão de trânsito diz respeito à administração da demanda, particularmente da frota, onde temos medidas que agem com o objetivo principal de reduzir ou redistribuir a frota geográfica e temporalmente. Entre as medidas principais estão as restritivas, como o rodízio de veículos ou a cobrança de pedágios urbanos (ou aumentos nos preços de combustíveis), e as de incentivo, como os relacionados com o escalonamento de horários das atividades ou à flexibilização dos horários de expediente de forma que trabalhadores possam sair às ruas em horários fora dos períodos de pico.

Nos últimos anos testemunhamos a implantação de algumas ferramentas de suporte ao controle de tráfego, como câmeras com processamento de imagens que são conectadas a sistemas centralizados que realizam checagem de documentação, radares para aferição de velocidade média de vias importantes e sensores como laços indutivos (BIKOWITZ, 1985) que são dispositivos eletromagnéticos enterrados no pavimento responsáveis por registrar a passagem de veículos e auxiliar o

controle ativo por meio de semáforos ajustados em tempo real (“inteligentes”) por sistemas centralizados dotados de métodos como o SCOOT (ROBERTSON; BRETHERTON, 1991). Em São Paulo, a maior cidade do país, obstáculos variados impedem que recursos estejam plenamente disponíveis ou, quando disponíveis, facilmente acessíveis e manipuláveis para os operadores de tráfego. Vale aqui a ressalva de que a discussão dos exatos fatores não cabe ao escopo desta dissertação ou à responsabilidade do pesquisador.

Adicione-se a isso o fato de que muitas medidas têm seus efeitos reduzidos no longo prazo, pois como já brevemente citado, o comportamento da frota é constantemente alterado no sentido de que uma vez que novas estruturas passam a facilitar o fluxo de tráfego em determinada região urbana, esta passa a atrair mais usuários do que antes, criando uma nova situação de equilíbrio que poderá usar toda a nova capacidade viária (LOPES, 2003). Assim, algo que se pode fazer de forma mais ágil para adaptar a resposta das malhas viárias às constantes mudanças, redistribuições de demanda e consequentes novas situações de equilíbrio, é o ajuste frequente (pelos próximos meses, ou anos, dependendo da frequência de reajuste adotada pelo órgão responsável) dos planos de sincronização semafórica, ou seja, uma das partes mais simples e das mais relevantes dentro do controle de tráfego, que é dizer a cada semáforo quando ele vai fechar e quando ele vai abrir.

Pode-se ver que muitas ações são custosas ou impopulares, o que torna a melhoria da gestão viária um objetivo importante de atuação, ponderando a possibilidade de garantir a eficácia de cada tipo de ação. Logo, devido à necessidade de um controle que contemple o comportamento dinâmico das relações entre a malha viária e a frota ao longo dos anos e considerando-se as restrições tecnológicas e de políticas de implementação dos controladores, foi pensado que, partindo-se da criação de um novo simulador de trânsito, o controle semafórico poderia ser o problema a ser o tema tratado neste mestrado.

1.2 Estrutura da Dissertação

Nesta dissertação, o leitor já encontrou uma breve introdução sobre conceitos relacionados à engenharia de tráfego (mais idéias serão abordadas ao longo do texto, mas de forma difusa) e obterá conhecimento introdutório relativo a um rápido

panorama sobre o contexto da simulação de tráfego no mundo e as diferentes abordagens computacionais, bem como alguns exemplos mais relevantes.

Em seguida ao leitor serão apresentados bem rapidamente diversos métodos estatísticos e heurísticos de otimização (com maior ênfase dada aos últimos) utilizados atualmente nos mais diversos problemas de engenharia, e receberá uma maior ambientação na parte que abrange os AGs, heurística escolhida para este trabalho.

Mais a frente as duas partes anteriores se fundem em um capítulo que apresenta o leitor ao ambiente de AGs aplicados à simulação de trânsito, e a progressão do uso destes métodos ao longo, aproximadamente, dos últimos quinze anos.

O anti-penúltimo capítulo tem por motivo apresentar o leitor à ferramenta criada ao longo do mestrado, sendo exposta a sua arquitetura e diagrama de blocos, o seu modo de funcionamento, e mais detalhadamente os conceitos e lógicas aplicados ao seu motor de simulação.

Antes de terminar, será mostrado o estudo de caso realizado para testar a ferramenta prototipada, com exposição de resultados e comparações que levarão aos indicativos dos próximos desenvolvimentos em uma eventual continuidade do projeto. Por fim é apresentada a conclusão do trabalho.

Um CD-ROM segue anexado à obra contendo arquivos para a reprodução/processamento desse estudo de caso da seção 6, e com arquivos relativos à prototipagem para que, no caso de haver maior interesse, ao leitor seja aberta a manipulação do modelo e permitida a reprodução do experimento.

2 SIMULADORES DE TRÂNSITO

Neste capítulo pretende-se dar ao leitor um panorama acerca das iniciativas de criação e evolução dos modelos e simuladores de trânsito ao redor do mundo, bem como apresentar conceitos gerais e específicos sobre o tema.

2.1 *Conceitos Gerais*

Nesta parte serão explicados os conceitos básicos e presentes na maioria dos simuladores de trânsito para maior entendimento durante a leitura do resto do texto. De início é importante mencionar o que significa exatamente simular trânsito, pois este também pode ser analisado diretamente sob a vista de processos analíticos.

Como foi visto em Burghout (2004) para se realizar estudos de tráfego existem dois caminhos principais a serem usados. Por um lado existem métodos analíticos baseados em processos numéricos que comumente fazem uso de conjuntos de equações diferenciais para encontrar os resultados de um problema de tráfego pontual. Como exemplo disso podemos citar o estudo de um único corredor de via expressa, que por sua relativa simplicidade, necessita de pouco poder computacional. Por outro lado temos a simulação, que envolve um grande conjunto de descrições mais elementares dos ambientes alvo de estudo para que se possa acompanhar o comportamento dinâmico de um sistema de tráfego ao longo do tempo. Essa descrição tem a vantagem de servir para observar comportamentos, principalmente emergentes, que sejam consequência do encadeamento complexo das diversas diretrizes que regem seus elementos. Mas por outro lado requer um poder computacional muito maior. Nas últimas décadas houve um grande aumento de soluções relacionadas com simulação de trânsito, devido aos avanços no poder computacional, com o que a capacidade exigida para tais simulações passou a estar disponível a um menor custo, até então um impeditivo para a simulação.

Em uma simulação de trânsito temos basicamente dois lados que são descritos das mais diversas maneiras, o lado da oferta, composto pelos elementos virtuais correspondentes à infra-estrutura viária, como as próprias vias e seus sistemas de controle e informação, e o lado da demanda, composto pelos agentes navegantes virtuais que representam veículos e seus condutores.

Classicamente, os ambientes de simulação muitas vezes envolveram análises de curtos períodos de tempo (cerca de quinze minutos) que correspondem a situações de oferta e demanda muito pontuais. Essa falta de abrangência dificulta a compreensão do comportamento das malhas viárias como sistemas dinâmicos, pois não contempla as relevantes variações que podem ocorrer em longos intervalos de até algumas horas, e os efeitos decorrentes do carregamento ou descarregamento da rede prévios ao horário da janela de tempo estudada. Logo ferramentas mais novas lidam com longos períodos de simulação que apresentem situações de demanda temporalmente variáveis, sendo essa uma forma de traduzir mais fielmente o comportamento dinâmico de redes metropolitanas, evidenciado pelas leituras de nível de congestionamento diárias como a representada abaixo, relativa à cidade de São Paulo. Aqui podemos ver que a janela de tempo de uma simulação que contemple todos os efeitos de carregamento, saturação (estado em que a demanda supera a oferta) e descarregamento de uma grande malha viária pode ter a dimensão de aproximadamente quatro horas, contenha ela o horário de pico da manhã ou o horário de pico da noite.

Figura 1: Gráfico de nível de congestionamento medido por porcentagem de vias monitoradas com lentidão, retirado em 03/2011, do portal da CET-SP
(<http://cetsp1.cetsp.com.br/monitransmapa/agora/graficolimite.asp>)

Tradicionalmente, em processos de simulação estas demandas são geradas através da modelagem de matrizes de origem-destino (OD) que governam a escolha das rotas dos veículos ou fluxos simulados partindo-se de pólos geradores de tráfego, sejam eles de liberação de veículos, de recebimento de veículos ou ambos e em gradações distintas, dependendo do horário do dia a ser estudado. Esses pólos são estimados estudando-se o perfil socioeconômico de cada região da rede estudada, através da realização de séries de pesquisas de origem e destino em importantes pontos de acesso às malhas de estudo, que visam amostrar os pontos iniciais e finais dos trajetos percorridos na rede, e também a partir de contagens de fluxo de tráfego em pontos relevantes da rede (CASCETTA, 1988). No caso de São Paulo, um método utilizado para revisão de planos semaforicos é mais pontual e leva em conta a medida direta de fluxo de tráfego em todos os segmentos de cada sub-rede estudada (EJZENBERG, 2005. VILANOVA, 2005).

2.2 Tipos de Simuladores

Como em qualquer outro tópico de simulação, há divergências quanto ao que se julga devido representar ou não em um simulador, e é natural então que se encontre uma grande variedade de modelos, cada um voltado para um tipo de resultado. Logo, o que se aceita é que existem numerosos meios de se representar computacionalmente tanto os lados da oferta quanto da demanda e os modos como eles se relacionam (BURGHOUT, 2004). Isto se deve ao fato de que é virtualmente impossível tratar, em um único ambiente, de todos os aspectos físicos e, mais dificilmente, os sociais, que afetam o comportamento de um sistema de tráfego. Em decorrência disto a arquitetura de cada uma das ferramentas desenvolvidas serve a propósitos distintos.

Quando decide criar um novo modelo de simulação, o especialista se depara com a realidade de que, por motivos de dificuldade de codificação, escassez de modelos e restrições de tempo e de poder computacional, é impossível tratar de todos os detalhes envolvidos na retratação fiel da realidade dentro de um computador. Por consequência o projetista deve determinar qual o foco de seu estudo. Após essa escolha, é necessário determinar quais dados da realidade se

deseja parametrizar para que a simulação seja funcional e traga o tipo de leitura esperada para esse estudo.

Devido a isso, ao longo do tempo, naturalmente surgiram diversos tipos de abordagem que, por consequência, levaram ao estabelecimento de convenções e consensos conceituais necessários para que os especialistas de controle de tráfego e simulação possam mais facilmente identificar, correta e rapidamente, o intuito por trás de um ambiente de simulação, bem como o tipo de resultado que se pode esperar do mesmo.

Predominantemente isso diz respeito a classificar cada modelo de acordo com o nível e o tipo de detalhamento, e o grau de agregação ou desagregação da oferta e da demanda. Do lado da demanda, em um sistema de tráfego, por exemplo, isso diz respeito a se os veículos vão ser individualizados e se eles possuem alguma capacidade de percepção e decisão sobre sua movimentação. Do lado da oferta, a classificação depende de quais variáveis da movimentação de um veículo serão consideradas, da modelagem ou não de diversos tipos de sinalização e sistemas de informação e gestão de tráfego, ou da escolha por diferentes formas de se parametrizar uma via e suas faixas (BOXILL; YU, 2000 e BURGHOUT, 2004). Essa classificação, geralmente é dicotomizada entre simuladores microscópicos e simuladores macroscópicos.

Vale a colocação de que muitas vezes, mesmo sob essa ótica, as regras não são, e nem devem ser, tão claras, de modo as classificações são nebulosas, flexíveis e não é incomum que sejam desenvolvidos modelos híbridos entre as categorias que serão explicadas mais à frente nesta seção. Estes modelos híbridos normalmente visam aumento de desempenho computacional por meio de simplificações ao mesmo tempo em que mantém o detalhamento nos aspectos ou porções da simulação onde cada atitude for mais necessária, e podem então apresentar a mistura dos conceitos de duas formas diferentes. Primeiramente temos simuladores possuindo um único modelo de simulação que siga um conjunto de regras provindas de mais de um paradigma (BOURREL; LESORT, 2007), um tipo de abordagem que muitas vezes é denominado de simulação mesoscópica (que é o caso do simulador deste trabalho), ou programas onde em um único ambiente são usados vários modelos ao mesmo tempo em sub-redes diferentes (YANG;

MORGAN, 2006) ou na mesma rede, mas em momentos diferentes (HALE, 2008), de forma que os resultados da execução de um paradigma em um momento, etapa ou camada da simulação podem ser usados para alimentar ou calibrar o outro modelo, e vice-versa.

Finalizando, abaixo, na Figura 2 podemos ver um mapa que foi elaborado para que haja um panorama gráfico aproximado da simulação de tráfego no mundo. O mapa foi compilado a partir de toda a bibliografia lida e é em boa parte devido ao artigo de Boxill e Yu (2000) que fornece a leitura mais abrangente encontrada na área e que apresenta foco na análise de boas ferramentas que auxiliem a análise de sistemas ITS.

Figura 2: Panorama mundial da simulação de tráfego

Rapidamente, se o leitor parar para refletir, lembrará que a grande São Paulo é atualmente a quarta maior aglomeração urbana do mundo e talvez se surpreenda com o fato de que, apesar de podermos pressupor que haja aqui uma grande expertise contida em seus especialistas e lembrando que para desenvolvimento e aquisição de software não existem fortes barreiras de entrada, as iniciativas na área são poucas e recentes (SIRI de VILANOVA, 2009; e ITSUMO, de BAZZAN et al., 2010). Enquanto isso países como a Holanda e Alemanha, que nem figuram na lista de países com as maiores metrópoles possuem uma grande e incomparável concentração de modelos dos mais variados tipos.

2.2.1 Simuladores Macroscópicos

Estes são os modelos mais antigos e os que mais se assemelham a soluções analíticas por tratarem os veículos de forma altamente agregada, pois fazem uso geralmente de conjuntos de equações diferenciais que traduzem o fluxo destes em uma forma contínua. Dependendo do modelo, esses simuladores podem ter seu comportamento semelhante a processos físicos como aqueles envolvidos em mecânica de fluídos (HELBING; TREIBER, 1998), abstraindo desse modo a identidade de cada veículo e também eventualmente elementos topológicos da via como número e largura de faixas. Enquanto isso existem outros modelos que se assemelham a autômatas celulares (DAGANZO, 2004), onde os segmentos são divididos em diversas células, ou subsegmentos, que transmitem a cada segundo quantidades maiores ou menores de veículos para as células à jusante seguindo funções de densidade espacial.

Independente da abordagem, normalmente o intuito dos modelos macroscópicos é retratar, de forma estática, o equilíbrio que ocorre entre a oferta da infraestrutura viária e a demanda da frota circulante para que o planejador de transportes consiga analisar o volume de tráfego em cada via e quais são os trajetos percorridos na malha simulada.

Este tipo de simulador oferece algumas vantagens provindas da simplicidade trazida ao se representar o movimento, ou viagens, de diversos veículos de forma agregada. A principal vantagem dos simuladores macroscópicos reside no fato de que as pesquisas de origem e destino que mapeiam os itinerários nas malhas viárias estudadas, os modelos de geração de viagem e os dados de fluxo de tráfego usados para alimentá-los são mais facilmente obteníveis devido à ampla disponibilidade dada pelos sistemas de monitoramento (BURGHOUT, 2004) e por tradicionais procedimentos de coleta de dados. Além disso, o tratamento destes dados se limita à qualidade das amostragem e não envolve a necessidade de se contemplar detalhados fatores culturais comportamentais dos motoristas para inserção no modelo.

Por outro lado pela carência de detalhes comportamentais, esta categoria demonstra limitações no momento em que é necessário realizar estudos mais

específicos e que sejam dependentes dos elementos abstraídos (em sua maior parte as interações entre os veículos), e que sirvam para observar efeitos comportamentais em trechos da malha viária como pontos de convergência, bifurcações, pistas de acesso, inclinações e curvas com raios variados (BURGHOUT, 2004). Em casos como esses a modelagem é explícita, traduzida em reduções de capacidade das vias ou atrasos inerentes.

Como exemplos importantes, temos o TRANSYT-7F (HALE, 2008) que foi concebido especificamente para otimização de sinalização semafórica em redes americanas (geometria predominante definida entre ruas de sentido Norte-Sul e ruas de sentido Leste-Oeste) e que é o mais frequentemente usado para realização de benchmark de aplicações análogas, o EMME (©INRO, Canada) e o VISUM (©PTV-AG, Alemanha), ferramenta comercial completa de estimativa de fluxos de tráfego, direcionada para realização de estudo e planejamento das malhas viárias e roteamento de sistemas de transporte público em cidades inteiras. Bem como outros simuladores macroscópicos o VISUM faz uso de procedimentos iterativos variados para determinação do equilíbrio em horários de pico entre a demanda de tráfego e a oferta viária simuladas.

2.2.2 Simuladores Microscópicos

No outro extremo temos estes simuladores em que normalmente a intenção é a de atacar problemas mais pontuais em sistemas de tráfego, começando por tratar cada veículo ou partícula de forma independente e representá-lo como um agente ativo no sistema, e assim permitir inclusive retratar seus traços comportamentais, como agressividade / tranquilidade dos motoristas. Tais atributos podem ser parametrizados, e a partir deles os agentes reagem de diferentes formas aos diversos sinalizadores contidos na infraestrutura bem como aos outros veículos. Juntamente a isso, a lógica de movimentação por sua vez segue algoritmos bem definidos que calculam aceleração e desaceleração do veículo e, em alguns casos, inclusive a troca de faixas a cada passo de simulação. Normalmente esses modelos são denominados pela expressão “carro-seguidor” (ROTHERY, 1992 e ZHANG, 2005) pelo fato de as decisões de cada veículo virtual serem baseadas primeiramente na posição, velocidade e aceleração do veículo à frente ou na

distância que resta até a próxima parada e, em alguns casos, baseadas de forma complementar no comportamento dos veículos das outras faixas.

Do lado da oferta, ou da infraestrutura, é possível configurar diversos parâmetros topológicos das vias, como largura de faixas, número delas, sinalizações e regulamentações (como proibições de ultrapassagem ou troca de faixas) e elementos geométricos como detalhes sobre curvatura que influem na movimentação do veículo. Além disso, este tipo de simulação pode ser integrado a outros tipos de ferramentas de controle de tráfego, incluindo avançados sistemas de integração viária que permitem leituras online sobre o status da rede e que influem nos processos decisórios sobre caminhos a serem seguidos por parte dos veículos simulados (BOXILL; YU, 2000).

Então, estes simuladores possuem a grande vantagem de permitir uma visualização fiel dos detalhes do comportamento da rede, mas, por requererem um custo computacional alto para poder levar em consideração um elevado número de fatores de infraestrutura e comportamentais, e ainda por cima estabelecer inter-relações detalhadas entre os veículos da rede, é mais usual que sejam abordadas situações de demanda estática e em redes limitadas, o que não compromete seu propósito. Frequentemente estas arquiteturas são alimentadas com dados provenientes de simulações macroscópicas, separando-se uma pequena parte de interesse da malha viária simulada, para que se possa obter uma representação mais transparente daquilo que foi computado, a saber, os tamanhos das filas e a velocidade dos veículos.

Assim, como nos simuladores macroscópicos, o grande problema da família de ambientes microscópicos está justamente atrelado ao que lhe confere suas vantagens, pois neste caso o problema geralmente está na dificuldade em se conseguir os dados necessários para calibrar todos os detalhes. Coletar informações mais específicas como taxas de troca de faixa, distanciamento mínimo entre veículos, estabelecer distribuições estocásticas de comportamento e ainda repetir toda essa parametrização para cada tipo de veículo a mais a ser incluído na simulação envolve custos muito altos para que as simulações apresentem fidelidade e assim possam ser realmente comparadas com cenários reais (BURGHOUT, 2004).

Como exemplos importantes temos o CORSIM (analisado em OWEN et al., 2000), que é o simulador atualmente utilizado para avaliar em maiores detalhes as redes otimizadas pelo TRANSYT-7F e o VISSIM (©PTV-AG, Alemanha; FELLENDORF, 1994) que, normalmente alimentado pelo VISUM, possui grande foco em avaliação de sistemas atuados e inteligentes, apresentando grande e versátil aplicabilidade para estudos relacionados à implementação de priorização semafórica do transporte público e mudanças infra-estruturais como criação de faixas de circulação exclusivas para o mesmo. No que interessa a maioria dos estudos relevantes em engenharia de tráfego, os dois simuladores têm seus desempenhos comparados a partir de um teste de sensibilidade em que diversos cenários de demanda são aplicados a uma mesma sub-rede viária, concluindo-se que, apesar de as abordagens serem um tanto diferentes, os dois aplicativos produzem resultados bem semelhantes (BLOOMBERG, DALE; 2007).

2.2.3 Simuladores Mesoscópicos

As ferramentas de simulação macroscópica normalmente são indicadas para planejamento estratégico, já que retratam o comportamento da rede como um todo e de forma estática. Elas passam ao especialista qual é o equilíbrio que há entre a malha viária e a sua frota de veículos. Entretanto, essa representação não permite a avaliação da malha viária como um sistema dinâmico onde há crescimento, dissipação e propagação de filas.

Na outra extremidade, a da simulação microscópica, por conta de inúmeros detalhes e interações, se mostra uma arquitetura computacionalmente pesada. Isso reflete diretamente na limitação do tamanho das sub redes viárias que pode representar, bem como nas curtas janelas de tempo analisadas.

