



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

GEORGE ALLAN MENEZES GOMES

**VISUALIZAÇÃO INTERATIVA DE DINÂMICAS DE TRÁFEGO ATRAVÉS DE
DADOS DE TRAJETÓRIAS**

FORTALEZA

2018

GEORGE ALLAN MENEZES GOMES

VISUALIZAÇÃO INTERATIVA DE DINÂMICAS DE TRÁFEGO ATRAVÉS DE DADOS DE
TRAJETÓRIAS

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação.
Área de Concentração: Computação Gráfica

Orientadora: Profa. Emanuele Marques dos Santos, Dra.

Coorientador: Prof. Creto Augusto Vidal, Dr.

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

G614v Gomes, George Allan Menezes.

Visualização interativa de dinâmicas de tráfego através de dados de trajetórias / George Allan Menezes Gomes. – 2018.

127 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2018.

Orientação: Profa. Dra. Emanuele Marques dos Santos.

Coorientação: Prof. Dr. Creto Augusto Vidal.

1. Visualização. 2. Dinâmica de Tráfego. 3. Dados de Trajetória. I. Título.

CDD 005

GEORGE ALLAN MENEZES GOMES

VISUALIZAÇÃO INTERATIVA DE DINÂMICAS DE TRÁFEGO ATRAVÉS DE DADOS DE
TRAJETÓRIAS

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação. Área de Concentração: Computação Gráfica

Aprovada em:

BANCA EXAMINADORA

Profa. Emanuele Marques dos Santos,
Dra. (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Creto Augusto Vidal, Dr. (Coorientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. José Antônio F. de Macêdo
Universidade Federal do Ceará (UFC)

Prof. Dr. Ascânio Dias Araújo física
Universidade Federal do Ceará (UFC)

Prof. Dr. João Luiz Dihl Comba
Universidade Federal do Rio Grande do Sul
(UFRGS)

À minha irmã Andréa Marla (*In Memoriam*).

AGRADECIMENTOS

Muitas pessoas colaboraram nesta conquista, tantas que certamente cometerei lapsos ao mencionar apenas algumas. Fica a certeza que só cheguei até aqui porque fui incentivado e apoiado por muitas pessoas importantes.

Em primeiro lugar, agradeço a Deus por mais uma conquista em minha vida.

A minha esposa Lia, pela compreensão, suporte, paciência e principalmente pelos momentos felizes durante todo este percurso. Com certeza, esse trabalho não seria possível se não fosse seu apoio e amor incondicional.

Aos meus pais, Raimundo Gomes e Risetete, pelo sacrifício que tiveram que fazer para me educar, sempre me incentivando a continuar minha formação em busca de um futuro melhor. Cheguei onde estou graças a eles.

Aos meus filhos Arthur e Alice que tornaram meus dias mais felizes quando enfrentei alguma dificuldade. Por serem crianças, muitas vezes não entendiam o porquê da minha ausência quando precisei me dedicar a esse trabalho.

A minha irmã Taise, por ter me apoiado quando eu precisei, os meus mais sinceros sentimentos de gratidão. A minha irmã Andréa, que apesar de ter partido, sempre estará em nossos corações.

A todos os meus familiares, em especial aos meus sogros Elias e Célia pelo suporte para a conclusão deste trabalho.

Meus sinceros agradecimentos aos meus orientadores. A professora Emanuele pela orientação dedicada e qualificada durante todos esses anos de doutorado. Ao meu coorientador professor Creto que participou de toda minha formação acadêmica, me orientando na graduação, no mestrado e, agora, no doutorado.

Muitos amigos me ajudaram nesse percurso, em especial, ao meu amigo Melo Júnior pelo apoio, pelas discussões e pelas ideias, que tiveram uma grande contribuição para atingir esse meu objetivo de vida. Eu sou muito grato também ao meu amigo Henrique Pequeno pelo auxílio incondicional em todo o doutorado. Aos amigos Ricardo Brauner e Gilvan pela ajuda técnica na definição deste trabalho que foi fundamental durante o início dessa jornada.

Aos professores Antônio Macedo e Ticiane Linhares pela colaboração na principal publicação deste trabalho.

A meus amigos do Instituto UFC Virtual que direta ou indiretamente me apoiaram nessa jornada. Em especial, agradeço a Gabriel Paillard, Natal Chicca, Rafael Carmo, Ernesto

Trajano, Carlos Eduardo, Inga Sabóia, Henrique Silva, Clemilson Santos, Windson Viana e Alysson Diniz pela inestimável apoio ao longo do desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Aos colegas, funcionários e professores do Departamento de Computação pelo apoio e aprendizado durante o doutorado.

A todos, meu muito obrigado.

“Por isso não desanimamos. Embora exteriormente estejamos a desgastar-nos, interiormente estamos sendo renovados dia após dia, pois os nossos sofrimentos leves e momentâneos estão produzindo para nós uma glória eterna que pesa mais do que todos eles.”

(Coríntios 4:16-17)

RESUMO

O acelerado processo de urbanização no mundo causa sérios problemas de trânsito. A crescente disponibilidade de tecnologias de aquisição de localização GPS tem permitido a coleta contínua de grande quantidade de dados de movimento. Esses dados são uma fonte valiosa para ajudar as agências de trânsito a identificar eventos anormais que exigem atenção imediata para melhor direcionar o tráfego. Neste contexto, a análise automatizada em combinação com a visualização interativa auxiliam o entendimento do problema, a análise de suas causas e potenciais soluções, e a tomada de decisão eficaz. As abordagens tradicionais agregam o movimento, empregando o conceito de discretização em janela de tempo e explorando um conjunto de dados inteiro. No entanto, essas agregações, frequentemente, apresentam inconsistências, no tempo e no espaço, com a dinâmica real do tráfego. Nesta tese, uma nova abordagem é apresentada para descobrir padrões de mobilidade global e local em tempo real. Diferentemente de outras abordagens, o método proposto acompanha a evolução do movimento dos objetos em tempo real. De fato, na literatura, nenhuma outra abordagem captura e acompanha as evoluções das dinâmicas de tráfego de maneira incremental. Para avaliar a eficácia do método proposto, experimentos extensivos foram realizados utilizando conjuntos de dados reais e simulados. O método proposto também foi submetido à avaliação de um especialista do domínio quanto a seus benefícios e limitações. Além disso, são apresentados testes de desempenho com resultados muito encorajadores para apoiar a abordagem proposta na visualização do fluxo total de tráfego de uma grande cidade. Os resultados demonstram que a técnica proposta escala linearmente com o tamanho do conjunto de dados e é capaz de lidar com grandes volumes de dados e com fluxos que apresentem altas taxas de amostragem.

Palavras-chave: Visualização. Dinâmica de Tráfego. Dados de Trajetória.

ABSTRACT

Urbanization is accelerating worldwide, giving rise to serious traffic problems. With the increasing availability of location acquisition technologies, massive movement data are collected continuously in a streaming manner. These data are a valuable source to help transit agencies to identify abnormal events that require immediate attention to better direct traffic. In this regard, visual analytics can help by combining automated analysis with interactive visualization for effective understanding, reasoning, and decision-making. Traditional approaches aggregate movement by employing the concept of time-window discretization and exploring an entire dataset. However, they can present inconsistencies in time and space with the real traffic dynamics. In this thesis, we present a novel approach to discover global and local mobility patterns in real time. Different from other existing approaches, our method tracks the evolution of the objects' movement in real time. We believe that no other approach captures and keeps track of how the hot routes evolve in an incremental manner. Moreover, we conducted extensive experiments on real-world and simulated datasets to evaluate the effectiveness of our method. We also present the benefits and limitations of our visualization proposal based on domain expert feedback. Finally, we present performance tests with very encouraging results to support our approach in visualizing the total traffic flow of a big city. The results demonstrate that our method scales linearly with the size of the dataset, and is able to deal with large datasets and with streams of high-sampling rates.

Keywords: Visualization. Traffic Dynamics. Trajectory data.

LISTA DE FIGURAS

- Figura 1 – (a) Um problema no cruzamento destacado prejudicou o fluxo de tráfego. Esse comportamento de tráfego foi constante dentro do período de análise. (b) As abordagens tradicionais marcam toda a rota de P_s até P_e , pois o que importa é o fluxo total de tráfego agregado dentro de uma janela de tempo. (c) No entanto, apenas parte da rota é quente considerando-se a dinâmica do movimento dos objetos. 24
- Figura 2 – (a) No início de um intervalo de dez minutos, um incidente bloqueou a rua no final do caminho P_e . (b) Após dois minutos, há um engarrafamento ao longo da rua, e o bloqueio na estrada é removido. (c) O congestionamento está se dissipando. (d) Após quatro minutos, o fluxo de tráfego na rua torna-se normal novamente e permanece assim até o final do intervalo de tempo. . . 25
- Figura 3 – (a) Uma via com tráfego intenso, mas com fluxo constante e sem engarrafamentos. (b) Depois de dois minutos, um veículo decidiu fazer uma conversão à direita no local do retângulo pontilhado. Alguns veículos logo atrás mudaram de faixa. (c) Uma onda de tráfego deslocando-se em sentido oposto ao do fluxo. (d) O fluxo de tráfego normaliza-se novamente e permanece assim até o final do período de análise. 26
- Figura 4 – Uma demonstração da visualização de ondas de tráfego proposta neste trabalho. (a) O tráfego está denso, mas fluindo normalmente. Uma onda de tráfego (retângulo pontilhado) movendo-se da direita (b) para a esquerda (c) pode ser identificada por cor, comprimento dos rastros dos veículos e tamanho dos marcadores. As escalas de cor à esquerda e comprimento à direita codificam a velocidade para auxiliar a identificação. Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/uycpCRSPyLo>>. 30
- Figura 5 – Rotas quentes descobertas pela abordagem proposta, usando um conjunto de dados brutos reais. (a) Um engarrafamento propagando-se para cima ao longo da via. (b) A evolução das rotas quentes no retângulo pontilhado central em cinco momentos registrados a cada meia hora. (c) À direita, as rotas quentes estão se formando em ambos os sentidos devido a um gargalo de tráfego. Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/mDUy956-vSA>>. 31

Figura 6	– Exemplo de trajetória com erros de GPS. Os pontos p_5 e p_9 podem ser identifi- cados usando média e desvio padrão para detectar velocidades extremamente altas para a trajetória atual.	36
Figura 7	– Exemplo de compressão de dados de trajetórias utilizando o algoritmo de Douglas-Peucker (DOUGLAS; PEUCKER, 1973). A nova trajetória é com- posta apenas pelos pontos $p_1 \rightarrow p_4 \rightarrow p_9 \rightarrow p_{12}$	36
Figura 8	– Representação estruturada de uma trajetória através de seus pontos de parada: casa, trabalho e supermercado ($p_1 \rightarrow p_2 \rightarrow p_3$).	37
Figura 9	– (a) Exemplo de trajetória segmentada por tempo fixo de duração de duas horas em três segmentos (s_1, s_2, s_3). (b) Mesma trajetória segmentada por limiar de tempo de 30 minutos entre dois pontos consecutivos (s_1, s_2).	38
Figura 10	– (a) Exemplo de trajetória com as leituras de GPS originais com dois erros – pontos vermelhos. (b) A trajetória foi corrigida com a remoção dos erros após o <i>map-matching</i> . (c) Compactação da trajetória com sua representação pelos vértices (esquinas) da rede viária, ou seja, quantidade de pontos reduzida.	39
Figura 11	– (a) Análise visual de rotas de ciclistas em Ottawa, Canadá. (b) Mapa de fluxo de movimento de pessoas gerado a partir de dados do Twitter.	41
Figura 12	– (a) Análise de densidade de tráfego utilizando dados de táxis da cidade de Nova York. (b) Análise das rotas mais percorridas (em vermelho) por táxis da cidade de Pequim, China.	43
Figura 13	– Exemplos de padrões em grupos de objetos que se movem juntos durante um certo período. (a) <i>Flock, convoy e swarm</i> . (b) <i>Gathering</i>	44
Figura 14	– Exemplo de análise utilizando clusterização de trajetórias de navios na área do Mar do Norte, Países Baixos. (a) As trajetórias inteiras de navios são representadas por linhas no mapa. (b) As trajetórias são agrupadas (<i>clus- ters</i>) de acordo com os destinos. (c) Um dos <i>clusters</i> é selecionado. (d) A clusterização por similaridade de rota foi aplicado ao <i>cluster</i> selecionado.	45
Figura 15	– (a) Exemplo de rotas quentes encontradas na cidade de San Joaquin, EUA. (b) Outro exemplo com as 20 rotas mais quentes na cidade de Atenas, Grécia.	48
Figura 16	– (a) Representação das principais rotas de aproximação dos aeroportos de Lon- dres através de clusterização de trajetórias. (b) Análise visual de trajetórias de táxis da cidade de Nova York.	50

Figura 17 – (a) Visualização de um mapa de calor com as regiões de maior volume de tráfego, as regiões laranjas. (b) Mapa de densidade de trajetórias do tráfego aéreo dos EUA. (c) Mapa de densidade de tráfego de embarcações em torno de Roterdã.	51
Figura 18 – (a) Análise da mobilidade de pessoas em Londres utilizando dados do Twitter. (b) visualização de padrões de intercâmbio entre diferentes regiões em Singapura.	51
Figura 19 – Visualizações de trajetórias de migração nos EUA gerados pelos métodos apresentados por: (a) Holten e Van Wijk (2009); (b) Cui <i>et al.</i> (2008); (c) Lambert <i>et al.</i> (2010); (d) Ersoy <i>et al.</i> (2011).	52
Figura 20 – (a) Padrões de oscilação de tráfego. (b) Padrões e comportamentos anormais no tráfego de um cruzamento. (c) Dinâmica de tráfego baseada em campo vetorial. (d) Visualização de um grande volume de dados de trajetórias de táxis.	53
Figura 21 – (a) Os pontos de uma trajetória representados em um mapa sem considerar o tempo. (b) A mesma trajetória é representada usando um STC. (c) Exemplo de um STC com poluição visual. (d) Uma adaptação do STC para visualizar velocidade.	54
Figura 22 – Diferentes codificações de velocidade e tempo para representações 2D curvas e retas de trajetórias. Tanto velocidade constante como velocidade variada (seções lentas e com alta velocidade) são mostradas.	55
Figura 23 – (a) Uma área de interesse delimitada por uma <i>viewport</i> em três camadas de mapa com diferentes escalas. (b) O espaço da <i>viewport</i> foi discretizado como uma grade com 17 linhas e 18 colunas. Para uma célula da grade (r, s) no instante atual t_k , a soma dos pesos de contribuição $w_{ik}(r, s)$ dos objetos que cruzaram essa célula é calculada pela Eq. (4.2).	58
Figura 24 – Este modelo diminui o peso de um objeto exponencialmente do laranja (mais recente) para o branco (expirado) através de um percentual fixo em intervalos regulares.	59

- Figura 25 – Uma visão geral da arquitetura proposta. Os dispositivos GPS representam os objetos ativos em movimento que podem ser rastreados pelo Gerenciador de Fluxo de Dados (GFD). Esses dispositivos enviam suas posições atuais para o GFD, o qual mantém uma estrutura de dados sincronizada dos objetos em movimento (TOM). O componente Renderizador de Dinâmicas de Tráfego (RDT) usa uma GPU para processar a TOM e guardar o resultado em um buffer que é renderizado em uma tela. 61
- Figura 26 – A metodologia da abordagem proposta: na fase de Preparação de Dados, o sistema limpa os dados brutos e os sincroniza para serem visualizados. Na fase de Visualização, esses dados são renderizados em imagens, e alguns atributos dessas imagens são realçados para facilitar a análise visual. 62
- Figura 27 – Processo de sincronização de tempo. (a) Um mapa exibindo os registros originais das localizações dos objetos P e Q . (b) Os pontos originais distribuídos em cinco unidades de tempo do processo de sincronização (t_0 a t_4). (c) As novas localizações para os objetos P e Q correspondentes a cada instante t_i são calculadas por interpolação espacial. (d) Os novos pontos exibidos no mapa, com a distância entre o ponto original q_1 e o ponto interpolado q'_1 destacado como um exemplo. As trajetórias são exibidas apenas para fins ilustrativos, elas são renderizadas pelo RDT na próxima etapa, Renderização. 64
- Figura 28 – Uma visão gráfica geral da etapa de renderização. (a) O *buffer* de acumulação da iteração anterior é envelhecido através de uma função de decaimento. (b) O *buffer* ativo é gerado incrementando as trajetórias dos objetos ativos em movimento (gerado por *incrementarRotas* no Algoritmo 2). (c) Em seguida, os dois *buffers* são compostos produzindo o *buffer* de acumulação para a próxima iteração. O pseudo-código do *renderizarRotas* está listado no Algoritmo 1. 66
- Figura 29 – Os efeitos de suavização e filtragem. (a) Usando apenas as contribuições dos objetos para visualizar as temperaturas das rotas. (b) As mesmas rotas após a aplicação de uma função de suavização. (c) Usando a função de suavização e um filtro que realça as rotas que atingiram a temperatura \bar{T} . (d) Apenas as rotas com temperatura \bar{T} são visualizadas em uma cor preta (sem suavização). 69

Figura 30 – Exemplo de aplicação da abordagem proposta para encontrar rotas quentes. (a) Trajetórias originais geradas pela etapa de Renderização. (b) Trajetórias suavizadas e com a marcação das rotas quentes realizada na etapa de Processamento Visual destacando com a cor preta toda célula com opacidade igual a 1,0.	74
Figura 31 – (a) Representação de uma via expressa com 4 faixas no nível de zoom 18. (b) A via expressa em (a) é exibida como uma rua de pista única no no nível de zoom 15 na região delimitada pelo retângulo pontilhado.	76
Figura 32 – Rotas quentes encontradas pela abordagem de referência NETSCAN (a) e a abordagem proposta (b). Este exemplo utilizou um subconjunto do T-Drive no horário de maior movimento de uma manhã, das 8h30 às 8h40, com um limiar de densidade de 10 táxis. (c) A diferença de imagem (b - a) mostra um pequeno ruído, destacado por círculos pontilhados na cor laranja, no resultado da abordagem proposta.	78
Figura 33 – O resultado da abordagem proposta usando a função de decaimento em três instantes distintos 8h34 (a), 8h37 (b) e 8h40 (c) para mostrar a evolução das rotas quentes.	80
Figura 34 – Três exemplos de gargalos de tráfego. No primeiro exemplo, é possível perceber na mesma via expressa carros mais rápidos (a) sem rastros e carros mais lentos (b) formando uma rota quente. (c) O segundo exemplo mostra uma faixa de pedestre prejudicando o fluxo de tráfego em uma avenida. (d) Um gargalo de origem desconhecida apareceu repentinamente durante a análise.	82
Figura 35 – O impacto de diferentes níveis de zoom na qualidade das rotas quentes. (a) Rotas quentes de boa qualidade com erros de GPS minimizados no nível de zoom 15. Uma região no mapa (o retângulo pontilhado) é representada nos níveis de zoom 16 (b), 17 (c) e 18 (d) para comparar os resultados.	83
Figura 36 – (a) Uma avenida com oito faixas, quatro em cada sentido. Nenhuma delas é uma rota quente no nível de zoom 18. (b) As oito faixas dessa avenida (retângulo pontilhado) foram incorretamente agregadas em uma rota quente no nível de zoom 16. (c) O tráfego flui da esquerda para a direita. A avenida é quente nesse sentido (retângulo pontilhado). (d) O tráfego flui no sentido oposto e a avenida não é quente.	84

- Figura 37 – Comparação dos efeitos de diferentes valores de parâmetros. Dois conjuntos de dados do T-Drive foram utilizados: o original na coluna da esquerda e uma réplica na coluna da direita. A réplica contém três cópias do conjunto de dados original. (a) Algumas rotas quentes em Pequim. (b) A maioria das estradas está marcada como quente para os mesmos parâmetros. (c) Alterando o nível de zoom para 17. (d) A réplica mostrou rotas mais quentes. (e) Com cinco minutos de tempo de análise, nenhuma via foi marcada como quente. (f) Para o mesmo período, a réplica apresentou rotas quentes bem definidas. (g) A abordagem alternativa baseada em velocidade foi usada. (h) Os resultados obtidos foram muito semelhantes, mesmo com um conjunto de dados três vezes maior. 86
- Figura 38 – Exemplo de um experimento para investigar a influência dos carros autônomos na redução de ondas de tráfego. A parte superior mostra o local do experimento, e na inferior, os gráficos utilizados no estudo. 88
- Figura 39 – Exemplo de aplicação da abordagem proposta para visualizar ondas de tráfego em uma via circular. (a) Um quadro da animação quando o fluxo de veículos é estável. (b) Após um aumento na densidade de veículos, o fluxo se desestabiliza, aparecendo uma onda de tráfego. Os veículos no centro da onda (círculo pontilhado) estavam totalmente parados nesse instante. 90
- Figura 40 – Renderização do rastro de um objeto, quadro a quadro f_i . Neste exemplo, o objeto estava a 60 km/h no quadro f_0 , depois desacelerou nos quadros f_1 a f_4 . 91
- Figura 41 – O resultado da etapa de Processamento Visual aplicada ao rastro de um objeto. (a) O rastro suavizado. (b) O rastro cortado proporcionalmente à velocidade do objeto e com cor modificada. (c) a e b juntos. 92
- Figura 42 – Captura de tela de uma visualização do fluxo de tráfego de 1.800 veículos em movimento ao mesmo tempo. A imagem mostra as principais funcionalidades da ferramenta desenvolvida, como o (a) painel de personalização, onde o usuário pode definir o tamanho do marcador, o limite de velocidade, a largura do rastro, o filtro de processamento visual e outros parâmetros; (b) a exibição de data e hora; (c) o controle de tempo; (d) a escala da velocidade do rastro; (e) codificação da cor; e controle de zoom (f). Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/uycpCRSPyLo>>. 94

Figura 43 – O cenário de uma via circular. (a) Tráfego denso e estável sem congestionamentos. Os veículos estão se movendo a uma velocidade próxima ao limite. (b) Um carro freia abruptamente. (c) Veículos que seguem logo atrás desse carro mudam de faixa, propagando a perturbação para as outras faixas. (d) Então, uma onda de tráfego é formada e se move no sentido oposto ao fluxo.	96
Figura 44 – O cenário de uma via de acesso. (a) Os veículos acessam a via principal e provocam uma onda de tráfego. (b) A via de acesso gerou duas ondas de tráfego. Neste caso, a escala de cor foi configurada para representar um intervalo de velocidade menor de 0 a 10 km/h para realçar as diferenças de velocidade.	97
Figura 45 – (a) O cenário de um bloqueio de pista. A faixa interna é obstruída sem nenhuma sinalização de bloqueio no mapa. (b) O cenário de um aclave. Os veículos pesados tiveram sua velocidade reduzida significativamente na seção do aclave prejudicando o fluxo de tráfego.	98
Figura 46 – Distribuição do número de objetos ativos ao longo do tempo.	104
Figura 47 – Evolução dos objetos ativos que entram e saem da visualização ao longo do tempo.	104
Figura 48 – Os resultados de desempenho. Um total de 16 conjuntos de dados foram testados com três resoluções diferentes: HD (1280 x 720), Full HD (1920 x 1080) e Ultra-Wide Full HD (2560 x 1080). As linhas de referência pontilhadas representam a velocidade de animação de 30fps e de 12fps. . . .	105

LISTA DE TABELAS

Tabela 1 – Notação usada na etapa de Renderização.	67
Tabela 2 – Notação usada na etapa de Processamento Visual.	70

LISTA DE ALGORITMOS

Algoritmo 1	–	$\text{renderizarRotas}(n_t, \tau, B_w)$	67
Algoritmo 2	–	$\text{incrementarRotas}(O_a)$	68
Algoritmo 3	–	$\text{suavizarFiltrar}(B_w)$	71
Algoritmo 4	–	$\text{incrementarRotas}(O_a)$	81

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Motivação	21
1.2	Caracterização do Problema	23
1.3	Objetivos	27
1.4	Metodologia	28
1.5	Contribuições	29
1.6	Publicações	31
1.7	Organização	32
2	CONCEITOS PRELIMINARES	33
2.1	Dados de Trajetórias	33
2.2	Pré-processamento de Trajetórias	34
2.3	Considerações Finais	39
3	TRABALHOS RELACIONADOS	40
3.1	Aplicações de Análise de Dados de Trajetória	40
3.2	Clusterização de Objetos em Movimento	42
3.3	Clusterização de Trajetórias	44
3.4	Descoberta de Rotas Quentes	46
3.5	Visualização de Dados de Trajetórias	48
3.5.1	<i>Padrões de Mobilidade Global</i>	49
3.5.2	<i>Padrões de Mobilidade Local</i>	52
3.6	Considerações Finais	55
4	ABORDAGEM PROPOSTA	57
4.1	Definição do Problema	57
4.2	Arquitetura	60
4.3	Metodologia	62
4.3.1	<i>Pré-processamento</i>	62
4.3.2	<i>Sincronização</i>	63
4.3.3	<i>Renderização</i>	65
4.3.4	<i>Processamento Visual</i>	68
4.4	Considerações Finais	71

5	ANÁLISE VISUAL DE DINÂMICA DE TRÁFEGO	73
5.1	Visualizando Rotas Quentes	73
5.1.1	<i>Representação do Mapa e Níveis de Zoom</i>	75
5.1.2	<i>Resultados e Discussão</i>	77
5.2	Visualizando Ondas de Tráfego	87
5.2.1	<i>Codificando Tempo e Velocidade</i>	88
5.2.2	<i>Avaliação por Especialista</i>	93
5.3	Considerações Finais	100
6	ANÁLISE DE DESEMPENHO	101
6.1	Tecnologias Empregadas	101
6.2	Conjuntos de Dados Analisados	103
6.3	Análise dos Resultados de Desempenho	104
6.4	Considerações Finais	106
7	CONCLUSÕES E TRABALHOS FUTUROS	108
7.1	Resultados Alcançados	108
7.2	Trabalhos Futuros	110
	REFERÊNCIAS	112

1 INTRODUÇÃO

Esta tese apresenta uma abordagem para visualização de dinâmicas de tráfego através de dados de trajetórias com o objetivo de encontrar padrões de mobilidade local e global. Inicialmente, a motivação e a contextualização para o desenvolvimento deste trabalho são discutidas na Seção 1.1. A Seção 1.2 apresenta a descrição do problema abordado. Em seguida, os objetivos a serem alcançados para resolução do problema são detalhados na Seção 1.3. Na Seção 1.4, a metodologia que norteou a elaboração deste trabalho é apresentada. Depois, na Seção 1.5, são descritas as principais contribuições desta tese, incluindo as publicações geradas na Seção 1.6. Por último, a organização dos demais capítulos é descrita na Seção 1.7.

1.1 Motivação

Atualmente, devido ao grande número de veículos nas ruas todos os dias, há um grande interesse no monitoramento em tempo real do tráfego nas áreas metropolitanas para identificar eventos anormais que requerem atenção imediata e que precisam ser resolvidos com o mínimo de transtorno. Além disso, em situações de emergência, como inundações, incêndios, furacões e terremotos, as autoridades precisam monitorar continuamente as rotas de evacuação para direcionar a população a seguir por vias alternativas, pois as pessoas geralmente abandonam suas residências precipitadamente seguindo rotas semelhantes (WOLSHON, 2001). Mesmo em condições normais, há uma necessidade constante de monitorar o tráfego para tomada de decisões em tempo real. Geralmente, as agências de trânsito combinam vigilância por vídeo com métodos de visão computacional para auxiliar os operadores humanos de controle de tráfego no monitoramento de trânsito (VISHNU *et al.*, 2017; ROY *et al.*, 2017; KANRAR; MANDAL, 2017). No entanto, a detecção e o rastreamento de objetos em movimento continuam sendo um grande problema para a maioria das aplicações de visão computacional (KALIRAJAN; SUDHA, 2015). Na vigilância por vídeo em ambientes abertos, por exemplo, o problema de rastreamento visual é particularmente desafiador, primeiro por ser oneroso e limitado a uma área supervisionada e, segundo, por apresentar problemas de oclusão, vibração de câmera e possíveis variações na iluminação de fundo. Como resultado, o monitoramento das condições de tráfego em uma cidade inteira por meio dessa abordagem exige um grande esforço humano e um custo elevado (ZHENG *et al.*, 2014).

O crescimento acelerado do uso de tecnologias para tornar as cidades mais inte-

ligentes facilita o monitoramento contínuo dos indivíduos. Os dispositivos com sensores de localização, como GPS, RFID e celulares, permitem coletar e monitorar em tempo real fluxos de dados (do inglês, *data stream*) de trajetórias, gerando grandes volumes de dados, que podem ser usados para análise em diferentes tipos de aplicação, como no planejamento urbano, para identificar padrões de fluxo de tráfego e as causas de congestionamentos de trânsito (XIA *et al.*, 2018; KONG *et al.*, 2018; PARK *et al.*, 2018; DING *et al.*, 2018a; KWON *et al.*, 2018). Assim, explorar e extrair informações úteis em fluxo de dados de trajetória foi extensivamente estudado na última década (DUAN *et al.*, 2016; SACHARIDIS *et al.*, 2008; LANGE *et al.*, 2011; MAO *et al.*, 2016). O 511.org¹, por exemplo, coleta dados de sensores RFID e GPS para apoiar a tomada de decisões e fornecer, em tempo real, informações de tráfego da área da Baía de São Francisco para seus usuários, através de aplicativos para *smartphones* e por meio de painéis distribuídos ao longo das principais vias. Além disso, os serviços baseados em localização, como o Google Maps e o Waze, mostram dados de tráfego de forma *online* para que indivíduos planejem suas rotas de tráfego ou encontrem pontos de interesse.

Considerando os aspectos globais na análise de dados de trajetórias, a maioria das abordagens para monitorar tráfego visam descobrir as rotas com grandes volumes de tráfego, também conhecidas como rotas quentes (SACHARIDIS *et al.*, 2008; LI *et al.*, 2007). A visualização de como essas rotas quentes evoluem permite que as agências de trânsito identifiquem eventos anormais em um estágio inicial para direcionar melhor o tráfego. No entanto, o monitoramento de dezenas de milhares de objetos em movimento em tempo real é uma tarefa muito desafiadora, pois requer processamento em alta velocidade de um vasto volume de dados em fluxo. Existem várias abordagens para a descoberta de rotas quentes, como as propostas em (ALMEIDA *et al.*, 2017; GUI *et al.*, 2016; WU *et al.*, 2014; SACHARIDIS *et al.*, 2008; LI *et al.*, 2007). Além disso, os trabalhos de clusterização de trajetórias em fluxos de dados (MAO *et al.*, 2016; LANGE *et al.*, 2011; YU *et al.*, 2013; SILVA *et al.*, 2016; LI *et al.*, 2010b) poderiam ser usados para capturar rotas quentes por meio de grupos de subtrajetórias semelhantes de objetos em movimento que estão presentes em uma sequência de intervalos de tempo recentes – também conhecidos como janelas de tempo (do inglês, *time windows*). Embora a clusterização de dados seja importante para extrair conhecimento potencialmente útil, agregar trajetórias em uma janela de tempo prejudica a análise de toda a dinâmica dos objetos em movimento, dinâmica essa que é vital para os sistemas de monitoramento de tráfego.