Como meio de resolver a ausência de um comportamento dinâmico em uma arquitetura e a limitação espaço-temporal no escopo de análise da outra, foram desenvolvidos diversos modelos de simulação com abordagens intermediárias. Isso foi feito para que se pudesse ganhar abrangência, sem perder todos os detalhes, e assim manter uma alta velocidade de execução

Esses modelos seriam os mesoscópicos e seu intuito é normalmente o de representar a movimentação das frotas pelas malhas viárias calculando e propagando as filas que são geradas em cada um de seus links. Algumas plataformas de simulação já foram desenvolvidas e serão retratadas abaixo.

Um dos exemplos seria o DynaMIT (Akiva, 1998), um modelo onde a demanda é alocada de forma semelhante à dos modelos macroscópicos. Nele os links são quebrados em diversos segmentos, sendo que cada segmento tem seu comportamento dividido entre uma parte móvel e uma parte onde é simulado o enfileiramento dos veículos. A parte móvel possui velocidade fixa para todos os veículos até o ponto onde esta começa a decair linearmente devido à proximidade do veículo em movimento com o fim da fila. Ao mesmo tempo, na parte do enfileiramento, modelos de bloqueio e de formação e dissipação de filas são implementados de acordo com o estado da fila no segmento.

Outra plataforma é o Metropolis (Palma et. al, 2007) onde a movimentação dos veículos também é determinada de forma agregada em função da quantidade e densidade de veículos adiante no link. De forma semelhante há o DYNEMO (Nökel, 2002), e o Mezzo (Burghout, 2004) onde também são aplicadas equações de velocidade-densidade para reger o deslocamento dos veículos, bem como modelos para cálculos de filas.

Para simulação mesoscópica, vários modelos podem ser aplicados, tanto que, especificamente para esta classe de simuladores, há uma maior flexibilidade sobre a abordagem que deve ser tomada. A única restrição aparente se dá pelo fato de comumente seguir-se o conceito de que a movimentação dos carros deve ser feita de forma agregada e desvinculada de parâmetros comportamentais ou reativos individuais, implementações que fazem aumentar o custo computacional e o tempo de processamento em simuladores microscópicos.

Como já citado, o modelo de simulação retratado nesta obra é mesoscópico. Esta abordagem foi selecionada por permitir simplificações na prototipagem e por ser uma arquitetura rápida que viabilize o uso de algoritmos genéticos.

2.3 Coletas Manuais de Dados para Simulações

Nesta seção encontram-se referências e explicações acerca de práticas mais usuais da Companhia de Engenharia de Tráfego (CET) da cidade de São Paulo, as quais são consideradas no estudo de caso detalhado no capítulo 6.

Uma vez tendo-se elaborado um modelo de simulação coerente, uma das maiores dificuldades validação e utilização do simulador reside em se conseguir dados de entrada com alta qualidade para a realização dos estudos. Em Vilanova (2006) os especialistas afirmam que sem isso qualquer simulador estaria fadado ao insucesso. Normalmente isso está altamente relacionado com a disponibilidade de pessoal para realizar monitoramentos e contagens presenciais. Embora isso seja predominantemente terceirizado, recomenda-se que todo especialista participe em parte dos processos de observação para que ele mesmo ganhe alguma sensibilidade experimental mais fina no momento em que for necessário calibrar os dados, pois caso contrário haveria um distanciamento prejudicial entre o programador da simulação e a realidade.

No que concerne a elaboração deste trabalho existem três parâmetros experimentais, os principais das leituras de campo, que serviram de base para a construção do GenPolis:

Tempo de Percurso: O intervalo de tempo médio necessário para que um veículo percorra livremente o segmento de rua que está sendo parametrizado

Fluxo: É a quantidade de veículos-equivalentes (uma unidade padrão que denota, para fins de simulação, um carro de passeio médio) que passam pelo segmento por hora, sendo geralmente medido através de contagem direta e com observador fixo. Vale notar que para cada segmento é necessária a observação ao longo de alguns ciclos, ou cerca de quinze minutos, para que a medição possa abstrair as oscilações referentes às mudanças de fase. Assim, esta variável deve representar um fluxo homogêneo, abstraindo pequenas variações (em pequenos intervalos) que podem ser consideradas como ruídos, mas permitindo observar as variações de menor frequência (ao longo do dia) que representem diferentes perfis de trânsito e que portanto devem ser incorporadas a uma boa simulação.

Fluxo de Saturação: É o ritmo máximo de vazão medido em uma via, também dado em veículos equivalentes por hora (veq/h). Usualmente sua leitura envolve observar o comportamento do segmento após algum tempo do início de um verde

(cerca de cinco segundos), momento em que os veículos começam a ser liberados. Isso somente pode ser feito se os segmentos à jusante estão vazios o suficiente para receber livremente a demanda de veículos enviada pelo trecho monitorado.

Entretanto, a contagem por si só não é simples e envolve dificuldades relacionadas a alguns ajustes e adaptações para que a leitura sirva melhor aos estudos e simulações.

Como já dito, os fluxos são medidos em veículos-equivalentes. Isso se deve não só ao fato de ser necessário simplificar a medida meramente para fins de registro, mas também ao fato de que normalmente simuladores de trânsito não contemplam todos os inúmeros tipos de veículos que podemos encontrar e contar na rua. Logo, para que a medição fique mais precisa, tanto para a aferição do fluxo quanto do fluxo de saturação, é necessário o uso de um fator multiplicativo para ajustar a contagem a cada um dos tipos e subtipos específicos de veículos e transpô-los para veículos-equivalentes, sendo que esse fator também pode depender da inclinação física do segmento observado, se a facilidade de movimentação do veículo contado for afetada por ela (efeito mais comumente observado em veículos pesados).

Nesses casos, ao mesmo tempo em que um carro de passeio médio em uma rua plana tem valor unitário, alguns caminhões em atividades maiores podem sofrer a aplicação de fatores multiplicativos grandes para denotar quantos carros médios poderiam ter passado no seu lugar e ocupado a mesma capacidade da via. Vale ressaltar aqui também que motos não são contadas, ou consideradas (fator multiplicativo nulo), pela sua facilidade de locomoção e pelo fato de muitas vezes ocuparem os espaços disponíveis entre as faixas.

Além disso, no caso mais comum do foco do estudo e simulação ser em um período de pico de saturação da malha, uma dificuldade que surge é a de que a medição direta não é completa. Isso se deve ao fato de que, para períodos de tráfego intenso, ainda é necessário observar e estimar a demanda represada, que constitui a frota de veículos que deixou de passar pela linha de retenção devido a algum bloqueio mais a frente causado pela saturação da via.

Além disso há a demanda suprimida, que é impossível de ser aferida por ser constituída pelos veículos que deixaram de sair às ruas devido a notícias, previsões de congestionamentos ou uma simples sazonalidade que seja popularmente

conhecida. Esse contingente suprimido é representado principalmente pelos motoristas rotineiros dos horários de pico que ou saem mais cedo para o trabalho ou atrasam a sua volta pra casa no fim do dia para não ficarem presos no trânsito.

Vale ressaltar a preocupação que existe em relação à demanda suprimida, que é a de que, considerando determinada malha viária, na hipótese dos problemas críticos de fluxo de tráfego e atraso serem solucionados com o auxílio da simulação, este contingente deixaria de ser uma demanda suprimida, parcial ou totalmente, para de fato interar o fluxo de tráfego efetivo da região. Logo, na medida em que o fluxo aumenta e os atrasos/congestionamentos diminuíssem, ainda assim deve-se ter em mente que há o risco da rede voltar a ficar congestionada rapidamente devido à demanda que não teria sido considerada no ajuste.

Finalmente, respeitadas as sazonalidades (como diferentes épocas do ano, períodos de férias ou de natal) para fins de agrupamento e classificação das medidas, o processo de contagem para cada estudo deve ser feito de modo que uma leitura, amostra ou registro envolva diversas coletas de dados simultâneas, no mesmo horário, de todos os segmentos de rua contemplados no estudo. Para a realização das medidas, são escolhidos os movimentos mais críticos da malha viária, sendo que, no caso de estudos relativos a horários de pico o ideal é selecionar os períodos em que a rede a ser simulada se encontra mais carregada. Entretanto por aspectos de natureza prática, é muito difícil organizar o contingente desejado para realizar de forma simultânea todas as medições necessárias nos segmentos das sub-redes escolhidas, o que acaba restringindo o tamanho das redes modeladas e criando a necessidade de se estabelecer estimativas e às vezes algumas extrapolações (como foi o caso da atividade de teste, detalhada no fim deste trabalho).

Mais a frente, na seção 6.2., o leitor receberá as impressões do que foi de fato realizar a coleta desses dados para o estudo de caso. Assim, em caso de intenção de reprodução, ele saberá como lidar com as dificuldades de se realizar a atividade sozinho e tentar mitigar as distorções para levantar um conjunto de parâmetros que crie uma simulação minimamente aproximada da realidade.

3 MÉTODOS DE OTIMIZAÇÃO

Neste capítulo o leitor obterá uma visão panorâmica acerca dos principais tipos de algoritmos existentes usados para buscas e otimizações em problemas de engenharia (com ênfase dada à descrição de procedimentos evolutivos), sendo que em seguida será dada atenção especial aos AGs. A maioria dos algoritmos tem a inspiração provinda de mecanismos naturais sejam eles físicos, químicos ou biológicos e pode-se encontrar descrições mais detalhadas para sua implementação, bem como uma grande variedade de referências em Brownlee (2011). Em parte também foi usado como texto base a obra de Yang (2008).

3.1 Tipos de Heurísticas

3.1.1 Algoritmos Estocásticos

Estes são procedimentos de busca global que em sua grande maioria são usados para achar vizinhanças no espaço de busca que possam servir como bons pontos de partida para processos de otimização local.

Como exemplos principais temos: a Busca Aleatória, que, como o nome diz, consiste em amostrar o espaço de busca da forma mais livre possível; *Stochastic Hill Climbing*, baseado no *Hill Climbing* simples, um método de gradiente onde uma solução candidata é modificada ligeiramente e a modificação só é aceita se resultar em melhora; *Scatter Search*, usa métodos de otimização local para gerar um portfolio de soluções variadas que reúnem informação sobre as boas vizinhanças do espaço de busca. *Tabu Search* mantém armazenados na memória os movimentos realizados no espaço de busca para evitar que o procedimento repita tentativas em regiões já testadas, isso pode servir para fazer o algoritmo convergir a vizinhanças melhores.

3.1.2 Algoritmos Físicos

Aqui temos as lógicas baseadas em processos físicos e da natureza, e que são normalmente tidos como algoritmos estocásticos que misturam busca global com busca local.

Seriam bons representantes dessa classe: Arrefecimento Simulado, ou *Simulated Annealing* que é inspirado nos processos metalúrgicos de aquecimento seguido de resfriamento que reorganizam a estrutura cristalina do material e criam uma liga mais resistente, onde de forma análoga o algoritmo passa por um “aquecimento”, um momento de maior liberdade de movimento pelo espaço de busca (exploração), seguido de um resfriamento, onde há uma movimentação menor que resulta em ajustes mais finos; Busca Harmônica que é inspirado na adaptabilidade do improviso de músicos de jazz, onde cada atributo da solução candidata representa um músico, com um tom e um alcance sonoro, que a cada iteração busca se adaptar ao toque dos outros, e vice-versa.

3.1.3 Algoritmos de Enxame (*Swarm Algorithms*)

Nestes algoritmos é usada a idéia de auto-organização em sistemas coletivos descentralizados, ou em outras palavras, são procedimentos que mimetizam fenômenos de inteligência coletiva, aonde a informação é guardada das mais diferentes maneiras.

Dentre os principais temos: Otimização por Colônia de Formigas, que simula o estilo de busca por comida desses animais, que inicia pelo percorrimto de caminhos aleatórios no espaço de busca até o momento em que se encontra comida, momento a partir do qual a formiga que o encontra passa a soltar um feromônio que reforça aquele traçado, passando a ser seguido por outras formigas, que também soltam feromônios, de forma a especializar a busca na direção de vizinhanças para as quais o caminho estará tão forte e carregado de feromônios quanto maiores forem os seus destaques e benefícios – de forma análoga o algoritmo percorre o espaço de busca probabilisticamente, percorrendo um parâmetro por vez para construir esse “caminho”, armazenando os movimentos e dando força para os mais recentes e bem sucedidos; Algoritmo de Abelhas é um método que, de forma semelhante à colônia de formigas, envia exploradores em regiões diversas do espaço, dessa vez sem o uso de traçados mas com localização direta de pontos promissores onde serão feitas buscas locais com o uso de mais abelhas.

3.2 Algoritmos Evolucionários

Aqui é feita uma breve e conveniente recapitulação acerca de processos evolutivos naturais para permear os conceitos aplicados nas implementações inspiradas nesse mecanismo natural. Após essa parte será dada atenção exclusiva aos AGs, visto que sua descrição intuitiva serve como base para diversos outros algoritmos dessa classe, como programação genética e estratégias evolutivas, procedimentos que podem ser vistos com maior detalhe em Brownlee (2011).

3.2.1 Procedimentos Evolutivos

Desde o surgimento da vida, uma série infindável de mudanças espontâneas transformou os primeiros aglomerados membranosos de proteínas na imensa biodiversidade, em suas mais complexas formas, que dá rosto à superfície do planeta atualmente. Além disso, o ecossistema gerado por essa história de mais de três bilhões de anos carrega em suas moléculas de DNA, seja a de uma ameba ou de um ser humano, todas as suas páginas. Para escrever essa história e criar seres cada vez mais ou menos adaptados ao ambiente e, no caso dos primeiros, mais aptos a dar prosseguimento à sua linhagem, a evolução fez uso de dois tipos de procedimentos complementares.

O primeiro, e mais primitivo, é a mutação, que ocorre da forma mais espontânea possível, seja pela alteração acidental do código genético ou pela ativação de uma parte adormecida dele que se reflita no fenótipo do ser vivo (constituição e estruturação física, anatômica e fisiológica em geral). A mutação se apresenta como um ruído que dá certa variabilidade inédita às espécies e que as faz percorrer, assim, caminhos evolutivos inovadores.

O outro procedimento é a recombinação que fez com que a evolução ocorresse de forma mais acelerada, pois permitiu aos seres vivos, mesmo os mais simples, atuarem de forma ativa na sua evolução ao escolherem, ou serem levados a escolher instintivamente, seus parceiros, permitindo às espécies agregar nas futuras gerações as boas características que, isoladas, estavam presentes em diferentes indivíduos, sendo que muitas apareciam através de mutações (Figura3).

Figura 3: Curva de evolução em função do tempo quando há escolha de parceiros para reprodução (azul) e quando não há essa escolha (verde)

Através desses procedimentos a vida encontrou inúmeras soluções para todo tipo de problema que lhe fosse apresentada, sendo elas desde aparatos morfológicamente muito especializados para determinados ecossistemas, até os generalistas que permitiram, mesmo de forma ineficiente, que as espécies se espalhassem pelo mundo. Há até mesmo soluções dadas através de estruturas especializadas e generalistas ao mesmo tempo, como a mão, que é a melhor solução especializada, até hoje, em manipulação de objetos, mas é generalista o suficiente para permitir a manipulação de uma infinidade deles (NETTO, 2010).

3.2.2 Introdução ao Algoritmo Genético

Seja qual for a forma que qualquer espécie tomou após os processos evolutivos, as principais exigências do sistema são a sobrevivência e a capacidade de se destacar e propagar os genes. Não há restrições enumeráveis acerca de como as soluções podem se apresentar e o caminho que devem tomar quando esse processo envolve diversos passos.

Disso, o conceito prático principal que podemos extrair é o da emergência, que seria o surgimento espontâneo, após inúmeras tentativas e erros da natureza, de formas adaptadas e bem sucedidas das espécies como a solução para a sobrevivência nos diversos ambientes. Este é um processo que não possui um

coordenador ativo, ele baseia-se totalmente na aleatoriedade e nas “exigências” do sistema em que cada ser vivo está inserido.

Foi a partir dessa relação, dentro dos processos evolutivos, dos inúmeros caminhos possíveis e imprevisíveis entre o surgimento do problema e a emergência da solução, que se pensou na hipótese da aplicação desses conceitos para resolver problemas da vida prática, e com isso surgiram os programas que são classificados pelo nome desta seção.

Nos seres vivos mais avançados como o homem, o código de DNA das células diploides possui encadeadas cerca de seis bilhões das seguintes quatro possibilidades de pares proteicos: Adenina-Timina, Timina-Adenina, Guanina-Citosina e Citosina-Guanina. Isto configura o DNA como um código composto por palavras de dois bits. Se lembrarmos que na memória de um computador oito bits são um byte, a codificação binária do genoma diploide ocuparia um pouco mais do que dois CDs inteiros (cerca de 1,5 Gigabytes). Embora haja muitas redundâncias e genes inativos, grosso modo é esse o volume de memória que define, direta ou indiretamente, todas as nossas características.

Existe nesse volume de informações uma parcela que conterà as informações para a constituição geral de nossos braços ou pernas pelo simples fato de sermos humanos, e outra parte que conterà as informações relativas às características dos membros do indivíduo especificamente, podendo eles serem mais compridos ou curtos, flexíveis ou robustos e rígidos. No que diz respeito ao fenótipo e à resolução de problemas computacionais não seria necessário usar códigos com extensões da ordem de megabytes para codificar os fenótipos tanto gerais quanto específicos de um braço ou de cada braço humano, de modo que os fenótipos podem ser classificados e codificados em poucos bytes para formar uma espécie de DNA simplificado que descreva o braço de uma pessoa de forma resumida.

Assim como podemos realizar essa conversão e simplificação de um braço humano, muitos problemas práticos, em aplicações como as de engenharia, gestão e economia que normalmente são complexos e envolvem um sem número de variáveis, pode ser codificado de forma binária e ter assim descrito o seu DNA. A partir do momento da codificação desses problemas, mecanismos que mimetizem os procedimentos de mutação e reprodução podem trazer boas resoluções.

Para resumir o resultado da inspiração nos mecanismos naturais de evolução cabe começar por enumerar os passos de um procedimento padrão na área de AGs (RENNARD, 2000):

Inicialização:

1. Codificação do problema, normalmente em uma série binária;
2. Criação aleatória (ou não) de uma população inicial de soluções alternativas.

Ciclo evolutivo:

3. Atribuição de uma pontuação de acordo com o índice de mérito (fitness) de cada alternativa (indivíduo da população);
4. Seleção de uma determinada quantidade de soluções de forma proporcional aos seus níveis de adequação;
5. Mutações esporádicas e recombinações entre os genótipos selecionados para criação da próxima geração.

Como uma ilustração do que consiste a procura por uma boa solução com o uso de AGs, imagine um grande território montanhoso coberto totalmente por uma neblina espessa e sobre o qual não há conhecimento, adicione não linearidades a este sistema na forma de túneis, montanhas, torres e buracos (passo 1) e, agora, é necessário sair andando à procura de algum dos maiores picos (mutação), avaliando o caminho traçado usando somente um altímetro (passo 3). Uma única pessoa, por mais rápido que se locomovesse, demoraria muito para achar estes picos, sendo que muito provavelmente ficaria parada em um pico local, então, para conseguir cobrir uma área extensa seria necessário espalhar uma grande quantidade de pessoas pelo terreno para explorá-lo (passo 2). Permita então a algumas pessoas, em sua maioria as que estão em grandes altitudes, se comuniquem para trocar informações sobre as regiões em que se encontram (recombinação) e avaliem as melhores alternativas baseadas também no que as outras sabem. Por fim recolha certa quantidade de pessoas, na sua maioria as que se encontram em altitudes mais baixas e substitua-as por pessoas que sejam colocadas em posições relativas às conclusões de cada uma das conversas que foram permitidas. Geralmente os

espaços amostrais que compreendem todas as alternativas possíveis para um problema complexo são muito grandes e as frequentes não linearidades complicam muito a realização de um traçado bem definido das inúmeras regiões existentes e, conseqüentemente, das que contém as melhores configurações de parâmetros. Logo a geração de uma população inicial que explore diferentes regiões se configura como um paralelismo implícito que reduz muito o tempo de procura pelas soluções, mas há outros fatores que reforçam este argumento.

Ao descrever uma pessoa, como podemos seccioná-la em partes, ela pode ser alocada em várias populações diferentes, de modo que alguém que tenha braços longos e pernas curtas representa não só as pessoas que compartilham essas características, mas também as que têm apenas braços longos, ou as que têm apenas pernas curtas. Remeter isso à ilustração anterior implica dizer que é permitido a cada uma das pessoas ocupar diversas regiões ao mesmo tempo. Se voltarmos à codificação de um problema, suponhamos que, como exemplo, isso seja feito a partir de palavras de 8 bits, temos então que cada conjunto de dois ou mais bits constitui um bloco de construção, por exemplo: 10*****, *00*****, 1*0**1**. Como podemos ver abaixo na Figura 4, durante um cruzamento em um AG, dois genótipos trocam seus códigos através de um corte feito em um ponto aleatório, no caso de um código de 8 bits podemos fazer cortes do primeiro até o sétimo bit, sendo que cada bit da linhagem gerada tem uma pequena chance de trocar de valor no fim do processo (mutação).

Figura 4: Processo de recombinação (crossover) com mutação em um AG

Então podemos perceber que existe um paralelismo implícito na exploração do espaço amostral na medida em que um único genótipo representa diversas regiões ao mesmo tempo, pois mesmo que através das recombinações as linhagens seguintes de genótipos passem a se localizar (como código inteiro) em regiões muito distantes, alguns dos seus blocos de construção, principalmente os que têm grande participação na adequação do indivíduo ao problema, possuem chances muito

maiores de se propagar por diversas gerações seguidas e permitir a exploração de uma boa região, sendo estes intuitivamente os pares de bits seguidos, como 10*****, *00**** que são muito mais persistentes que 1*0**1** [HOLLAND, 1992].