¹ O 511.org disponibiliza um número telefônico (511) e um serviço Web (<www.511.org>) para obter informações atualizadas sobre o trânsito da baía de São Francisco.

Na análise de aspectos locais de mobilidade, pesquisadores investigam como o comportamento de um indivíduo pode afetar diretamente o fluxo de tráfego, o qual apresenta instabilidades à medida que a densidade de veículos aumenta. Neste caso, como aqueles provocados por um indivíduo ao frear bruscamente ou ao mudar de faixa, podem causar uma reação em cadeia (TREIBER; KESTING, 2013). Essas perturbações podem causar ondas de tráfego que venham a se tornar engarrafamentos. Inúmeros trabalhos apresentam estudos investigando oscilações no fluxo de tráfego (ver o trabalho de Li *et al.* (2014) para uma visão geral). Nesses estudos, a onda de tráfego é o fenômeno mais investigado e um dos principais tópicos na teoria da dinâmica do fluxo de tráfego (STERN *et al.*, 2018; JIANG *et al.*, 2017; NI, 2015a; TREIBER; KESTING, 2013). No entanto, uma grande parte desses estudos empregam observações empíricas e modelos teóricos para analisar locais específicos. Assim, eles não podem ser aplicados facilmente na análise, em tempo real, de grandes volumes de dados de trajetórias em uma abordagem computacional sistemática.

Devido ao vasto volume de dados envolvidos, analisar dados de trajetória não é uma tarefa fácil e, muitas vezes, requer visualização exploratória para detectar e descrever padrões, tendências e relações dentro dos dados (ANDRIENKO; ANDRIENKO, 2006). Em muitos casos, é necessário que os usuários incorporem seus conhecimentos sobre o domínio e sua inteligência humana no processo de análise exploratória (ANDRIENKO *et al.*, 2017). A análise de dados de trajetória é um desses casos, e um processamento totalmente automatizado não seria eficiente. A visualização permite que os usuários usem sua capacidade de pensamento criativo e crítico, fazendo associações e gerando ideias, enquanto os computadores processam, agregam e extraem conjuntos de dados que seriam grandes demais para um ser humano lidar com eficiência sozinho (ANDRIENKO *et al.*, 2017). Neste sentido, esta tese propõe uma nova abordagem para visualizar padrões de mobilidade na análise de dinâmicas de tráfego. A seguir, são apresentados os cenários que caracterizam o problema que este trabalho pretende resolver.

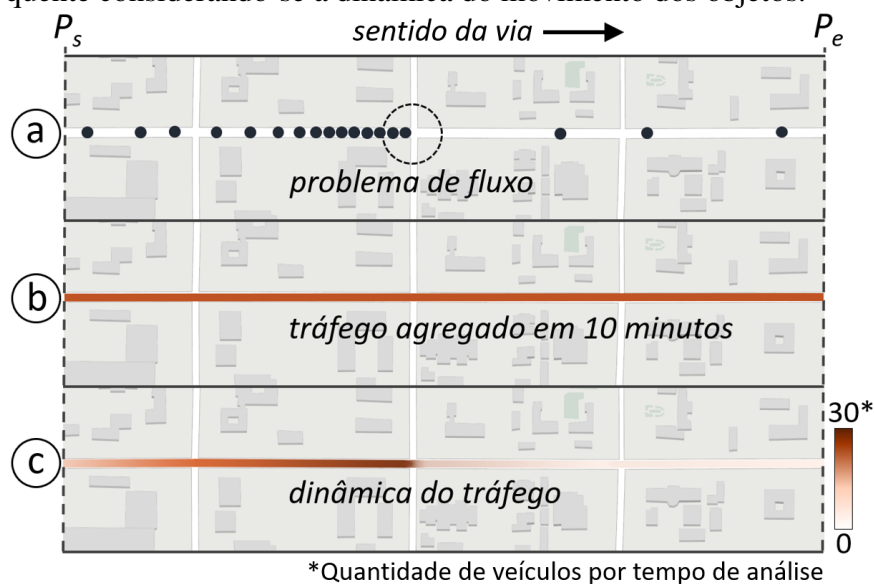
1.2 Caracterização do Problema

Nesta seção, são apresentados três cenários que comparam a abordagem proposta com técnicas tradicionais para caracterizar o problema tratado nesta tese. Os dois primeiros cenários abordam aspectos de mobilidade global, enquanto o terceiro aborda aspectos de mobilidade local. Neste trabalho, a análise de aspectos globais refere-se a encontrar padrões de mobilidade que abrangem um grande espaço geográfico. Por exemplo, encontrar as rotas mais quentes em

uma área de mapa de uma cidade inteira. Neste caso, agregações e simplificações do espaço são necessárias, como representar uma avenida com várias faixas de rolamento como uma via de pista única. Além disso, a representação dos indivíduos ou dos grupos geralmente não é necessária, mas, sim, o resultado da agregação do movimento deles. Na análise dos aspectos locais de mobilidade, o foco é no comportamento dos indivíduos. Assim, a representação do movimento dos indivíduos é fundamental para identificação de padrões. Além disso, o espaço deve ser representado de forma mais fiel, evitando simplificações ou agregações. Por exemplo, as ondas de tráfego surgem, na maioria dos casos, quando os veículos mudam de faixa de rolamento. Logo, para identificar a origem de uma onda de tráfego, o comportamento dos indivíduos e as faixas da via devem ser representadas.

No primeiro cenário, suponha que um grupo de veículos percorra um caminho entre o ponto inicial P_s e o ponto final P_e em dez minutos. Primeiro, a Figura 1 mostra que o fluxo de tráfego é prejudicado por um problema no cruzamento marcado (Figura 1a) – talvez um semáforo dessincronizado, uma rua com preferência errada ou um acidente. Neste cenário, o comportamento do tráfego é constante dentro do intervalo, e as faixas da via são representadas como uma única pista. Para simplificar a análise visando encontrar padrões de mobilidade, abordagens tradicionais agregam o movimento ocorrido dentro do intervalo de dez minutos. Neste exemplo, para encontrar a rota quente, essas abordagens marcam todo o caminho como

Figura 1 – (a) Um problema no cruzamento destacou prejudicou o fluxo de tráfego. Esse comportamento de tráfego foi constante dentro do período de análise. (b) As abordagens tradicionais marcam toda a rota de P_s até P_e , pois o que importa é o fluxo total de tráfego agregado dentro de uma janela de tempo. (c) No entanto, apenas parte da rota é quente considerando-se a dinâmica do movimento dos objetos.

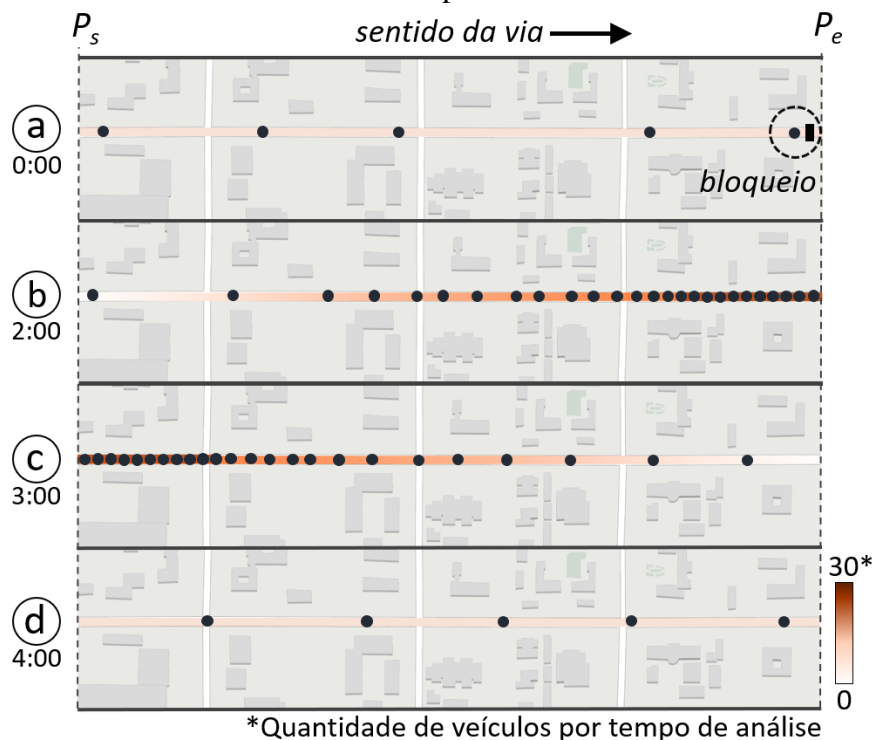


*Quantidade de veículos por tempo de análise

quente após o período (Figura 1b), sem levar em consideração como os objetos em movimento estavam distribuídos no espaço. No entanto, considerando a dinâmica do tráfego, somente a primeira parte da rota deve ser marcada como quente (Figura 1c). Apesar deste cenário mostrar um local específico, encontrar rotas quentes auxilia na análise de aspectos globais de tráfego. Para fins ilustrativos, esse local foi selecionado para demonstrar o problema.

O segundo cenário é representado na Figura 2. No início do intervalo de tempo, um carro colide contra uma árvore e bloqueia a rua no final do caminho P_e (Figura 2a). Depois de dois minutos, há um engarrafamento na via. Nesse instante, o carro quebrado é removido (Figura 2b). Assim, a via é liberada e o fluxo de tráfego é restabelecido (Figura 2c). Dois minutos depois, o fluxo de tráfego retorna ao normal (Figura 2d). Novamente, as abordagens tradicionais classificam toda a rota como quente durante toda a janela de tempo de dez minutos, não importando como os objetos em movimento estavam distribuídos no tempo. Então, essas abordagens não descrevem a dinâmica real da situação. Além disso, dependendo do limiar de densidade usado na maioria das abordagens, a rota não seria classificada como quente, pois, durante a maior parte do tempo, o fluxo de tráfego era normal e o número total de veículos seria

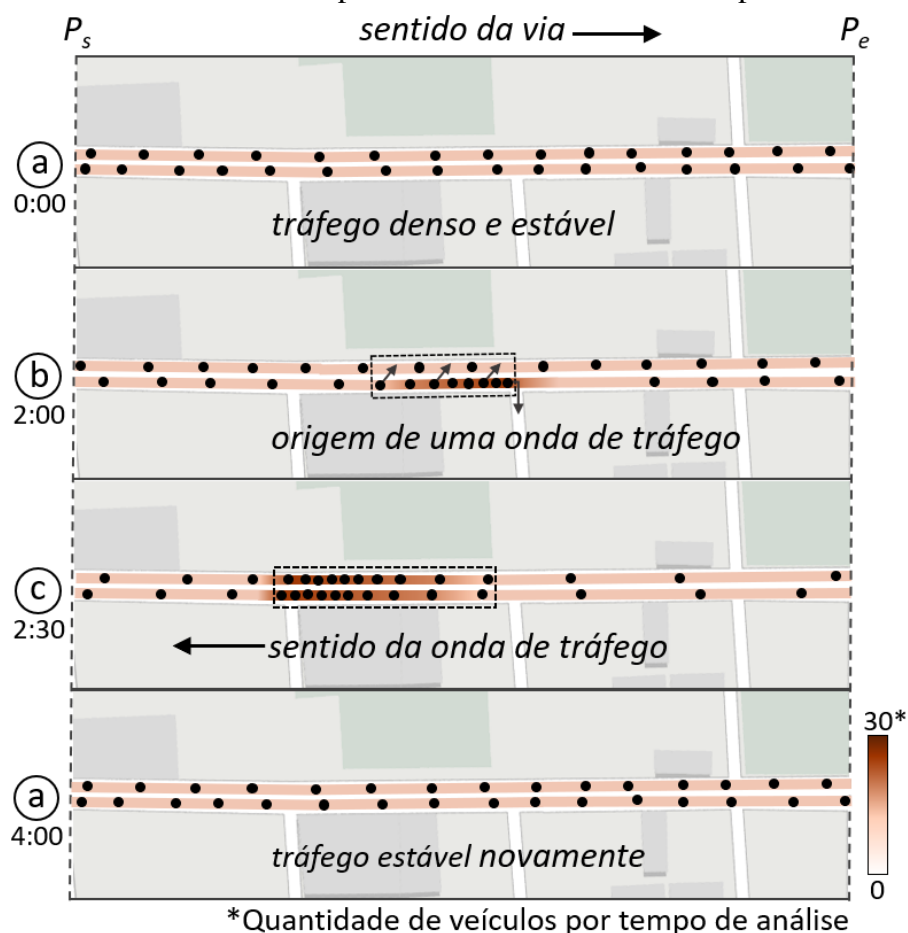
Figura 2 – (a) No início de um intervalo de dez minutos, um incidente bloqueou a rua no final do caminho P_e . (b) Após dois minutos, há um engarrafamento ao longo da rua, e o bloqueio na estrada é removido. (c) O congestionamento está se dissipando. (d) Após quatro minutos, o fluxo de tráfego na rua torna-se normal novamente e permanece assim até o final do intervalo de tempo.



insuficiente para considerar a rota como quente. Também, um evento poderia ser erroneamente considerado irrelevante se sua duração fosse dividida em duas janelas de tempo consecutivas. Assim, esse tipo de problema não seria detectado por uma aplicação de monitoramento de tráfego usando essas abordagens.

Por último, o terceiro cenário retrata um problema no fluxo de tráfego provocado por uma ação comum de um indivíduo saindo de uma via com trânsito denso. A Figura 3a representa uma via com tráfego intenso, mas com fluxo constante e sem engarrafamentos. Depois de dois minutos, um veículo decidiu fazer uma conversão à direita no local do retângulo pontilhado na Figura 3b. Para isso, ele precisou reduzir significativamente sua velocidade. Os veículos logo atrás também tiveram que reduzir a velocidade para manter uma distância de segurança e começaram oportunamente a mudar de faixa para desviar desse trecho lento (Figura 3b). Depois

Figura 3 – (a) Uma via com tráfego intenso, mas com fluxo constante e sem engarrafamentos. (b) Depois de dois minutos, um veículo decidiu fazer uma conversão à direita no local do retângulo pontilhado. Alguns veículos logo atrás mudaram de faixa. (c) Uma onda de tráfego deslocando-se em sentido oposto ao do fluxo. (d) O fluxo de tráfego normaliza-se novamente e permanece assim até o final do período de análise.



Fonte: o autor.

de 30 segundos, o fluxo se desestabilizou, provocando uma onda de tráfego em ambas as faixas. De fato, a mudança de faixa é a principal causa da propagação da onda de tráfego para todas as faixas da via (NI, 2015a). A Figura 3c mostra a onda de tráfego já formada depois de 30 segundos, movimentando-se no sentido contrário ao da via. Aos quatro minutos, o fluxo volta ao normal e permanece assim até o final do período de dez minutos. Cada acesso à via neste cenário comporta-se como um gerador estacionário de ondas de tráfego em um fluxo denso (TREIBER; KESTING, 2013). Identificar esse comportamento é essencial para compreender a dinâmica do tráfego e intervir para melhorar a eficiência do trânsito, como criar recuos de acesso, ou impedir que veículos saiam ou entrem em determinados trechos críticos da via. As abordagens tradicionais que agregam o movimento dentro de uma janela de tempo novamente falham em identificar esse problema. Assim como no cenário anterior, esse evento pode não ser percebido se estiver dividido em duas janelas de tempo consecutivas, ou se a definição do limiar de densidade estiver incorreta, já que não há uma variação significativa na densidade do fluxo durante toda a janela de tempo. Diferente do primeiro cenário, que demanda uma análise da distribuição dos objetos no espaço, e do segundo, que é dependente de uma análise de distribuição no tempo, este exige uma análise conjunta da distribuição dos objetos no tempo e no espaço simultaneamente. Para isso, são necessárias a representação do espaço com a identificação de todas as faixas da via e a representação do movimento individual de cada veículo.

Concluindo, agregar movimento empregando discretização do tempo em janelas pode apresentar inconsistências no tempo e no espaço com o fluxo de tráfego real. Não foi encontrada na revisão bibliográfica desta tese nenhuma abordagem para capturar e acompanhar progressivamente a evolução de um fluxo de tráfego identificando padrões de mobilidade global e local. No entanto, essa é uma tarefa bastante desafiadora, pois representar em tempo real o movimento de objetos sem usar janelas de tempo requer algoritmos otimizados e uso de tecnologias modernas para monitorar todo o fluxo de tráfego de uma cidade inteira.

1.3 Objetivos

O objetivo principal deste trabalho é propor uma nova abordagem, que pode ser usada por usuários de domínio e gestores urbanos, para visualizar de forma interativa e dinâmica padrões de mobilidade local e global. Essa abordagem utiliza uma estratégia incremental para descobrir e manter esses padrões, recebendo, como entrada, um fluxo de dados de trajetória. Para a obtenção deste objetivo, foram definidos os seguintes objetivos específicos:

- Realizar uma revisão bibliográfica sobre a análise de dados de trajetórias, investigando os conceitos fundamentais, as técnicas de pré-processamento de trajetórias, e as aplicações desenvolvidas;
- Descrever uma arquitetura com seus principais componentes, seguindo um modelo de dados compacto, incremental e baseado em grade. Nesse modelo, dados históricos não são armazenados, viabilizando, assim, uma aplicação de monitoramento de dinâmicas de tráfego;
- Especificar uma metodologia sistemática para visualizar os padrões de mobilidade a partir de dados brutos de trajetórias. Para isso, são descritas as etapas de pré-processamento, sincronização, renderização, e processamento visual;
- Propor e aplicar uma visualização para análise de padrões de mobilidade global, monitorando a evolução das rotas com tráfego pesado, ou seja, as rotas quentes;
- Analisar oscilações de tráfego através da visualização de ondas de tráfego, caracterizadas como um padrão local de mobilidade;
- Avaliar a eficácia e o desempenho das visualizações propostas através de experimentos extensivos em conjuntos de dados reais e simulados com a intenção de verificar a possibilidade da aplicação da abordagem proposta no monitoramento de todo o fluxo de tráfego de uma cidade inteira;
- Apresentar os benefícios e as limitações da abordagem proposta com base nos resultados dos experimentos e na avaliação por um especialista do domínio.

1.4 Metodologia

A elaboração deste trabalho seguiu a seguinte metodologia:

1. Estudo da fundamentação teórica necessária para o desenvolvimento desta tese;
2. Realização de uma revisão bibliográfica segundo uma análise sistemática de trabalhos relacionados;
3. Definição conceitual da abordagem proposta, especificando os componentes da arquitetura e as etapas da metodologia;
4. Desenvolvimento de protótipos para definição dos algoritmos de processamento gráfico em GPU e das tecnologias para processamento em tempo real de grandes volumes de dados;
5. Desenvolvimento dos componentes da arquitetura e seus algoritmos;
6. Aplicação da proposta para visualizar padrões de mobilidade global com dados de trajetó-

rias reais, comparando os resultados com uma abordagem tradicional para encontrar rotas quentes;

7. Avaliação experimental para identificar os benefícios e limitações da abordagem proposta;
8. Desenvolvimento de uma aplicação para visualização de ondas de tráfegos com o objetivo avaliar a eficácia dessa aplicação com um especialista do domínio;
9. Avaliação da eficiência através de uma análise de desempenho verificando a adequação do método proposto para visualizar o fluxo total de tráfego de uma cidade inteira com as tecnologias empregadas;
10. Análise dos resultados das aplicações e ajustes na proposta.

1.5 Contribuições

As principais contribuições desta tese são:

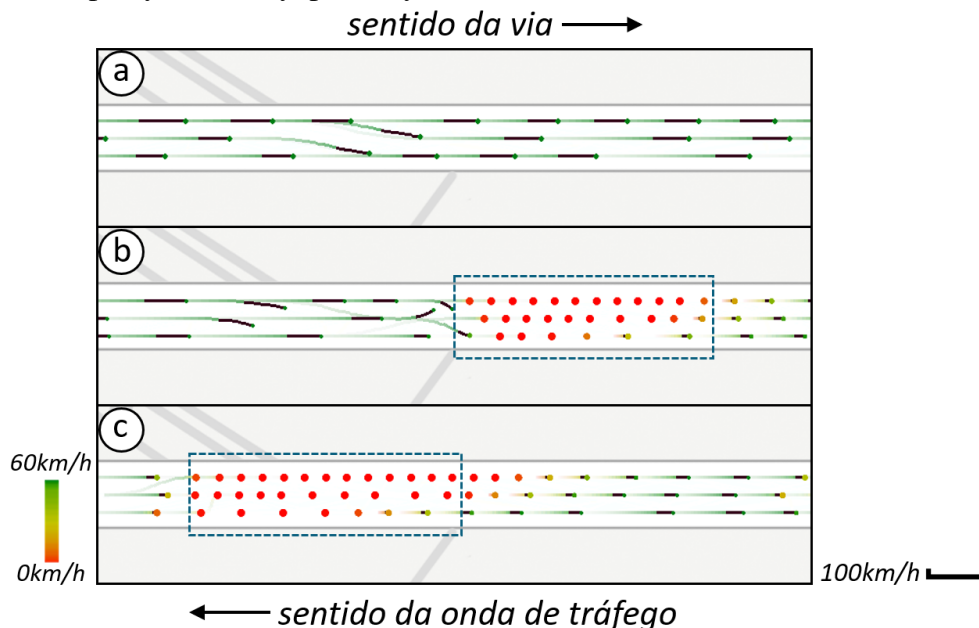
1. Uma nova abordagem para descobrir e monitorar padrões de mobilidade local e global, acompanhando a evolução desses padrões. Diferentemente da maioria das abordagens, a abordagem proposta considera que um objeto em movimento pode aparecer, desaparecer ou atualizar sua posição atual ao longo do tempo. Os experimentos realizados mostram resultados mais consistentes com a dinâmica real do tráfego;
2. Um método para facilitar a análise dos padrões de mobilidade, visualizando o movimento dos objetos no espaço sem utilizar janelas de tempo, isto é, o cálculo é realizado continuamente ao longo do tempo. Para isso, são empregados recursos visuais, como rastros, formas, animações e codificações de cores e comprimentos, para perceber comportamentos, como velocidade, direção, aceleração e localização;
3. Uma metodologia passo-a-passo sistemática para visualizar dinâmicas de tráfego a partir de dados brutos – latitude, longitude e tempo. A abordagem proposta utiliza principalmente dois parâmetros: o tempo de análise contínuo e o limiar de densidade. Ambos os parâmetros podem ser ajustados em tempo real e as alterações são visualizadas instantaneamente. Diferentemente de outras abordagens, não é necessário que o conjunto de dados inteiro seja mapeado para uma rede viária por *map-matching* (discutido posteriormente no Capítulo 2), e nem uma análise estatística prévia para a definição dos parâmetros.
4. Um modelo de dados compacto para representação de trajetórias o qual não cresce com a quantidade de objetos em movimento processados. O tamanho de um fluxo de dados é potencialmente ilimitado. As abordagens tradicionais armazenam parte dos dados que

chegam e o resultado, o qual depende do tamanho do conjunto de dados armazenado. A abordagem proposta usa uma representação baseada em grade de tamanho fixo, a qual é independente do tamanho da entrada. Nela, os padrões de mobilidade são guardados, mas não é possível identificar nem rastrear os objetos que geraram esses padrões. Nesse modelo, os dados dos objetos são descartados após serem processados para economizar poder computacional e espaço de memória, requisito essencial para aplicações em tempo real;

5. Uma abordagem que é executada em paralelo na Graphics Processing Unit (GPU) e mantém taxas de exibição interativa de quadros (*frame rates*), mesmo com grandes conjuntos de dados. Experimentos extensivos em conjuntos de dados reais e simulados foram realizados para avaliar a eficácia e o desempenho do método proposto. Os resultados demonstram que a abordagem proposta é eficientemente escalonada linearmente com o tamanho do conjunto de dados.

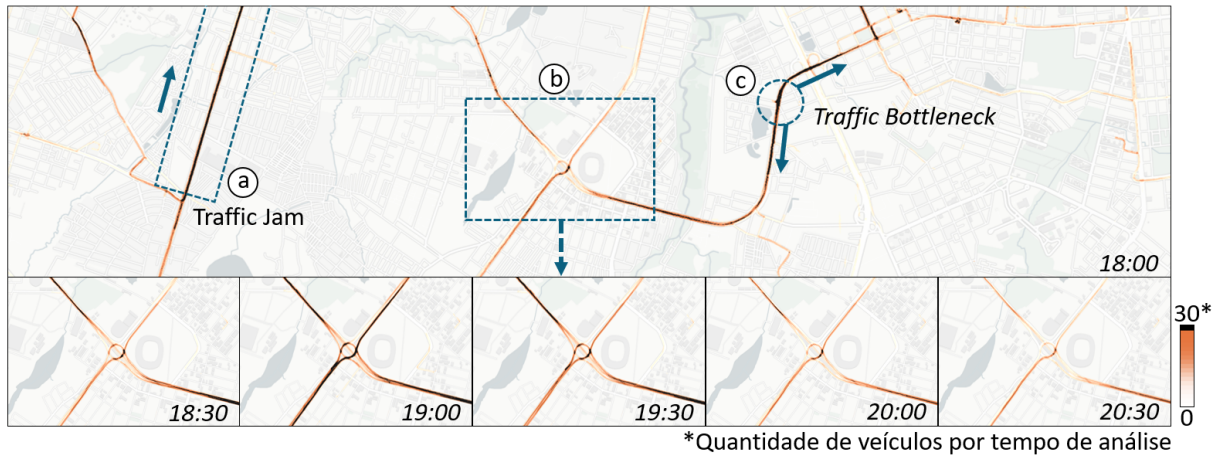
Como demonstração, a Figura 4a apresenta um exemplo da visualização de ondas de tráfego utilizando a proposta deste trabalho. Através da codificação da velocidade com

Figura 4 – Uma demonstração da visualização de ondas de tráfego proposta neste trabalho. (a) O tráfego está denso, mas fluindo normalmente. Uma onda de tráfego (retângulo pontilhado) movendo-se da direita (b) para a esquerda (c) pode ser identificada por cor, comprimento dos rastros dos veículos e tamanho dos marcadores. As escalas de cor à esquerda e comprimento à direita codificam a velocidade para auxiliar a identificação. Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/uycpCRSPyLo>>.



Fonte: o autor.

Figura 5 – Rotas quentes descobertas pela abordagem proposta, usando um conjunto de dados brutos reais. (a) Um engarrafamento propagando-se para cima ao longo da via. (b) A evolução das rotas quentes no retângulo pontilhado central em cinco momentos registrados a cada meia hora. (c) À direita, as rotas quentes estão se formando em ambos os sentidos devido a um gargalo de tráfego. Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/mDUy956-vSA>>.



Fonte: o autor.

uma escala de cor, com uma escala de comprimento dos rastros dos objetos e com o tamanho dos marcadores, neste exemplo, é possível identificar uma onda de tráfego deslocando-se da direita (Figura 4b) para a esquerda (Figura 4c). Para a análise de aspectos globais, a Figura 5 apresenta um exemplo de rotas quentes descobertas usando um conjunto de dados brutos reais. A Figura 5a mostra um engarrafamento propagando-se para cima ao longo da via. A Figura 5b apresenta a evolução das rotas quentes em cinco momentos registrados a cada meia hora. À direita (Figura 5c), as rotas quentes estão se formando em ambos os sentidos devido a um gargalo de tráfego.

1.6 Publicações

Ainda como resultado desta tese foram geradas as seguintes publicações:

1. GOMES, G. A.; SANTOS, E.; VIDAL, C. A. Interactive visualization of traffic dynamics based on trajectory data. In: IEEE. **Graphics, Patterns and Images, 2017 30th SIBGRAPI Conference on**. [S.l.], 2017. p. 111-118.
2. GOMES, G. A.; SANTOS, E.; VIDAL, C. A.; SILVA, T. L. C.; MACEDO, J. A. F. Real-time discovery of hot routes on trajectory data streams using interactive visualization based on GPU. **Computers & Graphics**, Elsevier BV, v. 76, p. 129-141, nov 2018.

1.7 Organização

Este capítulo apresentou o contexto motivacional para a escrita desta tese, bem como a caracterização do problema, os objetivos, a metodologia e as contribuições. O restante deste trabalho está organizado em seis capítulos. O Capítulo 2 discute a fundamentação teórica com os principais conceitos necessários para a compreensão da proposta deste trabalho. No Capítulo 3, a revisão bibliográfica com a identificação dos principais trabalhos relacionados é apresentada. O Capítulo 4 define formalmente o problema e descreve a arquitetura e a metodologia projetadas neste trabalho. Em seguida, no Capítulo 5, são descritas as aplicações da abordagem proposta para visualizar padrões de mobilidade global, através da análise de rotas quentes, e padrões de mobilidade local, através da visualização de ondas de tráfego. No Capítulo 6, o desempenho é analisado para avaliar a eficiência das visualizações propostas. Por último, as considerações finais e os possíveis trabalhos futuros são apresentados.

2 CONCEITOS PRELIMINARES

Este capítulo introduz os principais conceitos relacionados com a análise de dados de trajetória. Inicialmente, na Seção 2.1, apresentam-se os fundamentos relacionados aos dados de trajetórias. Na Seção 2.2, discutem-se as etapas para o pré-processamento de trajetórias. Por último, na Seção 2.3, tecem-se algumas considerações finais.

2.1 Dados de Trajetórias

O rápido crescimento populacional nos grandes centros urbanos impõe uma série de desafios que precisam ser tratados através de estratégias de planejamento eficientes. Alguns dos problemas mais críticos são: o congestionamento de trânsito, o alto consumo de energia, a poluição atmosférica, e o atendimento precário às demandas por serviços essenciais (DEAR; SCOTT, 2018; ZHENG *et al.*, 2014). Superar esses desafios era quase impossível há algumas décadas por conta da complexidade e da dinâmica dos grandes centros urbanos (GILBERT; GUGLER, 1992). Atualmente, a disseminação dos dispositivos móveis equipados com GPS (*Global Positioning System*) e dos serviços baseados em localização tornou economicamente e tecnicamente viável a coleta de grande quantidade de dados de trajetórias de objetos em movimento (ANDRIENKO *et al.*, 2017).

Os dados de trajetórias podem ser usados para análise em diferentes tipos de aplicação, tais como: no planejamento urbano, para identificar padrões de fluxo de tráfego e as causas de congestionamentos de trânsito (XIA *et al.*, 2018; KONG *et al.*, 2018; PARK *et al.*, 2018; DING *et al.*, 2018a; KWON *et al.*, 2018; GONZALEZ *et al.*, 2008); na análise de mobilidade sustentável (WANG; MORIARTY, 2018; LIU *et al.*, 2018); na otimização de rotas para empresas de transportes (KANCHARLA; RAMADURAI, 2018; LUKMAN *et al.*, 2018); e nos serviços personalizados de localização (RAJALAKSHMI; GOYAL, 2018; YANG; LIN, 2018).

Neste trabalho, estudam-se apenas os dados de trajetórias georreferenciados, que são denominados aqui: dados de trajetórias espaciais, ou, simplesmente, dados de trajetórias. Uma trajetória espacial é um caminho gerado por um objeto ao mover-se em um espaço geográfico (ZHENG; ZHOU, 2011a), normalmente representado por uma série de pontos ordenados cronologicamente, por exemplo, $p_1 \rightarrow p_2 \rightarrow \dots \rightarrow p_i \rightarrow \dots \rightarrow p_n$, onde n é a quantidade de pontos de uma trajetória. Um ponto p_i é definido por uma tupla formada por: suas coordenadas geoespaciais, sua marcação de tempo (*timestamp*) e seus atributos. Mais formalmente,

$p_i \in (S_d \times T \times A_1 \times \dots \times A_m)$, onde S_d define as coordenadas espaciais do ponto (por exemplo, se $d = 3$, S_d representa latitude, longitude e altitude), T define o tempo, e $A_j : 1 \leq j \leq m$ são atributos quantitativos ou qualitativos. Alguns dos atributos podem ser derivados diretamente das coordenadas geoespaciais e do *timestamp*, como, por exemplo, a velocidade, a direção, a aceleração e a distância percorrida. Outros atributos, por sua vez, são especificados de forma independente, como é o caso do tamanho e do formato do objeto. Essa definição mostra a complexidade do problema para análise de dados de trajetórias: aspectos espaciais e temporais, bem como atributos numéricos ou categóricos. Andrienko *et al.* (2011) apresentam uma classificação de atributos que podem ser extraídos de dados de trajetórias espaciais. Os três componentes, i.e., espaço, tempo e atributos, levam a uma incontestável riqueza de informação que torna a análise de dados de trajetórias uma tarefa interessante, desafiadora e produtiva.

A análise de dados de trajetórias objetiva perceber as inter-relações entre as componentes espacial (S) e temporal (T), de um lado, e as componentes de atributo (A), do outro lado, ou seja, a dependência funcional $S \times T \rightarrow A$ que define o comportamento dos atributos em relação ao espaço e ao tempo (TOMINSKI *et al.*, 2012). Um exemplo é analisar o comportamento da variação de velocidade de carros, ao longo de um dia, em um determinado cruzamento $s \in S$, ou seja, analisar a relação funcional $s \times T \rightarrow A$. Outro exemplo é analisar o comportamento da variação de velocidade ao longo de uma rodovia em um tempo específico $t \in T$, ou seja, analisar a relação funcional $S \times t \rightarrow A$. Assim, a análise de dados de trajetória deve permitir caracterização, pesquisa e comparação de comportamentos $T \rightarrow A$ e $S \rightarrow A$.

Portanto, entender toda a variedade de aspectos envolvidos na análise de dados de trajetórias espaciais é desafiador, pois esses dados podem conter informações interessantes não apenas no nível de trajetórias individuais, mas também no nível de conjunto de trajetórias. Por exemplo, trajetórias que cruzam regiões específicas no espaço ou que pertencem a intervalos específicos no tempo. Para grandes volumes de dados, determinar onde as informações interessantes podem ser encontradas não é uma tarefa fácil, nem tampouco é fácil identificar quais conjuntos de trajetórias precisam ser analisados detalhadamente.

2.2 Pré-processamento de Trajetórias

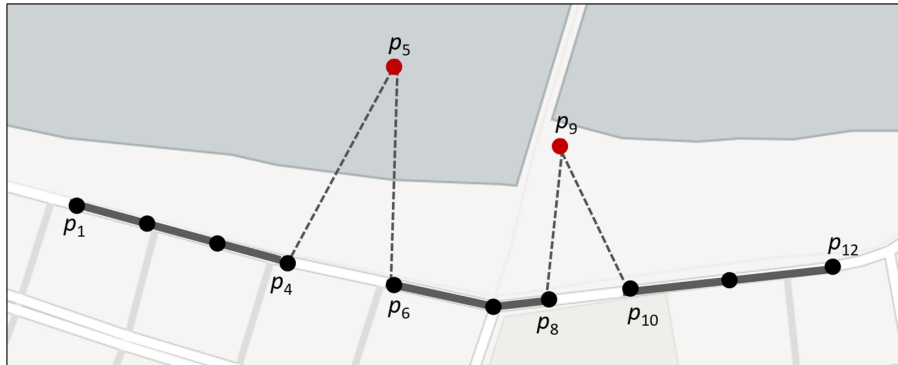
Antes de começar a analisar dados de trajetórias, é necessário lidar com uma série de problemas que podem afetar diretamente a qualidade da visualização desses dados. Como mencionado anteriormente, o movimento contínuo de um objeto é representado de forma

aproximada por uma trajetória espacial através de um conjunto de amostras discretas de pontos de localização (ZHENG; ZHOU, 2011a). Quanto maior for a taxa de amostragem de pontos, mais precisa será a trajetória. Entretanto, altas taxas de amostragem resultam em uma enorme quantidade de dados, que elevam os custos computacionais. Além desse problema, o uso de sensores na captura dos pontos da trajetória acarreta erros de localização em parte dos dados, e, assim, as trajetórias geradas com dados brutos apresentam imprecisões causadas pelos erros de localização embutidos nos dados capturados. Se esses erros não forem eliminados ou minimizados, a visualização desses dados pode apresentar poluição visual, que prejudica a análise dos dados. Para eliminar ou minimizar essas imprecisões e permitir uma análise confiável das trajetórias, é necessário realizar o tratamento prévio das trajetórias, conhecido como pré-processamento de trajetórias. Além disso, eliminar dados redundantes nesse momento, como pontos muito próximos, ajuda na redução da poluição visual e diminui a carga de processamento e armazenamento de dados para a visualização de grandes volumes de dados. Assim, no pré-processamento de trajetórias, são realizadas as etapas de correção, compressão, detecção de paradas, segmentação, *map-matching* e enriquecimento semântico de trajetórias.

O objetivo da etapa de correção é remover os pontos que apresentam erros, sem perder as principais características da trajetória (PATIL *et al.*, 2018; YANG *et al.*, 2018; ZHANG *et al.*, 2017). As causas dos erros de captura são variadas: baixo número de satélites disponíveis, condições atmosféricas ruins, falha de amostragem, reflexão e bloqueio de sinais. Uma revisão das metodologias atuais de registro de localização por GPS e as possíveis causas de erros é encontrada em (HE *et al.*, 2017). Um exemplo de erro em uma trajetória é mostrado na Figura 6. Nesse exemplo, os pontos p_5 e p_9 podem ser efetivamente identificados usando média e desvio padrão para detectar pontos que têm velocidade extremamente alta para a trajetória atual. Além disso, erros podem ser propositalmente gerados por motivo de privacidade. Nesses casos, as pessoas que não querem revelar suas localizações expõem deliberadamente apenas uma aproximação de posicionamento (BAMBA *et al.*, 2008). Além de remover ou reduzir erros de posicionamento, em alguns casos, é necessário encontrar amostras de posições que estejam faltando. Essa ausência de dados também é considerada um erro e pode acontecer devido à perda de sinal, por exemplo, quando o dispositivo com GPS entra em um túnel. Uma das abordagens para corrigir esse tipo de erro utiliza interpolação polinomial para estimar as amostras ausentes (HOREMUŽ; ANDERSSON, 2006).

A etapa de compressão de trajetórias objetiva reduzir o tamanho de uma trajetó-

Figura 6 – Exemplo de trajetória com erros de GPS. Os pontos p_5 e p_9 podem ser identificados usando média e desvio padrão para detectar velocidades extremamente altas para a trajetória atual.

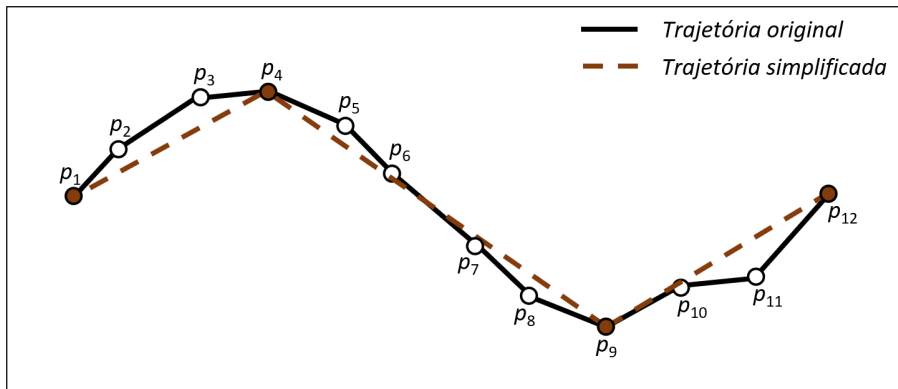


Fonte: Elaborado pelo autor.

ria para evitar sobrecargas no processamento, na comunicação e no armazenamento de dados (ZHENG, 2015). Assim, reduzindo a quantidade de dados brutos, é possível gerar representações mais compactas que descrevem o movimento de um objeto. Além disso, muitas aplicações não precisam de tal precisão de localização. Do ponto de vista geométrico, os algoritmos de compressão buscam a simplificação de linhas para remover pontos de uma trajetória sem distorcer sua geometria. A Figura 7 ilustra um exemplo de compressão de trajetórias utilizando o algoritmo de Douglas-Peucker (DOUGLAS; PEUCKER, 1973). Esse problema também é investigado pela comunidade de cartografia e de computação gráfica (DING *et al.*, 2018b; QIAN; LU, 2017; DU *et al.*, 2017; LIN *et al.*, 2017).

A detecção de paradas identifica locais onde os indivíduos permaneceram parados por algum tempo, como supermercados e atrações turísticas. Uma sequência de pontos de paradas pode fornecer uma visão abstrata para a compreensão do movimento de um indivíduo,

Figura 7 – Exemplo de compressão de dados de trajetórias utilizando o algoritmo de Douglas-Peucker (DOUGLAS; PEUCKER, 1973). A nova trajetória é composta apenas pelos pontos $p_1 \rightarrow p_4 \rightarrow p_9 \rightarrow p_{12}$.



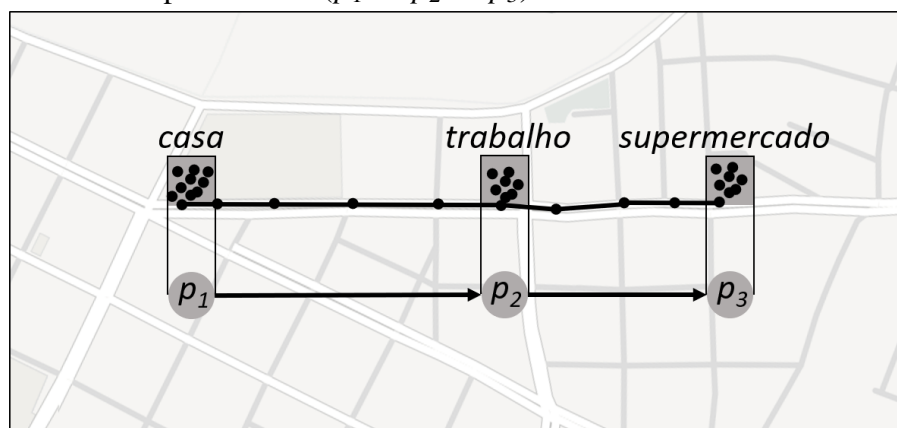
Fonte: Adaptado de (ZHENG, 2015).

ao invés de analisar a sequência original de todos os pontos de uma trajetória. Assim, uma trajetória com milhares de pontos GPS poderia ser representada apenas pelos lugares mais importantes. Por exemplo, a Figura 8 representa a detecção dos pontos de parada *casa* → *trabalho* → *supermercado*. A detecção de pontos de parada pode ser calculada por critérios espaciais e temporais, como velocidade, densidade e direção (BERMINGHAM; LEE, 2018; HWANG *et al.*, 2017; ZHENG, 2015; ANDRIENKO *et al.*, 2013).

Na etapa de segmentação, as trajetórias são subdivididas para serem usadas em processo de análise posterior. Geralmente, dados brutos de trajetórias espaciais possuem o registro do movimento de um objeto em um grande período de tempo, ou seja, a duração de uma trajetória pode ser de dias ou até de meses. Assim, para tornar o processo de análise de dados mais eficiente, é necessário dividir as trajetórias em segmentos, nos quais o objeto em movimento mantém um comportamento homogêneo, como mesmo local, velocidade semelhante, mesma direção e sentido (BUCHIN *et al.*, 2010). A classificação (BARBOSA *et al.*, 2018; WU; CHAN, 2017) e a clusterização (ZHANG *et al.*, 2018; CAI *et al.*, 2017) de trajetórias são exemplos de processos de análise que demandam segmentação. Em geral, as trajetórias são segmentadas por três critérios: intervalo de tempo, geometria e semântica (HWANG *et al.*, 2018; AHEARN; DODGE, 2018; DAMIANI *et al.*, 2018; SOARES JUNIOR *et al.*, 2018). A Figura 9 mostra dois exemplos de segmentação de trajetórias por intervalo de tempo: segmentação por tempo fixo e segmentação por limiar de tempo entre pontos consecutivos.

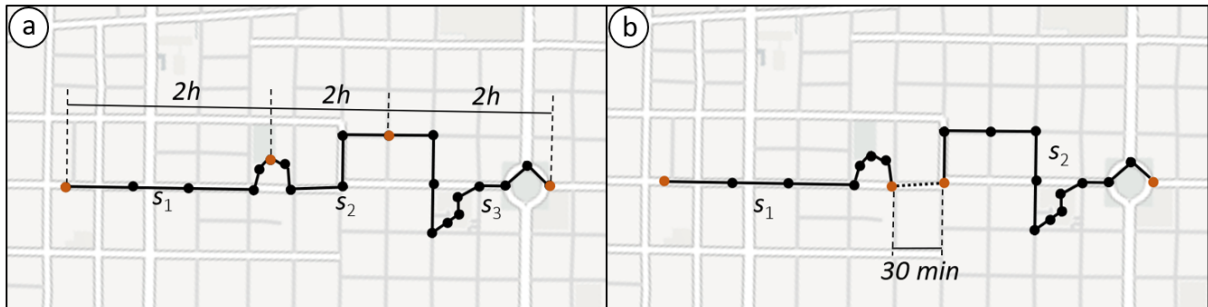
A etapa de *map-matching* converte dados brutos de GPS em um modelo lógico representativo, como uma rede viária de uma cidade representada por um grafo. Geralmente indivíduos em movimento, tais como pessoas e veículos, são limitados a transitar dentro de uma

Figura 8 – Representação estruturada de uma trajetória através de seus pontos de parada: casa, trabalho e supermercado ($p_1 \rightarrow p_2 \rightarrow p_3$).



Fonte: Elaborado pelo autor.

Figura 9 – (a) Exemplo de trajetória segmentada por tempo fixo de duração de duas horas em três segmentos (s_1, s_2, s_3). (b) Mesma trajetória segmentada por limiar de tempo de 30 minutos entre dois pontos consecutivos (s_1, s_2).

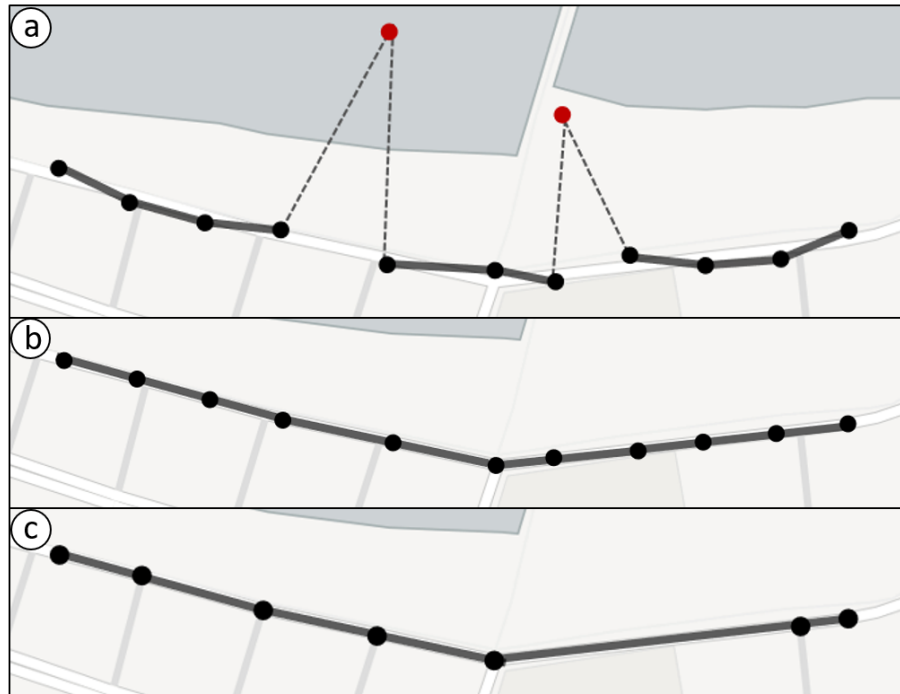


Fonte: Adaptado de (ZHENG, 2015).

rede restrita (GüTING *et al.*, 2005). Por exemplo, os trens transitam em uma rede ferroviária, os ônibus públicos têm suas próprias rotas planejadas, e os carros se deslocam dentro de uma rede viária. Dadas uma base de dados de pontos GPS e uma rede restrita, os algoritmos de *map-matching* buscam transformar esses pontos de modo a posicioná-los sobre os segmentos de rede por onde o objeto efetivamente passou (CHAMBERS *et al.*, 2018; CHEN *et al.*, 2018; ALGIZAWY *et al.*, 2017). Como as redes restritas são geralmente mais precisas do que as leituras de GPS (Figura 10a), o *map-matching* permite melhorar a precisão do ponto observado (Figura 10b). Além do mais, o *map-matching* pode ser utilizado para auxiliar na etapa de compressão dos dados (Figura 10c), pois permite representar a trajetória apenas pelos vértices ou segmentos da rede (SUI; YANG, 2018). Entretanto, o *map-matching* não é um problema fácil de ser resolvido por causa de várias características da rede restrita. Por exemplo, mapear pontos em segmentos paralelos próximos ou em cruzamento de segmentos é complicado (ZHENG, 2015). Além disso, esse mapeamento é um processo bastante dispendioso computacionalmente.

Por último, o enriquecimento semântico visa atribuir contexto aos dados das trajetórias, segundo um modelo de conhecimento, para análise posterior dos dados enriquecidos. Geralmente, o processo de enriquecimento é a última etapa do pré-processamento das trajetórias. Nesta etapa, os dados são correlacionados com informações geográficas – como pontos e regiões de interesse – e podem ser relacionados com anotações de redes sociais ou com outros conjuntos de dados relevantes para o domínio da aplicação desenvolvida. Por exemplo, na Figura 8, a classificação dos pontos (casa, trabalho e supermercado) pode ser considerada um enriquecimento semântico dessa trajetória. Vários trabalhos foram desenvolvidos para transformar dados brutos de movimento em trajetórias semanticamente enriquecidas (NOGUEIRA *et al.*, 2018; SANTIPANTAKIS *et al.*, 2017; BOGORNY *et al.*, 2013; BAGLIONI *et al.*, 2009).

Figura 10 – (a) Exemplo de trajetória com as leituras de GPS originais com dois erros – pontos vermelhos. (b) A trajetória foi corrigida com a remoção dos erros após o *map-matching*. (c) Compactação da trajetória com sua representação pelos vértices (esquinas) da rede viária, ou seja, quantidade de pontos reduzida.



Fonte: Elaborado pelo autor.

2.3 Considerações Finais

Este capítulo apresentou os principais conceitos relacionados com a análise de dados de trajetórias. Os dados de trajetórias oferecem conhecimento sem precedentes para entender objetos em movimento. A demanda por uma ampla gama de aplicações exige pesquisas sistemáticas sobre novas tecnologias de computação para a descoberta de conhecimento a partir desses dados. As pesquisas associadas a esse tema têm sido consideradas cada vez mais relevantes, e, por isso, têm atraído a atenção de várias outras áreas, incluindo biologia, sociologia, economia e geografia. Os trabalhos mais relacionados à abordagem proposta são discutidos no próximo capítulo.

3 TRABALHOS RELACIONADOS

Este capítulo apresenta os trabalhos mais relevantes relacionados com a visualização interativa de dinâmica de tráfego de veículos utilizando dados de trajetórias. Na Seção 3.1, apresentam-se algumas aplicações relacionadas com a análise de dados de trajetórias. Em seguida, na Seção 3.2, apresentam-se os estudos sobre a clusterização de objetos em movimento. Em seguida, na Seção 3.3, relacionam-se os trabalhos sobre clusterização de trajetórias. Na Seção 3.4, apresentam-se os métodos para descoberta de rotas quentes. Depois, na Seção 3.5, discutem-se os trabalhos de visualização de dados de trajetória. Por último, tecem-se algumas considerações finais.

3.1 Aplicações de Análise de Dados de Trajetória

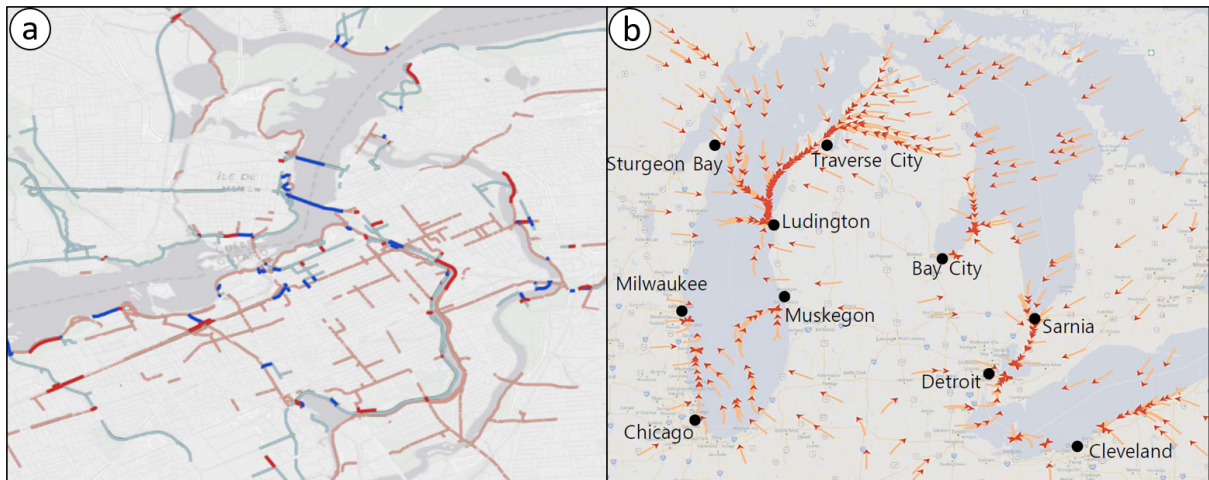
Nesta seção, alguns exemplos que analisam dados de trajetórias são apresentados de modo geral para contextualizar as possibilidades de aplicações desses dados no estudo de mobilidade de pessoas e veículos. Especificamente, os trabalhos de visualização de dados de trajetórias são apresentados na Seção 3.5, no qual são discutidos aspectos técnicos relacionados à abordagem proposta nesta tese.

Frequentemente, pessoas registram, passiva ou ativamente, seus movimentos no mundo real sob a forma de trajetórias espaciais. Na forma ativa de registro, os indivíduos gravam suas trajetórias com a finalidade de memorizar uma viagem ou compartilhar experiências. Os dados oriundos dessa forma de registro permitem a investigação de padrões de movimento de pessoas em uma cidade. Por exemplo, dados compartilhados por ciclistas foram usados para auxiliar no planejamento urbano (CONROW *et al.*, 2018; HE; SHIN, 2018; TU *et al.*, 2018; ZHANG; MI, 2018). Boss e seus coautores (2018) apresentaram uma técnica para investigar padrões de mobilidade em dados compartilhados através do aplicativo Strava¹ com o intuito de planejar ciclofaixas e ciclovias. A Figura 11a ilustra uma análise visual de dados coletados de ciclistas da cidade de Ottawa, no Canadá (BOSS *et al.*, 2018). No Flickr², uma série de fotos georreferenciadas pode representar uma trajetória espacial na qual cada foto corresponde a uma marcação de localização e de tempo. Os dados dessa plataforma podem ser usados para encontrar pontos ou regiões de interesse da população de uma cidade (BELCASTRO *et al.*, 2018; WANG

¹ Strava é um aplicativo usado para rastrear atividades de ciclismo, corrida e natação via GPS.

² Flickr (www.flickr.com) é um dos maiores sites de compartilhamento de imagens, com mais de 13 bilhões de fotos.

Figura 11 – (a) Análise visual de rotas de ciclistas em Ottawa, Canadá. (b) Mapa de fluxo de movimento de pessoas gerado a partir de dados do Twitter.



Fonte: a: (BOSS *et al.*, 2018), b: (KIM *et al.*, 2018)

et al., 2018b; HAUSMANN *et al.*, 2017).

Da mesma forma, o conjunto de *check-ins* de um usuário em uma rede social baseada em localização pode ser considerado uma trajetória, quando classificado cronologicamente. Por exemplo, dados do Twitter são comumente usados para investigar padrões no fluxo de pessoas em centros urbanos (KIM *et al.*, 2018; LIN *et al.*, 2018b). A Figura 11b ilustra um mapa de fluxo de movimento de pessoas gerado a partir de dados do Twitter (KIM *et al.*, 2018). Os dados de *check-in* de um usuário em uma rede social, por serem considerados uma trajetória de amostragem baixa, já foram usados no estudo de incerteza de trajetórias (FURTADO *et al.*, 2017) e na previsão de localização em trajetórias esparsas (WU *et al.*, 2018).

Na forma passiva de registro, as trajetórias são capturadas independentemente da ação do indivíduo. Por exemplo, uma pessoa com um telefone celular gera, involuntariamente, várias trajetórias espaciais (XU *et al.*, 2018a; XU *et al.*, 2018b). De maneira análoga, as transações efetuadas com um cartão de crédito em vários estabelecimentos comerciais registram, também de forma involuntária, a trajetória espacial do titular do cartão (LIN *et al.*, 2018a; TERROVITIS *et al.*, 2017), uma vez que cada transação contém um *timestamp* e um identificador do comerciante do local onde a transação ocorreu.

Os veículos de transporte equipados com dispositivos GPS geram dados de trajetórias espaciais que podem ser usados para fornecer uma visão de vários aspectos de mobilidade em centros urbanos. Por exemplo, a Taxi Limousine Commission³ registra e disponibiliza trajetórias de uma das maiores frotas de táxis do mundo, com mais de 13 mil táxis da cidade de Nova York.

³ A TLC (www.nyc.gov/html/tlc) disponibiliza informações de viagens de mais de 240 milhões de passageiros por ano. Os dados foram disponibilizados mensalmente de 2009 até 2017.

Recentemente, vários estudos foram realizados explorando e investigando esses dados (SUN *et al.*, 2018; JEONG *et al.*, 2018; HAGAN *et al.*, 2018; DORAISWAMY *et al.*, 2018). Por exemplo, Tseng e seus coautores (2018) examinaram os benefícios da utilização de táxis elétricos na cidade de Nova York em comparação com os táxis convencionais com motores de combustão interna. A Figura 12a ilustra um exemplo de análise de densidade de tráfego de táxis também em Nova York (DORAISWAMY *et al.*, 2018).

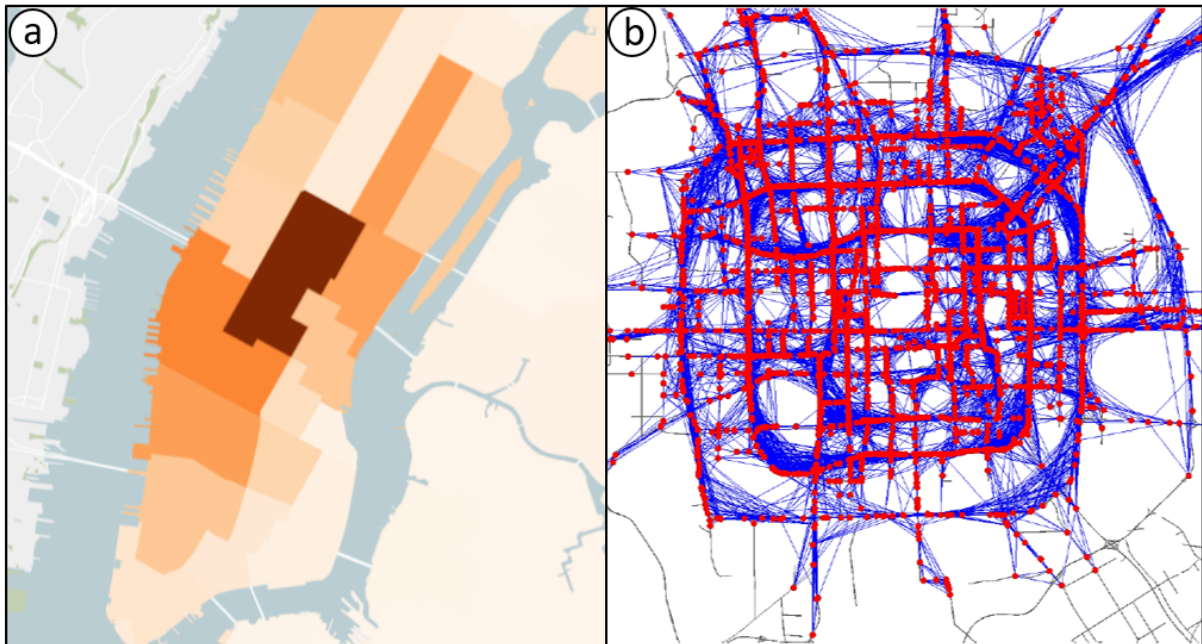
Além dessa, a base de dados T-Drive (YUAN *et al.*, 2011), que disponibiliza trajetórias espaciais de mais de 10 mil táxis da cidade de Pequim, foi também bastante estudada nos últimos anos (LEAL *et al.*, 2018; CAO *et al.*, 2018; HAN *et al.*, 2018; GOMES *et al.*, 2017). A Figura 12b mostra em vermelho as rotas mais utilizadas por táxis da cidade de Pequim, China (YUAN *et al.*, 2010). Atualmente, novos serviços de mobilidade sob demanda, como Uber e Lyft, estão crescendo em popularidade entre os viajantes urbanos (CRAMER; KRUEGER, 2016). Reconhecendo os possíveis impactos desses serviços, foram realizados vários estudos para analisar questões relativas a: qualidade e oferta (ALEMI *et al.*, 2018; HUGHES; MACKENZIE, 2016), planejamento urbano (JIANG *et al.*, 2018; DAVIDSON *et al.*, 2017), e impacto nos serviços tradicionais de táxis (JIANG; ZHANG, 2018; CRAMER; KRUEGER, 2016) e nas economias locais (ROGERS, 2015; HALL; KRUEGER, 2017). Muitos desses estudos utilizaram dados de trajetórias disponibilizados pelas próprias empresas prestadoras de serviço.

Por fim, na literatura, há vários outros trabalhos que tratam da análise de trajetórias de outros meios de transportes, como transporte público em geral (DABIRI; HEASLIP, 2018; KYAW *et al.*, 2018), caminhões (KANCHARLA; RAMADURAI, 2018; LUKMAN *et al.*, 2018), aeronaves (ANDRIENKO *et al.*, 2018; LI *et al.*, 2018) e navios (GUO *et al.*, 2018; ASSAF *et al.*, 2018).