Essa propagação de blocos fortes, em especial dos que contribuem na adequação dos genótipos, juntamente ao paralelismo implícito que ele reforça é o que permite a rápida convergência de um AG a bons resultados, independente da vertente da área à qual essa lógica é aplicada. Esses bons resultados aparecem após os dois momentos básicos na convergência do algoritmo, primeiramente temos o momento de exploração no sentido de uma busca e descobrimento geográfico macro, em que há uma grande variedade nos genótipos que permite ocupar o espaço de forma relativamente homogênea, através de uma enorme gama de blocos de construção. Num segundo momento, depois que o algoritmo acha uma região com bons níveis médios de adequação, ele passa a ter novas populações mais concentradas nas regiões que possuam uma variedade reduzida de blocos, mas todos de boa qualidade, sendo que esse momento também configura uma exploração, mas mais semelhante a uma mineração profunda. Essa sequência encontra-se representada na Figura 5.

Figura 5: Fases de um Algoritmo Genético. Busca e Descobrimento / Exploração (esquerda), Clusterização (centro), Mineração ou Garimpo / Ajuste (direita)

Entretanto é importante citar que assim como muitas lógicas de busca tradicionais baseadas em estatística, os AGs apresentam curvas de convergência que diminuem seu ritmo com o tempo, estabilizam em um determinado nível e não há garantia de que tenham atingido o máximo ou mínimo global, mas essa sua capacidade de auto-ajuste e esse aprendizado, por assim dizer, por meio da adaptação pode levantar uma série de opções já excelentes que, devido ao grande espaço de busca em que se encontram, demorariam muito tempo para serem

descobertas por especialistas humanos, que, por sua vez, de posse de tais soluções, podem fazer os ajustes necessários para aprimorá-las. Isso também pode ocorrer de forma inversa, pois algumas vezes o AG evolui a partir de alguns pontos que foram pré-definidos pelos especialistas, trazendo resultados na forma de ajustes ou eventualmente de quebras de paradigma. Essas duas abordagens fazem do AG, assim como muitas ferramentas computacionais, um instrumento complementar que agiliza muito a atividade criativa e inovadora humana.

Além do paralelismo implícito, algo que contribui intrínseca e significativamente para a convergência de um AG é a competição. Populações cada vez maiores de genótipos promovem a evolução de forma mais rápida do que em algoritmos de busca heurística e estatística, pois, assim como na vida real, o ritmo do desenvolvimento das habilidades de um agente quando feito de forma isolada não se compara ao ritmo observado no desenvolvimento coletivo multi-agente, onde há troca de informações e onde a quebra de vícios e paradigmas é mais frequente devido aos contrastes e às comparações que permitem uma autocrítica mais rotineira, coisas que por fim resultam no cultivo das melhores práticas (conceito tratado no curso de Vida Artificial, NETTO, 2010 e ilustrado na Figura 6).

Figura 6: Curva de adaptação em sistemas com um agente face à curva de adaptação em sistemas multiagente

Os tempos de convergência dos AGs são muito dependentes do conjunto de parâmetros de controle e dos métodos de seleção, cruzamento e mutação que são escolhidos, e para cada aplicação. Na próxima seção serão tratados alguns aspectos relacionados a essas escolhas para engenharia de tráfego, mas em relação aos AGs em geral uma avaliação já pode ser encontrada em Cavicchio (1972), onde são medidos os tempos de convergência para diferentes conjuntos de parâmetros e o nível de melhora trazido pelos mesmos.

4 AGs APLICADOS À ENGENHARIA DE TRÁFEGO

Nesse capítulo é apresentada uma revisão referenciando alguns trabalhos importantes que envolveram estudos de AGs em controle de tráfego ao longo dos últimos quinze anos.

Em Lebdeh e Benekohal (1997) o foco da aplicação é a prevenção da formação de bloqueios em que um sinal verde não libera veículos devido ao enchimento do segmento à frente (jusante). Neste trabalho é indicada a necessidade de controle semafórico dinâmico para situações saturadas, pois o comportamento do tráfego em uma janela de tempo depende daquele que foi registrado na janela de tempo anterior. Além disso, o autor relata que execuções diferentes de AGs (com variadas sementes de números aleatórios) levam a valores de convergência semelhantes.

Mais tarde, em 2001, os mesmos pesquisadores usaram um AG em diferentes cenários de priorização de fluxo em uma via arterial e fizeram uso de diversos índices de mérito para poder comparar os resultados de uma forma mais rica, chegando à conclusão de que a técnica poderia ser utilizada para controle em tempo real independente da priorização requerida pelo engenheiro de tráfego responsável.

No estudo realizado por Lo et al. (1999), um AG otimiza o tráfego simulando janelas de 45 minutos. Nele é indicado que em situações não congestionadas o recomendável seria otimizar um semáforo de cada vez, de forma serial, enquanto que em situações congestionadas seria mais interessante otimizar todos os semáforos paralelamente, sendo que ambas as estratégias apresentaram resultados 20 a 40% melhores que aqueles alcançados até então pelo software Transyt-7F, que usava o método de *Hill-Climbing* para isso, alternativa altamente dependente das condições iniciais e que é mais indicada para otimizações que visam ajustes finos.

Em Park et al.(1999) um algoritmo genético é usado para a otimização de todas as variáveis da temporização semafórica. Para isso um simulador mesoscópico próprio foi criado e o desempenho do AG foi comparado àquele do TRANSYT-7F. O aplicativo foi testado em três níveis diferentes de demanda: baixa, na qual o programa apresentou resultados estatisticamente melhores; média onde

performou de forma semelhante; alta, cenário em que, estatisticamente, apresentou maior vantagem. O estudo de caso desse trabalho possuía 4 semáforos e 13 aproximações, e o AG foi executado com 40% de crossover, 3% de mutação e 250 gerações de 10 indivíduos, sendo que as simulações representavam uma janela real de 15 minutos. À época, as otimizações demoraram 23 minutos em um computador dotado de processador pentium 133MHz.

Mais tarde, em 2000, para avaliar o impacto do uso de diferentes funções objetivo, Park et al. mediram o desempenho do AG em situações de saturação, ao classificar os cromossomos de acordo com três diferentes critérios: maximização de fluxo (throughput); minimização de atraso e minimização de atraso com função de penalidade. Ao fim do experimento os cientistas chegaram ao indício de que o critério mais simples, o do atraso, é o mais eficiente.

Em outro estudo, de Park et al (2001), são tidas como pressupostas as discrepâncias entre resultados de simulações macroscópicas e microscópicas, no que é discutido que as macroscópicas apesar de serem mais simples de se parametrizar e de mais rápida execução não representam o comportamento estocástico dos usuários da rede. Usando na experiência o simulador microscópico CORSIM o GA aplicado apresenta planos melhores do que aqueles saídos do Transyt-7F em 95% das vezes e que conseguem lidar com variações estocásticas na demanda.

Outro trabalho interessante liderado por Park, em 2004, envolveu o uso de AGs para a otimização de diversos planos fixos ao longo de um dia de operação, mas determinando também os pontos de transição entre eles e o número de transições. Vale destacar aqui que o número otimizado de planos diários para o qual a ferramenta convergiu é 7, valor semelhante àquele praticado pela CET-SP na rede que foi o estudo de caso desta obra.

Em Lebdeh (2002), é tomado como premissa o conceito de que as velocidades máximas permitidas em cada segmento devem se dar em função do nível de congestionamento e, pressupondo que todos os condutores respeitem este limite, os dados de ocupação e tamanho das filas são usados para otimização em tempo real dos tempos de verde e defasagens enquanto é aplicado um algoritmo genético para a determinação de melhores conjuntos de velocidades máximas. O

foco da aplicação é a redução do número de paradas para uma consequente redução no consumo de combustível e menos interações decorrentes das manobras de “stop and go”, resultando em uma melhora no fluxo de tráfego na medida em que as ondas de choque nas filas são suavizadas.

Como uma guia para o uso de AGs em sinalização semafórica, em Girianna (2004) há uma explicação bem clara acerca dos cálculos de todas as variáveis de controle da temporização semafórica. Após isso é apresentado um GA simples paralelo que possui a função de maximizar fluxo administrando a temporização semafórica de forma dinâmica, fazendo uso de algumas restrições para os tempos de verde. O AG foi testado em cenários de congestionamento que variaram entre os cruzamentos críticos estando nos pontos de saída ou nos pontos de entrada da rede, para estudar o comportamento do conjunto de defasagens para o qual o GA converge. Os resultados mais uma vez foram animadores, pois o GA soube lidar com variações temporais de volume, evitando que os cruzamentos mais carregados ficassem congestionados. O autor indica ainda a aceleração obtida pelo uso de uma arquitetura paralela e sugere o uso de técnicas complementares “gananciosas” de busca local que usem os resultados do AG como bons pontos de partida para um ajuste fino.

Em um trabalho particularmente útil para planejadores de tráfego, Varia e Dhingra (2004) empregam AGs não só para otimização de sinalização semafórica, mas para manipular a distribuição da demanda pela rede através da otimização da matriz origem-destino. No modelo é considerando uma situação hipotética onde para todos os condutores é informado o melhor caminho a se tomar de uma forma que cada trajeto traçado contribua para o desafogamento dos outros trajetos e vias, de uma maneira integrada. Esse trabalho usou um modelo mesoscópico de pelotões e só não otimiza as defasagens

Ceylan e Bell (2004a) contemplaram o efeito dinâmico de longo prazo do comportamento da rede ao adicionarem ao AG um procedimento periódico em que o comportamento do tráfego simulado frente aos melhores planos encontrados é reavaliado, de forma a estabelecer a nova situação de equilíbrio entre oferta e demanda e assim remodelar as matrizes de origem-destino antes de prosseguir com a otimização. Em trabalho semelhante naquele ano (CEYLAN E BELL, 2004b), os

pesquisadores aferiram o aumento de capacidade da rede causado pelo mesmo método. Mais a frente, em 2006, os mesmos pesquisadores utilizaram um algoritmo de redução de espaço de buscas juntamente à uma otimização dada pela combinação de um AG e um HC com a intenção de estabelecer um bom ponto inicial para a otimização de sinalização semafórica do Transyt.

Cantarella et al. (2006) compararam, frente a tempos fixos de processamento, diversos métodos heurísticos de otimização, a saber o Hill Climbing, Tabu Search, um AG, e métodos híbridos. O autor identificou que os métodos AG e Tabu Search puros são os que convergem mais rapidamente, e constatou que o Hill Climbing foi o único que obteve performance consideravelmente pior, mas alertando que pode ao mesmo tempo ser o melhor se os outros forem configurados de maneira errada.

Também para contemplar as mudanças nos fluxos de tráfego, dois trabalhos, de Sun et al. (2006) e Teklu et al. (2007) aplicaram ciclos simultâneos de otimização e reavaliação de fluxos, usando diferentes AGs, com os primeiros indicando que é necessário realizar diversas otimizações e os últimos que muito ainda deve ser feito para que se saiba selecionar os melhores parâmetros do AG.

Para aumentar a rapidez de execução do AG, Stevanovic et al. (2007), executaram um AG usando computação distribuída, e avaliaram como função objetivo uma composição de atrasos e paradas (o total dos atrasos da rede em horas é somado ao número de paradas multiplicado por um fator). Com maior agilidade puderam avaliar diferentes conjuntos de parâmetros do AG recomendando então o uso de números de gerações no mínimo três vezes maior que o tamanho da população, probabilidade de cruzamento de 40 a 70% e de mutação de 1 a 4%.

Na mesma busca por boas configurações, Kesur (2009) avaliou outros elementos como os tipos de codificação, cruzamento e mutação, chegando aos indicativos de que o AG funciona melhor com o uso de código gray para uniformizar as mutações binárias. Além disso, em detrimento do método de corte simples, os resultados favoreceram o uso do cruzamento misturado, em que os cromossomos dos pais são quebrados em diversos pedaços, cada um contendo um parâmetro, com os filhos recebendo esses parâmetros de um ou outro pai de forma aleatória.

Usando um modelo macroscópico que simula a dinâmica das filas e o bloqueio de segmentos (spillback), Liu e Chang (2011a), aplicam um AG para diversas situações de demanda, focando em cenários de saturação que apresentem bloqueios e consequente propagação de filas de um segmento a outro. Para efeito de comparação, em uma mesma rede de teste, foram gerados planos otimizados pelo modelo apresentado e pelo TRANSYT-7F, sendo cada um posteriormente simulado repetidas vezes no simulador microscópico CORSIM (para eliminar efeitos estocásticos da arquitetura microscópica). Após o processo concluiu-se que em situações de alta demanda e congestionamento da rede o plano originado pelo modelo proposto apresentou níveis de atraso total 18% menores do que aqueles originados pelo TRANSYT-7F.

Também usando modelos macroscópicos de fluxo de tráfego, Liu et al (2011b) simulam cenários de bloqueio parcial de vias expressas e consequente redirecionamento da demanda para vias arteriais vizinhas otimizando com AGs a sinalização da via arterial para esses casos. Os autores estabeleceram como os índices de mérito o fluxo total de veículos e o atraso total das viagens na via arterial para os veículos desviados, realizando diversas otimizações e atribuindo pesos diferentes à minimização do atraso e à maximização do fluxo em cada uma. É concluído que essa mudança de pesos entre objetivos de controle não impacta sistemas pouco congestionados, dispensando nesses casos a necessidade de estabelecimento de estratégias de controle específicas. Entretanto em sistemas congestionados a otimização fica sensível às mudanças de valores desses pesos.

Em Medina et al. (2010) foi realizado estudo considerando diversos níveis de situações saturadas e não saturadas, levando à constatação de que o AG traz ganhos evidentes em redes mais carregadas.

Pohlmann e Friedrich (2010) ajudam a quebrar a barreira do AG como aplicação apenas offline e combinam uma ferramenta estimadora de tráfego a sistemas de monitoramento em tempo real e um modelo rápido de simulação por transmissão celular para desenvolver uma aplicação que otimiza a sinalização semafórica a cada quinze minutos. A ferramenta foi aplicada em uma rede real com bons resultados, mas indicando que o estimador apresenta problemas na hora de lidar com o tráfego mais volátil dos períodos de pico.

5 O SIMULADOR GENPOLIS

Neste capítulo será explicado o modelo de simulação desenvolvido, a princípio, para este mestrado. O capítulo inicia com uma breve explicação para a escolha do desenvolvimento total de um modelo próprio passando por explicações de sua conceituação, arquitetura de funcionamento e detalhes pertinentes da sua programação (motor).

5.1 *Porque Fazer um Simulador Dedicado*

Todo modelo de simulação criado deve levar em conta os fatores relevantes à solução de problemas enfrentados pelo programador e pelos especialistas que o idealizam. Logo, para apresentar eficiência, essa ferramenta deve ser projetada de forma a atender às necessidades específicas da realidade presente na região onde é concebida (VILANOVA, 2009). Quando não é feito dessa forma específica e o intuito é a concepção de um simulador generalista, devem ser atendidos múltiplos conjuntos de requisitos provenientes de diversas culturas e cenários, o que pode aumentar em muito a complexidade do modelo e sobrecarregar a lógica e o motor da simulação, tornando-o mais lento e pesado.

Num primeiro momento, o que interessa ao programa deste mestrado é o alinhamento do paradigma dos AGs à prática local, que é a de estabelecer planos mais generalistas que durem por períodos longos a partir de observações de campo onde basicamente são feitas coletas de dados agregados típicos de simuladores macroscópicos, como fluxo, fluxo de saturação e tempos de percurso.

Considerando que cada simulação deverá compreender janelas temporais de algumas horas e que para cada execução de um AG é necessário realizar milhares de simulações, a ferramenta deve ser rápida o suficiente para que a execução ocupe um intervalo de tempo plausível de, no máximo, poucas horas até chegar a uma boa solução e permitir a análise e ajuste pelo engenheiro de tráfego. Além disso, o programa possuirá, enquanto protótipo dentro do tempo e recurso envolvidos em um mestrado, diversas simplificações para que ele atenda, neste momento, somente os requisitos de parametrização usuais da Companhia de Engenharia de Tráfego de São Paulo (CET-SP), mas que não comprometam sua verossimilhança.

Em segundo plano, deve-se pensar em um simulador que seja rápido o suficiente para uma futura aplicação do AG à resolução de problemas em tempo real. Apesar do estado tecnológico da rede paulistana ainda estar longe de permitir a aplicação de um sistema como esse, num futuro em que isso seja possível, se o protótipo seguir evoluindo, com interface gráfica fácil e intuitiva, mais flexibilidade de controle semaforico e um AG mais completo, acredita-se que o AG possa trazer ganhos maiores ainda, como observado em alguns dos artigos citados anteriormente.

Foi levado em consideração então que não seria o melhor caminho importar um simulador comercial já que as soluções aplicadas para a cidade de São Paulo são mais simples e são tipicamente planos fixos de temporização que perduram por horários de pico ou entropico inteiros, e, além disso, no que diz respeito ao próprio controle semaforico, ainda estamos lidando com uma grande predominância de semáforos cujo funcionamento não passa da carta de tempos configurada, salvo as raras exceções dos semáforos inteligentes que estão totalmente funcionais.

5.2 *Resumo Descritivo do Modelo*

O GenPolis é um simulador de trânsito mesoscópico, projetado para simular rapidamente malhas complexas de vias concorrentes e semaforizadas que possuam segmentos mais curtos. Nele, toda a rede é simulada simultaneamente em passos equivalente a um segundo real, que é a prática mais usual. Isso significa que, a cada ciclo de cálculos da rede toda (passo), o simulador apresenta uma foto virtual do estado da rede estudada. O projeto possui um construtor preliminar responsável por preparar a malha viária virtual automaticamente como grafo através de dados externos.

O ambiente de simulação é composto basicamente de três entidades:

- *Segmento de Rua (Link)*: Se assemelha a uma esteira, movimentando todos os veículos com igual velocidade;
- *Cruzamento (Nó)*: Pode conter um semáforo ou não e é responsável por distribuir os veículos dos segmentos à montante para os segmentos conectados à jusante;

- *Veículo*: Partícula básica de carga do sistema.

Do lado da herança microscópica temos apenas que o simulador descreve seus veículos de forma unitária, enquanto que pelo lado mesoscópico predominante, o comportamento dos veículos é regido por determinações agregadas de velocidade fixa (em função do tempo de percurso medido), não permite ultrapassagens e as conversões são baseadas em probabilidade, sem a existência de caminhos pré-traçados. Logo os fluxos das malhas virtuais serão compostos de forma estatística e consolidada sem se preocupar em atribuir qualquer tipo de identidade a cada um dos veículos que percorre a malha.

Os veículos são descritos dessa forma unitária no protótipo pois, além de ser um modo intuitivo e simples, ainda se contempla a possibilidade de se criar um simulador microscópico com velocidades variáveis e o modelo carro-seguidor a partir dele no futuro. Ensaio e maiores detalhes acerca disso são tratados na última seção do trabalho. Ao mesmo tempo, como já foi dito, os dados necessários para a constituição de uma simulação que esteja de acordo com o padrão de coleta da CET-SP são somente agregados e apontam para o uso de simuladores macroscópicos ou mesoscópicos, de modo que a constituição de um simulador microscópico aumentaria desnecessariamente a necessidade de leituras de campo e calibrações relativas a parâmetros como distanciamento entre veículos e ritmo de frenagem/aceleração, sendo que tal processo envolveria uma equipe de engenheiros de tráfego, muito mais do que apenas um único pesquisador.

Além disso, vale ressaltar aqui que não é do escopo deste projeto simular corredores mais longos e vias expressas, de modo que ainda não se busca saber se este simulador pode representá-los fielmente, ou mesmo otimizá-los de forma plausível e eficiente, pois geralmente as soluções encontradas para estas vias dependem mais de aspectos topológicos como número e largura de pistas, traçado e velocidades máximas.

5.3 Arquitetura da Ferramenta

Nesta seção serão descritos os blocos que constituem a ferramenta para que seja passada uma idéia inicial sobre o fluxo de informação.

5.3.1 City Engine e OpenStreetMaps

A princípio, devido ao caráter inicialmente mais experimental desse estudo a geração da malha viária do simulador teria formato semelhante ao das redes do Transyt-7F e poderia ser feita de forma padronizada (“Blank”) ou, a partir desse padrão, seria constituída uma malha com uma conformação aleatória aparentemente mais realista, como podemos ver abaixo na Figura 7.

**Figura 7: Conceito inicial de rede viária para o simulador.
Rede padrão (esquerda) e rede aleatória (direita)**

Este modelo de vias perpendiculares foi escolhido inicialmente pois além da relativa simplicidade de implementação, parte-se do pressuposto simplificado de que enquanto dirigimos realizamos geralmente três escolhas: prosseguir, virar à esquerda ou virar à direita.

Entretanto, foram avaliadas e escolhidas ferramentas computacionais comerciais que trouxeram ganhos em vários sentidos, pois ao mesmo tempo que pouparam uma boa parte do tempo de programação para criação da malha padronizada/aleatória devido à possibilidade de automatização, ainda assim possibilitaram a criação de uma malha bem mais verossímil. A ferramenta escolhida para isso foi o software City Engine, que normalmente é usado para planejamento urbano e para a criação de ambientes virtuais destinados principalmente a jogos, demonstrações e filmes, e que possibilita a importação de malhas reais e seus respectivos dados geográficos disponíveis no site OpenStreetMap (<http://www.openstreetmap.org/>).

Figura 8: Captura de tela do recurso de exportação de redes viárias do site www.openstreetmap.org

As ruas enquadradas na imagem (Figura 8) são traduzidas para uma lista de segmentos que são especificados como uma coordenada geográfica inicial e outra final e que respeitam o sentido da rua. Essa lista é lida pelo programa City Engine (Figura 9), aonde vários parâmetros podem ser editados através de um campo específico, de forma que a malha possa então ser compactada e alterada de forma prática até ficar adequada para a simulação.

Figura 9: Captura de tela do software City Engine para planejamento urbano, num momento em que uma rede extraída pelo openstreetmap é editada para ser simulada. A rede em questão é a região da Avenida Paulista, em São Paulo

5.3.2 Construtor

Uma vez editada, a malha pode ser exportada do City Engine na forma de uma lista de segmentos com identificações numéricas para ser então utilizada no GenPolis. No protótipo, atualmente essa lista é lida pelo programa Construtor (Figura 10) e é combinada com outro arquivo de texto que contém uma tabela os parâmetros de cada segmento, provenientes das leituras de campo, a saber: o fluxo, fluxo de saturação e tempo de percurso, e o parâmetro topológico que indica o número de faixas, e que juntamente com o comprimento (calculado pelo programa), serve para efetuar o cálculo do número de vagas, ou seja, a capacidade de repesamento de veículos do segmento.

Dessa forma esse programa transforma a lista de segmentos e a tabela de parâmetros em um modelo de malha viária conexo onde não só temos os segmentos, mas um grafo provido de conexões entre segmentos e os consequentes nós.

Todos esses dados podem ser verificados visualmente e então salvos e traduzidos para uma tabela que deverá ser lida pelo construtor interno do Simulador.

Figura 10: Captura de tela apresentando a rede apresentada na Figura9, após edição no software City Engine e compilada no construtor do GenPolis

5.3.3 Simulador

Cada execução do construtor cria um arquivo de texto com o grafo e todos os parâmetros pertinentes à malha viária a ser simulada. Caso seja necessário esse arquivo é manipulado e tratado para ser aberto e lido pelo Simulador, que recria o grafo na memória do computador e prepara a simulação.

Em cada simulação o simulador é responsável por transformar os dados binários de um cromossomo recebidos do AG em um plano semafórico fixo para a rede executada. Isso também pode ser feito diretamente através de um arquivo de texto onde o tempo de ciclo e os tempos de abertura e fechamento dos sinais são descritos explicitamente. Após esses cálculos o simulador é carregado seguindo um ritmo inicial fixo até que se atinja certa estabilidade, para daí iniciar a avaliação e a simulação efetiva (maior detalhamento na seção 5.4.4.).

Durante a execução do AG o simulador é executado sem qualquer tipo de visualização, mas naturalmente foi criada uma interface simples em OpenGL para que o usuário possa checar visualmente o funcionamento e comportamento da rede, enquanto ao lado o console passa informações consolidadas de funcionamento e do estado de congestionamento da rede, como observado na Figura 11.

Para maior entendimento: os pontos luminosos são veículos em circulação, as barras vermelhas são as filas e os segmentos pintados em rosa são aqueles em que o fim da fila calculado se encontra na extremidade montante do segmento (bloqueio, ou *queue spillback* em inglês), impedindo a entrada de mais veículos neste.

Figura 11: Captura de tela de simulação de teste. Valendo ressaltar que esta rede se assemelha somente geometricamente à região da Av. Paulista pois aqui os parâmetros utilizados para as vias foram todos padronizados, sendo que existem semáforos em todos os cruzamentos para se estressar o modelo

5.3.4 Algoritmo Genético

O AG consiste de uma camada superior de controle que passa sequencialmente ao simulador cada um dos cromossomos gerados ao longo da evolução para interpretação e criação dos planos de sincronização semafórica a serem testados. Ao final de cada simulação o AG registra o índice de mérito conferido àquele cromossomo para posterior classificação, seleção cruzamento e mutação, que serão realizados para constituição da geração seguinte de planos candidatos.

5.3.5 Fluxograma da Ferramenta

Abaixo, na Figura 12 temos uma ilustração que resume o fluxograma do GenPolis.

Figura 12: Fluxograma do GenPolis

5.4 Detalhes do Funcionamento (Modelo de Simulação)

5.4.1 Movimentação dos Veículos Pelo Segmento

Diretamente falando, para o GenPolis a velocidade de deslocamento dos veículos é dada em função do tempo de percurso medido para cada segmento simplesmente pela equação:

$$V_{segmento} = \frac{Comprimento_{segmento}}{TempoPercurso_{segmento}} \quad (Eq.1)$$

Essa velocidade então é constante e inalterável na simulação, para todo veículo, durante todo o percurso e independente da condição de congestionamento em que se encontra o segmento. Dessa forma, todos os veículos seguem da entrada do segmento para a linha de retenção ininterruptamente e, uma vez que chegam nesse ponto, no caso de haver um sinal vermelho, eles são empilhados como em uma fila vertical, retratada na Figura 13 (de forma análoga ao que era implementado no TRANSYT antigamente). Mas no caso do GenPolis, as desvantagens dessa

abordagem foram eliminadas pois a fila é calculada e dissipada de forma explícita usando métodos que serão descritos mais a frente.

Figura 13: Em um segmento/link, veículos em movimento (verdes), veículos empilhados (vermelhos)

Vale observar que, durante o desenvolvimento, modelos mesoscópicos de controle de velocidade foram implementados e avaliados, mas, uma vez que foram desenvolvidos mecanismos específicos para o cálculo da extensão e dissipação de filas preferiu-se deixá-los para um estudo futuro. No Apêndice A encontra-se uma descrição desse modelo de controle.

5.4.2 Servidor de Filas

Como já visto, quando um veículo termina de percorrer o segmento, no ponto de saída do segmento ele incrementa a pilha de veículos que podem ser transferidos para os segmentos à jusante. Assim, existem dois métodos elegíveis para se distribuir a demanda (Ilustrados na Figura 14):

i) Roteamento por porcentagem de conversões do segmento:

Os veículos que são passados adiante dessa forma são sorteados de acordo com as porcentagens de veículos que seguem em frente e de veículos que fazem cada uma das conversões possíveis, como em uma roleta em que aos movimentos com maiores probabilidades é destinada uma fatia maior. Este tipo de roteamento é aplicado aos segmentos mais relevantes da rede, sendo normalmente os que possuem semáforos na linha de retenção (segmentos de aproximação) e aqueles cujos movimentos possuem relevância direta para outros segmentos que sejam aproximações de semáforos.

ii) Roteamento por fluxo dos segmentos à jusante:

Os veículos que são passados adiante dessa forma são sorteados de acordo com os fluxos medidos dos segmentos candidatos (jusantes), como em uma roleta em que aos segmentos com maiores fluxos é destinada uma fatia maior (VILANOVA,

2009). Este tipo de roteamento é aplicado aos segmentos da rede que não possuem semáforos e não se encontram no meio do caminho entre dois semáforos próximos, ou seja, o perfil de envio de veículos não é importante, só sendo necessário que os segmentos à jusante sejam carregados ao longo da simulação com quantidades de veículo condizentes com as leituras de campo.¹

Figura 14: Métodos de roteamento por i:porcentagens de conversão no segmento montante (esquerda) e ii:fluxo total dos segmentos à jusante (direita)

Por último, a cada ciclo de simulação, se o segmento possui um ou mais veículos para serem transferidos, é incrementado em um contador de transmissões do segmento um valor igual ao fluxo de saturação, mas que agora é convertido em veículos equivalentes por segundo. Nos ciclos em que este contador possui valor maior do que um, o segmento entra em um ciclo de decremento seguido de passagem de veículo até o contador voltar a ter um valor menor do que um, e nos momentos em que não há veículos para serem transferidos, a esse contador é

¹ Essa diferenciação existe, pois na versão protótipo do GenPolis, desde que a primeira lista de segmentos é extraída do OpenStreetMaps, não há uma ordem bem definida de quais segmentos da malha escolhida são registrados primeiro e quais são registrados por último. Então quando essa lista é passada para o construtor, a montagem das tabelas de relação link_montante → link_jusante não segue uma regra. Assim, um segmento que tenha dois outros segmentos ligados a ele após a linha de retenção, terá os dois segmentos referenciados na sua lista de segmentos à jusante, mas não se pode dizer automaticamente nesse momento qual deles é o de “seguir em frente” e qual é o de conversão. Desse modo, a segunda regra de roteamento é adotada como padrão no construtor, pois os dados de fluxo já estão referenciados no arquivo que possui a tabela de consulta. Finalmente, para que a primeira regra de roteamento seja implementada, é necessário, atualmente, abrir o arquivo e atribuir essas porcentagens manualmente, baseando-se também, nas leituras de conversão realizadas ao mesmo tempo que as leituras de fluxo.

atribuído valor nulo. Este tipo de truncamento foi escolhido pelo fato deste simulador não trabalhar com pacotes de múltiplos veículos (ou parcelas) como normalmente é feito em simuladores mesoscópicos, mas sim elementos unitários.

Quando um veículo é passado para outro segmento, ele vira o veículo apontado como o veículo do fundo pelo segmento receptor, e o veículo apontado por ele como o veículo de trás é designado como o veículo à frente da fila do segmento transmissor.

5.4.3 Cálculo Explícito de Filas e de Bloqueio de Segmento

Uma das variáveis mais importantes que devem ser calculadas é o tamanho da fila ou, mais importante ainda, a posição do final dela e, eventualmente, a ocorrência de bloqueio do segmento caso essa posição seja o ponto de entrada deste. Já de início é importante ilustrar como a fila mais simples, a que cresce atrás de um período de sinal vermelho que acabou de iniciar em uma rede não saturada, é calculada no GenPolis.

Já foi dito que os veículos são empilhados no ponto de saída do segmento enquanto o atravessam por inteiro. Observe a Figura 15 para melhor entendimento do raciocínio que vem a seguir. No simulador, em todo passo de simulação e para cada segmento, os veículos empilhados são contados (Figura 15, parte A) e o correspondente comprimento horizontal dessa primeira parcela da fila é calculado, tendo-se por consequência uma medida preliminar da posição do fim da fila (Figura 15, parte B). Num segundo momento do cálculo é executado um ciclo que inicia por observar se o veículo que está em movimento e com posição mais avançada já se encontra em uma posição mais adiantada do que a posição preliminar do fim da fila (Figura 15, parte B). Em caso positivo, o programa acrescenta o comprimento desse veículo à posição do fim da fila (Figura 15, parte C) e repete essa rotina para os outros veículos em movimento e mais avançados até não haver mais veículos a serem checados ou até um veículo ser checado como em posição mais atrasada em relação ao fim calculado da fila (Figura 15, parte C e parte D). Em uma analogia, estes veículos em movimento que são capturados nessa rotina são aqueles que, na realidade já estariam parados na fila caso o cenário fosse o de uma simulação microscópica, com a implementação de uma fila horizontal.

Figura 15: Exemplo de cálculo do fim da fila realizado em um passo de simulação

Uma vez entendido isso, adiciona-se somente a observação de que, como o link do GenPolis é apenas uma esteira homogênea, independente do número de faixas da sua contraparte real, cada veículo ocupa um espaço proporcional ao seu tamanho real, dividido pelo número de faixas, o que significa que se um carro ocupa seis metros de sua faixa de rolagem, em um link de uma faixa do GenPolis ele ocupa esses mesmos seis metros, mas em um link de duas faixas ocupa três metros e em um link de três faixas, ocupa dois metros.

Propagadores de Espaço

Em uma fila, normalmente, temos partículas que se movimentam para frente, do fim até o início dela. Eventualmente, quando o elemento da frente da fila sai, ele libera um espaço que deverá se propagar para trás na medida em que o elemento que estava logo atrás anda para frente para ocupar o espaço vago, seguido pelo elemento que está atrás dele e assim sucessivamente até o fim da fila. Logo, o que se adotou para este simulador foi o uso de propagadores de espaço que andam para trás no segmento e que funcionam como uma memória que continua considerando, durante mais algum tempo, os espaços previamente ocupados pelos veículos que saíram do segmento. Em todo passo de simulação, para cada link é criado um propagador, que possui armazenada a contagem de veículos que foram liberados no segundo em que ele foi criado. Dessa forma, qualquer link onde houver

fila formada possuirá não só veículos como também uma sequência de propagadores em movimento contrário.

Pensando de forma intuitiva, quando estamos parados em uma rua com o semáforo fechado e em algum ponto intermediário de uma fila, normalmente não iniciamos o movimento no mesmo instante em que o semáforo abre, pois devemos esperar que os espaços se propaguem até o ponto onde estamos. Assim, é como dizer que estes espaços propagados são os nossos semáforos efetivos, semáforos móveis que estão constante e exatamente nos dizendo quando andar e quando parar no momento em que passam pela nossa posição, podendo essa sequência de sinais ser mais ou menos bem comportada. Pensando nisso, foi determinado que cada propagador carregaria consigo não só a informação de quanto espaço foi liberado no momento da sua criação, mas também o próprio fato mais fundamental de se o segmento pôde ou não liberar veículos no referido passo de simulação.

Assim, num primeiro momento um propagador é criado como semáforo móvel vermelho se o semáforo do nó está vermelho para o segmento e como um semáforo verde se o semáforo do nó está verde para o segmento. Normalmente quando o semáforo do nó está verde, o segmento pode liberar veículos, porém isso não significa que ele vai conseguir. Supondo que o segmento transmissor só possa liberar um único veículo neste ciclo e que o segmento sorteado para o qual esse veículo será transmitido seja um que se encontra lotado, ou bloqueado, o segmento não liberará o veículo e o propagador será criado como um semáforo móvel vermelho no lugar de um verde. Ou seja, o semáforo móvel é vermelho se o segmento tenta liberar veículos sem sucesso, apesar do semáforo do nó estar verde à sua frente.

Abaixo temos na Figura 16 uma ilustração exemplo para facilitar a visualização do processo que ocorre em alguns passos de simulação em seguida ao instante t_0 , o passo de simulação que foi retratado na Figura 15. Nessa passagem, num primeiro momento, o semáforo fica verde, alguns veículos são liberados e propagadores 1 e 2 relativos aos espaços desses veículos são criados (t_0+1 e t_0+2), ocupando cada um 6 metros de fila. Esse semáforo é fechado logo depois (t_0+3), mas o espaço ocupado pelos veículos liberados continua contribuindo para o tamanho da fila enquanto mais veículos (8 e 9) passaram pelo fim da fila em t_0+1 .

Com isso o fim da fila fica mais longe da linha de retenção apesar do segmento possuir menos veículos dentro dele, bem como acontece no mundo real.

Figura 16: Exemplo de propagação de espaços com bloqueio e liberação de segmento

Após algum tempo, supondo que em t_1 mais um veículo tenha entrado no link com o semáforo fechado, os propagadores ainda estão percorrendo a fila e o resultado é o de que o segmento fica bloqueado por algum tempo (t_1+1 e t_1+2) até os propagadores chegarem ao ponto de entrada, finalmente liberando espaço para mais veículos entrarem.

Essa intermitência depende da possibilidade de andar ou não a cada segundo, principalmente ao longo do tempo em que o veículo está em uma fila

congestionada, sendo feita no intuito de que se possa calcular precisamente a propagação de uma ou mais ondas de choque pela fila principal do segmento, e assim simular a existência de diversas sub-filas, como podemos observar na figura mais a frente, que contém uma imagem retirada do simulador.

A velocidade de propagação destes sinais é fixa e refere-se ao tempo necessário, quando o segmento está lotado, para um espaço se propagar do começo (primeiro veículo) ao fim da fila (último veículo) (Eq.5). Essa velocidade é calculada primeiramente a partir do tempo de esvaziamento do link, que por sua vez é calculado diretamente a partir do número de veículos que ele pode represar dividido pelo seu fluxo de saturação em veículos por segundo (Eq.2). Entretanto, o tempo de esvaziamento da fila envolve o escoamento do primeiro ao último veículo de um segmento lotado para o segmento seguinte e somente usar esse fator criaria o erro de que o final da fila só seria liberado quando todos os veículos já tivessem saído do segmento.

Dessa forma, deve ser subtraído do resultado obtido anteriormente (Eq.4) o tempo que o último veículo leva para atravessar o segmento inteiro (Eq.3), pois assim obtemos o intervalo de tempo entre o primeiro veículo ser liberado e o último veículo começar a andar a partir do ponto de entrada do segmento. (conceito fechado em reunião com VILANOVA em 2011). Abaixo, temos o processo resumido do cálculo dessa velocidade de propagação.

$$t_{\text{esvaziamento}} = n_{\text{carMax}} * \frac{3600}{F_{\text{sat}}} \quad (\text{Eq.2})$$

$$t_{\text{percurso}} = L_{\text{link}} / V_{\text{livre}} \quad (\text{Eq.3})$$

$$t_{\text{propag}} = t_{\text{esvaziamento}} - t_{\text{percurso}} \quad (\text{Eq.4})$$

$$V_{\text{propag}} = L_{\text{link}} / t_{\text{propag}} \quad (\text{Eq.5})$$

Com n_{carMax} sendo o número máximo de veículos admitidos pelo link e L_{link} o comprimento do link. Valendo observar que, em segmentos de entrada da rede essa lógica não deve ser implementada, pois teoricamente eles possuem capacidade infinita de represamento por possuírem segmentos ligados à montante que sejam passivos de serem bloqueados. Segmentos que não possuem veículos empilhados não geram propagadores.

A cada passo de simulação um propagador é criado e ele e todos os outros propagadores existentes são deslocados à velocidade de propagação na mesma rotina em que os veículos são deslocados, ou seja, antes da rotina de cálculo do fim da fila.

Com essa modelagem desenvolvida para o GenPolis o intuito é o de se respeitar explicitamente os fluxos de saturação medidos e os tempos de dissipação de fila, que são os dados que vão impactar no comportamento da posição do fim da fila, variável essencial para a determinação da disponibilidade que cada segmento tem de receber veículos ou não a cada passo de simulação.

Sequência do Cálculo da Fila

Concluindo, podemos dividir a população de elementos móveis contida no segmento em três, aquela composta pelos veículos que estão na pilha esperando pela sua liberação, aquela composta pelos veículos que estão percorrendo o link e aquela composta pelos propagadores de espaço, Cada um destes conjuntos terá uma parcela de responsabilidade no cálculo do ponto em que se encontra o fim da fila a cada passo de simulação.

Inicialmente, tendo sido exatamente contabilizados todos os veículos empilhados o fim da fila F é calculado da seguinte maneira:

$$F = n_{carPilha} * tam_{carro} / n_{pistas} . \quad (Eq.6)$$

Com n_{pistas} sendo o número de pistas do segmento, lembrando que esta divisão é feita, pois no GenPolis, o link, independente do número de pistas que a contraparte real possui, constitui um único canal homogêneo de passagem de veículos onde não há diferenciação de pistas.

Em seguida, todos os contadores armazenados nos propagadores são somados e multiplicados pela relação tam_{carro} / n_{pistas} , e, por último é executada a rotina de adição dos veículos ainda em movimento, e que logicamente não estão na pilha. Por fim a cada passo de simulação o trem de propagadores tem sua extensão controlada seguindo duas regras:

i. Após a movimentação dos propagadores o propagador que estiver localizado após o fim da fila atual é eliminado

ii. Quando um propagador verde é o propagador eliminado, o programa entra em um ciclo de eliminação de propagadores até encontrar um propagador vermelho.

A regra *i* traduz o fato de que não há necessidade de se haver um semáforo móvel em propagação após o fim da fila se o sentido de sua existência é a própria fila, e a regra *ii* é análoga ao fato de que uma sequência de semáforos móveis verdes que chega ao fim da fila representa que o novo fim da fila está mais à frente, pois o espaço referente a eles é um espaço vago.

Abaixo, na figura 17 podemos ver parte de uma tela do simulador em funcionamento seguido de legendas que ajudam a clarificar o funcionamento do cálculo das filas e propagação de espaços

- A** – Link Bloqueado (fim da fila no final do segmento);
- B e F** – Espaço se Propagando;
- C** – Carros em Movimento;
- D** – Fim da última Fila do segmento;
- E** – Segunda Sub-Fila;
- G** – Primeira SubFila

Figura 17: Detalhe legendado do ambiente de simulação

5.4.4 Manipulação de Demanda

Nesta seção será explicado como o simulador é alimentado, ou em outras palavras como são inseridos os veículos na malha virtual.

Curva de Inserção e Pontos de Entrada

Para podermos simular uma situação de trânsito real, e como já fora abordado no início deste trabalho, é necessário que se pense em atribuir ao protótipo uma demanda dinâmica que permita simular os efeitos de carregamento e descarregamento da rede. Logo para este projeto foi escolhido um sistema de inserção de veículos que represente não somente o momento de pico de tráfego mas sim todo o período em que se dá o aumento e a diminuição da demanda, fazendo com que o perfil de entrada de veículos na malha virtual lembre situações reais como as observadas na Figura 1, e então, a situação típica a ser simulada é aquela cuja curva de inserção (em número de veículos por segundo) se assemelha a uma gaussiana, sendo que esse processo de inserção faz uso de técnica de truncamento semelhante àquela relativa à liberação de veículos por meio dos links. Em outras palavras, se de acordo com a curva, em um ciclo de simulação devem ser inseridos 2.5 veículos, e no próximo ciclo, 2.7, no primeiro ciclo serão adicionados 2 veículos e no próximo 3, sobrando 0.2 para futuros truncamentos (Figura 18).

Figura 18: Perfis contínuo (vermelho) e discreto (azul) de inserção de demanda.