3.2 Clusterização de Objetos em Movimento

O enorme volume de dados de trajetórias espaciais disponível atualmente cria oportunidades para analisar os padrões de mobilidade de objetos em movimento, que podem ser representados por uma trajetória individual contendo um determinado padrão ou um grupo de trajetórias compartilhando padrões semelhantes. Um método bastante difundido na literatura para descobrir padrões de mobilidade é a clusterização, que é o processo de agrupar um conjunto de objetos físicos ou abstratos em classes similares (HAN *et al.*, 2011). Assim, esse processo fornece um resumo dos padrões de distribuição e correlações em um conjunto de dados.

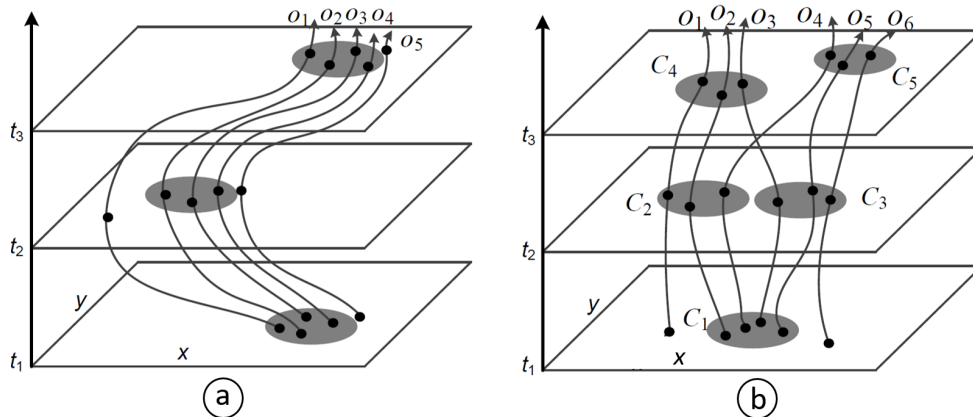
Figura 12 – (a) Análise de densidade de tráfego utilizando dados de táxis da cidade de Nova York. (b) Análise das rotas mais percorridas (em vermelho) por táxis da cidade de Pequim, China.



Fonte: a: (DORAISWAMY *et al.*, 2018), b: (YUAN *et al.*, 2010)

Os estudos sobre clusterização de objetos em movimento visam encontrar grupos de objetos que se movem juntos durante um certo período de tempo (LI *et al.*, 2004). Muitas abordagens foram propostas nos últimos anos para investigar esse problema. Yuan *et al.* (2017) revisam e resumem o desenvolvimento e as tendências da pesquisa sobre clusterização de objetos em movimento. De acordo com esse trabalho, as técnicas mais conhecidas são: *flock*, *convoy*, *swarm*, *traveling companions*, e *gathering*. *Flock* é um grupo de objetos que viajam juntos dentro de um círculo com raio pré-definido (LEVIS; LIEBCHEN, 2018; DORFMAN *et al.*, 2015; GUDMUNDSSON *et al.*, 2004). O problema dessa abordagem é que, nem sempre, um círculo consegue representar corretamente um grupo, podendo haver perdas de elementos. Para resolver esse problema, a técnica *convoy* emprega agrupamento baseado em densidade, e não na restrição de forma do grupo em movimento (BOUKHERS *et al.*, 2016; ORAKZAI *et al.*, 2016; JEUNG *et al.*, 2008). Enquanto *flock* e *convoy* exigem que os objetos estejam juntos durante um determinado intervalo de tempo, *swarm* descobre objetos que permaneceram juntos durante um período de tempo que não necessariamente é contínuo (ZHENG *et al.*, 2014; LI *et al.*, 2010a). A abordagem *traveling companions* é uma evolução das técnicas anteriores, em que os grupos de objetos em movimento são encontrados de forma incremental e *online* (PUNTHEERANURAK *et al.*, 2018; NASERIAN *et al.*, 2018; TANG *et al.*, 2012). Por último, o método *gathering* identifica um evento ou um incidente que envolve a congregação de indivíduos. Aplicações que

Figura 13 – Exemplos de padrões em grupos de objetos que se movem juntos durante um certo período. (a) *Flock, convoy e swarm*. (b) *Gathering*.



Fonte: (ZHENG, 2015)

utilizam esse método podem identificar comemorações, desfiles, protestos, engarrafamentos e outros eventos públicos (HACHEM *et al.*, 2017; ZHENG *et al.*, 2014; ZHENG *et al.*, 2013). Assim, essa técnica identifica situações não triviais em mobilidade urbana.

A Figura 13 ilustra os padrões encontrados pela clusterização de objetos em movimento. A Figura 13a mostra um grupo de objetos (o_2, o_3, o_4) como um *flock* do instante t_1 ao t_3 . Embora o_5 acompanhe o grupo, não pode ser incluído devido ao tamanho fixo do círculo empregado. Por outro lado, um *convoy* pode incluir o_5 , uma vez que o *convoy* (o_2, o_3, o_4, o_5) é um grupo encontrado por densidade. Os cinco objetos também formam um *swarm* durante o período de tempo não contínuo entre t_1 e t_3 . A Figura 13b mostra um exemplo de *gathering* (C_1, C_2, C_4). Já (C_1, C_3, C_5) não é um *gathering*, pois C_5 está muito distante de C_2 e C_3 (ZHENG, 2015).

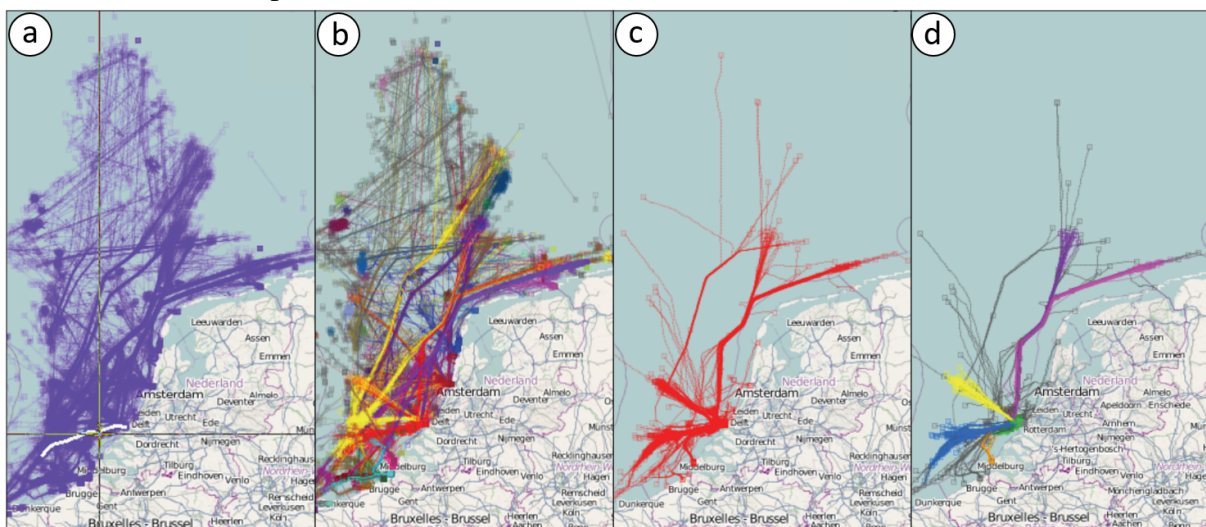
3.3 Clusterização de Trajetórias

Os trabalhos de clusterização de trajetórias geralmente almejam encontrar caminhos representativos ou características compartilhadas por trajetórias de objetos em movimento. A clusterização de trajetórias é um tema de pesquisa que tem sido bastante estudado nos últimos dez anos (BIAN *et al.*, 2018; ZHENG, 2015). Assim, diversas abordagens foram propostas para descobrir *clusters* através da similaridade entre trajetórias ou subtrajetórias (ANDRIENKO *et al.*, 2018; YAO *et al.*, 2017). Porém, como os dados de trajetória são multidimensionais – espaço, tempo e atributos –, as técnicas tradicionais de clusterização são ineficientes para analisá-los. Além disso, as trajetórias podem apresentar propriedades bem diferentes entre si – como comprimento, forma, taxa de amostragem, número de pontos –, tornando ainda mais

difícil encontrar similaridades. Portanto, a comparação entre trajetórias deve seguir estratégias específicas, considerando o propósito da clusterização. Desta forma, a medida de similaridade entre trajetórias é um dos fatores mais importantes de uma técnica de clusterização de trajetórias. Essa medida geralmente é calculada por uma função de similaridade que pode considerar a distância entre: um ponto e uma trajetória (CHEN *et al.*, 2010; ZHENG; ZHOU, 2011b), duas trajetórias completas (ZHANG; ZHANG, 2018; CHOONG *et al.*, 2017), ou segmentos de trajetórias (DEVOGELE *et al.*, 2017; LEE *et al.*, 2007).

A Figura 14 ilustra uma análise que faz uso da clusterização de trajetórias de navios. Para obtenção dessas trajetórias, os navios foram rastreados por radar pela guarda costeira dos Países Baixos na área do Mar do Norte (ANDRIENKO; ANDRIENKO, 2013). Na Figura 14a, as trajetórias inteiras dos navios são representadas por linhas sobre o mapa da região. Em seguida, na Figura 14b, um algoritmo de clusterização baseado em densidade foi aplicado às trajetórias dos navios. Neste caso, cada um dos *clusters* resultantes consiste em um grupo de trajetórias com pontos finais espacialmente próximos. Desta forma, é possível descobrir os principais destinos desses navios. Depois, um dos *clusters* com trajetórias que terminam em Roterdã, na Holanda, é selecionado na Figura 14c. Por último, na Figura 14d, uma clusterização por similaridade de rota é aplicada ao *cluster* selecionado anteriormente, ou seja, encontrando *clusters* dentro de um *cluster*. Assim, é possível conhecer as rotas típicas dos navios que vão para Roterdã.

Figura 14 – Exemplo de análise utilizando clusterização de trajetórias de navios na área do Mar do Norte, Países Baixos. (a) As trajetórias inteiras de navios são representadas por linhas no mapa. (b) As trajetórias são agrupadas (*clusters*) de acordo com os destinos. (c) Um dos *clusters* é selecionado. (d) A clusterização por similaridade de rota foi aplicado ao *cluster* selecionado.



Fonte: (ANDRIENKO; ANDRIENKO, 2013)

Devido à complexidade computacional de comparar trajetórias ou partes delas, a maioria dos métodos de clusterização de trajetórias é desenvolvida para ser executada de forma *offline*, isto é, realizando a análise de dados em todas as trajetórias previamente coletadas (ZHENG, 2015). Alguns trabalhos propõem métodos para serem executados *online*, ou seja, os *clusters* são incrementalmente encontrados durante o registro das trajetórias. Li e seus coautores (2010b), por exemplo, propuseram um algoritmo de clusterização incremental com o objetivo de reduzir o custo computacional e o armazenamento de trajetórias recebidas de forma *online*. Por sua vez, Mao et al. (2016) propõem uma clusterização de trajetórias *online* utilizando janelas deslizantes no tempo. Geralmente, os métodos de clusterização processam trajetórias dentro de um período de tempo – também conhecido como janela de tempo ou *time window* –, e os resultados são apresentados ao final desse período. Mesmo nos métodos *online*, o resultado da clusterização não é apresentado em tempo real, mas apenas no final de cada janela de tempo.

Os métodos discutidos nesta seção são voltados para o movimento de objetos em espaço livre, ou seja, sem restrição de rede. Embora existam algumas publicações discutindo a clusterização de trajetórias em uma rede viária, esse problema pode ser resolvido pela combinação de algoritmos de *map-matching* e algoritmos de clusterização em grafos (ZHENG; ZHOU, 2011b). Assim, primeiro, usam-se algoritmos de *map-matching* para projetar trajetórias em uma rede viária e, em seguida, emprega-se clusterização em grafos para encontrar subgrafos comuns entre as trajetórias na rede viária. Nessa linha, Kharrat *et al.* (2008) apresentam o NETSCAN, uma técnica para encontrar similaridades entre trajetórias restritas a uma rede viária.

3.4 Descoberta de Rotas Quentes

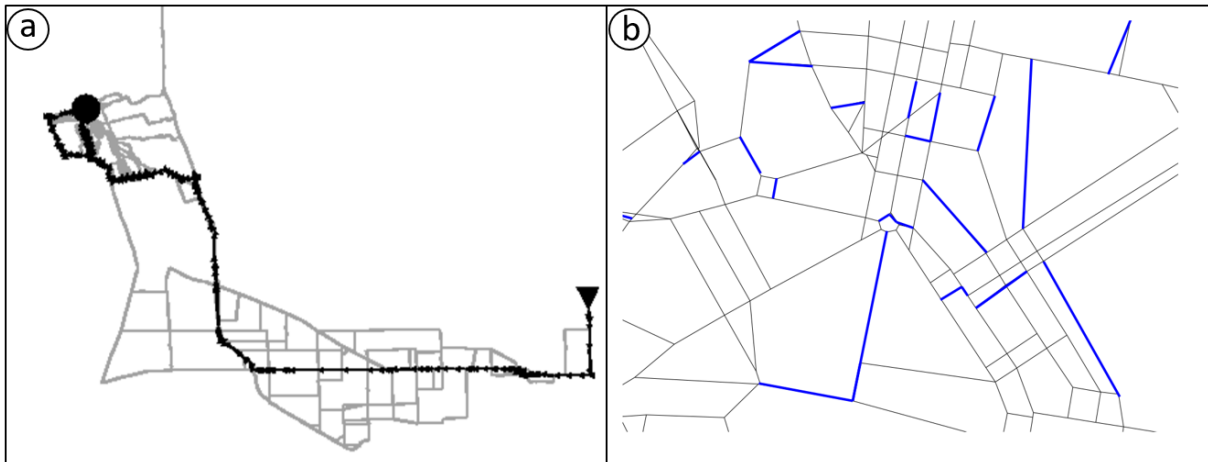
Dentro do escopo de análise de trajetórias de veículos, no qual está inserido este trabalho, o problema de descobrir rotas com tráfego pesado – igualmente conhecidas como rotas quentes ou caminhos quentes – vem sendo investigado em trabalhos recentes (ALMEIDA *et al.*, 2017; GUI *et al.*, 2016; WU *et al.*, 2014; SACHARIDIS *et al.*, 2008; LI *et al.*, 2007). Monitorar essas rotas ajuda as agências de trânsito a identificar eventos atípicos para tomada de decisão no direcionamento de tráfego. Esse é um problema bastante desafiador devido à natureza complexa dos dados. Se os veículos viajassem em *clusters* organizados, poderia ser simples usar um algoritmo de clusterização para encontrar as rotas quentes. No entanto, no mundo real, os veículos se movem de forma imprevisível. Variações na velocidade, no tempo, na rota e em outros fatores fazem com que eles viajem em grupos bastante irregulares e dispersos. Por

exemplo, dois veículos viajando da mesma origem para o mesmo destino e utilizando a mesma rota terão trajetórias com diferenças na duração e na velocidade média. Condições de tráfego aleatórias, como um acidente de trânsito ou semáforos dessincronizados, podem causar ainda mais variações.

A clusterização de objetos em movimento permite encontrar uma rota quente através das trajetórias de veículos que se movem juntos (LI *et al.*, 2007). A clusterização de trajetórias também permite descobrir rotas quentes, através da similaridade entre as trajetórias ou subtrajetórias. No entanto, devido à complexidade dos dados, as duas abordagens não funcionam bem para resolver o problema, pois a maioria desses métodos adota clusterização por densidade baseada na distância euclidiana entre pontos de trajetórias. Como os dispositivos GPS têm frequências de amostragem diferentes, bem como o problema de perda de sinal, os conjuntos de dados de trajetória geralmente têm densidades diferentes, o que complica o uso da maioria dos métodos de clusterização por densidade baseada em distância euclidiana (GUI *et al.*, 2016). Além disso, para encontrar uma rota quente, não é preciso identificar e nem acompanhar os *clusters* e seus veículos integrantes. Uma rota pode ser considerada quente mesmo se nenhum *cluster* percorrê-la. Isso acontece sempre que um número suficiente de veículos percorre a rota dentro de um determinado intervalo de tempo, não importando se eles viajam de modo síncrono ou não por essa rota. Além disso, essas abordagens de clusterização geralmente encontram muitos *clusters* pequenos e perdem o fluxo geral do movimento dos veículos (LI *et al.*, 2007). Portanto, não é trivial usar clusterização de dados de movimento para encontrar rotas quentes.

A maioria das abordagens encontradas na literatura foi proposta para trabalhar com dados históricos, isto é, aplicações que utilizam base de dados de trajetórias previamente coletadas. Li e seus coautores (LI *et al.*, 2007) propuseram um algoritmo baseado em densidade chamado FlowScan, que encontra rotas quentes medindo o volume de tráfego em uma rede viária restrita. A Figura 15a mostra rotas quentes encontradas por esse algoritmo na cidade de San Joaquin, EUA. Um algoritmo *online* também foi desenvolvido por Sacharidis *et al.* (2008) para encontrar e manter os caminhos quentes enquanto os dados de trajetórias são recebidos. A Figura 15b ilustra, em um dado período, as 20 rotas mais quentes encontradas por esse algoritmo na cidade de Atenas, Grécia. Almeida *et al.* (2017) propõem um método automatizado para incentivar a formação de comboios de veículos nas principais rodovias de uma cidade através da integração de semáforos. Para isso, um algoritmo é usado para simular o movimento de ondas verdes de tráfego nas rotas mais quentes da cidade. Gui e seus colaboradores (2016) apresentam

Figura 15 – (a) Exemplo de rotas quentes encontradas na cidade de San Joaquin, EUA. (b) Outro exemplo com as 20 rotas mais quentes na cidade de Atenas, Grécia.



Fonte: a: (LI *et al.*, 2007), b: (SACHARIDIS *et al.*, 2008)

um método para encontrar rotas quentes a partir de segmentos de trajetória com velocidade baixa. Wu *et al.* (2014) propõem uma nova abordagem de inferência de rotas quentes através de um algoritmo que extrai a rede viária a partir de um conjunto de trajetórias, usando métodos de processamento de imagem. No entanto, nessas abordagens, não é possível visualizar as rotas quentes em tempo real, pois o resultado é exibido apenas no final de uma janela de tempo. Além disso, como o movimento é agregado empregando discretização de janela de tempo, esses métodos podem apresentar inconsistências de tempo e espaço com o fluxo de tráfego real, como mencionado nos cenários mostrados na Introdução, Capítulo 1. Por fim, o trabalho de Gomes *et al.* (2018), que é uma aplicação direta da abordagem proposta nesta tese, apresenta um método para monitorar como as rotas quentes evoluem em tempo real de uma forma incremental.

3.5 Visualização de Dados de Trajetórias

Muitas vezes, a visualização é fundamental para entender os resultados de processos automáticos de abstração de dados, como a clusterização. Embora certas abordagens de clusterização gerem representações visuais implicitamente – por exemplo, dendrogramas ou mapas –, para muitas outras técnicas, representações visuais precisam ser construídas em uma etapa de pós-processamento para a compreensão dos resultados. Além disso, em muitos casos, é necessário que os usuários incorporem seus conhecimentos sobre o domínio e sua inteligência humana no processo de análise exploratória. Nesses casos, a visualização atende essa demanda, possibilitando a interação do analista com os resultados, alterando parâmetros e visualizando suas alterações instantaneamente. O analista emprega sua capacidade de pensamento criativo e

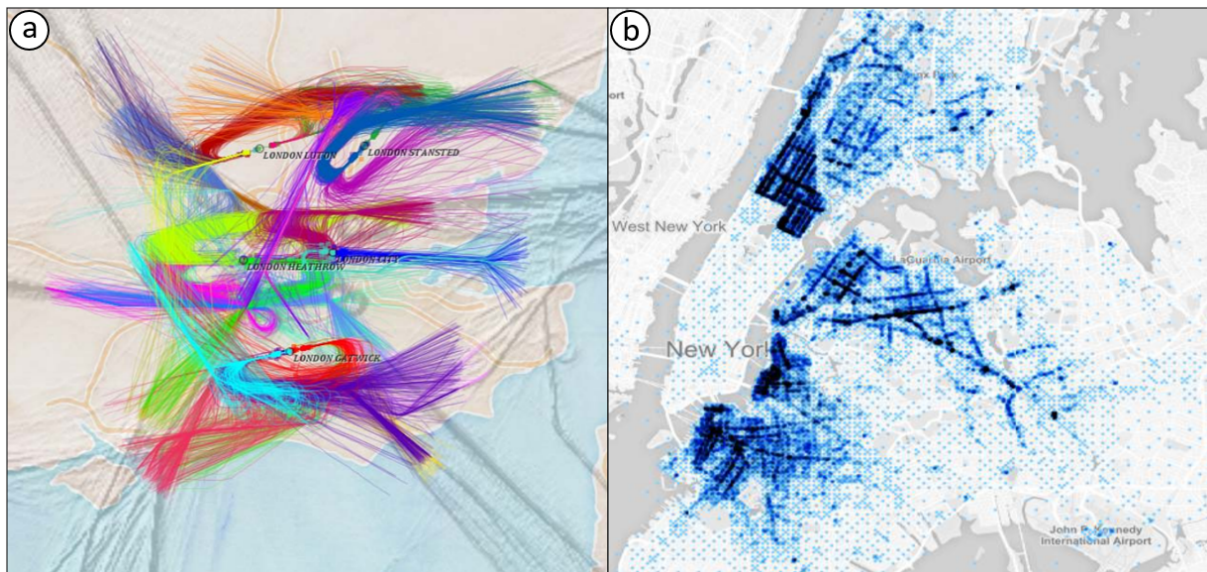
crítico, fazendo associações e gerando ideias, enquanto os computadores processam, agregam e extraem conjuntos de dados que seriam grandes demais para um ser humano lidar com eficiência sozinho (ANDRIENKO *et al.*, 2017). Assim, a visualização promove avanços no processo de análise, e extrai, de forma eficiente, conhecimento potencialmente útil. As ferramentas de visualização são particularmente importantes para analisar fenômenos e processos que estão acontecendo no espaço geográfico, pois a heterogeneidade do espaço e a variedade de propriedades envolvidas na análise não podem ser adequadamente representadas por um processamento totalmente automático (ANDRIENKO; ANDRIENKO, 2012).

A comunidade de visualização vem ajudando a análise de dados de trajetórias fornecendo meios eficazes de integrar seres humanos no processo de exploração de dados (CHEN *et al.*, 2015; COOK; THOMAS, 2005). Este trabalho segue a ideia de que a análise visual integra a análise automática de dados com a experiência humana, fornecendo, para esse fim, meios visuais interativos (WANG *et al.*, 2016; KEIM *et al.*, 2008). Recentemente, Cai *et al.* (2018) apresentaram as aplicações e as pesquisas em visualização de dados de trajetória espaciais. Uma revisão da literatura sobre aplicações de dados de trajetória na perspectiva de uma agência de transporte rodoviário foi apresentada por Markovic e seus coautores (2018). O estado da arte e as novas direções na pesquisa em análise visual de movimento e transporte foram sumarizados em (ANDRIENKO *et al.*, 2017). Além disso, Andrienko e seus coautores apresentaram também uma revisão de sistemas e técnicas para análise visual de dados de movimento (ANDRIENKO; ANDRIENKO, 2012) e uma lista de desafios na análise visual geoespacial (ANDRIENKO *et al.*, 2007a). Nesta seção, os trabalhos relacionados são apresentados considerando dois aspectos: padrões de mobilidade global e padrões de mobilidade local.

3.5.1 Padrões de Mobilidade Global

Várias abordagens foram propostas para analisar visualmente os aspectos globais dos dados de trajetórias (LORENZO *et al.*, 2016; SCHEEPENS *et al.*, 2016; HUANG *et al.*, 2016). Andrienko *et al.* (2018) propuseram um conjunto de técnicas e diretrizes de visualização para auxiliar a análise de dados de tráfego aéreo por meio de clusterização de trajetórias para identificar as rotas mais importantes (Figura 16a). Pahins *et al.* (2017) propuseram a Hashedcubes, uma técnica que permite a exploração visual, em tempo real, de grandes volumes de dados. A Figura 16b mostra o uso da Hashedcubes para analisar dados de trajetórias de táxis na cidade de Nova York.

Figura 16 – (a) Representação das principais rotas de aproximação dos aeroportos de Londres através de clusterização de trajetórias. (b) Análise visual de trajetórias de táxis da cidade de Nova York.

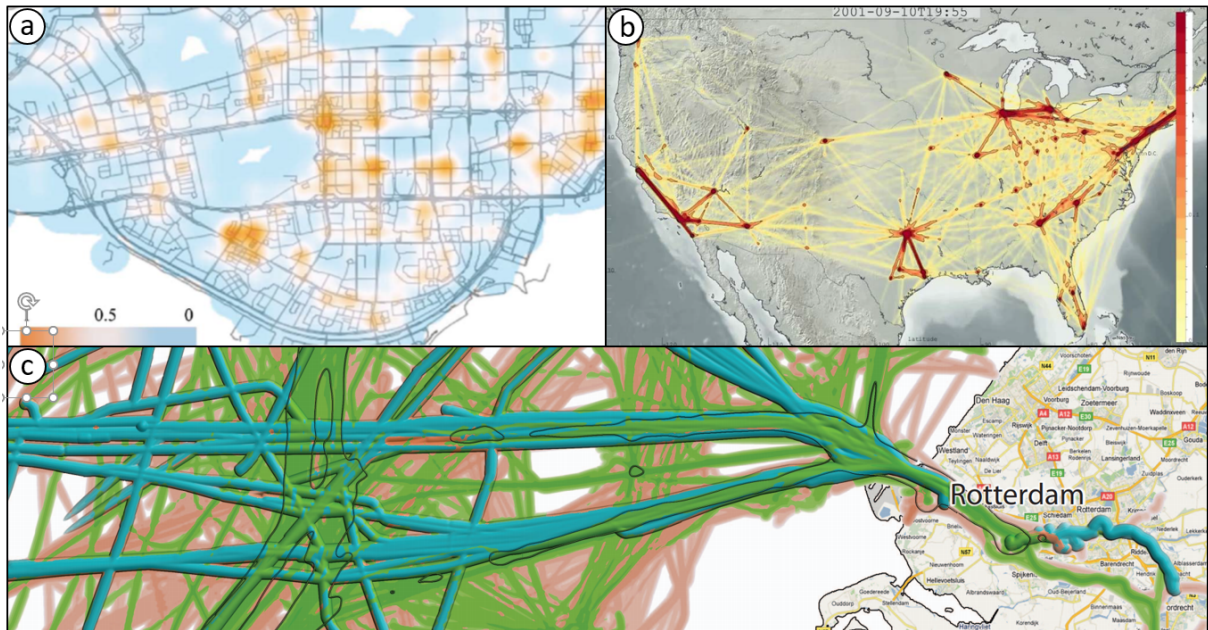


Fonte: a: (ANDRIENKO *et al.*, 2018), b: (PAHINS *et al.*, 2017)

Liu e seus coautores (2013) usam um mapa de calor para mostrar áreas com alto volume de tráfego. Como mostrado na Figura 17a, as regiões laranjas representam um grande volume de tráfego, enquanto as regiões azuis indicam baixo volume de tráfego. A técnica de Kernel Density Estimation (KDE) é comumente usada para gerar mapas de calor. Com essa técnica, é possível visualizar mapas de densidade de trajetórias (SCHEEPENS *et al.*, 2016; SCHEEPENS *et al.*, 2011a; LAMPE; HAUSER, 2011). Uma modificação do KDE, o Network KDE (NKDE), permite caracterizar eventos ao longo de redes viárias (XIE; YAN, 2008; BORRUSO, 2008). A Figura 17b mostra as rotas com maior volume de tráfego aéreo dos EUA encontradas por KDE, codificando a cor para indicar a densidade das trajetórias: regiões mais escuras significam maior densidade de tráfego (LAMPE; HAUSER, 2011). Outro exemplo é ilustrado na Figura 17c, um mapa de densidade de tráfego de embarcações em torno da cidade de Roterdã (SCHEEPENS *et al.*, 2011a).

Algumas abordagens representam o movimento entre regiões como um grafo, no qual as regiões são os nós do grafo e o fluxo entre um par de nós é representado por uma aresta direcionada e ponderada (WANG *et al.*, 2013). O movimento entre as regiões é simplificado e visualizado usando um mapa de fluxo. Landesberger *et al.* (2016) desenvolveram uma técnica baseada em grafos, chamada MobilityGraphs, que revela novos padrões de movimento não identificados em mapas de fluxo (Figura 18a). Zeng *et al.* (2013) apresentaram uma técnica para visualizar padrões de intercâmbio entre diferentes regiões de uma cidade. Uma metáfora visual

Figura 17 – (a) Visualização de um mapa de calor com as regiões de maior volume de tráfego, as regiões laranjas. (b) Mapa de densidade de trajetórias do tráfego aéreo dos EUA. (c) Mapa de densidade de tráfego de embarcações em torno de Roterdã.

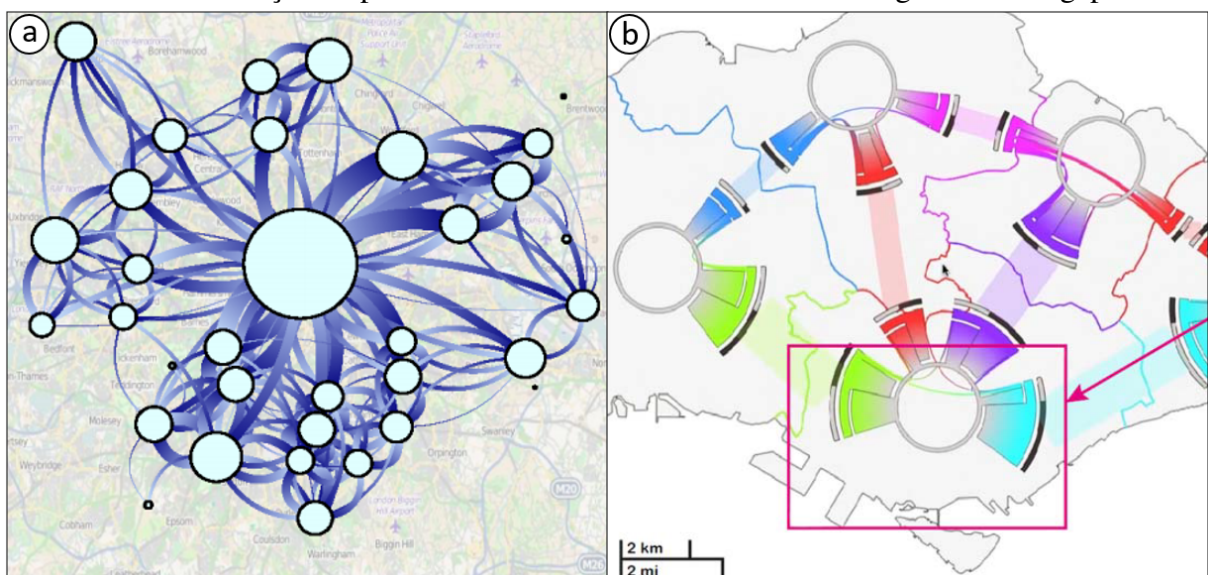


Fonte: a: (LIU *et al.*, 2013), b: (LAMPE; HAUSER, 2011), c: (SCHEEPENS *et al.*, 2011a)

radial foi projetada para representar os intercâmbios entre as regiões (Figura 18b).