O processo de cálculo da inserção no simulador consiste primeiramente em, ainda no construtor, somarmos todos os fluxos (veq/h) dos links de entrada, que é então multiplicado pelo número de horas reais que a simulação vai representar para termos então o número total de veículos a serem simulados. Os veículos poderão ser inseridos no simulador de duas maneiras distintas, primeiramente, e em quase

sua totalidade, pelos próprios links de entrada que constituíram essa soma e que representarão a grande maioria dos veículos que apenas passam por qualquer sub-rede viária real e, complementarmente, em uma parte menor, pelos links internos, sendo esses veículos por sua vez aqueles que representam os veículos que saem de pólos geradores de tráfego como estacionamento, shoppings, escritórios e vagas de rua homogêneas distribuídas pela rede. Normalmente esse segundo contingente é modelado através de links de entrada que se encontram dentro da rede e representam esse pólos geradores mencionados, mas para o estudo de caso desse trabalho, foi avaliado ser desnecessário simular tais links, pois a coleta de dados que seria destinada a eles ocuparia um tempo importante que foi destinado a executar as medidas de segmentos mais relevantes. Então a medida adotada para inserção interna foi a de 2% (segundo orientação de Luis Vilanova, especialista em trânsito da CET-SP, em 2011).

Para regular isso, um fator deve ser definido no início da simulação para que se crie uma proporção entre o número de veículos que estão apenas de passagem e o número de veículos que inicia seu movimento dentro da rede. Determinado esse fator, a inserção ao longo da simulação de cada veículo segue duas regras distintas:

- Quando é sorteado que o veículo deve ser inserido na malha através de um link de entrada, o link por onde esse veículo vai ser adicionado é sorteado através do uso de uma roleta em que cada link de entrada possui uma fatia proporcional ao seu fluxo.
- Quando é sorteado que o veículo deve ser inserido na malha através de um link interno, é feita a mesma coisa, mas nesse caso a fatia é proporcional ao peso do segmento na malha interna, que é dado pela multiplicação de seu fluxo pelo seu tamanho.

No caso dos links de entrada, a inserção é sorteada de acordo com o fluxo porque estes segmentos são tratados como se tivessem comprimento indeterminado, pois representam toda a demanda de origem externa ao sistema e que envolve inúmeros links não simulados que estariam contribuindo de forma virtual para alimentar o sistema. Enquanto isso, no caso dos links internos, os sorteios são feitos de acordo com o peso dos segmentos da malha interna, pois como essa inserção é feita de forma homogênea, deve-se considerar o tamanho de cada segmento além do fluxo deste.

É importante lembrar que para ambas as possibilidades de entrada ignora-se o fato do segmento estar bloqueado ou lotado e trata-se a inserção como um represamento que, no caso dos links de entrada representa os veículos que estão esperando para entrar nas vias externas à malha simulada e, no caso dos links de entrada a situação é semelhante à tentativa de saída de um estacionamento quando à sua frente a rua encontra-se com o trânsito parado.

Carregamento Inicial

Toda simulação precisa ter suas condições iniciais bem definidas para que não haja transitórios anômalos. No caso de uma simulação de trânsito, qualquer período que se deseje representar traduz uma situação real ao longo de um determinado período, e em situações reais os veículos e o trânsito não simplesmente aparecem na malha viária, pois no momento em que se começa a observar o trânsito, muitas coisas já estavam acontecendo, como veículos no meio de seus percursos e até mesmo congestionamentos. Então, ao invés de se regular manualmente essas condições até que se encontre um resultado satisfatório, a solução encontrada foi deixar a própria malha virtual criar as condições iniciais que recriem a situação do sistema viário simulado no instante em que a simulação de fato é iniciada, ou seja, é dado um tempo para o sistema entrar em regime, quando de fato é iniciada a simulação. Foi escolhido fazer isso, pois a princípio a rede deve possuir uma capacidade de auto regulação de sua população na medida em que, ao mesmo tempo em que determinadas quantidades de veículos entram na rede, outra quantidade de veículos sai dela.

Assim, para evitar transitórios abruptos, o que se faz neste sistema é usar o valor inicial da gaussiana escolhida como o ritmo de entrada padrão para cada segundo simulado e, ao longo dessa simulação prévia, após determinados períodos são feitas leituras do número de veículos (normalmente dois minutos) e comparadas com as leituras dos períodos anteriores. Quando a diferença percentual de contagem entre as últimas duas leituras é tal que não supera pequenos valores definidos (1% a 5%) previamente à simulação o sistema tem naquele momento a sua condição inicial satisfeita e começa então a simulação pretendida, com o relógio de inserção percorrendo a gaussiana de fato.

Retirada ou Recolocação de Veículos

Mesmo com dados confiáveis que remontam uma situação real, não há garantia de que a malha virtual alimentada por esses dados não ficará instável e apresente um de dois comportamentos extremos, a saber, o quase completo esvaziamento da rede durante a simulação ou o estabelecimento de um congestionamento total da rede, quando tanto uma situação quanto outra não correspondesse à realidade no momento da coleta de dados.

Nesse caso, para regular o sistema haverá um parâmetro que definirá uma proporção de reinserção ou de retirada de veículos e que servirá assim para ajustá-la ao comportamento da rede real, sendo que, no primeiro caso, esse parâmetro desafogaria a rede e, no segundo caso, iria fazer alguma manutenção para mantê-la com uma população de veículos mais próxima àquela observada na realidade. Até o fim do projeto não se julgou necessário realizar esse ajuste, entretanto isso não significa que isso não venha a ser feito no futuro.

5.4.5 Programação Semafórica

Nesta seção será retratado o processo utilizado pela CET para a determinação do funcionamento dos semáforos na cidade e que, conseqüentemente é o adotado para o mesmo fim neste simulador. O processo em questão é o método dos graus de saturação (VILANOVA, 2005a) que se baseia nos níveis de ocupação e uso das capacidades das vias para determinar tempos ótimos de ciclo e de verdes para cada movimento, fazendo uso de índices experimentais de ajuste para permitir ao sistema mais ou menos flexibilidade no que diz respeito às flutuações comportamentais da rede viária.

Descrição da Carta de Tempos

Considerando como exemplo um cruzamento com dois movimentos transversais, temos alguns períodos de tempo a serem considerados que vão além de simplesmente discretizar todo o ciclo como verde ou vermelho. Aqui vale explicar sequencialmente qual é o perfil do fluxo em vias concorrentes.

Partindo de um sinal vermelho em um segmento com um número considerável de veículos, quando este passa para verde existe um período de transitório em que o fluxo de veículos passa de nulo até o nível de saturação. Como

já dito, o fluxo de saturação é a máxima vazão que um segmento pode dar aos veículos no seu ponto de saída (linha de retenção). Então temos um perfil inicial de fluxo que se assemelha a uma rampa que parte do nível nulo e sobe até um patamar grampeado no valor do fluxo de saturação. Para fins de simplificação de cálculo de tempos de verde e de simulação, este período de transitório pode ser dividido em duas partes, sendo a primeira um tempo perdido do movimento com fluxo virtual zero e a segunda metade o tempo aproveitado com fluxo virtual igual ao de saturação.

Ao fim do movimento, no momento em que o sinal troca de verde para amarelo, existe um tempo de acomodação em que é esperado que alguns veículos ainda cruzem a linha de retenção apesar do sinal de bloqueio iminente. Dessa forma, neste momento do ciclo se observa uma rampa contrária à anteriormente descrita, onde se aplicam as mesmas simplificações relativas a tempo perdido e aproveitado. Dessa forma a simplificação leva a um perfil de fluxo que se assemelha a um pulso de duração igual ao tempo de verde dimensionado, diminuído dos tempos perdidos no início e no fim do movimento.

Além dos perfis de fluxo, ao serem considerados os dois movimentos concorrentes, para fins de segurança é necessário dimensionar um tempo de vermelho comum às duas vias para tentar fazer com que nenhum dos veículos que atravessam o cruzamento no transitório final do ciclo colidam com veículos do outro movimento que estarão no início do seu ciclo. A esse período é dado o nome de vermelho de limpeza.

Assim, ao final temos que é intrínseco a todo sistema de vias concorrentes certa quantidade de tempo sem movimentação efetiva, chamado de tempo morto, que é retratado por dois períodos (podendo ser mais do que dois no caso de cruzamentos com mais movimentos concorrentes) compostos cada um pela soma do tempo perdido no fim de um movimento, o tempo de vermelho de limpeza e finalmente o tempo perdido do início do outro movimento (VILANOVA, 2005c).

Para maior clareza, a Figura 19 consolida toda a descrição da carta de tempos de uma forma ilustrativa, valendo citar que na realidade as transições são mais abruptas e estão representadas desta forma para facilitar a visualização da carta.

Figura 19: Perfil de Fluxo em ambos sentidos de uma linha de retenção e suas respectivas cartas de tempos. Acima a linha azul se refere à carta de tempos A abaixo e a linha laranja se refere à carta de tempos B abaixo. As linhas contínuas se referem ao comportamento real enquanto as linhas pontilhadas se referem ao modelo com as devidas simplificações.

Cálculo da Carta de Tempos

Como já citado, o método utilizado para os cálculos de temporização de semáforos é o método dos graus de saturação. O grau de saturação de um link, que varia de 0 a 100%, significa o quanto da capacidade da via deverá ser utilizado para a passagem efetiva de veículos. Assim, quanto mais próximo de 100% esse índice menor é a janela de tempo que sobra para os veículos que passarem acima do nível esperado, e desse modo menor é o tempo de verde alocado e consequentemente menor é o tempo de ciclo calculado. Por outro lado, quanto menor o grau de saturação, mais tempo de verde do que o necessário será destinado à via, de forma que maior será o verde ocioso do sistema quando o fluxo de veículos estiver no nível esperado, mas que servirá para acomodar eventuais fluxos adicionais não contabilizados nas observações. Para São Paulo, de acordo com a experiência dos especialistas, os graus de saturação recomendados encontram-se na faixa de 80% a 90%. (VILANOVA, 2005), supondo então flutuações de fluxo que variam dos 11% aos 25% acima das medidas.

Para definir o tempo de ciclo de uma rede inteira, deve-se determinar qual a intersecção crítica desta, ou, em outras palavras, aquela que possuir os links críticos com o maior fluxo de veículos no sistema. Uma vez encontrada essa intersecção, deve-se definir o nível de saturação para seus links montantes, em seguida calcula-se a taxa de ocupação de cada um deles (Eq.7). Essas taxas serão então divididas

pelos respectivos graus de saturação (Eq.8) e os resultados serão somados para a obtenção da porcentagem total de verde demandada a cada ciclo de operação do semáforo (Eq.9).

$$Ocup_i = \frac{Fluxo_i}{FSat_i} \quad (Eq.7)$$

$$Pvd_i = \frac{Ocup_i}{GSat_i} \quad (Eq.8)$$

$$Pvd_{tot} = \sum Pvd_i \quad (Eq.9)$$

$$Tciclo = \frac{Tmorto}{1 - Pvd_{tot}} \quad (Eq.10)$$

$$Tvd_i = Tciclo * Pvd_i \quad (Eq.11)$$

De forma intuitiva, com o que temos até agora, se pensarmos no ciclo inteiro simplificado de um semáforo com dois movimentos, temos três períodos a considerar, o tempo de verde do primeiro movimento, o tempo morto e o tempo de verde do outro movimento. Assim, após calcularmos a porcentagem de verdes necessária para cada movimento e somarmos ambas, o que faltar, se faltar, para se atingir os 100% de capacidade do tempo de ciclo será a parte destinada ao tempo morto.

Uma vez calculado o tempo morto e tendo em mãos a porcentagem de tempo de ciclo referente a ele chega-se diretamente ao tempo de ciclo desejado para os graus de saturação selecionados (Eq.10) e assim que forem aplicadas as porcentagens teremos o tempo de verde efetivo para cada movimento do cruzamento (Eq.11) e, conseqüentemente, a carta de tempos simplificada. Por último, uma vez tendo os tempos perdidos e aproveitados do início e fim de cada movimento, é possível traçar a carta de tempos real incluída dos tempos de amarelo.

Para o caso da soma das porcentagens de verde ser maior que 100%, o cruzamento encontra-se saturado e é necessário grampear o tempo de ciclo no maior permitido (típicamente 120 segundos para o caso de São Paulo), subtrair o

tempo morto, e atribuir os tempos de verde ao tempo que restar de acordo com as porcentagens Pvd_i obtidas.

Tempos de Ciclo Máximo e Verde de Segurança

Ao se calcular os tempos de ciclo e as distribuições de tempos de verde deve-se tomar alguns cuidados baseados principalmente na experiência dos especialistas de trânsito. Esses cuidados têm cunho mais prático e são alguns dos que levam em consideração aspectos psicológicos no controle de trânsito.

Primeiramente, independente do que os cálculos forneçam como os tempos de verde para quaisquer cruzamentos, eles não podem ser muito curtos, isso se deve ao fato de que é necessário impedir situações em que motoristas que saibam do tempo curto de verde tentem chegar ao cruzamento e atravessá-lo com pressa e de forma arriscada, no intuito de evitarem ficar parados por mais um ciclo inteiro esperando o próximo semáforo verde curto. Assim, convencionou-se o período mínimo possível de qualquer sinal verde em teoria como sendo o de 12 segundos (segundo recomendações de VILANOVA, especialista da CET, em 2010).

Além disso, tempos de ciclo demasiadamente grandes podem ser arriscados na medida em que confundem e desconfortam motoristas e pedestres que esperam em um farol vermelho muito longo. Dessa maneira, recomenda-se (VILANOVA, 2005b) adotar um tempo máximo de ciclo da ordem de 120 segundos, não podendo este ultrapassar os 180 segundos.

5.4.6 Algoritmo Genético

Nesta seção será explicado com algum detalhamento o algoritmo genético usado no GenPolis.

Elementos do AG

Para que durante a seção de resultados se possa entender melhor o que será descrito, essa seção apresentará a analogia entre cada elemento do algoritmo genético e sua contraparte real no processo de seleção natural.

Cromossomo

No caso deste projeto, cada conjunto de valores que constitui o código genético de uma solução é uma linha de código binário salva em arquivos que reunirão várias delas. O código será composto pelos seguintes parâmetros:

- Graus de saturação de cada segmento que for aproximação de semáforo;
- Defasagens para cada nó da rede.

Neste projeto, como os graus de saturação recomendados (VILANOVA, 2005) situam-se entre 80% e 90%, optou-se por trabalhar com 16 níveis de discretização (4 bits) entre 76% e 91%. Enquanto isso, como o tempo de ciclo máximo adotado para o estudo de caso foi o de 120 segundos, optou-se trabalhar com 128 níveis de discretização de defasagem (7 bits). Na Figura 20 se encontra um cromossomo exemplo usado para a otimização da rede do estudo de caso, nela existem seis semáforos e treze aproximações, totalizando 42 bits para as defasagens (6 semáforos x 7 bits/semáforo) e 52 bits para os graus de saturação (13 aproximações x 4 bits/aproximação).

```

CODIGO 01
Defasagens
100010100100110001100000010000100000011111
Graus de Saturacao
0100011000001000001010001110101010101100011001111100

```

Figura 20: Exemplo de cromossomo usado pelo GenPolis na codificação de um plano de temporização semaforica para o estudo de caso

Espécime e Espécie

Cada solução que for parametrizada por um cromossomo, ou seja, cada configuração da rede semaforica que for simulada constituirá um espécime, um indivíduo, que será avaliado, selecionado e, dependendo do seu desempenho, poderá participar de um cruzamento e mutação. No GenPolis, os indivíduos usados para representar uma mesma malha viária virtual simulada constituirão a espécie, que, em outras palavras, será o estudo de caso.

Processo de Seleção

O tipo de seleção escolhida foi a baseada em janelamento onde os mais fortes serão mais favorecidos e para que assim aumente a pressão evolutiva e se permita rápidas convergências. Espera-se que haja grande diferenciação entre as alternativas da população inicial, de modo que aquelas que apresentarem desempenhos extremamente ruins já sejam eliminadas no começo. Existe alguma

dúvida se isso pode levar a convergências prematuras, mas por enquanto a prioridade é provocar a execução de um algoritmo genético um pouco mais rápido, dados os longos tempos de execução. Então outros métodos de seleção serão testados no futuro.

Técnica de Cruzamento Escolhida

Neste Algoritmo Genético a técnica de cruzamento escolhida será o corte duplo circular, representado na Figura 21, onde ao filho será atribuída uma metade do cromossomo de cada pai, sendo que o ponto de corte será definido aleatoriamente para cada cruzamento. Abaixo se encontram alguns exemplos ilustrativos desses cruzamentos.

Mutação

Cada novo indivíduo da nova geração poderá sofrer alterações simples e aleatórias em cada um dos pedaços do seu código. Então, cada bit do cromossomo terá uma probabilidade padronizada (normalmente definida entre 1% e 4%) de trocar de valor, efeito também representado na Figura 21.

Figura 21: Exemplos de corte duplo circular para cruzamento de cromossomos, com uma mutação no terceiro caso (bit em vermelho)

Índices de Mérito

Existem diversos dados passivos de serem eleitos para realizar a avaliação de cada indivíduo simulado. Apesar da alta correlação que há entre eles, avaliou-se executar o algoritmo genético usando todos, um de cada vez, para identificar qual o viés do conjunto de soluções encontrada para cada um. Nesta seção então são descritos cada um dos índices de mérito que o simulador permitirá aferir.

Soma total dos atrasos de percurso

Cada segmento possuirá um tempo de percurso livre que dependerá da velocidade máxima permitida nele e seu de comprimento. Logo qualquer veículo que demorar mais do que este tempo para cruzá-lo de ponta a ponta já estará somando um atraso que penalizará a simulação. Assim, a cada ciclo (segundo) de simulação a quantidade de veículos atrasados (empilhados) é somada a um contador até que a simulação acabe. Este valor será a consolidação do atraso de todos os veículos da simulação e servirá para classificar os cromossomos. Essa é a medida de desempenho mais importante e mais comumente utilizada devido ao seu desempenho, pois diz diretamente por quanto mais tempo os congestionamentos mantiveram os condutores na rua.

Número de paradas

Na prática, também deve ser considerado o custo psicológico pago pelo motorista, principalmente por aquele que se encontra em redes congestionadas, ao se realizar movimentos intermitentes de aceleração e frenagem. Além disso, essa movimentação incorre também no aumento da emissão de poluentes e portanto, pode ser considerada como medida complementar ao atraso se multiplicada por um fator arbitrário de até 30 (como usado no simulador SIRI, em Vilanova, 2009)².

No GenPolis o número de paradas de cada veículo é incrementado uma vez em que este entra na fila e sucessivas vezes para cada vez em que passar um tempo de ciclo inteiro e ele ainda estiver localizado no mesmo segmento, independente da posição.

² Durante o estudo de caso, após algumas simulações de teste com planos aleatórios o valor determinado para esse fator foi de 20 pois dessa forma a parcela do fitness relativa à soma das paradas se assemelhava à dada pela soma dos atrasos.

Penalização por bloqueio de transmissão (queue spillback)

Assim como com o contador de vias congestionadas, aqui as penalidades podem ser conferidas aos indivíduos a cada segundo dependendo da extensão dos segmentos que se encontram bloqueados, ou seja, cujo fim da fila se encontra na entrada deste, de modo a não permitir a entrada de mais veículos, fazendo a fila se propagar além do comprimento do próprio segmento.

5.4.7 Testes Iniciais de Funcionamento

Algumas simulações de teste foram rodadas com a rede da Figura 10 para avaliar a sensibilidade do simulador às curvas de demanda impostas e o comportamento das filas. Uma curva de demanda com um pico curto foi inserida com a janela simulada representando uma hora de tempo real (curva azul).

No geral o comportamento da rede mostrou-se dentro do esperado, apresentando uma defasagem de carregamentos e de pico e outra de descarregamento atrasadas em relação ao perfil de inserção de demanda escolhido.

Abaixo, na Figura 22 podemos observar uma análise de sensibilidade do modelo em que a rede recebe um mesmo curto perfil de demanda com diferentes amplitudes para que se observe o comportamento de carregamento e descarregamento da rede de teste.

Figura 22: Perfis de inserção de veículos em veículos equivalentes por segundo (acima e perfis de comprimento somado de filas em metros para cada perfil de demanda (abaixo/em vermelho))

Pelos gráficos é possível notar que existe uma mesma defasagem de aproximadamente 12 minutos entre o valor de pico de inserção de demanda e o valor de pico de soma de filas medido na rede. Além disso uma vez que a simulação passa pelo valor de pico e o ritmo de inserção de carros cai rapidamente, o ritmo de descarregamento da rede (a inclinação das curvas vermelhas) é o mesmo pois isso depende do fluxo de saturação nos pontos de saída da rede

5.4.8 Teste de Simulação Microscópica

Apesar do modelo como foi explicado já estar suficiente para a aplicação pretendida, no intuito de tentar deixar o simulador mais semelhante à realidade, foi desenvolvido um modelo microscópico incipiente baseado em atenuação da velocidade de locomoção em função de distância a barreiras efetivas ao movimento (tanto linha de retenção em farol vermelho quanto o veículo da frente). Essa lógica foi feita para simular de forma indireta as frenagens e acelerações onde foram usados os parâmetros descritos a seguir, valendo ressaltar que todo o cálculo dessa atenuação é feito frente à velocidade máxima do segmento e que o leitor pode conferir o código no CD-ROM que acompanha a obra:

Distância de Frenagem com Semáforo Vermelho (dFV): No caso do segmento estar com sinal vermelho à frente, esta é a distância à linha de retenção em que o veículo mais a frente no segmento inicia a frenagem

Atenuação de Velocidade com Semáforo Vermelho (aVV): Este é o fator, de 0 a 1, responsável pela atenuação da velocidade máxima (V_{Max}) do veículo no segmento em função da distância em que ele se encontra da linha de retenção (dLR). Quanto menor for essa distância, menor é a velocidade. Esse fator indica ao mesmo tempo qual a velocidade mínima de deslocamento admitida no segmento ($(1-aVV) \cdot V_{max}$), uma vez que sem isso o tempo de frenagem seria praticamente infinito pois o veículo nunca chegaria de fato à linha de retenção, pelo fato da lógica se assemelhar a uma soma de PG infinita de razões menores que 1. Assim, quando essa expressão de cálculo de deslocamento faz com que o veículo ultrapasse a linha de retenção, por uma regra simples ele é mantido logo atrás dela.