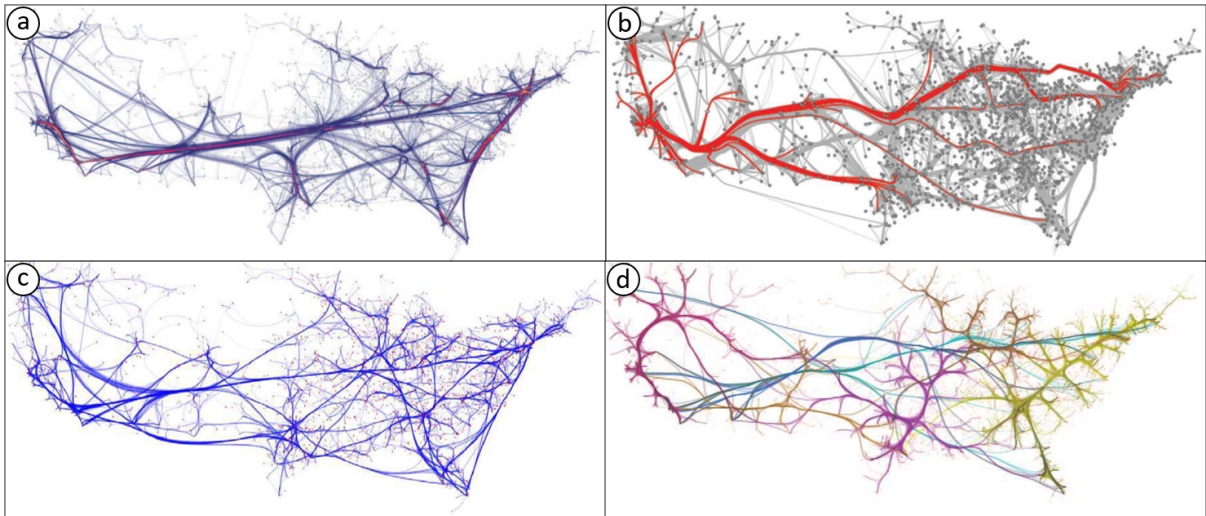
Quando o volume de dados se torna muito grande, geralmente as visualizações sofrem de poluição visual (*visual clutter*), pois uma grande quantidade de elementos gráficos, como linhas, podem facilmente sobrecarregar a visualização e prejudicar a identificação de padrões. Muitas técnicas de *edge-bundling* foram propostas para reduzir a poluição visual nas

Figura 18 – (a) Análise da mobilidade de pessoas em Londres utilizando dados do Twitter. (b) visualização de padrões de intercâmbio entre diferentes regiões em Singapura.



Fonte: a: (LANDESBERGER *et al.*, 2016), b: (ZENG *et al.*, 2013)

Figura 19 – Visualizações de trajetórias de migração nos EUA gerados pelos métodos apresentados por: (a) Holten e Van Wijk (2009); (b) Cui *et al.* (2008); (c) Lambert *et al.* (2010); (d) Ersoy *et al.* (2011).



Fonte: a: (HOLTEN; WIJK, 2009), b: (CUI *et al.*, 2008), c: (LAMBERT *et al.*, 2010), d: (ERSOY *et al.*, 2011)

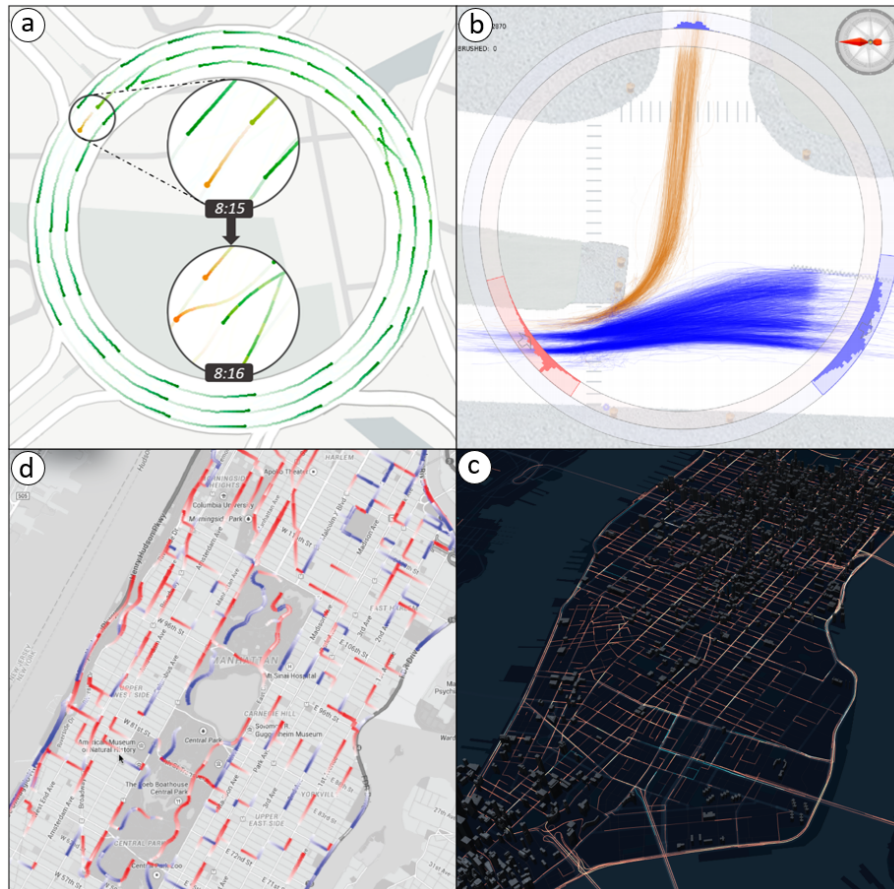
visualizações (ERSOY *et al.*, 2011; LAMBERT *et al.*, 2010; HOLTEN; WIJK, 2009; CUI *et al.*, 2008). Uma revisão sobre os trabalhos de redução de poluição visual pode ser encontrada no artigo apresentado por Ellis e Dix (2007). A Figura 19 apresenta a aplicação de quatro abordagens diferentes para resolver o mesmo problema: representar a migração nos EUA.

As abordagens apresentadas nesta seção geralmente agregam o movimento no espaço ou no tempo para facilitar a análise visual. Assim, a percepção das mudanças das posições espaciais dos veículos é perdida, ou seja, a própria essência do movimento se perde (ANDRIENKO *et al.*, 2007b).

3.5.2 Padrões de Mobilidade Local

Diferentes trabalhos analisam os padrões de mobilidade local. O trabalho de Gomes *et al.* (2017), que é a primeira aplicação da abordagem proposta nesta tese, propõe uma técnica para visualizar padrões de oscilação de tráfego sem utilizar discretização de tempo em janelas (Figura 20a). Guo *et al.* (2011), por exemplo, propuseram um método de visualização para investigar padrões e comportamentos anormais em tráfego local (Figura 20b). Apesar dessa abordagem permitir a visualização individual de carros em movimento através de animação, trata-se de uma aplicação para analisar um cruzamento de avenidas específico. Portanto, a técnica proposta por Guo *et al.* não pode ser utilizada para monitorar o tráfego de uma cidade inteira. Poco e coautores (POCO *et al.*, 2015) propuseram um método para visualização da dinâmica de

Figura 20 – (a) Padrões de oscilação de tráfego. (b) Padrões e comportamentos anormais no tráfego de um cruzamento. (c) Dinâmica de tráfego baseada em campo vetorial. (d) Visualização de um grande volume de dados de trajetórias de táxis.



Fonte: a: (GOMES *et al.*, 2017), b: (GUO *et al.*, 2011), c: (POCO *et al.*, 2015) e d: (UBER,)

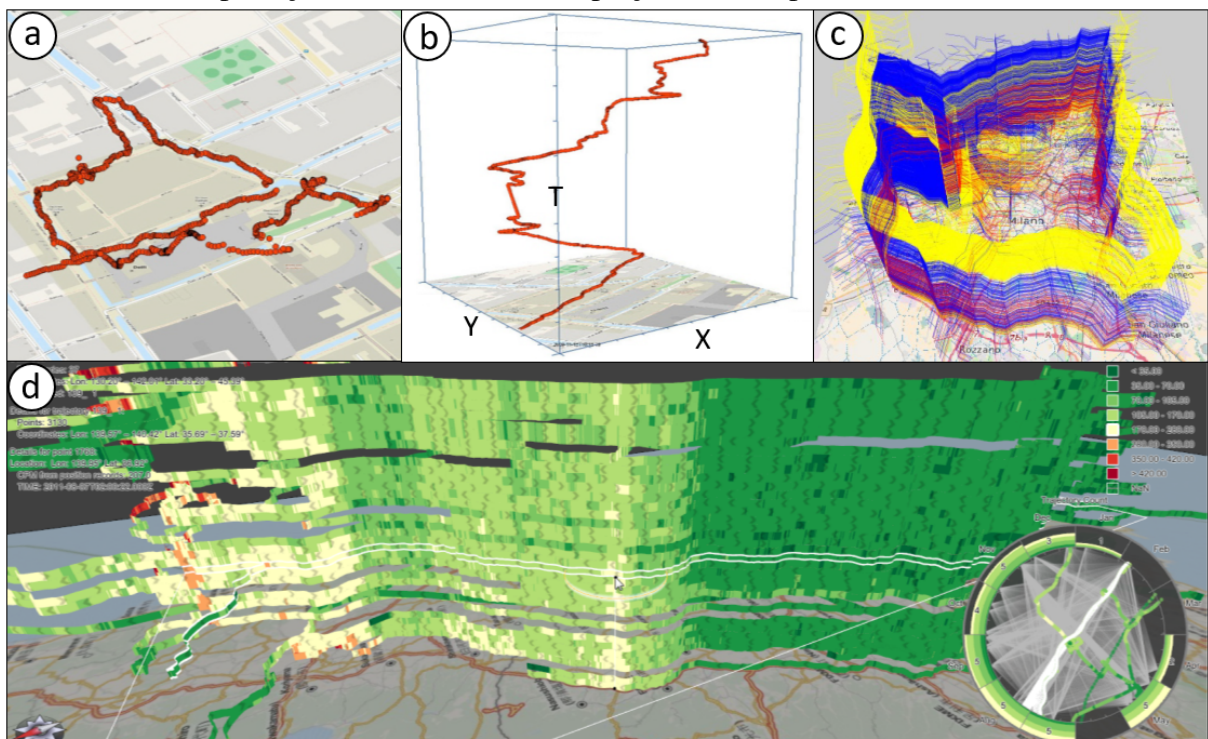
tráfego baseada no movimento de múltiplas partículas em um campo vetorial (Figura 20c). Eles apresentaram uma função para adaptar um campo vetorial calculado com dados de trajetórias de táxis sobre a rede viária da cidade de Nova York. Essa adaptação é uma tarefa cara e demorada. Assim, essa abordagem não é adequada para visualizar dinâmica de tráfego em tempo real. Recentemente, a Uber Engineering lançou uma API para análise visual exploratória de grandes bases de dados (UBER,). Apesar de uma demonstração de desempenho disponibilizada usando um grande volume de dados de trajetória de táxi (Figura 20d), essa API não pretende ser uma abordagem sistemática para análise visual de tráfego. No entanto, ela pode ser usada como uma ferramenta poderosa no processo de renderização de visualizações.

Como trajetórias são compostas por atributos espaciais e temporais, vários trabalhos representam as trajetórias em um sistema de coordenadas 3D, no qual o plano XY é usado para mapear informações geográficas espaciais e o eixo Z representa a dimensão do tempo. Essa técnica é conhecida como Space Time Cube (STC) e é muito utilizada para a análise visual

de dados de movimento (YANG *et al.*, 2017; MATASCI *et al.*, 2017; AMINI *et al.*, 2015; TOMINSKI *et al.*, 2012). Por exemplo, a Figura 21a representa no mapa todos os pontos de uma trajetória sem considerar o tempo. Na Figura 21b, a mesma trajetória é representada usando um STC. Desse modo, os pontos das trajetórias são conectados por segmentos de linha em ordem cronológica. No entanto, a técnica STC geralmente sofre de poluição visual (Figura 21c) além dos problemas de visualização 3D, como a oclusão (WARE, 2012) e a percepção inconsistente através de rotações (HANCOCK *et al.*, 2009). Por causa dessas limitações, o STC tem melhores resultados para analisar padrões de mobilidade local. Para minimizar esses problemas, um STC necessita de meios de interação com os dados. Ferramentas de navegação como *pan*, rotação e zoom para ajustar o ponto de vista são fundamentais para melhorar a percepção dos padrões espaço-temporais. Além disso, a técnica STC pode ser aprimorada para representar outros atributos que não sejam espaço-temporais. Por exemplo, Tominski *et al.* (2012) apresentaram uma técnica para visualizar várias trajetórias distribuídas ao longo do eixo Z como faixas empilhadas para analisar a evolução da velocidade dos objetos em movimento (Figura 21d).

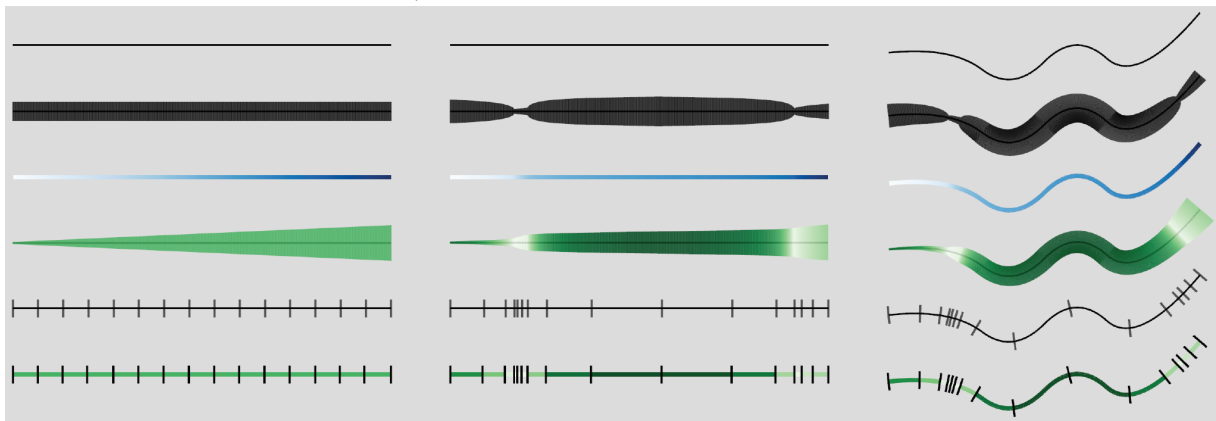
Para representação 2D, um trabalho bastante relevante para o contexto desta tese foi proposto por Perin *et al.* (2018) para avaliar a percepção do tempo e da velocidade em

Figura 21 – (a) Os pontos de uma trajetória representados em um mapa sem considerar o tempo. (b) A mesma trajetória é representada usando um STC. (c) Exemplo de um STC com poluição visual. (d) Uma adaptação do STC para visualizar velocidade.



Fonte: a e b: (KVELADZE *et al.*, 2012), c: (ANDRIENKO *et al.*, 2017) e d: (TOMINSKI *et al.*, 2012)

Figura 22 – Diferentes codificações de velocidade e tempo para representações 2D curvas e retas de trajetórias. Tanto velocidade constante como velocidade variada (seções lentas e com alta velocidade) são mostradas.



Fonte: (PERIN *et al.*, 2018)

uma representação 2D de uma trajetória. Esse trabalho avaliou, empiricamente, a percepção das pessoas sobre a variação do tempo e a codificação da velocidade em representações 2D de trajetórias. No estudo de percepção gráfica realizado nesse trabalho, nove codificações da literatura para caminhos retos e curvos foram avaliadas (Figura 22). No entanto, não há diretrizes na literatura sobre como representar o tempo e a velocidade nesses caminhos. Os resultados do estudo fornecem orientações claras sobre quais codificações usar e quais evitar. Resumindo, foi sugerido usar cor para codificar a velocidade e comprimento para codificar o tempo.

Por último, reduzindo o escopo para o monitoramento de tráfego, alguns trabalhos concentram-se na investigação de incidentes de tráfego. O Incident Cluster Explorer (ICE) é uma aplicação que investiga e visualiza incidentes rodoviários (PACK *et al.*, 2009). O T-Watcher é um sistema interativo de análise visual para monitorar e analisar situações complexas de tráfego em grandes cidades (PU *et al.*, 2013). O sistema Traffic Origins foi projetado para visualizar o impacto que os incidentes rodoviários têm em gerar congestionamentos (ANWAR *et al.*, 2014). Wang *et al.* (2013) apresentaram um sistema de análise visual para visualização de engarrafamentos, que são detectados automaticamente pelo sistema. No entanto, nenhuma dessas propostas foi projetada para ser utilizada no monitoramento de tráfego em tempo real.

3.6 Considerações Finais

Embora as técnicas de clusterização de dados sejam importantes para extrair conhecimento potencialmente útil, agregar trajetórias em uma janela de tempo prejudica a análise de toda a dinâmica do movimento dos objetos, o que é vital para os sistemas de monitoramento de

tráfego. Uma característica importante do tráfego de veículos no mundo real é a complexidade do movimento. Objetos viajam com velocidades e tempos diferentes, mesmo quando estão na mesma rota. Consequentemente, analisar o movimento dos objetos de forma contínua e interativa fornece uma compreensão muito melhor dos padrões de mobilidade encontrados em dados de trajetória. Neste sentido, a análise visual fornece subsídios para atingir esse objetivo. Assim, a principal contribuição deste trabalho em relação aos apresentados é apresentar uma nova abordagem para analisar de forma visual, contínua e interativa a dinâmica de tráfego, encontrando padrões de mobilidade local e global.

4 ABORDAGEM PROPOSTA

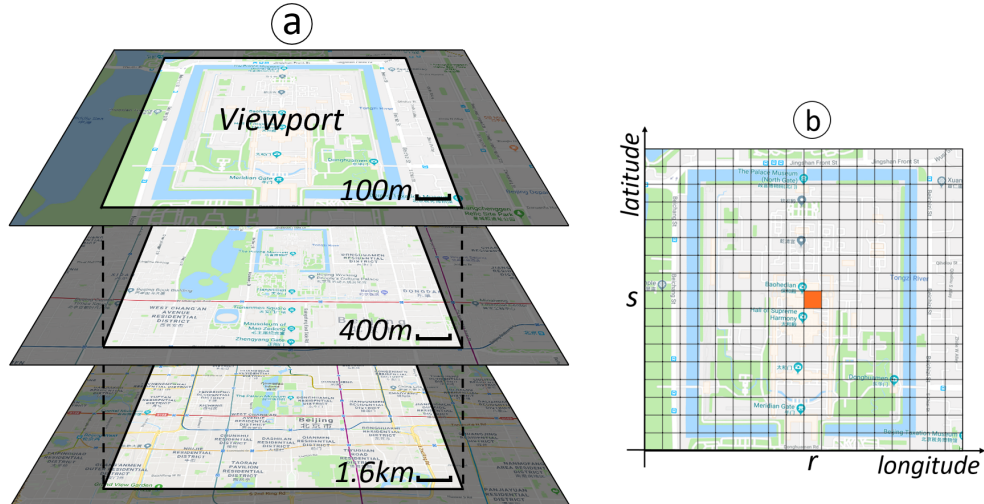
O objetivo principal deste trabalho é apresentar uma nova abordagem para análise de dinâmica de tráfego visualizando o movimento dos objetos no espaço e no tempo. Com esse objetivo, inicialmente é apresentada a definição do problema na Seção 4.1. Em seguida, a arquitetura com seus principais componentes é descrita na Seção 4.2. Na Seção 4.3, uma metodologia sistemática é definida, considerando desde o pré-processamento das trajetórias até a visualização dos dados. Por último, as considerações finais são apresentadas na Seção 4.4. É importante mencionar que neste capítulo são apresentados os aspectos técnicos da abordagem proposta. Não é descrito ainda como a abordagem deve ser aplicada para a análise de dinâmica de tráfego, que é discutida no próximo capítulo.

4.1 Definição do Problema

O problema que este trabalho propõe-se a resolver tem como entrada um fluxo de dados em tempo real (do inglês, *data stream*). A solução proposta visa descobrir padrões de mobilidade em uma área de interesse delimitada por uma *viewport* em um mapa (Figura 23a) e em um instante t_k que representa o tempo atual. Nesta solução, são considerados apenas os objetos que cruzam essa *viewport* no período de análise de $t_k - \Delta t$ até t_k , com um Δt suficientemente pequeno para abranger apenas as informações mais recentes do fluxo de dados. O conceito de período de análise definido neste trabalho é diferente da definição tradicional de janela de tempo usada nas abordagens apresentadas anteriormente, pois o período de análise é aplicado individualmente a cada objeto em movimento, enquanto a janela de tempo aplica o mesmo intervalo para todos os objetos.

Além disso, como a *viewport* é definida no espaço de tela, é importante mencionar que se houver mudanças na escala do mapa que está sendo exibido, o nível de zoom desse mapa ou foi aumentado (*zoom-in*, quando a escala aumenta - menor denominador) ou foi diminuído (*zoom-out*, quando a escala diminui - maior denominador). Assim, na Figura 23a, percorrer os mapas de cima para baixo (escalas 1:100, 1:400, 1:1600) corresponde a um *zoom-out*. Na Computação Gráfica, esse é o conceito comum de mapeamento de *Window* (em coordenadas de mundo) para *Viewport* (em coordenadas de tela). Neste trabalho, um mapa pode ser visualizado em diferentes escalas. Com esse objetivo, um mapa em uma determinada escala é associado a uma camada individual, também conhecida como nível de zoom. Todas as camadas são georreferenciadas,

Figura 23 – (a) Uma área de interesse delimitada por uma *viewport* em três camadas de mapa com diferentes escalas. (b) O espaço da *viewport* foi discretizado como uma grade com 17 linhas e 18 colunas. Para uma célula da grade (r, s) no instante atual t_k , a soma dos pesos de contribuição $w_{ik}(r, s)$ dos objetos que cruzaram essa célula é calculada pela Eq. (4.2).



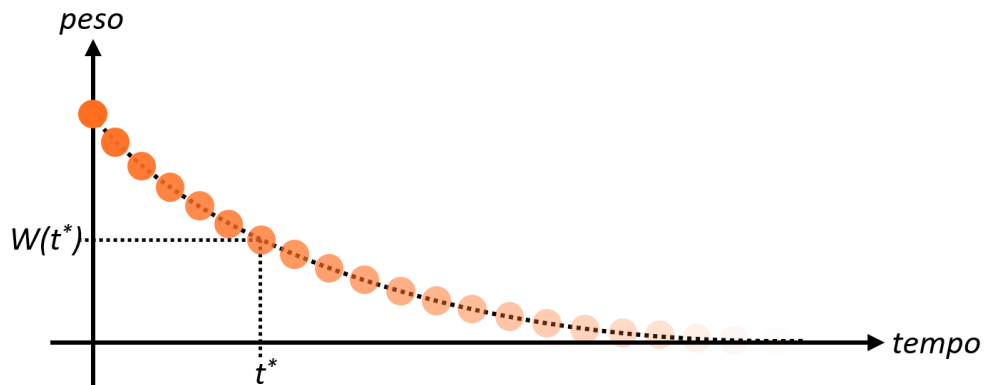
Fonte: o autor.

isto é, um determinado par [longitude, latitude] representa a mesma localização geográfica, independentemente do nível de zoom (Figura 23a).

Para descobrir padrões de mobilidade, esta abordagem utiliza o conceito de temperatura ao longo da trajetória de um determinado objeto em movimento, semelhante ao que acontece no caminho de um cometa. O objeto sempre tem a temperatura máxima em sua localização atual e deixa para trás um rastro de temperaturas que gradativamente vão diminuindo ao longo de seu caminho. Se um objeto em movimento cruzar o caminho de outro objeto em movimento, a temperatura nesse local será a contribuição resultante das temperaturas dos dois rastros que se cruzaram. Na Figura 24, a contribuição de temperatura no instante atual t_k em um determinado local ao longo do caminho de um objeto é a temperatura máxima desse objeto multiplicada por um fator de perda de temperatura (peso) que é em função do tempo decorrido, isto é, da diferença de tempo entre o instante atual e o instante em que o objeto passou por esse local.

O modelo proposto utiliza uma abordagem baseada em grade na qual a temperatura é calculada em cada pixel na *viewport*. Através do mapeamento *window-viewport*, a matriz de *pixels* na *viewport* é relacionada a uma grade de células sobre o mapa. Uma representação aproximada dessa grade é exibida na Figura 23b – cada célula da grade corresponde a um *pixel* na *viewport*. Assim, mudar o nível de zoom altera os tamanhos das células e a região coberta por cada célula. Portanto, a temperatura associada a um determinado *pixel* precisa ser recalculada. Se um dado objeto i estava na célula (r, s) em um instante t_{irs} no passado, a temperatura de seu

Figura 24 – Este modelo diminui o peso de um objeto exponencialmente do laranja (mais recente) para o branco (expirado) através de um percentual fixo em intervalos regulares.



Fonte: o autor.

rastro naquela célula no instante atual t_k é computada como

$$T_i(r, s, k) = T_i^{max} \cdot W(t_k - t_{irs}), \quad (4.1)$$

onde T_i^{max} é a temperatura máxima do objeto (a contribuição em sua localização atual no instante t_k), $t_k - t_{irs}$ é o tempo decorrido para o objeto i e $W(t^*)$ é o fator de perda de temperatura por um determinado tempo decorrido t^* (veja a Figura 24). Assim, a temperatura de uma dada célula (r, s) da grade no instante atual t_k é a soma das temperaturas de todos os objetos que ocuparam essa célula (Figura 23b), ou seja,

$$T(r, s, k) = \sum_{i=1}^n T_i(r, s, k), \quad (4.2)$$

onde n é quantidade desses objetos. Na abordagem proposta, o fator de perda de temperatura é representado como uma função exponencial de decaimento do tempo decorrido como

$$W(t^*) = W_0 \lambda^{t^*}, \quad (4.3)$$

onde t^* é o tempo decorrido e W_0 é o fator de perda de temperatura quando o tempo decorrido é zero. A Figura 24 representa essa função com círculos coloridos indicando o efeito de esmaecimento do laranja para o branco. Esse modelo foi adotado com pequenas diferenças na literatura (JIANG; GRUENWALD, 2006; CAO *et al.*, 2006; CHEN; TU, 2007; ISAKSSON *et al.*, 2012). A taxa de decaimento pode ser controlada especificando o fator de perda de temperatura no final do rastro como uma fração f de W_0 , ou seja,

$$W(t_{max}^*) = fW_0 = W_0 \lambda^{t_{max}^*} \Rightarrow \lambda = f^{1/t_{max}^*}. \quad (4.4)$$

Geralmente, o tempo decorrido máximo t_{max}^* é definido como igual ao período de análise Δt e $W_0 = 1,0$. Se o período de análise é discretizado em n_t partições iguais, a taxa de decaimento de uma partição para a próxima pode ser computada como

$$\lambda = f^{1/n_t}. \quad (4.5)$$

Por exemplo, se $n_t = 10$ e $f = 0,01$, então $\lambda \approx 0,63$.

Assim, com λ calculado pela Eq. (4.5), é possível determinar incrementalmente o fator de perda de temperatura atual do objeto i no tempo t_k , na célula (r, s) , a partir do fator de perda anterior como

$$W(t_k - t_{irs}) = W(t_{k-1} - t_{irs})\lambda. \quad (4.6)$$

Por último, é definida a equação da temperatura máxima T_i^{max} de um objeto i , ou seja, a temperatura quando o tempo decorrido é zero. Considerando que τ objetos em movimento estão simultaneamente em uma determinada célula da grade (todos com igual T_i^{max}), a temperatura da célula atinge o limite de temperatura \bar{T} , então

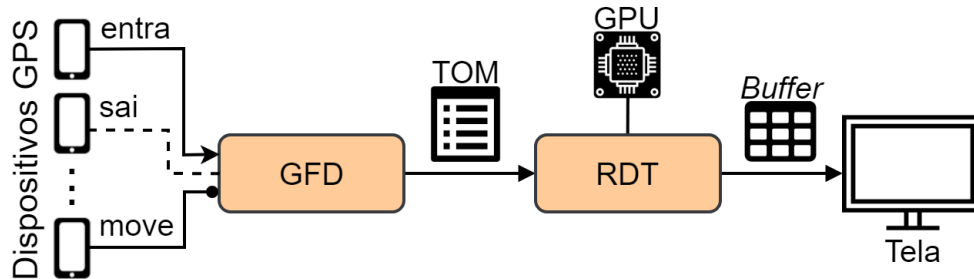
$$T_i^{max} = \bar{T} / \tau. \quad (4.7)$$

Conseqüentemente, quanto maior for τ , menor será T_i^{max} . A definição do limite \bar{T} é importante para representar a temperatura como um atributo na visualização dos dados, por exemplo, esse limite pode ser mapeado para o valor máximo de opacidade ou de um matiz de um pixel. A definição de τ depende da aplicação da abordagem proposta (veja Capítulo 5). Por exemplo, em uma análise de padrões de mobilidade local, na qual o foco é nos comportamentos individuais, o valor de T_i^{max} poderia ser igual a \bar{T} , logo, $\tau = 1$. Já na análise de padrões de mobilidade global, na qual o foco é o movimento agregado – vários rastros são agregados em uma mesma célula –, o valor de τ deve ser alto para que a contribuição de cada objeto no resultado final seja percebida.

4.2 Arquitetura

Assumindo que os objetos em movimento estão distribuídos geograficamente e podem enviar continuamente sua localização para um centro de monitoramento de tráfego, a arquitetura para a abordagem proposta é representada na Figura 25. A arquitetura é composta principalmente por dois componentes: o Gerenciador de Fluxo de Dados (GFD) e o Renderizador de Dinâmicas de Tráfego (RDT). A principal função do GFD é manter uma estrutura de dados

Figura 25 – Uma visão geral da arquitetura proposta. Os dispositivos GPS representam os objetos ativos em movimento que podem ser rastreados pelo Gerenciador de Fluxo de Dados (GFD). Esses dispositivos enviam suas posições atuais para o GFD, o qual mantém uma estrutura de dados sincronizada dos objetos em movimento (TOM). O componente Renderizador de Dinâmicas de Tráfego (RDT) usa uma GPU para processar a TOM e guardar o resultado em um buffer que é renderizado em uma tela.



Fonte: o autor.

atualizada com as posições atuais dos objetos ativos em movimento, representados pelos dispositivos GPS. Através de uma tabela indexada (TOM – Tabela de Objetos ativos em Movimento), o GFD insere novos objetos e suas posições atuais, remove aqueles que deixaram de ser rastreados, e atualiza as posições dos que estão se movendo e já estão na TOM. Em seguida, o RDT processa em GPU *buffers* gerados a partir da TOM e guarda o resultado em um outro *buffer* que pode ser renderizado em uma imagem em uma tela.

As abordagens encontradas na literatura armazenam os dados de trajetórias de uma janela de tempo para análise, mesmo aquelas voltadas para aplicações *on-line* (SILVA *et al.*, 2016; SILVA *et al.*, 2013; CHEN; TU, 2007; ISAKSSON *et al.*, 2012). Dependendo da quantidade de objetos rastreados e da taxa de amostragem de posicionamento, uma janela de tempo pode gerar um grande volume de dados a ser processado. Por exemplo, a frota de táxi de Pequim é composta por cerca de 70 mil táxis (GUI *et al.*, 2016), monitorar essa quantidade de veículos em tempo real é um grande desafio.

Na abordagem proposta, apenas as posições dos objetos ativos (TOM) e um *buffer* são temporariamente armazenados para serem visualizados, ou seja, os dados históricos não são mantidos. Portanto, diferentemente das abordagens encontradas na literatura, o modelo proposto é compacto e adequado para uma aplicação de monitoramento de dinâmicas de tráfego. Além disso, um centro de monitoramento de tráfego pode armazenar os *buffers* como imagens ou vídeos para manter um registro de ocorrências, como acontece em sistemas de vigilância. Na próxima seção, a descrição das funcionalidades dos componentes é apresentada com mais detalhes e inserida dentro do contexto da metodologia proposta.

Figura 26 – A metodologia da abordagem proposta: na fase de Preparação de Dados, o sistema limpa os dados brutos e os sincroniza para serem visualizados. Na fase de Visualização, esses dados são renderizados em imagens, e alguns atributos dessas imagens são realçados para facilitar a análise visual.



Fonte: o autor.

4.3 Metodologia

Este trabalho propõe uma metodologia sistemática dividida em etapas organizadas em duas fases (representadas pelas caixas amarelas na Figura 26): a fase de Preparação de Dados e a fase de Visualização. Na fase de Preparação de Dados, a etapa de pré-processamento (Seção 4.3.1) remove inconsistências e pré-processa atributos espaço-temporais, como velocidade e distância. Em seguida, a etapa de Sincronização (Seção 4.3.2) atualiza as posições dos objetos ativos para a mesma unidade de tempo discreta. Na fase de Visualização, a etapa de Renderização (Seção 4.3.3) recebe a TOM para ser renderizada e compõe *buffers* para a visualização. Por último, a etapa de Processamento Visual (Seção 4.3.4) realça os atributos das imagens geradas anteriormente para facilitar a análise visual. Os algoritmos da fase de Visualização são executados em paralelo em GPU.

4.3.1 Pré-processamento

O registro de localização por GPS geralmente apresenta imprecisões causadas por vários fatores (HE *et al.*, 2017). Assim, a etapa de Pré-processamento inicialmente remove: pontos com má precisão; pontos que estão fora da área geográfica de análise; pontos duplicados ou muito próximos; e segmentos de trajetória com velocidade irreal – muito além do limite. Esta etapa foi baseada no trabalho de Zheng (ZHENG, 2015), que lista um conjunto de ações para remover inconsistências em dados coletados por GPS. Na arquitetura proposta, esta etapa é executada nos dispositivos GPS antes de enviar a localização para o GFD. Para esse fim, foi adotada uma abordagem distribuída proposta por Sacharidis *et al.* (2008) para processar e filtrar atualizações de localização em tempo real utilizando dispositivos móveis.

Para minimizar o alto custo computacional de renderizar uma grande quantidade de elementos gráficos em uma visualização dinâmica, esta etapa pré-processa os dados da

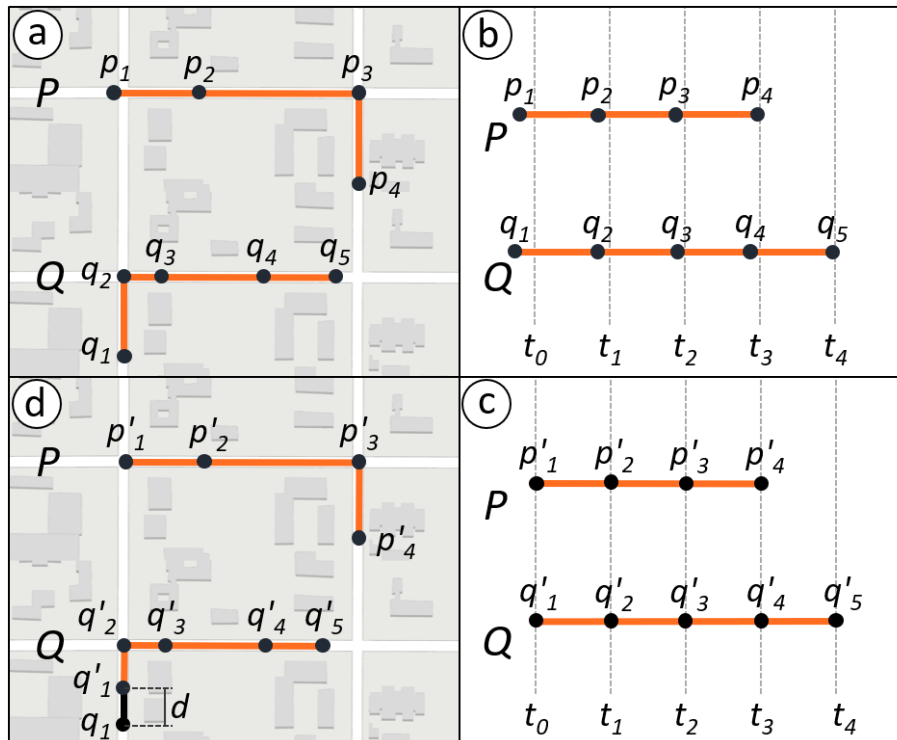
trajetória de modo que os atributos espaço-temporais, como velocidade e distância, sejam calculados antecipadamente nos dispositivos GPS. Desta forma, o processamento desses atributos é distribuído entre esses dispositivos. Em seguida, esses atributos são enviados junto com a localização de um objeto ativo para o GFD que os armazena na TOM. A TOM representa pontos e objetos ativos usando um modelo de dados estruturados. De acordo com esse modelo, para cada ponto recebido em um fluxo de dados, além de seus dados brutos (latitude, longitude, tempo), os atributos espaço-temporais (velocidade, distância e direção) são armazenados em mapas *hash* indexados por um *id*. Além desses, os pontos que vêm antes e depois de um determinado ponto também são guardados como atributos, a fim de facilitar a navegação entre eles. Esses atributos são usados na interpolação espacial na próxima etapa.

4.3.2 Sincronização

Além de atualizar as posições dos objetos ativos na TOM, o GFD sincroniza essas posições para a mesma unidade discreta de tempo. Isso é necessário porque as posições dos dispositivos GPS geralmente chegam com marcações de tempo (*timestamps*) diferentes devido a problemas de comunicação de rede – como atrasos, pacotes perdidos, sincronização de horário e assim por diante –, mesmo quando esses dispositivos enviam sua localização em intervalos regulares (AGUILERA *et al.*, 1997). Assim, para representar graficamente esses objetos em uma aplicação de monitoramento em tempo real, em um instante de tempo t_i , a aplicação deve saber onde cada objeto está localizado na tela, ou seja, para cada objeto, a localização correspondente ao instante t_i é calculada por interpolação espacial e armazenada na TOM. Essa etapa de sincronização de dados é semelhante a um processo de normalização de trajetória (ZHENG; ZHOU, 2011b).

A Figura 27 mostra o processo de sincronização de tempo para dois objetos: P e Q . Nesse cenário, os objetos enviam suas localizações em um intervalo regular de um segundo, ou seja, em uma taxa de frequência de um segundo. Cinco iterações correspondendo a cinco unidades de tempo de sincronização são mostradas (t_0 a t_4) nessa ilustração. A Figura 27a mostra um mapa com os pontos originais p_i e q_i enviados ao GFD pelos objetos P e Q , respectivamente. As trajetórias dos objetos são exibidas apenas para fins ilustrativos, elas são renderizadas pelo RDT em outro processo (veja a Seção 4.3.3). A Figura 27b ilustra a distribuição dos pontos originais no tempo, mostrando que no instante t_1 , P e Q estavam ativos e, portanto, ambos devem estar na TOM e, no momento t_4 , como apenas Q estava ativo, P foi removido da TOM.

Figura 27 – Processo de sincronização de tempo. (a) Um mapa exibindo os registros originais das localizações dos objetos P e Q . (b) Os pontos originais distribuídos em cinco unidades de tempo do processo de sincronização (t_0 a t_4). (c) As novas localizações para os objetos P e Q correspondentes a cada instante t_i são calculadas por interpolação espacial. (d) Os novos pontos exibidos no mapa, com a distância entre o ponto original q_1 e o ponto interpolado q'_1 destacado como um exemplo. As trajetórias são exibidas apenas para fins ilustrativos, elas são renderizadas pelo RDT na próxima etapa, Renderização.



Fonte: o autor.

A Figura 27c mostra os novos pontos p'_i e q'_i interpolados e sincronizados para t_i . Por último, a Figura 27d mostra os novos pontos sincronizados no mapa. Nessa última ilustração, a distância entre os pontos q_1 e q'_1 é destacada como exemplo. Posteriormente, no Capítulo 5 (experimento 4 da Seção 5.1.2), demonstra-se que as distâncias entre os pontos originais e interpolados não afetam significativamente a visualização final. Para fins ilustrativos, os pontos originais p_3 e q_2 estão perfeitamente posicionados nas esquinas. Como a rede viária não é utilizada, e dependendo da taxa de amostragem, as trajetórias dos objetos P e Q não são tão perfeitas na prática. Erros de trajetória, como um carro atravessando uma esquina, são comuns utilizando dados brutos GPS. No entanto, esses erros podem ser minimizados usando diferentes escalas de mapa. Esse tópico também é discutido no Capítulo 5 (experimento 4 da Seção 5.1.2).

4.3.3 Renderização

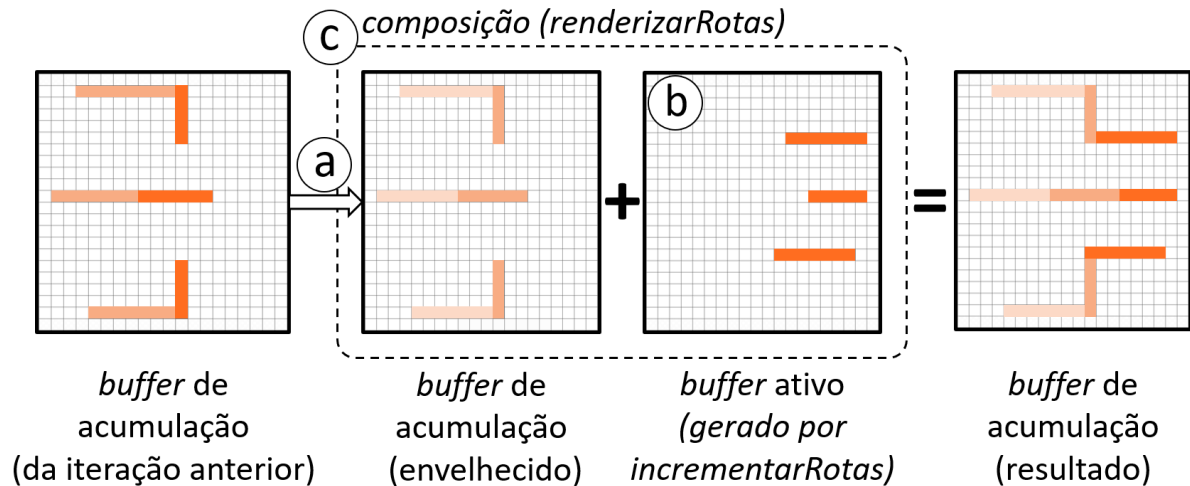
Usando a TOM fornecida pelo GFD com as posições sincronizadas dos objetos ativos, em conjunto com a abordagem baseada em grade e os conceitos explicados na Seção 4.1, esta etapa renderiza de forma incremental as trajetórias dos objetos em movimento que servirão de base para a análise visual das dinâmicas de tráfego. A Figura 28 apresenta uma visão esquemática da técnica de renderização desenvolvida. Nela, dois *buffers* são utilizados: um *buffer* de acumulação que representa as trajetórias acumuladas no período de análise; e um *buffer* ativo que representa uma atualização do movimento dos objetos ativos do instante t_{k-1} para o instante atual t_k . Essa ilustração descreve uma iteração do algoritmo para o instante t_k com três objetos ativos em movimento. Em cada iteração do algoritmo, o *buffer* de acumulação da iteração anterior é envelhecido através de uma função de decaimento (Figura 28a), e um *buffer* ativo é gerado incrementando a trajetória de cada objeto ativo (Figura 28b). A taxa de decaimento é calculada pela Eq. (4.5). Por último, os dois *buffers* são compostos (Figura 28c) produzindo um novo que será o *buffer* de acumulação para o tempo t_{k+1} . Este processo é executado continuamente ao longo do tempo. Para tornar este exemplo mais compreensível, cada objeto na Figura 28 tem a temperatura igual ao limite \bar{T} , quando $\tau = 1$, calculada pela Eq. (4.7).

As dimensões de ambos os *buffers* são compatíveis com o número de *pixels* na *viewport*. Se a mesma região de interesse (janela em coordenadas de mundo) for exibida em telas com resoluções diferentes, a quantidade de *pixels* será maior na *viewport* da tela de resolução mais alta. Portanto, altas resoluções implicam grandes *buffers* a serem processados em GPU, prejudicando o desempenho. Esse tópico é discutido no Capítulo 6.

Para gerar uma visualização suave ao longo do tempo, posições intermediárias também são interpoladas entre as localizações atuais e anteriores para cada objeto ativo. Isso permitirá uma melhor percepção do movimento dos objetos como uma animação. Por exemplo, para uma visualização a 30 quadros por segundo (*fps*, do inglês *frames per second*), 28 posições intermediárias são interpoladas por segundo e por objeto. Para conseguir isso, o processo *inbetweening*¹ e o princípio de animação Pose-to-Pose (THOMAS *et al.*, 1995) são utilizados para definir *quadros-chave* e manter a velocidade de visualização com quadros intermediários. Na abordagem proposta, os quadros-chave são as posições sincronizadas na TOM. Para ilustrar o esforço computacional empregado neste processo, considere que a TOM contenha 10.000

¹ *Inbetweening* é um processo utilizado em animação para criar a ilusão de movimento a partir da geração de quadros intermediários entre duas imagens, chamadas quadros-chave, para dar a aparência de que a primeira imagem evolui suavemente para a segunda imagem.

Figura 28 – Uma visão gráfica geral da etapa de renderização. (a) O *buffer* de acumulação da iteração anterior é envelhecido através de uma função de decaimento. (b) O *buffer* ativo é gerado incrementando as trajetórias dos objetos ativos em movimento (gerado por *incrementarRotas* no Algoritmo 2). (c) Em seguida, os dois *buffers* são compostos produzindo o *buffer* de acumulação para a próxima iteração. O pseudo-código do *renderizarRotas* está listado no Algoritmo 1.



objetos que estão sendo visualizados a 30fps, então 300.000 interpolações são realizadas por segundo, incluindo pontos sincronizados e intermediários. Além disso, 30 operações de filtragem e composição de *buffer* também são realizadas por segundo. Para visualizar o movimento desses objetos processando essa grande quantidade de interpolações e operações, é necessário utilizar uma GPU.

Esta etapa apresenta uma solução eficiente com complexidade linear que resolve esse problema. O pseudo-código para processar cada iteração é listado nos Algoritmos 1 e 2. A Tabela 1 resume as notações usadas nesses algoritmos.

Inicialmente, um *loop* de animação que chama a função *renderizarRotas* (Algoritmo 1) é executado em intervalos regulares. Todas as interpolações, as composições e a filtragem de *buffers* devem levar menos tempo que o intervalo regular para não prejudicar a visualização. Essa função executa todos esses cálculos a um custo computacional mínimo para renderizar um quadro de animação. Os parâmetros de entrada são: o período de análise representado por seu número de partições de tempo n_t ; o limiar de densidade τ ; e o *buffer* de acumulação B_w da iteração anterior (instante t_{k-1}). Na primeira vez que essa função é chamada, B_w está vazio. Para a saída, ela retorna o *buffer* B_w atualizado para o instante t_k que será usado como um quadro de animação na visualização. Primeiro, essa função recebe O_a do GFD (linha 1). Na linha 2, λ é calculado pela Eq. (4.5) com o parâmetro n_t . Em seguida, a opacidade do *buffer* B_w com

Tabela 1 – Notação usada na etapa de Renderização.

Notação	Descrição
O_a	O conjunto (obtido da TOM) de objetos em movimento que estão ativos durante um tempo de iteração; o fluxo de dados de entrada.
n_t	O número de partições do período de análise. Esse parâmetro pode ser definido pelo usuário.
τ	O limiar de densidade, isto é, o número mínimo de objetos em movimento em uma célula da grade para que ela atinja o limite máximo de temperatura \bar{T} . Esse parâmetro também pode ser definido pelo usuário.
B_w	O <i>buffer</i> de acumulação com todos os rastros gravados dos objetos que foram rastreados em algum momento dentro do período de análise.
B_a	O <i>buffer</i> ativo que representa uma atualização do movimento dos objetos dentro de um tempo de iteração. Este <i>buffer</i> é gerado incrementando a trajetória de cada objeto ativo.
λ	O fator λ controla o percentual de perda de temperatura dos objetos por iteração.
o	Um objeto em movimento.
o_w	A temperatura contribuída por um objeto.
$[lat, lng]$	As coordenadas de latitude e longitude.
t	Um segmento de trajetória.

Algoritmo 1: renderizarRotas(n_t, τ, B_w)**Entrada:**

- o número de partições n_t
- o limiar de densidade τ
- o *buffer* de acumulação B_w

Retorno :

B_w atualizado

- 1 $O_a \leftarrow GFD.retornarTOM()$
- 2 $\lambda \leftarrow calcularDecaimento(n_t)$
- 3 $B_w \leftarrow B_w \cdot \lambda$
- 4 $B_a \leftarrow incrementarRotas(O_a)$
- 5 $B_w \leftarrow compor(B_w, B_a)$
- 6 $B_w \leftarrow suavizarFiltrar(B_w)$
- 7 $mostrarMarcadores(O_a)$
- 8 **retorne** B_w

todas as rotas já desenhadas até o instante t_{k-1} é reduzida usando λ (linha 3). Na linha 4, a rota de cada objeto ativo é estendida por um segmento de linha em *incrementarRotas* (Algoritmo 2). A linha 5 compõe as rotas dos objetos ativos em B_a com as rotas acumuladas em B_w . Essa composição é executada em paralelo na GPU. A função *suavizarFiltrar* (apresentada na próxima etapa, Seção 4.3.4) aprimora a qualidade da visualização destacando e filtrando as densidades mais relevantes. Por fim, para ajudar na percepção do número de objetos monitorados e suas posições atuais, os marcadores são renderizados opcionalmente para cada objeto ativo (linha 7).

A função *incrementarRotas* no Algoritmo 2 recebe como entrada o conjunto de

Algoritmo 2: incrementarRotas(O_a)

Entrada:o conjunto de objetos em movimento O_a **Retorno :**o *buffer* ativo B_a

```

1  $B_a \leftarrow \emptyset$ 
2 para todo  $o \in O_a$  faça
3    $[lat, lng] \leftarrow encontrarLocal(o)$ 
4    $o_w \leftarrow 1/\tau$ 
5    $t \leftarrow criarSegmentoTrajetria([lat, lng], o, o_w)$ 
6    $B_a \leftarrow B_a + t$ 
7 fim
8 retorne  $B_a$ 

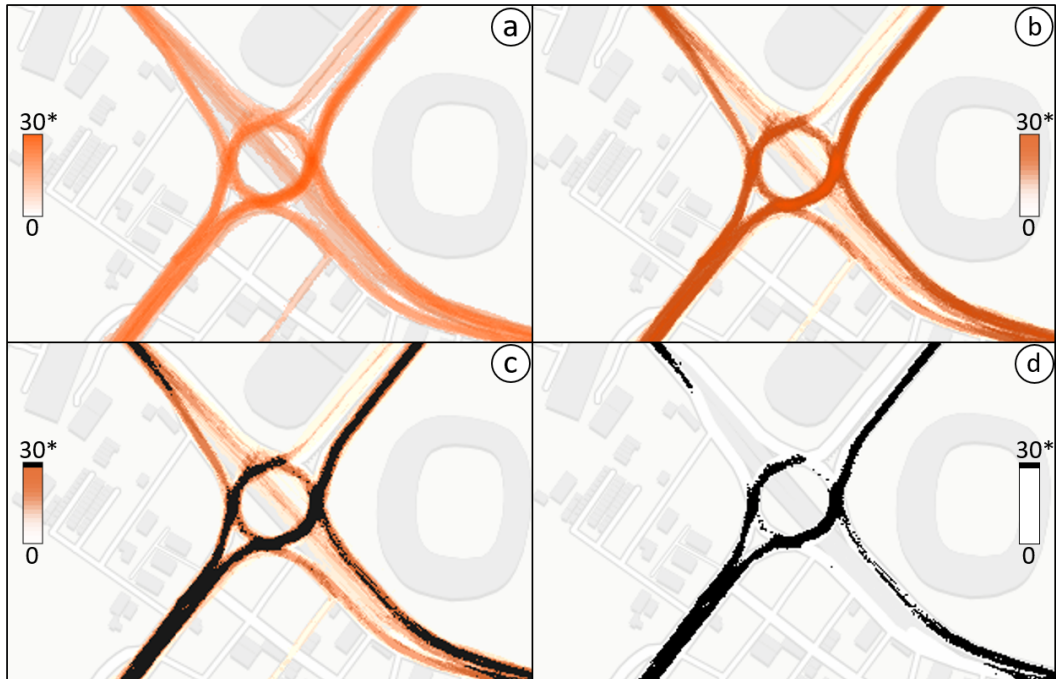
```

objetos ativos O_a e retorna como saída o *buffer* B_a . Na linha 1, ela começa limpando o *buffer* B_a . Em seguida, para cada objeto ativo o (linha 2), a função *encontrarLocal* (linha 3) interpola a posição atual $[lat, lng]$ de o no mapa. Para esta tarefa, cálculos loxodrômicos (WEINTRIT; KOPACZ, 2011) são utilizados considerando a projeção do mapa em um globo. Na linha 4, é calculada a temperatura o_w do objeto, que é inversamente proporcional ao limiar de densidade τ . Na linha 5, um segmento de linha é criado entre as posições anterior e atual de o . A função *criarSegmentoTrajetória* cria o segmento de linha e usa o_w para definir sua opacidade e cor representando a temperatura do objeto. Essa função usa primitivas baseadas em textura, como linha e círculo, que aumentam o desempenho do sistema em cerca de dez vezes quando comparadas às primitivas geométricas padrões do WebGL, mesmo utilizando GPU. Assim, é possível obter um alto desempenho de renderização com taxas de visualização interativas. Em seguida, esse segmento de trajetória é adicionado a B_a (linha 6). No final, B_a consolida todos os segmentos como uma única textura, representando uma atualização do movimento dos objetos no período entre os instantes t_{k-1} e t_k .

4.3.4 Processamento Visual

A rota de um objeto em movimento e seus atributos espaço-temporais, como velocidade e direção, são essenciais para compreender melhor a dinâmica do fluxo de tráfego. Esta etapa destaca esses atributos por meio de um conjunto de operações de processamento de imagens, como filtros, suavização e contraste. Desta forma, com um conjunto pré-definido de aprimoramentos visuais, é possível destacar visualmente padrões de mobilidade, como ondas de tráfego (TREIBER; KESTING, 2013) e rotas quentes (LI *et al.*, 2007). Esses padrões são

Figura 29 – Os efeitos de suavização e filtragem. (a) Usando apenas as contribuições dos objetos para visualizar as temperaturas das rotas. (b) As mesmas rotas após a aplicação de uma função de suavização. (c) Usando a função de suavização e um filtro que realça as rotas que atingiram a temperatura \bar{T} . (d) Apenas as rotas com temperatura \bar{T} são visualizadas em uma cor preta (sem suavização).



*Quantidade de veículos por tempo de análise

Fonte: o autor.

discutidos em aplicações no Capítulo 5.

Ocasionalmente, as rotas que atingem o limite de temperatura \bar{T} não podem ser facilmente percebidas ao visualizar apenas a temperatura resultante das rotas dos objetos. Por exemplo, na visualização das contribuições dos objetos na Figura 29a, é difícil identificar quais rotas atingiram a temperatura \bar{T} . Para resolver isso, esta etapa aplica uma função de suavização que enfatiza as rotas com alta densidade e suaviza aquelas que apresentam baixa densidade (Figura 29b). Além disso, é possível realçar essas rotas com uma cor diferente (Figura 29c), ou até mesmo filtrar para visualizar apenas as rotas com temperatura \bar{T} (Figura 29d).

Esta etapa é executada em paralelo diretamente na GPU por um programa *shader*² representado pela função *suavizarFiltrar* no Algoritmo 3. A Tabela 2 resume as novas notações usadas nesse algoritmo. Essa função utiliza os recursos de mapeamento de texturas das GPUs modernas para otimizar o processo. Ela recebe como entrada o *buffer* de acumulação B_w ; a chave booleana *suavizar* e o tipo *filtro*; e retorna como saída o *buffer* B_w atualizado. Para cada

² Em computação gráfica, um *shader* é um tipo de programa de computador que desempenha uma variedade de funções especializadas em processamento de imagens, efeitos especiais ou pós-processamento de vídeo. Geralmente, esses programas são codificados para ser executados em uma GPU.

Tabela 2 – Notação usada na etapa de Processamento Visual.

Notação	Descrição
<i>suavizar</i>	Um chave booleana que, se for verdadeira, ativa a suavização dos rastros dos objetos.
<i>filtro</i>	O tipo de filtro aplicado aos <i>buffers</i> .
<i>fg</i>	Uma cor de fragmento que representa um pixel de um <i>buffer</i> com as componentes de cor RGBA. Esses fragmentos são gerados e processados em paralelo pela GPU.

fragmento *fg* em B_w (linha 1), o algoritmo verifica se seu componente alfa (RGBA) é zero (linha 2). Se verdadeiro, o fragmento é descartado na linha 3 e não prossegue para os próximos estágios da *pipeline* gráfica. Geralmente, o *buffer* B_w é muito esparso devido à restrição de movimento em uma rede viária e ao envelhecimento causado pela função de decaimento. Essa função aproveita essa característica e descarta diretamente na GPU todos os fragmentos com opacidade zero, resultando em melhora significativa do desempenho. Caso contrário, se o componente alfa do fragmento não for zero (linha 4), a suavização e a filtragem poderão ser aplicadas a esse fragmento. Assim, nas linhas 5-7, se *suavizar* for verdadeira, o algoritmo atualiza o alfa do fragmento, aplicando uma função *Sigmoid* para o intervalo de opacidade [0,1]. Na abordagem proposta, a *Sigmoid* foi utilizada, mas qualquer outra função de suavização, como uma Gaussiana, pode ser aplicada.

Em seguida, um filtro marca em preto as células que atingiram a temperatura \bar{T} , ou seja, a cor dos fragmentos com opacidade máxima é alterada. Relembrando o Algoritmo 2, em que a temperatura do objeto foi usada como opacidade para criar um segmento de trajetória, o Algoritmo 3 verifica se o componente alfa do fragmento é 1 e se o *filtro* é do tipo 1. Se verdadeiro, a cor do fragmento é configurada para preto com o máximo de alfa (Figura 29c) para distinguir as células que atingiram o limite \bar{T} (linha 9). Caso contrário, na linha 10, se o *filtro* é do tipo 2, então apenas as células com temperatura \bar{T} são mostradas. Esse filtro é semelhante ao filtro tipo 1, mas os fragmentos com um componente alfa menor que 1 são descartados (linhas 11-14, Figura 29d). Um valor diferente para o tipo de filtro resultará em não aplicação de filtragem. Concluindo, este algoritmo contém as principais funcionalidades que são executadas em paralelo na GPU. Outros recursos também podem ser incluídos, como correção de cor, redução de ruído e diferentes escalas de cores. Nos experimentos realizados neste trabalho (Capítulo 5), o processamento para suavizar e filtrar um *buffer* afetam o desempenho de renderização em apenas 5%, aproximadamente.

Algoritmo 3: suavizarFiltrar(B_w)

Entrada:

- o *buffer* de acumulação B_w
- a chave lógica *suavizar*
- o tipo *filtro*

Retorno :

```

   $B_w$  atualizado
1 para todo  $fg \in B_w$  faça em paralelo
2   se ( $fg.a = 0$ ) então
3     | descartar
4   senão
5     | se ( $suavizar = 1$ ) então
6       |  $fg.a = 1 / (1 + \exp(-(fg.a - 0.5) / 0.1))$ 
7     | fim
8     | se ( $filtro = 1$  e  $fg.a = 1$ ) então
9       |  $fg = cor(0, 0, 0, 1.0)$ 
10    | senão se ( $filtro = 2$ ) então
11      | se ( $fg.a < 1$ ) então
12        | descartar
13      | senão
14        |  $fg = cor(0, 0, 0, 1.0)$ 
15      | fim
16    | fim
17  | fim
18 fim
19 retorne  $B_w$ 

```

4.4 Considerações Finais

Este capítulo descreveu em detalhes a abordagem proposta para análise visual de dinâmica de tráfego. Inicialmente, a definição do problema foi apresentada. Em seguida, a arquitetura e seus principais componentes foram descritos. E por último, uma metodologia foi detalhada descrevendo as etapas desde o pré-processamento dos dados de trajetórias até a visualização desses dados.

As abordagens atuais para a análise de dinâmica de tráfego normalmente empregam discretização de tempo em janelas e não podem ser empregadas para monitoramento em tempo real. A solução apresentada neste capítulo emprega uma abordagem incremental baseada em grade, que usa uma estrutura de dados indexada para rastrear os objetos ativos e suas posições atuais. A visualização é mantida em um *buffer* ativo e em um *buffer* de acumulação que são iterativamente atualizados e renderizados usando *inbetweening*. A solução é conceitualmente simples, mas requer um esforço de implementação significativo e uma combinação não trivial de

recursos para lidar efetivamente com dados em fluxo e atender a requisitos de qualidade visual e escalabilidade. Essa solução pode monitorar cenários dinâmicos com mais eficiência e requer um pequeno número de parâmetros: um período de análise e um limiar de densidade.

Apesar deste trabalho ter sido projetado para aplicações de monitoramento de tráfego em tempo real, a arquitetura e a metodologia proposta podem ser adaptadas para análise de dados históricos. Para isso, o GFD carregaria todos os dados de trajetória de uma vez e depois atualizaria a TOM em períodos regulares, simulando um fluxo de dados para o RDT. Nesta configuração, a etapa de pré-processamento também deve ser executada no GFD, ou seja a remoção de inconsistências nos dados GPS e o processamento dos atributos espaço-temporais, que antes eram realizados nos dispositivos GPS, agora podem ser realizados no GFD. Além disso, outras operações de pré-processamento de trajetórias podem ser incluídas nesta etapa, como: compressão e simplificação de trajetórias para reduzir o volume de dados; mapeamento (*map-matching*) para reduzir os erros GPS e melhorar a qualidade da visualização; e o enriquecimento semântico para analisar outros aspectos de mobilidade.

O próximo capítulo apresenta as aplicações da abordagem proposta na análise de padrões de mobilidade global e local.