A partir dessas duas variáveis é possível calibrar o movimento de frenagem frente a um farol vermelho com a seguinte equação, que a cada passo de simulação indica quanto o veículo se desloca (D):

$$D = V_{max} * (1 - a_{VV} * (1 - \frac{dLR}{dFV})) \quad (\text{Eq.12})$$

Fator de Segurança (fSG): Este parâmetro é um fator que, multiplicado pela velocidade máxima do segmento estabelece uma distância segura que, assim como o dFV, determina a distância inicial de frenagem, mas em relação ao veículo da frente

Atenuação de Velocidade com Carro à Frente (aVC):: Assim como o parâmetro aVV, este é responsável por uma atenuação de velocidade, mas em função da distância do veículo à frente (dVF) e o estabelecimento de velocidade mínima de deslocamento ((1-aVC)*Vmax).

Da mesma forma o deslocamento para um passo de um segundo de simulação é calculado da seguinte maneira:

$$D = V_{max} * (1 - a_{VC} * (1 - \frac{dVF}{fSG * V_{max}})) \quad (\text{Eq.13})$$

Distância de Largada (dL): em uma fila quando o veículo apontado encontra-se parado, este parâmetro define a que a distância o veículo à frente dele precisa estar, logo após iniciar seu movimento, para que este possa iniciar seu deslocamento.

A partir da variação desses fatores foi possível experimentar a sensibilidade que um simulador microscópico pode ter aos seus parâmetros. Foi notado que pequenas alterações faziam grande diferença no comportamento das filas. Então após um longo processo de calibragem, a simulação, com captura de tela observável na Figura 23 ficou visualmente plausível com a seguinte parametrização:

$$dFV = 100; \quad aVV = 0.8; \quad fSG = 4; \quad aVC = 0.99; \quad dL = 4$$

Assim pôde-se observar de uma forma mais verossimilhante a formação de filas em linhas de retenção, a propagação delas (como ondas de choque) após a abertura do sinal verde e os efeitos do acúmulo de várias filas uma após a outra em situações de congestionamento.

Apesar deste não ser o convencional modelo veículo-seguidor (ROTHERY, 1992 e ZHANG, 2005) onde o deslocamento se dá explicitamente com o uso de aceleração e frenagem, pôde-se observar que, pelo fato de com o uso dessa formulação não ser necessário o uso de propagadores, a execução da simulação ficou consideravelmente mais leve e rápida. A partir desse ponto, para que o simulador se tornasse microscópico de fato seria necessário alterá-lo para representar várias faixas e incluir procedimentos para troca de faixas.

Figura 23: Captura de tela da simulação com modelo microscópico testado

6 ESTUDO DE CASO

Após a programação do simulador e do algoritmo genético, em reunião com especialistas da CET-SP, foi escolhida uma rede que possuísse semáforos com apenas duas fases para ser usada como estudo de caso.

Na mesma reunião foi proposto o uso de ambos os modelos, o mesoscópico, que já estava pronto e o microscópico que estava sendo experimentado. A decisão foi a de usar apenas o modelo mesoscópico por ser mais simples e por já representar a propagação de filas enquanto que seria muito difícil fazer a calibragem do mesmo efeito no modelo microscópico.

6.1 Rede Escolhida

O estudo de caso selecionado foi o de um conjunto de seis semáforos de duas fases em uma via coletora de mão única na zona Sul de São Paulo, a Avenida Padre Antônio José dos Santos, no bairro do Brooklin, em São Paulo (Figuras 24 e 25). Os semáforos da rede ficam nos seguintes cruzamentos:

• Av. Padre Antônio José dos Santos	X	Av. Portugal
• Av. Padre Antônio José dos Santos	X	Rua Nova York
• Av. Padre Antônio José dos Santos	X	Rua Califórnia
• Av. Padre Antônio José dos Santos	X	Rua Ribeiro do Vale
• Rua Ribeiro do Vale	X	Rua Indiana
• Av. Padre Antônio José dos Santos	X	Rua Guaraiúva

A escolha foi feita primeiramente por esta via ter, à época do estudo, a maior rede sincronizada de duas fases que a CET colocou à disposição para estudo de caso, o que favorece o uso do AG por representar uma explosão combinatória. A via também foi escolhida por ser uma coletora, o que a evidencia como sendo pressupostamente mais turbulenta e congestionada. Ela é usada por uma demanda que precisa principalmente acessar rapidamente a Avenida Santo Amaro, em ambos os sentidos ou seguir em direção ao Aeroporto de Congonhas, e assim possuir alta capacidade, possuindo três faixas em toda esta extensão. Além disso, esta via possui intersecção semaforizada com a rua Ribeiro do Vale, muito usada como via de acesso à Avenida dos Bandeirantes em ambos os períodos de pico.

Outro motivo importante foi o fato dessa via já ser bem conhecida e usada, e frequentemente saturada, pelas pessoas por ser uma alternativa aos congestionamentos da Av. dos Bandeirantes (sentido Rodovia dos Imigrantes) ou da Marginal Pinheiros logo após a ponte do Morumbi, perto da Av. Eng. Luis Carlos Berrini, importante centro comercial da cidade.

Figura 24: Imagem de Satélite da Região Simulada

Figura 25: Mapa Correspondente da Região Simulada

6.2 Coleta de Dados

Esta sessão passa por mais práticas acerca da coleta de dados realizada para o estudo de caso. De início vale dizer que a atividade foi realizada apenas pelo pesquisador com suporte eventual de outras pessoas, de modo que algumas extrapolações e compactações tiveram de ser realizadas, mas sem comprometer a qualidade das leituras de forma a inviabilizar o estudo. Mais especificamente, a compactação da rede se deu por meio de um reconhecimento inicial para que se pudesse montar uma tabela de parâmetros físicos de cada segmento, como placas de pare, valetas, pistas estreitas e aclividade/declividade, de forma a determinar pequenos grupos de segmentos que possuíssem características semelhantes. Isso foi feito no intuito de replicar alguns ou todos os parâmetros de alguns segmentos para outros, de forma a reduzir o número de medidas necessárias para se construir a rede. No caso do fluxo de saturação, com essa compactação, foi necessário realizar a medida de apenas 32% dos segmentos da rede enquanto que no caso do tempo de percurso essa taxa foi de 45%.

Revisando, os parâmetros que foram obtidos são os mesmos aplicados nas otimizações da CET-SP, a citar: fluxo de saturação, o parâmetro que diz qual a máxima vazão instantânea que um segmento pode ter em sua linha de retenção, medida em veículos/hora; fluxo, que diz a frequência com a qual os veículos passam pela linha de retenção em um determinado período, dado também em veículos/hora; e o tempo de percurso ideal que é usado para se determinar a velocidade média livre dos veículos que percorrem o segmento. Em seguida, ao leitor serão apresentadas algumas orientações para que se façam as medidas de campo, na forma ideal, e na forma simplificada, no caso da atividade ser realizada sem suporte.

No Apêndice B podem ser conferidas as tabelas que resultaram das coletas de dados, bem como as compactações e extrapolações realizadas, caso o leitor queira reproduzir o experimento.

Primeiramente, para medição do tempo de percurso é recomendável ter a via observada totalmente livre, de maneira que o veículo possa entrar no segmento com sua velocidade normal, percorrê-lo por inteiro e atravessar um sinal verde no fim. Durante as medidas de campo preliminares realizadas para reconhecimento da rede, a cada ciclo de operação semafórica, com sorte foi possível obter apenas uma

leitura dessas, e isso se as defasagens entre o semáforo anterior e o do segmento observado favorecessem essa medida. Normalmente os engenheiros de tráfego possuem equipamentos que permitem a eles atuar no semáforo da forma como preferirem, o que possibilita manipular a rede para que os veículos façam esse percurso livre com uma frequência maior. Então, sem esse equipamento e estando sozinho é possível cronometrar os veículos pelo fato de uma medida isolada ser uma atividade tecnicamente simples, mas o processo de medição em grandes números (10 medidas limpas para cada segmento no mínimo) com muitos segmentos acaba por ser extremamente lento e, no caso, a rede em questão possui mais de sessenta segmentos para serem parametrizados.

Para contornar esse problema foi necessário realizar as cronometragens em horários de baixo fluxo, a partir das 22h, posicionando-se no assento de passageiro de alguns veículos (a citar, um compacto, um *station wagon* e um sedã grande), emprestando os referidos veículos a diferentes condutores em horários com baixo fluxo e pedindo para eles adotarem direções mais ou menos agressivas a cada passagem pela rede. Mesmo desse modo, para alguns segmentos foi necessário fazer aproximações por inspeção visual e pela sensação de velocidade, pois o semáforo fechava antes de se conseguir realizar todo o percurso do segmento e era necessário, durante uma frenagem, projetar o deslocamento livre do veículo, estimando quando este passaria livremente pela linha de retenção.

Alguns outros detalhes devem ser considerados para as medidas, como o tempo necessário para a passagem sobre uma valeta e, no caso do GenPolis que, para segmentos não semaforizados, não trata a questão das vias preferenciais e placas de pare, o tempo de percurso livre pode envolver as reduções e frenagens frente a placas de pare ou para a passagem de um ou dois veículos na via concorrente com fluxo preferencial.

Para medir o fluxo de saturação a recomendação é que se façam histogramas onde a quantidade de veículos que atravessou a linha de retenção no ciclo observado é registrada a cada cinco segundos. Assim, as diferenças de quantidades aferidas entre cada período de cinco segundos permitem o traçado de um perfil de vazão de veículos para intervalos de verde no segmento. A partir desse perfil, determina-se o intervalo onde esse fluxo está em um patamar e a média dos valores nesse período dá o fluxo de saturação daquela medida.

Entretanto essa medida precisa ser feita com duas pessoas, sendo que uma conta as passagens e outra registra manualmente o valor da contagem a cada cinco segundos em uma tabela. Além disso, para essa medida, o segmento observado precisa estar quase cheio e o segmento à jusante não, uma condição que se apresenta em períodos curtos logo antes e logo após horários de pico, devendo-se atender ao fato de que diversos fatores como saídas de estabelecimentos, entradas, embarques e desembarques demorados não incomumente comprometem a leitura de um ciclo inteiro de descarga. A leitura que se tirou desses problemas é a de que, uma vez estando sozinho, não é possível operacionalizar a medida por histograma. Além disso, conseguir registrar patamares consistentes de fluxo de saturação de todos os segmentos se torna uma tarefa muito difícil e pouco eficiente na medida em que as janelas horárias ideais para a coleta desse dado são curtas e que muitas medidas devem ser descartadas por conta de longos bloqueios de faixas.

Além disso, para o estudo de caso, somente foram analisadas vias de duas ou três faixas, logo a resolução da medida era baixa uma vez que entre um momento transição e os períodos de fluxo de saturação ou não se notava grande diferença na contagem de veículos passantes ou, mais notavelmente, nos primeiros cinco segundos já era possível dizer que o segmento estava em fluxo de saturação pois não havia diferenciação alguma com os períodos de alta vazão.

Logo, foi testado um procedimento alternativo consideravelmente mais simples, mas que requer certa familiarização de quem coleta os dados com a via observada, que é o de durante cerca de 15 minutos, por inspeção visual e considerando a velocidade média livre do segmento (baseada na impressão que se tem após algum tempo observando a via), determinar quando um pelotão que está prestes a passar se encontra compactado o suficiente de forma que se possa concluir que o segmento entrará em vazão máxima. Nesse momento o cronômetro é disparado e a contagem é iniciada, devendo durar o tempo que for necessário para o pelotão atravessar a linha de retenção ou até que os veículos que passam estejam a uma distância maior um do outro.

Existem duas considerações importantes acerca dessa medida. A primeira é o fato de que não devem interromper ou invalidar a medida pequenas obstruções típicas da via como rápidos desembarques, embarques, acessos e saídas de vagas na rua, estacionamento, ou supermercados e fatores de redução instantânea de

fluxo como condutores extremamente lentos, dentre eles, por exemplo, os indecisos entre seguir um caminho ou outro, ou os que demoram muito para atravessar o cruzamento por conta de cuidado excessivo com as valetas (comuns no estudo de caso). A outra consideração é a de que não é necessário realizar a medida apenas com o primeiro pelotão que deixa o segmento na abertura de um sinal verde, mas outros pelotões mesmo que menores (a partir de três ou quatro veículos por faixa) que chegarem mais tarde podem ser contados também. Em resumo, pode-se dizer que o bom senso aliado ao conhecimento da via dão àquele que realiza as contagens a capacidade de decidir quando que cada pelotão inicia a sua passagem, quando que ele termina e, dentro desse período, quando o ritmo de passagem é o maior.

A contagem de fluxo, por sua vez é simples, entretanto precisa ser feita em todos os links sem exceção durante todo o período de pico no caso das aproximações de semáforos e no mínimo 10 minutos de leitura para cada trecho menos importante ou não crítico (grosso modo os não semaforizados). Novamente aqui houve a restrição de que com apenas uma pessoa não é possível observar toda a rede logo foi necessário realizar algumas extrapolações.

Como a intenção da otimização é gerar um plano que vise um funcionamento coordenado da rede, é necessário que a parametrização com extrapolações vise traçar um comportamento geral da rede. Para isso foi necessário eleger o cruzamento crítico que seria mais acompanhado, no caso o cruzamento 5, da Padre Antônio José dos Santos com a Rua Ribeiro do Vale, por serem essas as coletoras representativas de dois movimentos principais do região, da demanda que acessa a Avenida Santo Amaro e demanda travessa que segue à Avenida dos Bandeirantes. Em determinada semana, durante terça, quarta e quinta-feira, para os dois períodos de análise, o pico da manhã e da tarde, das 06:00 às 10:00 e das 16:00 às 20:00 respectivamente, uma medida de fluxo de 10 minutos era feita em média a cada 15 minutos, sendo que essas medidas alternavam entre uma medida no cruzamento principal e outra medida em outro cruzamento até que todos os cruzamentos, semaforizados e não semaforizados, tivessem sido contemplados.

Com isso, ao final da semana para cada período de pico, temos três leituras de cada cruzamento ao lado da aproximação de um perfil temporal de demanda da rede baseado nas várias leituras de fluxo realizadas no cruzamento principal. Para a

determinação dos fluxos de cada segmento, inicialmente deve-se determinar o horário exato de pico que essa aproximação indica e usar o respectivo valor de fluxo médio nos segmentos do cruzamento principal. Em seguida, como as medidas dos outros cruzamentos foram feitas em horários variados, estas medidas devem ser multiplicadas por um fator de forma que se aproximem de um valor de pico esperado. Esse fator deve ser calculado para cada uma, obtendo-se, no perfil do cruzamento principal, a razão entre a demanda do horário de pico e a demanda do horário da medida a ser ajustada.

Não há evidências científicas de que qualquer desses métodos alternativos sejam válidos para parametrizar uma simulação, tampouco se pode garantir que sejam precisos (o que não é um fator crítico considerando a volatilidade de qualquer medida na engenharia de tráfego), mas uma vez que os métodos tradicionais não foram viabilizados, recorrer a esse procedimento permitiu dar uma vazão prática aos dados necessários para a continuidade do trabalho.

Como recomendação geral, a primeira coleta de cada parâmetro deve servir principalmente para o observador se habituar com os segmentos observados e o tipo de medida que estiver sendo feito, bem como para obter uma idéia preliminar da faixa de valores com que a rede do estudo opera, valendo lembrar que as alternativas aqui adotadas foram pensadas durante o estudo de caso para sua rede específica, e que podem não ter aplicação nenhuma em estudos de redes diferentes.

6.3 Calibragem

Uma vez com todos os dados coletados e organizados a rede foi parametrizada e simulada algumas vezes usando-se a temporização antiga para checagem do volume de veículos que passava por cada link virtual, com isso percebeu-se a necessidade do uso de um algoritmo de geração de números aleatórios com distribuição regular, o que levou à adoção do *Mersenne Twister* (Matsumoto, 1998). Dessa maneira foi possível observar que o comportamento da rede simulada estava dentro do que se esperava uma vez que os fluxos medidos em cada link de interesse (todas as aproximações) correspondiam aos parametrizados, com desvios de até 7% para cima ou para baixo, sendo que alguns links não semaforizados e com fluxos baixos apresentaram desvios maiores.

No Apêndice B encontram-se os ajustes realizados com maiores detalhes.

6.4 Otimizações

Após a coleta de dados, foi feita a parametrização manual da rede para cada período de pico (manhã e tarde) de forma que todos os segmentos de entrada da rede possuísem os fluxos de saturação aferidos, os tempos de percurso medidos, os fluxos calculados e as porcentagens de conversão observadas, bem como o perfil de demanda obtido.

Seguindo as recomendações encontradas no manual do software Transyt-7F (Hale, 2008), o tamanho da população do AG escolhido foi de 20 indivíduos e o número de gerações escolhido foi 200, sendo que cada cromossomo ou indivíduo era simulado três vezes para que se retirasse o fitness médio das execuções e se mitigasse o efeito da volatilidade, resultando em 12000 simulações. Assim no caso da rede escolhida, usando um computador provido de processador Intel Core i5 de 2.4 GHz e 4 GB de memória RAM DDR3 1333 MHz, com a demanda do pico da manhã a execução de cada simulação demora de 0,5 a 0,7s, o que implica em cada otimização demorar cerca de 100 a 140 minutos. A função de fitness escolhida para as otimizações foi a soma do atraso total (**A**) da simulação com o número de paradas (**P**) multiplicado por 20 (**Fit=A+20*P**). A escolha desse valor foi arbitrária, mas fica dentro da faixa de valores observada em outros trabalhos de otimização citados ao longo da obra.

Outras configurações como maiores tamanhos de população (até 100) e gerações (300 ou mais) foram testadas antes, mas o aumento na demora da execução, que chega a 10 horas, não compensa o pouco ganho no resultado da otimização, de cerca de mínimos de fitness apenas até 5% menores.

No caso da manhã, como não foi identificada saturação da rede, foi possível realizar otimizações com alguns níveis de demanda além daquele aferido. No caso então foram escolhidos quatro cenários sobre os quais foram realizadas seis otimizações, a começar por aquele representante da situação observada seguido pelos cenários com 110%, 120% e 130% da demanda aferida. No último cenário, de 130% de demanda, a rede apresentou saturação e os planos otimizados atingiram tempo de ciclo próximo ao máximo de 120 segundos permitido no estudo. Os resultados das otimizações encontram-se na Figura 26.

Uma vez com as seis otimizações realizadas para cada cenário, as cartas de tempo resultantes foram comparadas e o plano foi escolhido de forma arbitrária levando-se em consideração o fitness encontrado e a quantidade de semelhanças com os outros planos.

Figura 26: Curvas de evolução do melhor fitness de cada geração em função do número de gerações. Foram realizadas seis otimizações (cores distintas) para cada nível de demanda (de 100 a 130% do esperado), cada qual com uma semente de pseudo-aleatoriedade distinta.

Uma vez tendo todos os planos otimizados em mãos foi possível executar cada um deles com os diferentes níveis de demanda para testá-los fora da situação para os quais foram concebidos (por exemplo, o plano otimizado de 110% com uma demanda de 130%) e criar a tabela 2 que comparasse, frente aos diferentes níveis de demanda, os desempenhos simulados dos planos otimizados para cada cenário e o resultado do plano anterior para cada cenário também.

Com essa tabela foi possível escolher o plano que ainda apresentasse resultados aceitáveis fora do nível de demanda para o qual foi otimizado, ou seja, não necessariamente o melhor plano para os fluxos de pico observados, mas aquele mais robusto e mais confiável que ainda possa lidar com flutuações de demanda, para ser implantado na rede real.

Tabela 1: Plano antigo da manhã (esq.) e plano atual otimizado (dir.)

tCiclo	Plano Antigo				Plano Atual			
	Manhã							
	72				71			
	tG	tR	TG	pG	tG	tR	TG	pG
Pe. Antonio	0	35	35	49%	33	65	32	45%
Portugal	40	67	27	38%	70	28	29	41%
Pe. Antonio	0	37	37	51%	19	50	31	44%
Nova York	42	67	25	35%	55	14	30	42%
Pe. Antonio	0	39	39	54%	70	29	30	42%
Califórnia	44	67	23	32%	34	65	31	44%
Pe. Antonio	0	37	37	51%	58	17	30	42%
Ribeiro do Vale	42	67	25	35%	22	53	31	44%
Indiana	0	22	22	31%	2	29	27	38%
Ribeiro do Vale	27	67	40	56%	34	67	33	46%
Pe. Antonio	0	37	37	51%	41	3	33	46%
Guaraiuva	42	67	25	35%	8	36	28	39%

Tabela 2: Desempenhos de planos otimizados e do anterior para pico da manhã

Planos	Demanda	Atraso				20*Paradas				Fitness (A+20*P)			
		100%	110%	120%	130%	100%	110%	120%	130%	100%	110%	120%	130%
Planos	100% [56]	139	188	439	1228	222	257	367	657	361	445	806	1886
	110% [71]	153	188	353	620	217	243	304	378	371	430	657	997
	PA [72]	208	379	1521	3179	223	291	633	1114	432	670	2153	4294
	120% [90]	197	231	301	627	216	239	271	362	413	470	572	989
	130% [111]	252	293	375	500	217	239	268	312	469	532	625	813

Legenda: Plano Aplicado, Plano Anterior, [tempo de ciclo]

Ao observar a tabela, percebe-se que a rede não está saturada para as demandas de até 120%, pois não se observa para cada um dos planos uma grande diferenciação no número de paradas conforme a demanda aumenta, o que significa que ainda não é registrada com muita frequência a situação em que veículos demorem mais de um ciclo para atravessar algum semáforo. Já para a demanda de 130% onde foi identificada a saturação, o tempo de ciclo do plano otimizado é de 111 segundos e há um salto no fitness de cada plano frente esse cenário.