5 ANÁLISE VISUAL DE DINÂMICA DE TRÁFEGO

Este capítulo apresenta duas aplicações da abordagem proposta para análise visual de dinâmica de tráfego. A Seção 5.1 apresenta a visualização de rotas quentes para encontrar padrões de mobilidade global no fluxo de tráfego. Em seguida, com o objetivo de analisar padrões locais de oscilação de tráfego, a Seção 5.2 apresenta a visualização de ondas de tráfego. Por último, na Seção 5.3, tecem-se algumas considerações finais.

5.1 Visualizando Rotas Quentes

O problema que visa descobrir rotas que frequentemente apresentam tráfego pesado (rotas quentes) foi abordado recentemente por vários trabalhos (ALMEIDA *et al.*, 2017; GUI *et al.*, 2016; WU *et al.*, 2014; SACHARIDIS *et al.*, 2008; LI *et al.*, 2007). Além desses, os trabalhos de clusterização de trajetórias também podem ser usados para encontrar rotas quentes – como foi discutido na Seção 3.4 do Capítulo 3. Analisar como essas rotas quentes evoluem permite que as agências de trânsito identifiquem eventos anormais em um estágio inicial para direcionar melhor o tráfego. No entanto, nenhum desses trabalhos foi projetado para o monitoramento de tráfego em tempo real, pois eles encontram padrões de mobilidade ao final de uma janela de tempo através da agregação do movimento que ocorreu nesse período. Apesar de existirem abordagens *online*, como a proposta por Sacharidis (2008), elas ainda agregam dados de trajetórias dentro de uma janela de tempo, ou seja, não é uma aplicação em tempo real. Embora agregar movimento seja eficaz para extrair padrões de mobilidade, utilizar uma janela de tempo prejudica a análise de toda a dinâmica dos objetos em movimento, o que é vital para os sistemas de monitoramento de tráfego.

Assim, analisar os dados à medida que eles chegam fornece uma compreensão melhor dos padrões de rotas quentes, ou seja, caminhos que são compartilhados por um grupo de objetos em movimento. Neste sentido, a abordagem proposta neste trabalho monitora continuamente a dinâmica dos objetos em tempo real. Diferentemente de outras abordagens existentes, o método proposto não explora um fluxo de dados inteiro (KHARRAT *et al.*, 2008; HAN *et al.*, 2012; SACHARIDIS *et al.*, 2008; LI *et al.*, 2007) nem usa janelas de tempo (LI *et al.*, 2010b; SILVA *et al.*, 2016).

Para encontrar rotas quentes utilizando a abordagem proposta no capítulo anterior, apenas dois parâmetros são definidos diretamente pelo usuário: o período de análise representado

por n_t e o limiar de densidade τ . Note que n_t representa o número de iterações que a temperatura de um objeto i decai do valor máximo T_i^{max} a zero, e τ é utilizado para encontrar a T_i^{max} através da Eq. (4.7). Assim, uma determinada célula da grade é considerada quente se a temperatura dessa célula atinge o limite de temperatura \bar{T} , isto é, τ objetos em movimento estão simultaneamente nessa célula (todos com igual T_i^{max}). Portanto, quanto maior for τ , menor será T_i^{max} . A definição de τ para encontrar rotas quentes significativas depende do fluxo de veículos nas ruas. Por exemplo, o fluxo nas vias expressas é geralmente maior do que o fluxo nas estradas secundárias. Assim, em vias expressas o limiar de densidade deve ser mais alto para ter rotas quentes significativas. Além disso, τ também depende da representação amostral do fluxo. Por exemplo, quando os dados de trajetórias são registrados por câmeras, o fluxo de tráfego tem uma representação mais completa do que as trajetórias coletadas por aplicativos específicos, como aplicativos de táxi. Nas abordagens tradicionais de clusterização de trajetórias, o limiar de densidade é definido por uma análise estatística prévia. Nas aplicações de monitoramento em tempo real, esse processo não é possível, pois o conjunto completo dos dados ainda não está disponível. Assim, o método proposto utiliza análise visual interativa para descobrir valores adequados de τ , isto é, o usuário ajusta o parâmetro em tempo real e as alterações são visualizadas instantaneamente. Os resultados da aplicação da abordagem proposta nesta tese na análise visual de rotas quentes podem ser encontrados no artigo (GOMES *et al.*, 2018).

A Figura 30 mostra um exemplo de aplicação da abordagem proposta para visualização de rotas quentes. Para esse exemplo, foram utilizados dados brutos de GPS com trajetórias

Figura 30 – Exemplo de aplicação da abordagem proposta para encontrar rotas quentes. (a) Trajetórias originais geradas pela etapa de Renderização. (b) Trajetórias suavizadas e com a marcação das rotas quentes realizada na etapa de Processamento Visual destacando com a cor preta toda célula com opacidade igual a 1,0.



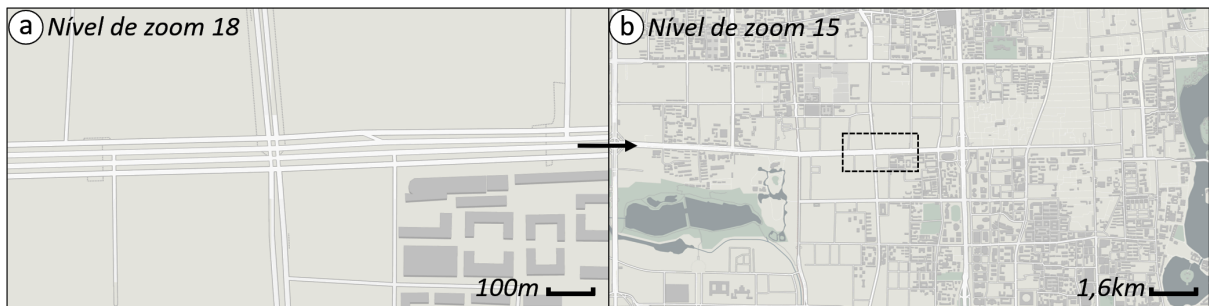
Fonte: o autor.

de aproximadamente 850 táxis. Os conjuntos de dados usados nos experimentos desta aplicação são detalhados na Seção 5.1.2. O limiar τ foi definido como 10, ou seja, a temperatura máxima T_i^{max} de um objeto i é representada na imagem como 0,1 de opacidade, onde $\bar{T} = 1,0$. O período de análise foi definido em 10 minutos com uma taxa de visualização de $30fps$, logo, $n_t = 18.000$ ($10min \times 60s \times 30fps$). A Figura 30a mostra as trajetórias originais geradas pela etapa de Renderização. A marcação das rotas quentes acontece na etapa de Processamento Visual através do Algoritmo 3 destacando com a cor preta toda célula com opacidade igual a 1,0. Diferentemente dos trabalhos existentes que retornam apenas as rotas quentes, o método proposto codifica a variação de densidade através de uma escala de densidade (0 a 10). Assim, com essa variação, podem-se perceber rotas que são significativas, mas que ainda não se tornaram quentes. Por exemplo, é possível perceber a formação de rotas quentes nos cruzamentos marcados por círculos pontilhados na Figura 30b. Com a suavização e o destaque das rotas quentes (Figura 30a), esses padrões ficam bem mais perceptíveis quando comparados às mesmas regiões na Figura 30a. A animação propicia um estímulo visual através do qual percebe-se a evolução dessas rotas que se formam ou que se dissipam. Um vídeo mostrando a visualização animada está disponível em <https://youtu.be/mDUy956-vSA>.

5.1.1 Representação do Mapa e Níveis de Zoom

Antes de apresentar os experimentos da abordagem proposta para encontrar rotas quentes, é necessário explicar o comportamento de dois pontos importantes na análise visual usando mapas: a representação do mapa e os níveis de zoom. Como a rede viária não é utilizada, a representação do mapa tem o papel de contextualizar a análise por meio de um conjunto de imagens (*tileset*) distribuídas em uma grade uniforme de células quadradas (*tiles*) em diferentes níveis predefinidos de zoom. Ao interagir com o mapa, navegando ou mudando o nível de zoom, mais imagens são carregadas. É importante mencionar que metadados ou geometrias do mapa não são utilizados no método proposto. Nos experimentos, foi utilizado um mapa com 18 níveis de zoom, que são comumente usados por diferentes aplicações, como o OpenStreetMap, o Mapbox, o Bing Maps e o Google Maps. Para tornar a navegação mais suave, o Google Maps, por exemplo, usa quatro níveis intermediários entre dois níveis consecutivos dos 18 níveis. Nessas aplicações, quanto maior o nível de zoom, maior será a escala do mapa. Por exemplo, no nível de zoom 18, podem-se ver todas as faixas de uma via expressa (Figura 31a); e no nível 15, as faixas não são distinguíveis, e a via é exibida como uma rua de pista única (Figura 31b).

Figura 31 – (a) Representação de uma via expressa com 4 faixas no nível de zoom 18. (b) A via expressa em (a) é exibida como uma rua de pista única no nível de zoom 15 na região delimitada pelo retângulo pontilhado.



Fonte: o autor.

Nessa ilustração, é possível perceber que uma pequena diferença de nível de zoom (3 de um total de 18 níveis) causa uma grande mudança na escala do mapa. Esse aspecto é discutido no experimento 5 da Seção 5.1.2.

Assim, mudanças no nível de zoom podem exigir alterações no limiar de densidade τ , especialmente nos níveis mais altos nos quais as faixas são distinguíveis. Nos experimentos deste capítulo, por exemplo, as faixas são distinguíveis nos níveis 16, 17 e 18. Assim, diferentes limiares são usados para obter representações significativas de rotas quentes. No entanto, como o usuário interage com o sistema de visualização, ele pode ajustar o limiar nesses níveis de zoom em tempo real e as alterações são visualizadas instantaneamente. Para os níveis restantes, o mesmo limiar pode ser usado, pois as faixas já estão combinadas em uma única pista. É importante destacar que estão fora do escopo deste trabalho os níveis de zoom nos quais as ruas não estão visíveis. Assim, nos experimentos deste capítulo, os níveis de zoom abaixo de 7 estão fora do escopo. Da mesma forma, os níveis altos de zoom (16, 17 e 18) nos quais uma faixa de uma via pode preencher uma grande área da tela não são interessantes para a aplicação de encontrar rotas quentes. Para o monitoramento de tráfego de cidades, regiões e bairros, os níveis de 12 a 15 são geralmente os mais relevantes.

Os níveis de zoom também afetam a percepção dos erros de GPS. Na abordagem proposta baseada em grade, os erros de GPS são minimizados em níveis baixos de zoom, pois há mais dados agregados por célula nesses níveis. Conseqüentemente, as rotas quentes têm uma aparência mais consistente em níveis menores de zoom. Além disso, a abordagem proposta apresenta melhores resultados quando o fluxo de dados tem uma alta taxa de amostragem, por exemplo, objetos em movimento enviando atualizações de posição a cada cinco segundos, no máximo. Como a rede viária não é utilizada, uma baixa taxa de amostragem pode causar inconsistências devido à esparsidade dos pontos das trajetórias, por exemplo, segmentos de

trajetória cruzando quarteirões. Em níveis menores de zoom, no entanto, essas inconsistências são minimizadas devido ao efeito de agregação em uma célula de grade. Esse aspecto é discutido no experimento 4 da Seção 5.1.2.

5.1.2 Resultados e Discussão

Seis experimentos foram desenvolvidos para avaliar os resultados da abordagem proposta para encontrar rotas quentes. Nesses experimentos, foram usados dois conjuntos de dados reais de táxi coletados por dispositivos GPS. O primeiro conjunto de dados é o T-Drive, que contém trajetórias de mais de 10.000 táxis em Pequim dentro de uma semana (YUAN *et al.*, 2011; YUAN *et al.*, 2010). Do T-Drive, foi selecionado um subconjunto de uma hora com cerca de 7.000 táxis no total e média de 1.500 objetos ativos, ou seja, táxis que se movem ao mesmo tempo. Embora seja um conjunto de dados com baixa taxa de amostragem (taxa de amostragem de um minuto), o T-Drive foi escolhido por conter dados reais coletados de um grande número de veículos. Assim, para simular uma tarefa de monitoramento de tráfego desses táxis, pontos intermediários foram interpolados e mapeados para uma rede viária (*map-matching*) para tornar esse subconjunto denso com uma alta taxa de amostragem (taxa de amostragem de um segundo). Neste caso, o *map-matching* foi utilizado apenas na preparação prévia desse subconjunto, ou seja, a rede viária não foi utilizada durante os experimentos. Também foi usado neste trabalho outro conjunto de dados reais, mas com uma alta taxa de amostragem original de um segundo a partir de 850 táxis rastreados na cidade de Fortaleza, Ceará. Desses, cerca de 730 táxis se movem ao mesmo tempo. Este segundo conjunto de dados é pequeno comparado ao T-DRIVE, mas possui uma grande quantidade de dados brutos, que são usados diretamente nos experimentos sem interpolação ou mapeamento. Nos experimentos, o fluxo de dados foi simulado a partir desses dois conjuntos de dados.

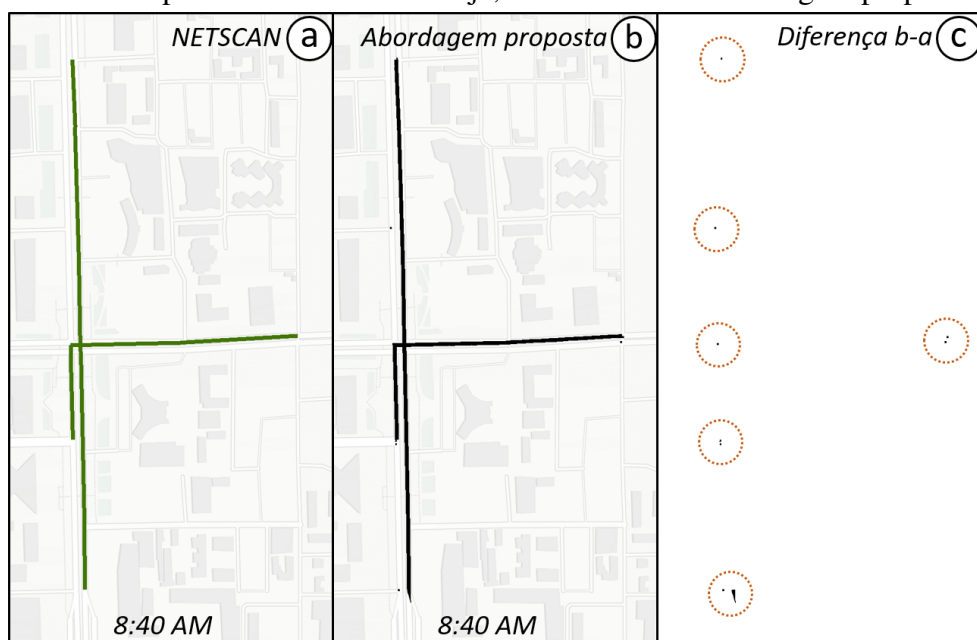
Experimento 1

O primeiro experimento foi criado para comparar a eficácia da abordagem proposta com o NETSCAN (KHARRAT *et al.*, 2008), um algoritmo tradicional de clusterização de trajetórias baseado em densidade para o movimento restrito em uma rede viária. Uma abordagem baseada em movimento restrito foi escolhida como referência, em vez de uma baseada em movimento livre no espaço euclidiano, para reduzir os erros de distância e similaridade entre as trajetórias. Neste experimento, o T-Drive foi utilizado, e ambas as abordagens foram configuradas

para retornar as rotas mais quentes nas quais o número de táxis que passam por elas está acima de um limiar de densidade de 10 objetos em 10 minutos. Além desses dois parâmetros, o NETSCAN também usa outro parâmetro ε , que é a diferença máxima de densidade entre segmentos vizinhos de vias. Para este experimento, o $\varepsilon = 10$, ou seja, o mesmo valor do limiar de densidade. Na abordagem proposta, a função de decaimento foi ignorada e a etapa de Processamento Visual utiliza o tipo *filtro*= 2, ou seja, apenas as células quentes (*fg.a*= 1) são retornadas. Assim, espera-se que ambas as abordagens retornem os segmentos de ruas percorridos por pelo menos 10 táxis em 10 minutos.

A Figura 32 mostra uma pequena região do mapa no nível de zoom 15 como exemplo para comparar os resultados das duas abordagens. O subconjunto T-Drive usado neste exemplo foi do horário de maior fluxo de uma manhã, das 8h30 às 8h40. As rotas quentes encontradas pelo NETSCAN e apresentadas na Figura 32a são as mesmas rotas quentes descobertas pela abordagem proposta apresentadas na Figura 32b. No entanto, se a diferença de imagem for calculada entre a Figura 32b e a Figura 32a, é possível ver ruído de imagem, o qual está destacado pelos círculos laranjas pontilhados na Figura 32c. Todavia, essa diferença não é significativa, especialmente considerando regiões maiores no mapa. Este exemplo representa

Figura 32 – Rotas quentes encontradas pela abordagem de referência NETSCAN (a) e a abordagem proposta (b). Este exemplo utilizou um subconjunto do T-Drive no horário de maior movimento de uma manhã, das 8h30 às 8h40, com um limiar de densidade de 10 táxis. (c) A diferença de imagem (b - a) mostra um pequeno ruído, destacado por círculos pontilhados na cor laranja, no resultado da abordagem proposta.



Fonte: o autor.

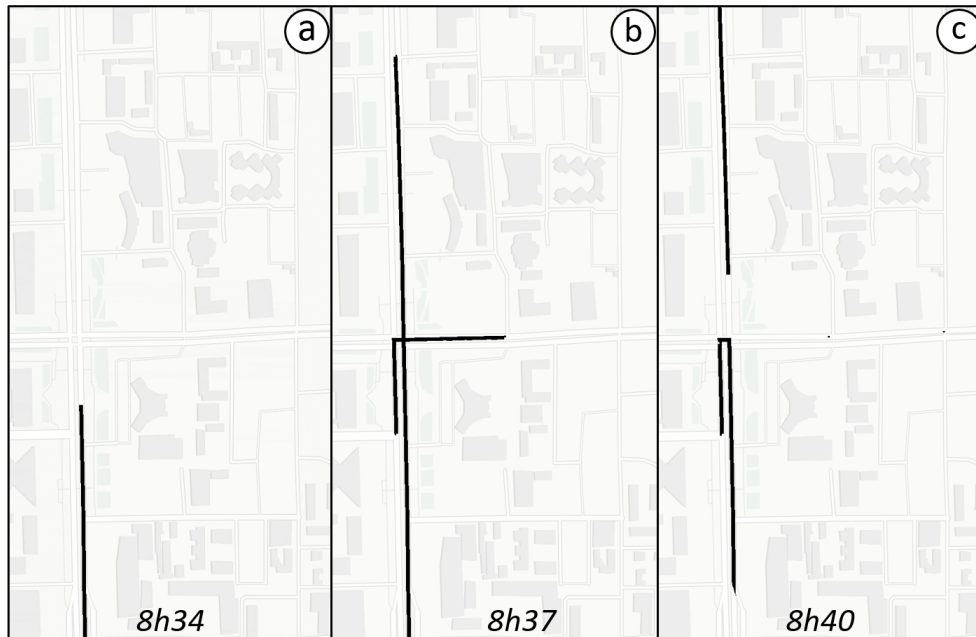
uma pequena região e duas grandes rotas quentes, e a diferença é apenas de 1,68% calculada pelo número de *pixels* remanescentes (Figura 32c) dividido pelo total de *pixels* verdes encontrados pelo NETSCAN (Figura 32a). Essa diferença pode ser explicada pela abordagem proposta ser baseada em grade, diferentemente da abordagem baseada em vetores do NETSCAN. A abordagem proposta não usa as informações da rede viária e, às vezes, nas interseções e limites de segmentos de vias, existem locais onde o limiar de densidade também é alcançado. Por exemplo, quando duas avenidas de quatro faixas se cruzam, mesmo que nenhuma das faixas seja quente, o método proposto pode marcar células quentes nas interseções das faixas. Isso também explica as diferenças na Figura 32c. Conceitualmente, essas células apresentam tráfego pesado desde que alcançaram o limiar de densidade. No entanto, a abordagem baseada em vetores do NETSCAN não identifica essas interseções.

Experimento 2

Neste experimento, a função de decaimento é aplicada para visualizar a evolução das rotas quentes. Ao contrário do NETSCAN, que retorna o resultado apenas no final de uma janela de tempo, a abordagem proposta apresenta continuamente as mudanças do resultado por meio de uma animação. Usando o método proposto com a função de decaimento, é possível perceber padrões que não são encontrados pelos métodos clássicos. Por exemplo, no cenário apresentado na Figura 2, depois que o bloqueio da rua é removido, a rota quente se comporta como uma onda de tráfego (TREIBER; KESTING, 2013) se movendo da direita para a esquerda até desaparecer.

Para este experimento, a mesma configuração do experimento anterior foi empregada: o mesmo subconjunto do T-Drive, a mesma região do mapa, o limiar de densidade de 10 objetos e o mesmo período de 10 minutos. A Figura 33 mostra o resultado da abordagem proposta em três momentos distintos 8h34, 8h37 e 8h40 para mostrar a evolução das rotas quentes. Analisando a animação como um todo, é possível perceber que uma rota quente aparece na parte inferior do mapa e está se movendo para cima como uma onda de tráfego (Figure 33b). Também é possível notar que existe uma outra rota quente que persiste por causa de um retorno na via (Figura 33c parte inferior). Comparando o resultado obtido neste experimento às 8h40 (Figura 33c) com o obtido pelo NETSCAN (Figura 32a), é possível observar que eles são bem distintos. Enquanto o resultado do NETSCAN representa todo o movimento agregado em 10 minutos, o resultado da abordagem proposta é mais consistente com a dinâmica do movimento dos objetos que está acontecendo às 8h40.

Figura 33 – O resultado da abordagem proposta usando a função de decaimento em três instantes distintos 8h34 (a), 8h37 (b) e 8h40 (c) para mostrar a evolução das rotas quentes.



Fonte: o autor.

Experimento 3

A velocidade é um atributo chave comumente utilizado na literatura para identificar engarrafamentos e gargalos de tráfego (POCO *et al.*, 2015; NI, 2015a; WANG *et al.*, 2013; TREIBER; KESTING, 2013). Este experimento visa identificar possíveis gargalos de tráfego usando a velocidade dos objetos na formação de rotas quentes. Uma rota quente é um bom indicador de um gargalo de tráfego, pois todos os gargalos de tráfego são rotas quentes, embora o oposto não seja verdadeiro. Por exemplo, uma via expressa com um alto fluxo de veículos pode ser considerada uma rota quente, mas certamente não é um engarrafamento.

Neste sentido, a abordagem proposta foi adaptada para lidar com esse problema. A ideia dessa abordagem alternativa baseia-se na teoria da transferência de calor (LIENHARD, 2013). De maneira similar, uma estrada se comporta como uma superfície termossensível, a qual recebe calor de objetos quentes que a tocam. A quantidade de calor transferida depende do tempo em que o objeto e a superfície permanecem em contato, ou seja, está diretamente relacionado à velocidade do objeto. Seguindo essa analogia, se a velocidade de um objeto é alta, seu peso de contribuição para a formação de uma rota quente é mínimo e vice-versa. Além disso, a função de decaimento representa o quanto a estrada perde calor.

Embora os resultados obtidos sejam bastante diferentes da abordagem original, a adaptação da abordagem proposta é muito simples. A instrução na linha 4 do Algoritmo 2 foi

Algoritmo 4: incrementarRotas(O_a)

Entrada:o conjunto de objetos em movimento O_a **Retorno :**o *buffer* ativo B_a

```

1  $B_a \leftarrow \emptyset$ 
2 para todo  $o \in O_a$  faça
3    $[lat, lng] \leftarrow encontrarLocal(o)$ 
4    $o_w \leftarrow speed/speedLimit$ 
5    $t \leftarrow criarSegmentoTrajetria([lat, lng], o, o_w)$ 
6    $B_a \leftarrow B_a + t$ 
7 fim
8 retorne  $B_a$ 

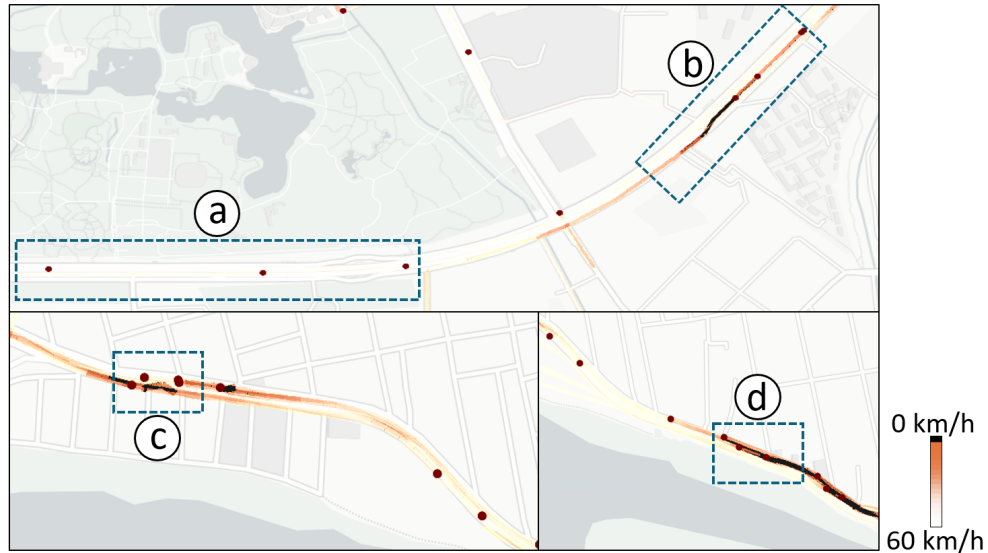
```

alterada para $o_w \leftarrow speed/speedLimit$, em que *speed* é a velocidade do objeto no instante atual t_k e *speedLimit* é a velocidade máxima permitida na via onde o objeto o está. O algoritmo modificado é apresentado no Algoritmo 4. Essa razão o_w representa a temperatura do objeto o e é usada como opacidade na função *criarSegmentoTrajetória*. Portanto, um fluxo constante de objetos com velocidade significativamente menor do que o limite de velocidade da via pode indicar um gargalo de tráfego. Por exemplo, em uma via expressa, os carros que se movem em baixa velocidade, mesmo que em pequeno número, indicarão um gargalo de tráfego. Por outro lado, um grande número de carros que se deslocam perto do limite de velocidade da estrada não constitui necessariamente uma rota quente nesta nova abordagem.

Assim, ao empregar a abordagem alternativa, este experimento utilizou os dois conjuntos de dados para identificar possíveis gargalos de tráfego. Para isso, um único limite de velocidade de 60 km/h foi definido para todo o mapa, mas o limite real de cada via poderia ser extraído de uma rede viária. A Figura 34 mostra três exemplos de gargalos de tráfego. Analisando a animação do primeiro exemplo usando o T-Drive, é possível perceber na mesma via expressa carros mais rápidos (Figura 34a) sem rastros e carros mais lentos (Figura 34b) formando uma rota quente. A causa desse gargalo é uma alça de saída à frente. O segundo exemplo mostra uma faixa de pedestre (Figura 34c), prejudicando o fluxo de tráfego em uma avenida. Por último, a Figura 34d apresenta um gargalo que apareceu repentinamente durante a análise. A causa desse último gargalo não foi identificada. O conjunto de dados de Fortaleza foi usado nos dois últimos exemplos.

Logo, o problema representado pelo cenário da Figura 1 pode ser solucionado empregando esta abordagem. Os veículos que reduziram sua velocidade para cruzar a rua deram

Figura 34 – Três exemplos de gargalos de tráfego. No primeiro exemplo, é possível perceber na mesma via expressa carros mais rápidos (a) sem rastros e carros mais lentos (b) formando uma rota quente. (c) O segundo exemplo mostra uma faixa de pedestre prejudicando o fluxo de tráfego em uma avenida. (d) Um gargalo de origem desconhecida apareceu repentinamente durante a análise.



Fonte: o autor.

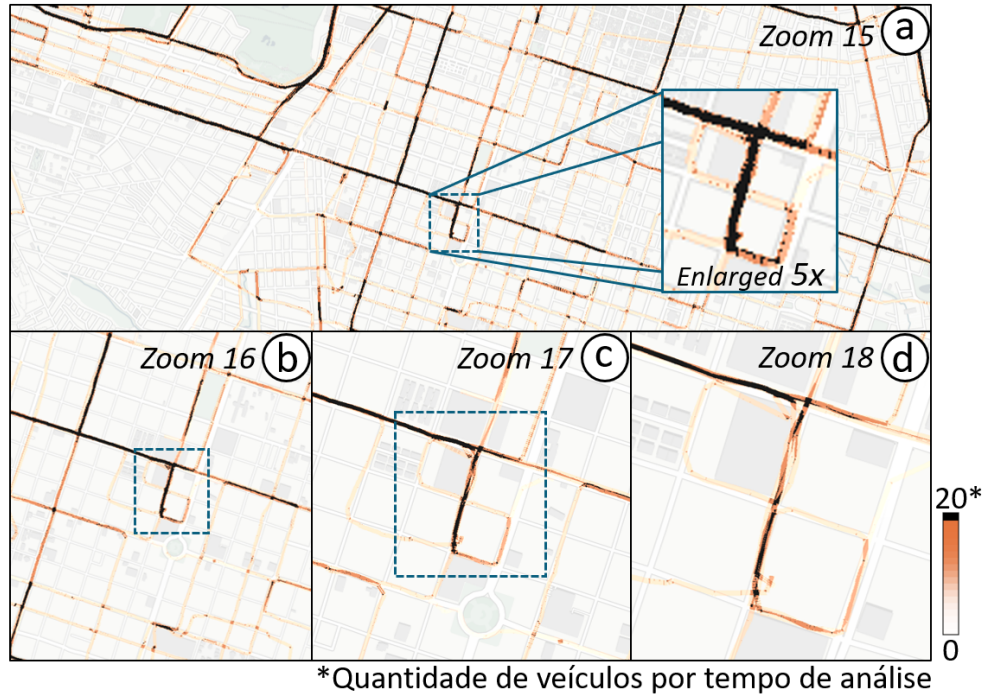
maior contribuição para a formação da rota quente.

Experimento 4

Este experimento foi realizado para verificar a eficácia da abordagem proposta ao usar um conjunto de dados brutos em diferentes níveis de zoom. Os dados coletados por dispositivos GPS geralmente contêm erros. Alguns deles são facilmente detectados e eliminados diretamente no dispositivo. No entanto, pequenos erros podem prejudicar a qualidade das rotas quentes, por exemplo, posições de veículos capturadas em calçadas de pedestre. Neste experimento, foi utilizado o conjunto de dados de Fortaleza com posições de táxis registradas a cada segundo. O objetivo é verificar a qualidade das rotas quentes em diferentes níveis de zoom. Em uma aplicação de monitoramento, o foco principal é a área visualizada nos níveis de zoom de cidade, ou seja, os níveis de 12 a 18.