Considerando que nos horários de pico, os níveis aceitáveis de flutuação de demanda são de cerca de 10% (informação verbal, CET-SP, 2011), é esperado que os semáforos tenham de lidar tipicamente com demandas de até 110%. Assim, o

plano escolhido foi aquele otimizado para a demanda de 110% (Tabela 1), visto que com esta última consideração este plano mantém níveis aceitáveis de atraso e paradas em demandas consideravelmente maiores, sem comprometer a rede quando o volume de tráfego é igual ao esperado.

Para o pico da tarde, como já foi identificada saturação da rede durante as leituras, somente se otimizou o nível de demanda aferido e, assim como nas otimizações para o pico matutino, foi escolhido o plano otimizado que possuía maior quantidade de semelhanças com os demais. Na Figura 27 há uma síntese com a curva de evolução do fitness, a tabela comparativa entre os desempenhos do plano otimizado aplicado e do anterior e as temporizações dos semáforos para cada um. Os outros planos otimizados, além do escolhido podem ser conferidos no Apêndice C.

6.5 Resultados

Após alguns ajustes, especificados no apêndice C, os planos otimizados encontram-se aplicados na rede real e, analisados por engenheiros da CET-SP, já apresentaram bons resultados em casos isolados de demanda atipicamente alta devido a redirecionamentos de tráfego causados por ocorrências em redes próximas, mas leituras mais conclusivas e acompanhamentos mais consistentes como os que foram feitos à época da coleta de dados ainda precisam ser feitos, pois a implantação foi realizada no período de férias escolares e portanto, a demanda média diária encontrava-se abaixo do esperado.

Mas de forma geral, como esperado, a ferramenta desenvolvida obteve resultados simulados mais expressivos no caso das redes congestionadas e se apresentou como boa alternativa para obtenção de soluções para temporização de semáforos já que os ganhos em relação aos planos anteriores se encontram na faixa de 15 a 30%, dependendo do nível de congestionamento da via.

Além disso, independente de qual plano anterior estivesse aplicado, a ferramenta foi capaz de encontrar bons planos otimizados e coerentes uma vez que seu espaço de busca é delimitado pela restrição dada com o intervalo de graus de saturação de 76% a 91% aceito pela experiência dos especialistas da CET-SP. Vale citar que mesmo dentro destas restrições as simulações apresentavam grande variação de valor de *fitness* entre os cromossomos testados, variação esta contemplada pelo algoritmo genético o que lhe permitiu descartar as piores alternativas e fazer emergir planos de semaforização que fossem pontos ótimos globais.

7 CONCLUSÕES

Apesar da situação de escassez de dados e recursos, e das consequentes aproximações e extrapolações das leituras, os procedimentos aqui relatados e o modelo de simulação adotado foram pensados para que todo o processo de desenho e parametrização de uma sub-rede sincronizada e da sua simulação e otimização com a aplicação de um AG fossem simplificados e possíveis de ser realizados por uma única pessoa, sem comprometer a qualidade do estudo. Nisso os resultados da aplicação real trouxeram bons indícios para que o desenvolvimento de um simulador mais avançado, juntamente à implementação de técnicas de AGs mais elaboradas e com melhores desempenhos, tragam simulações mais próximas à realidade e resultados melhores do que os obtidos com a versão atual.

Além disso, no futuro seria interessante realizar o mesmo estudo, mas dispondo de maior contingente para a obtenção de dados mais precisos que permitam traçados de perfis de demanda dinâmicos em cada aproximação ao invés de fazer isso para apenas um cruzamento principal, extrapolando essa curva para os outros cruzamentos.

Outro resultado interessante alcançado na elaboração do modelo foi o relativo à velocidade da simulação, que, mesmo sem qualquer otimização na codificação, é comparável àquela da aplicação *online* no trabalho de Pohlmann e Friedrich (2011), indicando que a arquitetura aqui desenvolvida favoreceria AGs para a otimização de planos semafóricos em tempo real mesmo sem o uso de computação paralela, que é uma ótima alternativa para a viabilização e aceleração desse tipo de aplicação, dado o caráter paralelo tanto explícito quanto implícito do AG.

Por fim, de modo geral foi importante e bastante imersivo realizar um trabalho que passou pela definição de um tipo especializado de estudo, no caso o de otimização semafórica realizado pela CET, e pela criação de um modelo próprio de simulação que atendesse especificamente a esse propósito. Dessa forma o trabalho acabou tendo uma integração forte entre a implementação e aplicação real, que provocaram aprendizado acerca tanto de questões computacionais quanto de questões práticas nos trabalhos de campo. Nisso o trabalho foi feito sempre mirando a simplicidade, com uma formulação voltada à coleta de dados agregados e que permitiu seguir uma metodologia voltada à realização de extrapolações e estimativas

importantes, o que levou a um modelo mais explícito de simulação e cálculo de filas que aparentemente dispensa processos exaustivos de calibragem e a um estudo de caso enxuto.

7.1 Futuros Desenvolvimentos

A partir deste trabalho, pode-se indicar uma série de melhorias a serem feitas no caso do intuito ser o desenvolvimento de uma ferramenta profissional ou semi-profissional de cálculo de filas e otimização semafórica.

Primeiramente, este protótipo só trabalha com semaforização de duas fases, sendo que grande parte dos semáforos de redes reais possuem mais fases, de pedestre, fases de retorno, de giro à esquerda e outras. Nisso o sequenciamento dessas fases é algo que também influencia na obtenção de bons resultados de temporização semafórica e que também contribuem consideravelmente para explosões combinatórias, que dificultam a obtenção de planos otimizados e favorecem o uso dos AGs.

Além disso, de forma geral, como já comentado, a elaboração de um modelo microscópico simplificado paralelo a partir de uma versão modificada do GenPolis, apesar de criar a necessidade de calibragem de parâmetros não agregados, traria ganhos na obtenção de uma simulação mais realista e transparente aos analistas, bem como ganhos em velocidade por eliminar o uso dos propagadores, que ocupam grande espaço da memória e do processamento destinados ao programa.

Para isso, alguns elementos precisam ser implementados, iniciando pelo modelo carro-seguidor, para que o comportamento das filas não se dê de forma explícita como no protótipo, mas implícita a partir dos parâmetros de simulação microscópica. Uma vez implementado, esse modelo seria necessário fazer a divisão dos segmentos por faixas para que se possam simular as filas de cada uma e não uma única fila agregada. Por último seria necessário aplicar um procedimento que simule as mudanças de faixa, bem como implementar os comportamentos de faixa, que determinam quais faixas de cada segmento que permitem ou não determinados tipos de conversão. Outro ponto interessante seria, uma vez desenvolvido este novo modelo, que as duas arquiteturas poderiam ser usadas para a realização de auto-calibrações através do uso do protótipo atual, que possui um funcionamento e formulação mais explícitos, para a calibragem do modelo microscópico.

Outro ponto importante é o desenvolvimento de um AG mais sofisticado e que atenda principalmente às indicações de Kesur (2009), como visto na seção de AGs aplicados à engenharia de tráfego, bem como a realização de um estudo dedicado somente a determinar quais as proporções entre os números de geração e tamanho de população que devem ser utilizados para que o AG faça melhor proveito do modelo aqui desenvolvido.

REFERÊNCIAS

- AKIVA, M.B.; BIERLAIRE M.; KOUTSOPOULOS H.; MISHALANI R. DynaMIT: a simulation-based system for traffic prediction. **DACCORD Short Term Forecasting Workshop**, artigo 12p, 1998
- AMIACH, M. Controle de tráfego em sistemas expressos. **Monografia (Bacharelado)**, Escola Politécnica da USP, 2010. 121p
- BAZZAN, A.L.C.; AMARANTE, M.B.; AZZI, G.G.; BENAVIDES, A.J. ITSUMO: an agent-based simulator for ITS applications. In: **IEEE ITSC2010 Workshop on Artificial Transportation Systems and Simulation**, artigo 7p. 2010
- BAZZAN, A.L.C.; AMARANTE, M.B.; AZZI, G.G.; BENAVIDES, A.J.; BURIEL, L.S.; MOURA, L.; RITT, M.P.; SOMMER T. Disponível em, <http://www.inf.ufrgs.br/~mrpritt/Publications/P34-ecms-2011.pdf>
Extending traffic simulation based on cellular automata: from particles to autonomous agents, Acesso em: março 2011. 7p.
- BOURREL, E. LESORT, J.B. Mixing microscopic and macroscopic representations of traffic flow hybrid model based on lighthill–whitham–richards theory, **Transportation Research Record: Journal of the Transportation Research Board**, v.1852, 2003. 8p
- BOXILL, S.; YU, L. An evaluation of traffic simulation models for supporting ITS development, **Relatório**. Texas Southern University, p21-67, 2000
- BROWNLEE, J. Clever algorithms – nature inspired programming recipes, **Creative Commons Attribution-Noncommercial-Share Alike 2.5 Australia License**, livro de referências, 436p, 2011
- BURGHOUT, W. Hybrid microscopic-mesosopic traffic simulation, **tese (doutorado)**, **Royal Institute of Technology – Suécia**, p1-62 (185), 2004
- CAVICCHIO, D., Reproductive adaptive plans, In: **Proceedings of the ACM annual conference**, v.1, 11p, 1972
- CANTARELLA, G.E.; PAVONE, G.; VITETTA, A. Heuristics for urban road network design: lane layout and signal settings, **European Journal of Operational Research** Volume 175, Issue 3, artigo, 14p, 2006
- CASCETTA E.; NGUYEN, S. A unified framework for estimating or updating origin/destination matrices from traffic counts, **Transportation Research Part B: Methodological**, v.22, Issue 6, artigo, 19p, 1988

CEYLAN, H.; BELL, M. G. H. Reserve capacity for a road network under optimized fixed time traffic signal control, **Journal of Intelligent Transportation Systems** v.8, Issue 2, artigo, 14p, 2004a

CEYLAN, H.; BELL, M. G.H. Traffic signal timing optimisation based on genetic algorithm approach, ncluding drivers routing, **Transportation Research Part B: Methodological Volume 38**, Issue 4, artigo, 14p, 2004b

CEYLAN H. Developing combined genetic algorithm—hill-climbing optimization method for area traffic control, **Journal of Transportation Engineering**, artigo, 9p, 2006

CEYLAN, H.; BELL, M. G.H. Genetic algorithm solution for the stochastic equilibrium transportation networks under congestion, **Transportation Research Part B: Methodological Volume 39**, Issue 2, artigo, 17p, 2005

DAGANZO, C. F. In traffic flow, cellular automata = kinematic waves, **Transportation Research Part B: Methodological Volume 40**, Issue 5, artigo, 12p, 2006

DALE J.; BLOOMBERG L. Comparison of VISSIM and CORSIM traffic simulation models on a congested network, **Transportation Research Record: Journal of the Transportation Research Board**, Volume 1727, artigo 9p, 2000

DNIT - Departamento Nacional de Infra-Estrutura de Transportes - Diretoria de Planejamento e Pesquisa. Coordenação Geral de Estudos e Pesquisas Instituto de Pesquisas Rodoviárias.Rio de Janeiro, Manual de estudos de tráfego, **manual**, 388p, 2006

EJZENBERG, S. Disponível em, <http://www.sinaldetransito.com.br>, **Reprogramação de semáforos – método de observação de campo**, artigo, 32p, 2005

FELLENDORF, M.; VISSIM: a microscopic simulation tool to evaluate actuated signal control including bus priority. In: **64th Institute of Transportation Engineering Annual Meeting**, 9p, 1994

FRANTZ, D. R. Non linearities in genetic adaptive search, **College of Engineering - Universidade de Michigan**, relatório, 181p, 1972

GASPAR, A.; COLLARD P. From GAs to artificial immune systems: improving adaptation in time dependent optimization, In: **Proceedings of the 1999 Congress on Evolutionary Computation**, artigo, 8p, 1999

GIRIANNNA M.; BENEKOHAL, R. F. Using genetic algorithms to design signal coordination for oversaturated networks, **Journal of Intelligent Transportation Systems**, v.8, Issue 2, artigo, 14p, 2004

GOLDBERG, D. E. Genetic and evolutionary algorithms come of age, In: **Communications of the ACM**, artigo, 7p, 1994

GREENSHIELDS, B. D. A study in highway capacity. **Highway Research Board, Proceedings 14**, artigo, 30p, 1935

GREFENSTETTE, J. J. Optimization of control parameters for genetic algorithms, In: **IEEE Transactions on Man and Cybernetics**, v.16, Issue 1, artigo, 7p, 1986

GUIMARÃES, José O. S. Computação evolutiva na resolução de equações diferenciais ordinárias não lineares no espaço de Hilbert, **Escola Politécnica da USP**, tese doutorado, 211p, 2009

HALE, D. K. Traffic network study tool, TRANSYT-7F, **United States Version, McTrans Center – University of Florida**, manual, 442p, 2008

HELBING, D.; TREIBER, M. Gas-kinetic-based traffic model explaining observed hysteretic phase transition, **Physical Review Letter 81**, artigo, 4p, 1998

KERNER, B. S. Introduction do modern traffic flow theory and control, **Editora Springer**, livro didático, 265p, 2009

KESUR, K. B. Advances in genetic algorithm optimization of traffic signals, **Journal of Transportation Engineering 135**, Issue 4, artigo, 14p, 2009

KRZANOWSKI, R.; RAPER, J. Spatial evolutionary modeling, **Oxford University Press**, livro didático, 265p, 2001

LEBDEH, G.A.; BENEKOHAL, R. Development of traffic control and queue management procedures for oversaturated arterials, **Transportation Research Record: Journal of the Transportation Research Board**, v.1603, artigo, 9p, 1997

LEBDEH, G. A., A.M.ASCE1, Integrated adaptive-signal dynamic-speed control of signalized arterials, **Journal of Transportation Engineering**, artigo, 5p, 2002

LEBDEH, G. A.; BENEKOHAL, R. F. Design and evaluation of dynamic traffic management strategies for congested conditions, **Transportation Research Part A: Policy and Practice**, V.37, Issue 2, artigo, 19p, 2003

LIU, Y.; CHANG, G. An arterial signal optimization model for intersections experiencing queue spillback and lane blockage, **Transportation Research Part C: Emerging Technologies**, v.19, Issue 1, artigo, 15p, 2011a

LIU, Y.; CHANG, G., YU, J. An integrated control model for freeway corridor under non-recurrent congestion, In: **IEEE Transactions on Vehicular Technology**, v.60, Issue 4, artigo, 15p, 2011b

LO, H. K; CHANG E.; CHAN, Y.C., Dynamic network traffic control, **Transportation Research Part A: Policy and Practice**, v.35, Issue 8, artigo, 24p, 2001

LOPES, Denise L. Viabilidade do uso de modelos sintéticos integrados de uso de solo e transportes: estudo de aplicação à cidade de são paulo, **Escola Politécnica da USP**, tese mestrado, 177p, 2003

MATSUMOTO, M.; NISHIMURA T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, **Journal ACM Transactions on Modeling and Computer Simulation (TOMACS) - Special issue on uniform random number generation** v.8, Issue 1, artigo, 25p, 1998

MEDINA, J. J. S.; MORENO, M. J. G.; ROYO, E. R. Traffic signal optimization in “La Almozara” district in saragossa under congestion conditions, using genetic algorithms, traffic microsimulation, and cluster computing, In: **IEEE Transactions on Intelligent Transportation Systems**, v.11, Issue 1, artigo 10p, 2010

MING, S. H. (2008), disponível em <http://www.sinaldetransito.com.br>, Dimensionamento do entreverdes - uma abordagem probabilística, **website instrucional Sinal de Trânsito – CET-SP**, 23p

MITCHELL, M., Genetic algorithms: **An overview**, **santa fe institute**, artigo, 17p, 1995

NEVES, R. P. O., A.L.I.V.E. artificial life in virtual environments, **Escola Politécnica da USP**, tese mestrado, 153p, 2003

NETTO, M. L. **Notas e conceitos expostos em aula durante a aula de Vida Artificial**, em 05/04/2010, 2010

NÖKEL, K.; SCHMIDT, M. Parallel DYNEMO: Meso-Scopic Traffic Flow Simulation on Large Networks, **Networks and Spatial Economics**, **2**, artigo, 17p

OH, S., RITCHIE S. G.; OH, C. Real-time traffic measurement from single loop inductive signatures, **Transportation Research Record: Journal of the Transportation Research Board**, v.1804, artigo 9p, 2002

OWEN, L. E.; ZHANG, Y.; RAO, L.; MCHALE, G. Traffic flow simulation using CORSIM, In: **Proceedings of the 2000 Winter Simulation Conference**, artigo 5p, 2000

PARK, B.; MESSER, C.; URBANIK, T. Initial evaluations of new TRANSYT-7F version 8.1 program, **Transportation Research Record: Journal of the Transportation Research Board**, v.1867, artigo, 6p, 1999a

PARK, B.; MESSER, C.; URBANIK, T. Traffic signal optimization program for oversaturated conditions genetic algorithm approach, **Transportation Research Record: Journal of the Transportation Research Board**, artigo, v.1683, artigo, 10p, 1999b

PARK, B.; MESSER, C.; URBANIK, T., Enhanced genetic algorithm for signal-timing optimization of oversaturated intersections, **Transportation Research Record: Journal of the Transportation Research Board**, v.1727, artigo, 10p, 2000

PARK, B.; ROUPHAIL N.; SACKS J., Assessment of stochastic signal optimization method using microsimulation, **Transportation Research Board**, v.1748, artigo, 6p, 2001

PARK, B.; SANTRA, P.; YUN, I.; LEE D. Optimization of time-of-day breakpoints for better traffic signal control, **Transportation Research Record: Journal of the Transportation Research Board**, v.1867, artigo, 7p, 2004

PARK, B.; WON, J. Microscopic Simulation Model Calibration and Validation Handbook, artigo e manual, **Transportation Research Record** v.1856, p15-52 (176p), 2006

PAPAGEORGIOU, M., Some remarks on macroscopic traffic flow modelling, **Transportation Research Part A: Policy and Practice** v.32, Issue 5, artigo, 7p, 1998

ROBERTSON, D. I.; BRETHERTON, R. D., Optimizing networks of traffic signals in real time, the SCOOT method, In: **IEEE Transactions on Vehicular Technology**, v.40, Issue 1, artigo, 5p, 1991

SOUZA, C., disponível em http://www.aedb.br/seget/artigos06/916_Copia%20de%20Placas.pdf, Reconhecimento de Placas de Automóveis Através de Câmeras IP, Universidade IMES São Caetano do Sul, 7p, 2001

STEVANOVIC, A.; MARTIN P., STEVANOVIC J. VisSim-based genetic algorithm optimization of signal timings, **Transportation Research Record: Journal of the Transportation Research Board**, v.2035, artigo, 10p, 2007

SUN D. S.; BENEKOHAL, R. F.; WALLER T. Bi-level programming formulation and heuristic solution approach for dynamic traffic signal optimization, **Computer-Aided Civil and Infrastructure Engineering**, v.21, Issue 5, artigo, 13p, 2006

TEKLU, F.; SUMALEE, A.; WATLING, D. A genetic algorithm approach for optimizing traffic control signals considering routing, **Computer-Aided Civil and Infrastructure Engineering**, v.22, Issue 1, artigo, 13p, 2007

TRENTINI, V. Reconhecimento automático de placas de veículos, In: **VI Workshop de Visão Computacional**, artigo 7p, 2010

TURLEY, C. Calibration procedure for a microscopic traffic simulation model, **Brigham Young University**, tese mestrado, 156p, 2007

VARIA, H. R.; DHINGRA, S. L Dynamic optimal traffic assignment and signal time optimization using genetic algorithms, **Computer-Aided Civil and Infrastructure Engineering** v.19, Issue 4, artigo, 14p, 2004

VILANOVA, L. disponível em , <http://www.sinaldetransito.com.br> **Programação de um semáforo usando o método do grau de saturação**, 14p, 2005a

VILANOVA, L. disponível em, <http://www.sinaldetransito.com.br> **Verde de segurança**, 3p, 2005b

VILANOVA, L. disponível em <http://www.sinaldetransito.com.br>, **Fundamentos da programação semaforica**, 20p, 2005c

VILANOVA, L. disponível em, <http://www.sinaldetransito.com.br>, **Redução de acidentes devido à reprogramação semaforica**, artigo, 10p, 2005d

VILANOVA, L. disponível em, <http://www.sinaldetransito.com.br>, **Levantamento expedito de dados para reprogramação semaforica**, artigo, 24p, 2006

VILANOVA, L. disponível em <http://www.sinaldetransito.com.br>, **SIRI - Um simulador mesoscópico para redes de semáforos**, artigo, 12p, 2009

XIE, F.; LEVINSON, D. M. Evolving transportation networks, **Editora Springer**, livro didático, 297p, 2011

YANG, X. Introduction to mathematical optimization – from linear programming to metaheuristics, **University of Cambridge**, livro didático, 161p, 2008

WHITLEY D. A genetic algorithm tutorial, **Statistics and computing Springer**, artigo, 21p, 1994

ZHANG H.M.; KIM, T. A car-following theory for multiphase vehicular traffic flow, **Transportation Research Part B: Methodological** **Volume 39**, Issue 5, artigo, 20p, 2005

REFERÊNCIAS COMPLEMENTARES

CAMPBELL, A. disponível em,
<http://sci.waikato.ac.nz/evolution/EvolutionOfLife.shtml>
 The Evolution of Life, University of Waikato

CASTRO, L.; ZUBEN, F. disponível em,
ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia707_02/topico11_02.pdf,
 Programação Evolutiva. Apostila, 2009

CHRISTIAN, J. disponível em,
<http://www.dcs.napier.ac.uk/~benp/summerschool/jdemos/herdy/EStheory.html>
 Evolutionary Strategy, 1999

Companhia de Engenharia de Tráfego CET-SP, disponível em,
<http://www.cetesp.com.br/internew/informativo/pico/pico.asp>
 Informativo sobre Operação Horário de Pico (Rodízio), 2011

DNIT – Departamento Nacional de Infraestrutura de Transportes, disponível em,
<http://www.dnit.gov.br/rodovias/operacoes-rodoviaras/postos-de-contagem/plano-nacional-de-contagem-de-transito>,
 Plano Nacional de Contagem de Trânsito, 2011

DORONSO, B. disponível em,
<http://neo.lcc.uma.es/cEA-web/ES.htm>
 Evolutionary Strategies – NEO, 2004

FERNANDEZ, J. disponível em,
<http://www.geneticprogramming.com/Tutorial/>
 The Genetic Programming Notebook, 2003

GIOVANELLI C. Trânsito-Só Perdemos com Ele, Veja São Paulo, 20p, 2011

GRAHAM, L., disponível em
http://www.stellaralchemy.com/lee/virtual_creatures.php
 Stellar Alchemy, 2009

HOLLAND, J. H. disponível em,
<http://econ2.econ.iastate.edu/tesfatsi/holland.gaintro.htm>
 Genetic Algorithms, 1992

KOPPLIN, J. disponível em
<http://www.computersciencelab.com/ComputerHistory/HistoryPt3.htm>
 An Illustrated History of Computers, 2002

MACEDO, L. disponível em,
<http://g1.globo.com/sao-paulo/noticia/2010/09/cet-reduz-velocidade-em-mais-de-10-vias-para-diminuir-acidentes-em-sp.html>,

CET reduz velocidade em mais de 10 vias para diminuir acidentes em SP, 2010

OBITKO, M. disponível em
<http://www.obitko.com/tutorials/genetic-algorithms/introduction.php>

Introduction to Genetic Algorithms, 1998

SKINNER, M. disponível em,
<http://geneticalgorithms.ai-depot.com/Tutorial/Overview.html>,

Genetic Algorithms Overview, 2001

RENNARD, J.P. disponível em
<http://www.rennard.org/alife/english/gavintrgb.html>

Introduction do Genetic Algorithms, 2000

ROSE, S., disponível em
<http://andabien.com/html/evolution-timeline.htm?=&id=9738234>

Andabien- Evolution Timeline

WALL, M. disponível em,
<http://lancet.mit.edu/~mbwall/presentations/IntroToGAs/P001.html>

Overview of Genetic Algorithms, 1996

Apêndice A: Controle de velocidade de um modelo mesoscópico

Inicialmente, o segmento é dividido em duas partes. A velocidade dos veículos é controlada por uma função agregada de velocidade, sendo que todo veículo que entra no segmento deveria percorrê-lo por inteiro de acordo com essa função até chegar ao servidor de filas. Dessa forma, ao se dispensar os complexos cálculos de proximidade e aceleração dos simuladores microscópicos, pretende-se obter uma velocidade de fluxo consolidada que seja semelhante às leituras de velocidade e densidade que observa-se em simuladores microscópicos (BURGHOUT, 2004). Abaixo podemos observar uma primeira forma dessa função de densidade d , proposta por Greenshields (1935).

$$V(d) = V_{\text{livre}} \left(1 - \frac{d}{d_{\text{congest}}} \right)$$

Em que V_{livre} é a velocidade máxima admitida ou regulamentada no segmento, e d_{congest} é a maior densidade possível de se observar no segmento.

Entretanto, no que concerne a simulação mesoscópica, essa primeira função causa duas distorções principais: primeiramente a velocidade máxima só acontece quando não há veículo algum ($d = 0$) no segmento, quando na realidade as vias podem possuir alguma quantidade de veículos em velocidade máxima até que seja alcançado um dado valor de densidade de tráfego; em segundo lugar temos, na outra ponta, a situação de congestionamento total, o que implicaria velocidade zero e teoricamente travaria o segmento criando tempos de percurso infinitos, o que seria algo inverossímil. Então, normalmente opta-se por usar uma velocidade máxima (aquela regulamentada no segmento real), outra mínima e dois valores de densidade de tráfego, assim, obtém-se a função de densidade-velocidade vista abaixo, que é àquela usada em outros simuladores mesoscópicos como o Mezzo (BURGHOUT, 2004) e o DYNAMIT (AKIVA, 1998) e que cria três intervalos de densidade para a parte móvel do segmento, um deles com uma velocidade máxima saturada, outro que apresenta uma transição contínua em função da densidade e um último com uma velocidade mínima saturada (BURGHOUT, 2004).

$$V(d) = \begin{cases} V_{\text{Livre}}, & \text{se } d < d_{\min} \\ V_{\min} + (V_{\text{Livre}} - V_{\min}) \left(1 - \left(\frac{d - d_{\min}}{d_{\max} - d_{\min}} \right)^a \right)^b, & \text{se } d_{\min} \leq d \leq d_{\max} \\ V_{\min}, & \text{se } d > d_{\max} \end{cases}$$

Em que $V_{\min}$ é a velocidade mínima, $d_{\min}$ e $d_{\max}$ são as densidades de saturação mínima e máxima e a e b são parâmetros de calibração, com a função apresentando o seguinte comportamento:

Apêndice B: Dados coletados e parâmetros do estudo de caso

Os dados foram dispostos aqui para que sejam lidos de maneira mais clara, caso o leitor deseje, eles podem ser extraídos do arquivo *parametrosRede.xlsx* contido na mídia anexada à obra

Compactações:

Tempo de Percurso					
De	Para				
IN02	IN01				
IN10	IN11				
PA02	PA01				
PA10	PA11				
PO02	PO01	PO03			
NY02	NY01	NY03			
CA02	CA01	CA03			
RV02	RV01	RV03	GU02	GU01	GU03
UR	PS				
MI02A	MI01A	MI03A			
MI02B	MI01B	MI03B			
NO02A	NO01A	NO03A	HW01A	HW02A	HW03A
NO02B	NO01B	NO03B	HW01B	HW02B	HW03B

Fluxo de Saturação										
De	Para									
PA04	PA03	PA02								
PA08	PA07	PA06	PA05							
PA10	PA11*									
IN07A	IN09A	IN08A	IN06A	IN05A	IN04A					
IN08b	IN09B	IN10B	IN07B	IN06B	IN05B					
PO02	MI01A*	MI02A*	MI03B*	MI02B*	NO02A*	NO03B*	NO02B*	HW02A*	HW03B*	HW02B*
	PS	UR	IN03	IN02	IN01	NO01A*	HW01A*	IN10A*	IN04B*	IN11B*
NY01	NY02	PO01	NY03	PO03						
CA02	CA01	CA03								
RV02	RV03									
RV01	NO01A*	HW01A*								
GU02	GU01	GU03								

*É usado o fluxo de saturação por faixa

Calibragens (todas para o pico da tarde):

- i) -130 (11%) do fluxo de NY01
- ii) -50 (5%) do fluxo de RV01
- ii) +200 (6,8%) do fluxo de saturação de PA10

Mapa de Tempos de Percurso

Mapa de Fluxos de Saturação

Perfil de Demanda da Rede (Pico da Manhã)

(baseado no cruzamento Av. Pe. Antonio x Rua Ribeiro do Vale)

Mapa de Fluxos (Pico da Manhã)

Mapa de Taxas de Conversão (Pico da Manhã)

Perfil de Demanda da Rede (Pico da Tarde)

(baseado no cruzamento Av. Pe. Antonio x Rua Ribeiro do Vale)

Mapa de Fluxos (Pico da Tarde)

Mapa de Taxas de Conversão (Pico da Tarde)

Apêndice C: Planos anterior e otimizados para cada nível de demanda

Relembrando que o estudo de caso foi feito em uma rede com cinco semáforos em linha de uma via coletora de mão única. Para facilitar a análise dos planos nas próximas figuras deve-se ter em mente que os veículos percorrem a rede na seguinte sequência de intersecções:

Av. Pe. Antonio José dos Santos X Rua Guaraiúva;
 Av. Pe. Antonio José dos Santos X Rua Ribeiro do Vale;
 Av. Pe. Antonio José dos Santos X Rua Califórnia;
 Av. Pe. Antonio José dos Santos X Rua Nova York;
 Av. Pe. Antonio José dos Santos x Av. Portugal.

tG instante início verde
 tR instante início vermelho
 TG tempo de verde
 pG porcentagem de verde

Plano escolhido para teste cruzado

Plano escolhido para teste real

Legendas

<u>tCiclo</u>	<u>72</u>			
Plano Antigo	tG	tR	TG	pG
Manhã				
Pe. Antonio	0	35	35	49%
Portugal	40	67	27	38%
Pe. Antonio	0	37	37	51%
Nova York	42	67	25	35%
Pe. Antonio	0	39	39	54%
Califórnia	44	67	23	32%
Pe. Antonio	0	37	37	51%
Ribeiro do Vale	42	67	25	35%
Indiana	0	22	22	31%
Ribeiro do Vale	27	67	40	56%
Pe. Antonio	0	37	37	51%
Guaraiuva	42	67	25	35%

Plano anterior para o pico da manhã

Deve-se atender ao fato de que o algoritmo genético aplicado encontrou planos com dois tipos de comportamento atrelados aos níveis de saturação.

Nos cenários do pico da manhã com demanda 100%, 110% a rede não estava ainda saturada e pode-se observar a formação de ondas verdes, com os semáforos abrindo em sequência (os tGs do plano percorrem o ciclo de forma progressiva ao longo da rede) para que os veículos do fluxo principal na avenida Pe. Antônio.

tCiclo	57				64				60				56				67				56			
	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG
Pe. Antonio	30	55	25	44%	30	57	27	42%	40	6	26	43%	30	55	25	45%	29	60	31	46%	30	55	25	45%
Portugal	4	25	21	37%	63	25	26	41%	12	35	23	38%	4	25	21	38%	66	24	25	37%	4	25	21	38%
Pe. Antonio	19	43	24	42%	16	44	28	44%	24	48	24	40%	19	42	23	41%	16	44	28	42%	18	40	22	39%
Nova York	49	14	22	39%	50	11	25	39%	54	19	25	42%	48	14	22	39%	50	11	28	42%	46	13	23	41%
Pe. Antonio	0	24	24	42%	0	27	27	42%	0	25	25	42%	0	23	23	41%	0	30	30	45%	0	22	22	39%
Califórnia	29	51	22	39%	32	58	26	41%	30	54	24	40%	28	50	22	39%	35	61	26	39%	27	50	23	41%
Pe. Antonio	34	55	21	37%	47	8	25	39%	37	59	22	37%	42	7	21	38%	54	16	29	43%	23	43	20	36%
Ribeiro do Vale	3	28	25	44%	13	41	28	44%	4	31	27	45%	12	36	24	43%	21	48	27	40%	48	17	25	45%
Indiana	47	10	20	35%	58	16	15	23%	48	8	20	33%	55	19	20	36%	66	24	25	37%	31	50	19	34%
Ribeiro do Vale	15	41	26	46%	21	52	31	48%	13	42	29	48%	24	49	25	45%	29	60	31	46%	55	25	26	46%
Pe. Antonio	15	38	23	40%	33	59	26	41%	17	42	25	42%	25	46	21	38%	40	0	27	40%	7	29	22	39%
Guaraiuva	43	9	23	40%	0	27	27	42%	47	11	24	40%	51	19	24	43%	5	34	29	43%	34	1	23	41%

Planos para demanda aferida, ou 100% (Pico da manhã)

tCiclo	74				71				74				77				71				71			
	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG
1 Pe. Antonio	35	68	33	45%	33	65	32	45%	33	68	35	47%	30	66	36	47%	33	65	32	45%	30	62	32	45%
Portugal	0	30	30	41%	70	28	29	41%	0	28	28	38%	72	25	30	39%	70	28	29	41%	67	25	29	41%
2 Pe. Antonio	23	55	32	43%	19	50	31	44%	19	51	32	43%	18	50	32	42%	20	50	30	42%	15	45	30	42%
Nova York	61	18	31	42%	55	14	30	42%	57	14	31	42%	56	13	34	44%	56	15	30	42%	51	10	30	42%
4 Pe. Antonio	0	33	33	45%	0	29	29	41%	0	32	32	43%	0	35	35	45%	0	31	31	44%	0	29	29	41%
Califórnia	38	68	30	41%	34	65	31	44%	37	68	31	42%	40	71	31	40%	36	65	29	41%	34	65	31	44%
5 Pe. Antonio	40	71	31	42%	58	17	30	42%	66	22	30	41%	53	8	32	42%	0	29	29	41%	28	57	29	41%
Ribeiro do Vale	2	34	32	43%	22	53	31	44%	27	60	33	45%	13	47	34	44%	34	65	31	44%	62	22	31	44%
0 Indiana	50	5	29	39%	2	29	27	38%	5	34	29	39%	59	11	29	38%	11	36	25	35%	38	65	27	38%
Ribeiro do Vale	10	44	34	46%	34	67	33	46%	39	73	34	46%	16	53	37	48%	41	5	35	49%	70	32	33	46%
3 Pe. Antonio	32	62	30	41%	41	3	33	46%	51	8	31	42%	37	70	33	43%	53	10	28	39%	15	43	28	39%
Guaraiuva	67	26	33	45%	8	36	28	39%	13	45	32	43%	75	31	33	43%	15	47	32	45%	48	9	32	45%

Planos para demanda 110% (Pico da manhã)

tCiclo	90				80				99				83				83				94			
	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG
Pe. Antonio	38	80	42	47%	33	71	38	48%	36	83	47	47%	21	60	39	47%	30	69	39	47%	26	70	44	47%
Portugal	86	33	37	41%	77	28	31	39%	89	31	41	41%	66	16	33	40%	75	25	33	40%	76	21	39	41%
Pe. Antonio	24	64	40	44%	21	56	35	44%	25	68	43	43%	7	44	37	45%	19	56	37	45%	13	55	42	45%
Nova York	70	19	39	43%	62	16	34	43%	74	20	45	45%	50	2	35	42%	62	14	35	42%	61	8	41	44%
Pe. Antonio	0	38	38	42%	0	36	36	45%	0	45	45	45%	0	37	37	45%	0	37	37	45%	0	41	41	44%
Califórnia	43	84	41	46%	41	74	33	41%	50	93	43	43%	42	77	35	42%	42	77	35	42%	46	88	42	45%
Pe. Antonio	38	75	37	41%	29	60	31	39%	10	53	43	43%	47	80	33	40%	47	79	32	39%	51	91	40	43%
Ribeiro do Vale	80	32	42	47%	65	23	38	48%	58	4	45	45%	2	41	39	47%	1	41	40	48%	2	45	43	46%
Indiana	53	87	34	38%	51	0	29	36%	35	77	42	42%	58	8	33	40%	55	2	30	36%	65	9	38	40%
Ribeiro do Vale	2	47	45	50%	5	45	40	50%	82	29	46	46%	13	52	39	47%	7	49	42	51%	14	59	45	48%
Pe. Antonio	31	69	38	42%	18	50	32	40%	92	38	45	45%	37	71	34	41%	37	70	33	40%	37	80	43	46%
Guaraiuva	74	25	41	46%	55	12	37	46%	43	86	43	43%	76	31	38	46%	75	31	39	47%	85	31	40	43%

Planos para demanda 120% (Pico da manhã)

Nos níveis mais altos, observados no cenário de 130% de demanda do pico da manhã e no pico da tarde, o que se nota é a alta superposição dos períodos de verde geralmente com poucos segundos de defasagem entre um semáforo e outro.

tCiclo	111				119				119				118				111				94			
	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG
Pe. Antonio	2	54	52	47%	1	57	56	47%	1	57	56	47%	28	83	55	47%	5	57	52	47%	0	44	44	47%
Portugal	60	108	48	43%	63	115	52	44%	63	115	52	44%	89	23	52	44%	63	0	48	43%	50	89	39	41%
Pe. Antonio	109	46	48	43%	11	65	54	45%	2	54	52	44%	14	64	50	42%	1	49	48	43%	89	35	40	43%
Nova York	52	104	52	47%	71	6	54	45%	60	116	56	47%	70	9	57	48%	55	107	52	47%	41	84	43	46%
Pe. Antonio	0	50	50	45%	0	52	52	44%	0	53	53	45%	0	51	51	43%	0	51	51	46%	0	42	42	45%
Califórnia	55	105	50	45%	57	113	56	47%	58	113	55	46%	56	112	56	47%	56	105	49	44%	47	88	41	44%
Pe. Antonio	59	106	47	42%	1	49	48	40%	100	34	53	45%	113	44	49	42%	31	78	47	42%	57	3	40	43%
Ribeiro do Vale	0	53	53	48%	54	114	60	50%	39	94	55	46%	49	107	58	49%	83	25	53	48%	8	51	43	46%
Indiana	68	0	43	39%	116	46	49	41%	16	64	48	40%	114	46	50	42%	48	93	45	41%	70	13	37	39%
Ribeiro do Vale	5	62	57	51%	51	110	59	50%	69	10	60	50%	51	108	57	48%	98	42	55	50%	18	64	46	49%
Pe. Antonio	29	81	52	47%	118	58	59	50%	79	11	51	43%	95	31	54	46%	29	80	51	46%	52	93	41	44%
Guaraiuva	86	23	48	43%	63	112	49	41%	16	73	57	48%	36	89	53	45%	85	23	49	44%	4	46	42	45%

Planos para demanda 130% (Pico da manhã)

tCiclo	100			
	tG	tR	TG	pG
Pe. Antonio	0	53	53	54%
Portugal	58	95	37	34%
Pe. Antonio	0	55	55	56%
Nova York	60	95	35	31%
Pe. Antonio	0	57	57	56%
Califórnia	62	95	33	31%
Pe. Antonio	0	52	52	56%
Ribeiro do Vale	57	95	38	31%
Indiana	0	25	25	31%
Ribeiro do Vale	30	95	65	56%
Pe. Antonio	0	50	50	56%
Guaraiuva	55	95	40	31%

Plano anterior para pico da tarde

tCiclo	120				120				120				120				120				120				Plano Antigo 100			
	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG	tG	tR	TG	pG
1 Pe. Antonio	13	80	67	56%	2	69	67	56%	31	99	68	57%	70	19	69	58%	43	114	71	59%	20	87	67	56%	0	53	53	54%
Portugal	86	8	42	35%	75	117	42	35%	104	26	42	35%	25	65	40	33%	0	38	38	32%	93	15	42	35%	58	95	37	34%
2 Pe. Antonio	6	68	62	52%	118	62	64	53%	17	82	65	54%	64	8	64	53%	51	111	60	50%	14	78	64	53%	0	55	55	56%
Nova York	76	1	45	38%	68	113	45	38%	87	12	45	38%	14	59	45	38%	1	46	45	38%	87	9	42	35%	60	95	35	31%
4 Pe. Antonio	0	57	57	48%	0	58	58	48%	119	59	60	50%	0	59	59	49%	0	59	59	49%	0	60	60	50%	0	57	57	56%
Califórnia	62	114	52	43%	63	114	51	43%	64	114	50	42%	64	114	50	42%	64	114	50	42%	65	##	49	41%	62	95	33	31%
5 Pe. Antonio	115	54	59	49%	87	27	60	50%	87	29	62	52%	91	30	59	49%	86	27	61	51%	52	##	66	55%	0	52	52	56%
Ribeiro do Vale	59	109	50	42%	32	81	49	41%	34	82	48	40%	35	85	50	42%	32	79	47	39%	3	49	46	38%	57	95	38	31%
0 Indiana	18	49	31	26%	5	39	34	28%	111	24	33	28%	106	19	33	28%	8	40	32	27%	75	##	34	28%	0	25	25	31%
Ribeiro do Vale	54	12	78	65%	44	119	75	63%	29	106	77	64%	24	100	76	63%	45	2	77	64%	##	69	75	63%	30	95	65	56%
3 Pe. Antonio	80	21	61	51%	60	1	61	51%	58	119	61	51%	58	118	60	50%	59	119	60	50%	16	76	60	50%	0	50	50	56%
Guaraiuva	26	74	48	40%	6	54	48	40%	4	53	49	41%	3	52	49	41%	4	53	49	41%	81	10	49	41%	55	95	40	31%

Planos Para Pico da Tarde

Uma vez implementado o plano destacado acima sofreu alguns ajustes por parte de especialista da CET-SP:

No cruzamento da rua Pe. Antonio com a rua Guaraiuva a defasagem foi alterada de 58 para 62

No cruzamento da rua Pe. Antonio com a rua Califórnia a defasagem foi alterada de 119 para 110