A Figura 35 mostra um exemplo do impacto do nível de zoom na qualidade das rotas quentes. As Figuras 35b a 35d descrevem a mesma região no mapa (o retângulo pontilhado na Figura 35a). No nível de zoom 15, a Figura 35a apresenta rotas quentes de boa qualidade pois os erros de GPS foram minimizados. No entanto, no nível de zoom 18, esses erros foram significativos e prejudicaram a formação das rotas quentes. Essas diferenças no resultado ocorrem devido à variação do tamanho da área representada por uma célula no *buffer* em

Figura 35 – O impacto de diferentes níveis de zoom na qualidade das rotas quentes. (a) Rotas quentes de boa qualidade com erros de GPS minimizados no nível de zoom 15. Uma região no mapa (o retângulo pontilhado) é representada nos níveis de zoom 16 (b), 17 (c) e 18 (d) para comparar os resultados.



Fonte: o autor.

diferentes níveis de zoom. Quanto menor o zoom, maior a área representada por uma célula no *buffer* e, conseqüentemente, mais dados são agregados nessa célula.

Concluindo, os resultados obtidos foram excelentes usando dados brutos a partir do nível de zoom 12 – com o qual é possível visualizar completamente uma grande cidade – até o nível de zoom 16, que é equivalente a um pequeno bairro. A abordagem proposta é adequada para um conjunto de dados com alta taxa de amostragem. Quanto menor a taxa de amostragem, mais difícil será a formação de rotas quentes e maior será o nível de zoom necessário para minimizar o problema. Além disso, quanto maior o número de objetos ativos, melhor a qualidade das rotas quentes, pois a temperatura máxima individual será baixa e, portanto, os erros são minimizados. Por exemplo, em uma aplicação com um grande número de objetos ativos, se o limiar de densidade $\tau = 100$, a temperatura individual inicial do objeto $T_i^{max} = 0,01$ – pela Eq. (4.7). Assim, essa pequena temperatura minimiza a percepção de possíveis erros de trajetória.

Experimento 5

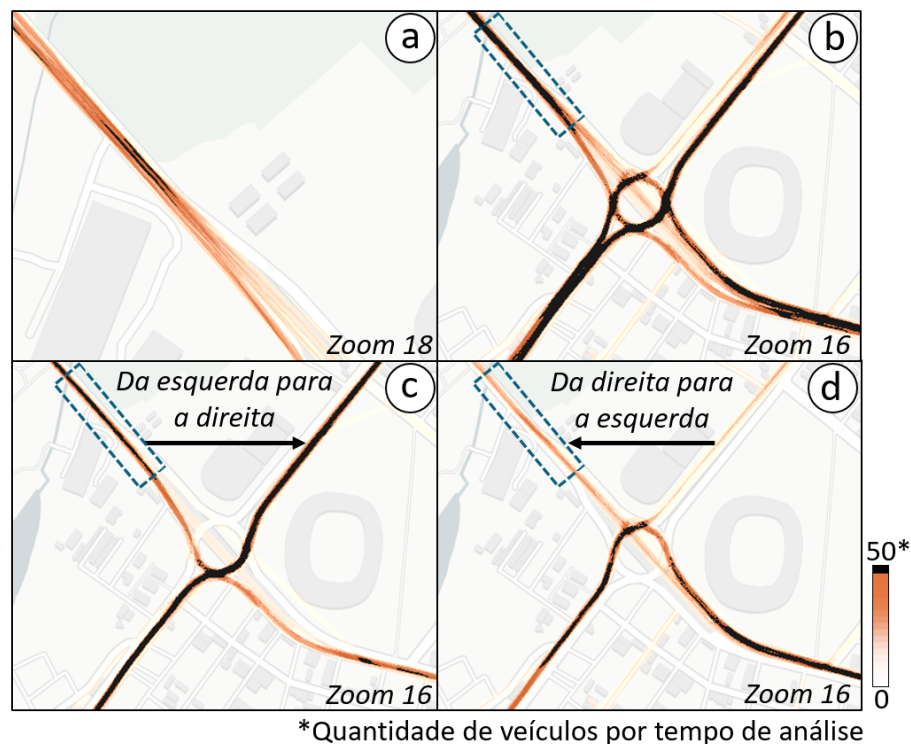
Como discutido no experimento anterior, a abordagem proposta agrega segmentos de trajetória nos níveis de zoom mais baixos. Este fato pode impactar negativamente a identificação

correta de rotas quentes. A Figura 36a ilustra uma avenida com oito faixas, quatro em cada sentido. Nenhuma delas é uma rota quente no nível de zoom 18. No entanto, no nível de zoom 16, como mostrado no retângulo pontilhado da Figura 36b, as oito faixas foram agregadas incorretamente em uma rota quente. Agregar faixas que têm o mesmo sentido de fluxo não é um problema; pelo contrário, uma rota deve considerar todas as faixas das vias que a compõem. O problema ocorre quando faixas de sentidos opostos são agregadas.

Assim, neste experimento, um método para filtrar o sentido do movimento dos veículos foi desenvolvido. A partir de um sentido definido pelo usuário, apenas os segmentos de linha com esse mesmo sentido são desenhados. O resultado também pode ser decomposto em duas camadas, uma para cada sentido da via, podendo ser visualizadas individualmente. A Figura 36c mostra o tráfego fluindo da esquerda para a direita. Neste caso, as quatro pistas da avenida foram agregadas corretamente nesse sentido de via, encontrando uma rota quente (o retângulo pontilhado da Figura 36c). Por outro lado, as pistas no sentido oposto não formaram uma rota quente (o retângulo pontilhado da Figura 36d).

Concluindo, a abordagem proposta pode ser usada para descobrir corretamente rotas

Figura 36 – (a) Uma avenida com oito faixas, quatro em cada sentido. Nenhuma delas é uma rota quente no nível de zoom 18. (b) As oito faixas dessa avenida (retângulo pontilhado) foram incorretamente agregadas em uma rota quente no nível de zoom 16. (c) O tráfego flui da esquerda para a direita. A avenida é quente nesse sentido (retângulo pontilhado). (d) O tráfego flui no sentido oposto e a avenida não é quente.



quentes. Mesmo sem usar o filtro de sentido do movimento, mas com o auxílio da visualização com a animação, com a representação individual dos veículos e com a ferramenta de zoom, é possível identificar o sentido das rotas quentes. Além disso, com a ajuda do filtro de sentido e outras ferramentas de consulta de dados, a identificação de rotas quentes torna-se ainda mais precisa.

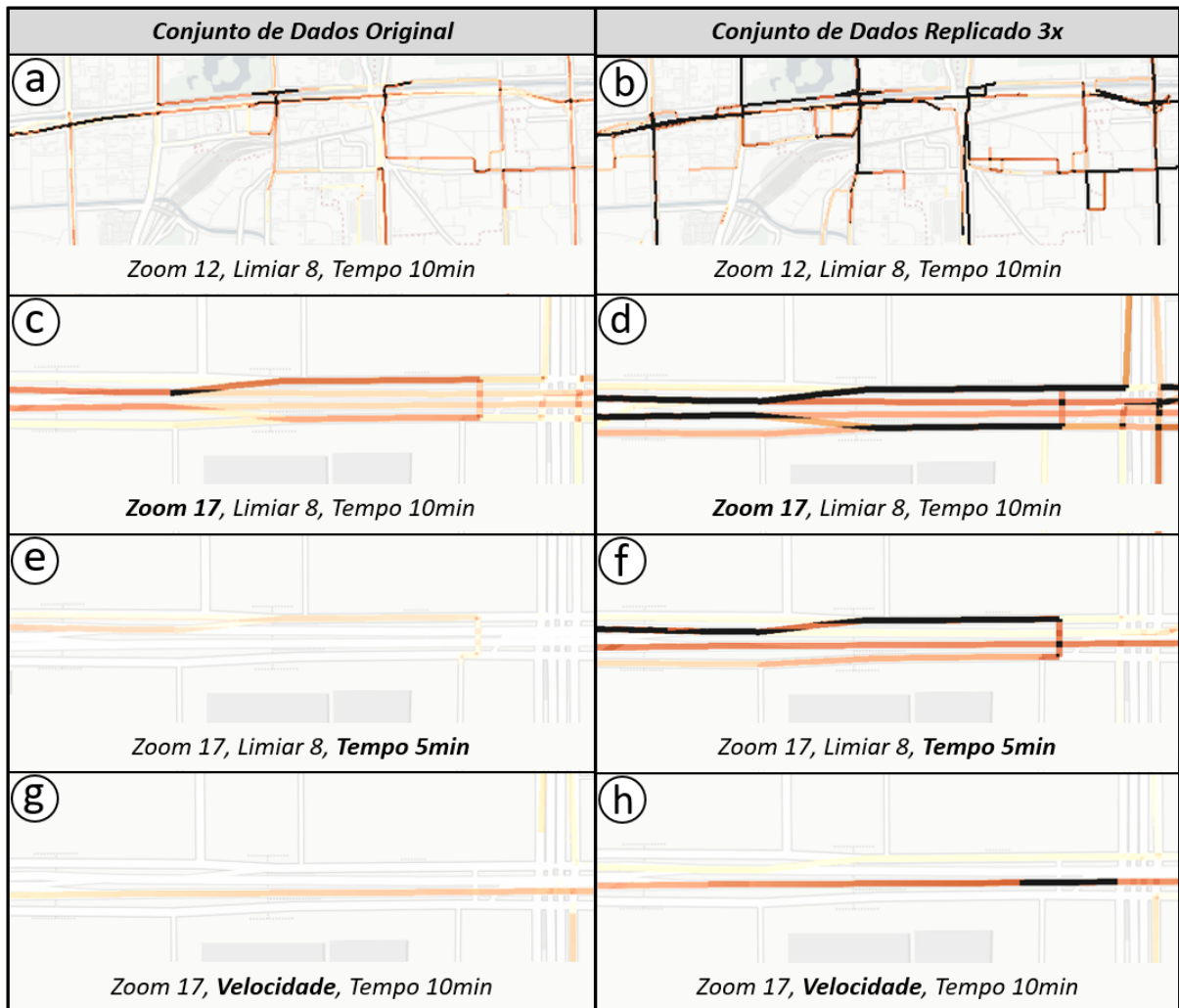
Experimento 6

Neste experimento, os resultados de diferentes valores de parâmetros são comparados e discutidos. Como mencionado anteriormente, o método proposto requer basicamente dois parâmetros: o limiar de densidade τ e o tempo de análise representado por n_t . Além desses parâmetros, há também outras variáveis indiretas: o tamanho do conjunto de dados, o nível de zoom e o tamanho da *viewport*. Como o tamanho da *viewport* não afeta a representação das rotas quentes, ou seja, ele limita apenas a área de visualização, seu efeito não é discutido aqui. No entanto, o tamanho da *viewport* tem impacto direto no desempenho, sendo discutido especificamente no Capítulo 6, Análise de Desempenho.

Dois conjuntos de dados do T-Drive foram utilizados neste experimento: o original e uma réplica. A réplica foi gerada usando três cópias do conjunto de dados original. O T-Drive original mantém um fluxo médio de cerca de 1.500 veículos ativos durante o intervalo de tempo, conforme mostrado na Figura 46a. Assim, a réplica desse conjunto de dados mantém uma média de cerca de 4.500 veículos ativos. Além disso, o fluxo de tráfego foi aumentado proporcionalmente na réplica, ou seja, o fluxo nas vias é três vezes maior. Além de experimentar diferentes valores de parâmetros, os efeitos causados pelo tamanho do conjunto de dados também foram analisados.

A Figura 37a mostra rotas quentes nas vias de Pequim no nível de zoom 12, com o limiar de densidade de 8 veículos, e o tempo de análise de 10 minutos. Como a réplica do T-Drive tinha uma quantidade três vezes maior de objetos ativos, as rotas quentes foram mais proeminentes no mapa empregando os mesmos parâmetros (Figura 37b). Neste caso, para obter rotas quentes mais significativas, o limiar poderia ser aumentado. Por outro lado, em outra situação, como mostrado na Figura 37d, usando os mesmos parâmetros de limiar e de tempo, mas mudando o nível de zoom para 17, foi possível visualizar rotas quentes na réplica do T-Drive enquanto que quase não havia rotas quentes no original (Figura 37c). Nesse caso, o limiar do conjunto de dados original deve ser reduzido para visualizar rotas mais quentes. Em

Figura 37 – Comparação dos efeitos de diferentes valores de parâmetros. Dois conjuntos de dados do T-Drive foram utilizados: o original na coluna da esquerda e uma réplica na coluna da direita. A réplica contém três cópias do conjunto de dados original. (a) Algumas rotas quentes em Pequim. (b) A maioria das estradas está marcada como quente para os mesmos parâmetros. (c) Alterando o nível de zoom para 17. (d) A réplica mostrou rotas mais quentes. (e) Com cinco minutos de tempo de análise, nenhuma via foi marcada como quente. (f) Para o mesmo período, a réplica apresentou rotas quentes bem definidas. (g) A abordagem alternativa baseada em velocidade foi usada. (h) Os resultados obtidos foram muito semelhantes, mesmo com um conjunto de dados três vezes maior.



Fonte: o autor.

seguida, o tempo de análise foi reduzido para cinco minutos, ou seja, a temperatura decaiu mais rapidamente. Na Figura 37e, não existiam rotas quentes no conjunto de dados original, mas ainda havia rotas quentes bem definidas no conjunto de dados maior (Figura 37f).

Nas Figuras 37g e 37h, a alternativa baseada em velocidade foi usada com um tempo de análise de 10 minutos. Relembrando que nessa abordagem o limiar de densidade não é utilizado. As faixas representadas nessas figuras eram de uma via expressa e o tráfego fluía bem,

então nenhuma rota foi marcada como quente no conjunto de dados original (Figura 37g). No caso da réplica, apenas um segmento de pista foi marcado como quente (Figura 37h) pois havia mais veículos com baixa velocidade marcando esse segmento. No entanto, os resultados obtidos por essa abordagem em ambos os conjuntos de dados foram muito semelhantes, especialmente em contraste com os resultados obtidos usando o limiar de densidade (Figuras 37c e 37d), que foram bastante diferentes. Para a abordagem alternativa, a velocidades dos veículos é mais importante do que a quantidade, como discutido anteriormente.

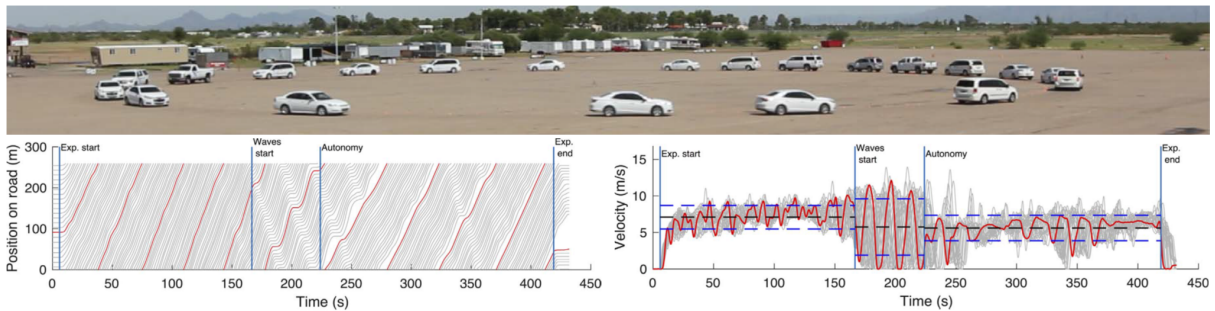
5.2 Visualizando Ondas de Tráfego

Analisar os padrões de mobilidade local é essencial para identificar como os comportamentos dos indivíduos afetam diretamente o fluxo de tráfego. Frequentemente, um fluxo de tráfego se comporta como uma solução supersaturada, na qual pequenos distúrbios, como um carro mudando de faixa, podem causar uma reação em cadeia. Essas perturbações podem causar ondas de tráfego que podem se tornar engarrafamentos. A onda de tráfego, também conhecida como tráfego *stop-and-go* ou *phantom intersection*, é um dos principais tópicos na teoria da dinâmica do fluxo de tráfego (STERN *et al.*, 2018; JIANG *et al.*, 2017; NI, 2015a; TREIBER; KESTING, 2013).

A dinâmica do fluxo de tráfego apresenta instabilidades à medida que a densidade de veículos aumenta. Neste caso, pequenas perturbações amplificam-se e crescem em ondas de tráfego que viajam no sentido oposto ao movimento (TREIBER; KESTING, 2013). Esse fenômeno é observável através do envolvimento de veículos em ciclos repetidos de aceleração e desaceleração, impactando a eficiência e sustentabilidade do tráfego de uma grande cidade. Pesquisadores realizaram inúmeros estudos investigando a oscilação do tráfego com observações empíricas e modelos teóricos, tentando entender a propagação das ondas de tráfego (ver o trabalho de Li *et al.* (2014) para uma visão geral). Nesses estudos, as principais causas das ondas de tráfego foram identificadas: semáforos, cruzamentos, rampas de acesso, mudanças de faixa, bloqueios, aclives, presença de caminhões e acidentes. No entanto, a maioria desses estudos envolve experimentos numéricos ou simulações que investigam lugares específicos, como um cruzamento ou uma seção de via. Por exemplo, recentemente Stern *et al.* (2018) realizaram um experimento para investigar a influência dos carros autônomos na redução de ondas de tráfego. A Figura 38 mostra o local do experimento com 21 veículos e os gráficos utilizados no estudo.

As descobertas desses trabalhos não podem ser facilmente aplicadas a grandes vo-

Figura 38 – Exemplo de um experimento para investigar a influência dos carros autônomos na redução de ondas de tráfego. A parte superior mostra o local do experimento, e na inferior, os gráficos utilizados no estudo.



Fonte: (STERN *et al.*, 2018)

lumes de dados de GPS de uma cidade inteira em uma abordagem computacional sistemática. Considerando ainda o monitoramento em tempo real, essa tarefa torna ainda mais difícil. Neste sentido, a análise visual pode ajudar combinando análise automatizada com visualização interativa na exploração de grandes volumes de dados. No entanto, a maioria das abordagens de análise visual existentes não aborda adequadamente a dimensão do tempo, pois se concentra apenas na visualização espacial ou usa alguma forma de simplificação da dimensão temporal. O espaço e o tempo são inseparáveis na análise visual da dinâmica de fluxo de tráfego. As singularidades em qualquer uma dessas dimensões não devem ser descartadas, pois podem revelar relacionamentos implícitos na dinâmica do movimento (ANDRIENKO *et al.*, 2010). Assim, para visualizar padrões de oscilação de tráfego, a dimensão do tempo é essencial. Além disso, codificar a informação de velocidade na visualização tem grande influência na percepção de ondas de tráfego. Neste sentido, os recursos da abordagem proposta para codificar tempo e velocidade são discutidos na próxima seção.

5.2.1 Codificando Tempo e Velocidade

Normalmente, uma trajetória representada em um mapa mostra a rota percorrida, mas omite informações contextuais que são dependentes da dimensão do tempo, como a velocidade de deslocamento, a velocidade permitida, os eventos anormais ou o tempo gasto para percorrer a rota. A literatura está repleta de exemplos que codificam informações sobre tempo e velocidade em representações de trajetórias (ver o trabalho de Perin *et al.* (2018) para uma visão geral). Pesquisadores e *designers* codificaram essas informações usando variáveis visuais como tamanho (BACH *et al.*, 2016; AIGNER *et al.*, 2011; BERTIN, 1983), brilho ou cor (BACH *et al.*, 2016; TOMINSKI *et al.*, 2012; SCHEEPENS *et al.*, 2011b; COLLBERG *et al.*, 2003;

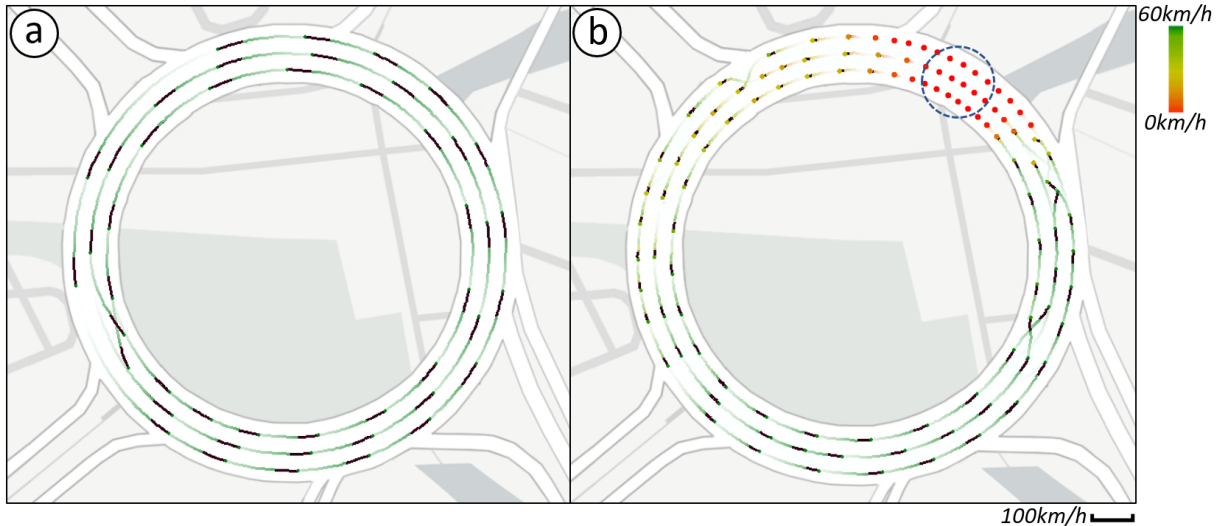
DYKES; MOUNTAIN, 2003), e comprimento de segmento (THUDT *et al.*, 2013; BERTIN, 1983). Perceber o tempo e a velocidade dos objetos é fundamental para compreender a dinâmica do fluxo de tráfego, especialmente as ondas de tráfego.

Como as trajetórias representam objetos em movimento, é natural analisá-las visualizando as características dinâmicas do movimento dos objetos através de uma animação. Essa visualização, por usar o paradigma da animação, promove a real compreensão do movimento coletivo dos objetos e a percepção dos impactos dos comportamentos individuais no fluxo de tráfego. Neste sentido, a animação pode contribuir para a percepção da dimensão tempo. No entanto, existem muitas armadilhas que devem ser consideradas no uso de animação (TVERSKY *et al.*, 2002). Por exemplo, Chevalier *et al.* (2014) apresentam um estudo demonstrando a dificuldade de acompanhar o movimento de um objeto dentro de uma nuvem de pontos sem utilizar recursos visuais adicionais. Esse mesmo problema pode acontecer na visualização de dados de trajetórias. Contudo, ele pode ser resolvido, por exemplo, utilizando a visualização animada de campos vetoriais (BECCARIO, ; HINT.FM,) com *path lines*. Essa técnica é bastante difundida (WANG *et al.*, 2018a; WEISKOPF, 2014) e foi inclusive utilizada para visualizar dados de trajetórias (POCO *et al.*, 2015).

Neste sentido, a visualização proposta nesta seção para analisar ondas de tráfego foi inspirada na visualização animada de campos vetoriais. Assim, o movimento de um veículo deixa um rastro que persiste na via, que é semelhante a uma *path line* na visualização de campo vetorial. No entanto, na visualização proposta, essa trilha desaparece lentamente no decorrer do tempo e seu comprimento representa a magnitude da velocidade do objeto. É importante mencionar que os campos vetoriais não são calculados na visualização proposta, ou seja, os dados das trajetórias são processados para visualizar o movimento dos objetos diretamente em tempo real.

A Figura 39 mostra uma aplicação da abordagem proposta para visualizar ondas de tráfego em uma via circular, que é muito utilizada para estudar esse padrão de mobilidade (STERN *et al.*, 2018; TREIBER; KESTING, 2013). Os dados de trajetórias utilizados nesta visualização foram gerados por um simulador de tráfego (TREIBER,). A ideia da visualização proposta é criar um efeito de rastro para cada objeto, de modo que seja possível perceber a dinâmica do movimento de um objeto pela magnitude de sua velocidade, por sua direção e por sua presença. Uma visão geral animada com vários rastros promove a percepção da evolução do fluxo dos veículos. A Figura 39a mostra um quadro da animação quando o fluxo de veículos era

Figura 39 – Exemplo de aplicação da abordagem proposta para visualizar ondas de tráfego em uma via circular. (a) Um quadro da animação quando o fluxo de veículos é estável. (b) Após um aumento na densidade de veículos, o fluxo se desestabiliza, aparecendo uma onda de tráfego. Os veículos no centro da onda (círculo pontilhado) estavam totalmente parados nesse instante.

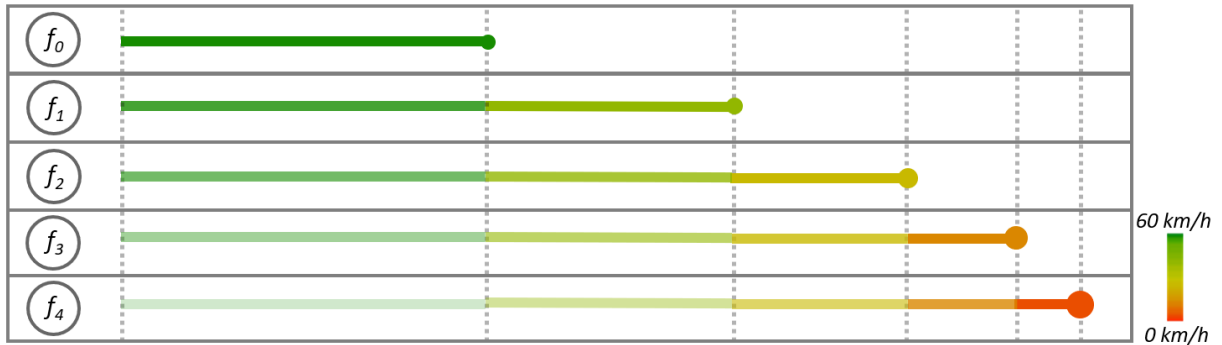


Fonte: o autor.

estável. Nela, é possível perceber que os veículos estão circulando com velocidade próxima a 60km/h pela cor verde do rastro, comparando com a escala de cor no canto direito superior, e pelo tamanho do segmento de cor preta, comparando com a escala de velocidade no canto direito inferior. Após um aumento na densidade de veículos, o fluxo se desestabiliza, aparecendo uma onda de tráfego (Figura 39b). Os veículos que estão no centro da onda (região circular pontilhada) estavam totalmente parados nesse instante. É possível perceber isso pela cor vermelha, pela ausência de rastro e pelo tamanho maior dos marcadores. Neste trabalho, a codificação de cor da matiz vermelha a verde foi empregada por ser muito utilizada em aplicativos de navegação, como o Google Maps e Waze, mas qualquer outra escala de cor pode ser empregada na abordagem proposta. A variação do tamanho dos marcadores com a velocidade busca atender o princípio de animação *squash and stretch* durante o movimento (THOMAS *et al.*, 1995). Portanto, a velocidade é codificada na visualização proposta utilizando variáveis visuais de cor, comprimento de rastro, tamanho de marcador e brilho (esmaecimento do rastro). Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/uycpCRSPyLo>>.

Para visualizar ondas de tráfego, a abordagem proposta nesta tese deve seguir algumas recomendações. Primeiramente, o nível de zoom deve ser suficientemente alto para identificar todas as faixas das vias para não haver agregação de rastros. Já que o foco é identificar padrões de mobilidade local, os níveis de zoom 16, 17 e 18 são os mais indicados. A segunda

Figura 40 – Renderização do rastro de um objeto, quadro a quadro f_i . Neste exemplo, o objeto estava a 60 km/h no quadro f_0 , depois desacelerou nos quadros f_1 a f_4 .



Fonte: o autor.

recomendação é que o valor do limiar de densidade τ seja igual a 1, isto é, cada objeto atinge o limite de temperatura individual (da Eq. 4.7, $T_i^{max} = \bar{T}$). Como não há agregação de rastros, como acontece na visualização de rotas quentes, o limiar de densidade não tem impacto no resultado desta visualização. Por último, o parâmetro n_t que representa o tempo deve ser suficientemente pequeno para perceber as variações de velocidade. Enquanto que na visualização de rotas quentes esse valor era de alguns minutos, na visualização de ondas de tráfego é de poucos segundos. Nesta visualização, o valor do parâmetro tempo é independente do tamanho do conjunto de dados, ao contrário da visualização de rotas quentes (como visto anteriormente no Experimento 6). Por exemplo, o rastro de um veículo pode ser visível por dois segundos, não importa quantos veículos estejam sendo visualizados.

Para compreender melhor a visualização proposta, a Figura 40 ilustra um exemplo de uma trajetória em linha reta de um objeto em cinco quadros de animação. Em cada quadro f_i , o rastro do objeto é estendido por um segmento de linha pela função *criarSegmentoTrajetória* (linha 5 do Algoritmo 2) usando a opacidade máxima, que representa o limite de temperatura \bar{T} . Essa função também define a cor do segmento de linha através da velocidade atual do veículo no local, da velocidade máxima permitida, e de uma escala de cor. Depois, a opacidade de todos os segmentos de linhas já desenhados é reduzida pelo fator de decaimento (linha 3 do Algoritmo 1). Existe uma relação direta entre o comprimento do rastro e a velocidade do objeto correspondente. Por exemplo, a velocidade do objeto no quadro f_5 é baixa, pois o comprimento do rastro entre f_4 e f_5 é curto. Assim, na Figura 40, é possível perceber uma desaceleração do objeto representado. Além do comprimento, da cor e da opacidade do rastro, a magnitude da velocidade também é codificada no tamanho do marcador (linha 7 do Algoritmo 1).

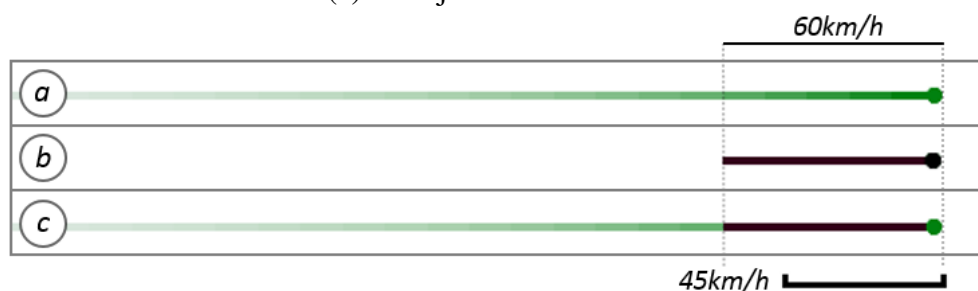
Como foi descrito anteriormente, a etapa de Processamento visual destaca os segmentos de trajetórias que atingem o limite de temperatura \bar{T} além de aplicar um filtro que destaca

os segmentos com alta densidade e suaviza os de baixa densidade. Nesta visualização, a temperatura individual dos veículos está máxima ($T_i^{max} = \bar{T}$), logo, os segmentos que representam o deslocamento recente no instante atual t_k são destacados por essa etapa. Como mostra a Figura 40, o comprimento de um segmento destacado está relacionado com a velocidade do objeto representado no instante atual t_k . Para permitir a percepção da magnitude da velocidade de um objeto, uma escala de referência de velocidade é exibida no mapa (Figura 41). A escala é calculada considerando o fator de decaimento λ , e o nível de zoom do mapa. Além disso, os mesmos filtros utilizados na visualização de rotas quentes podem ser aplicados nesta visualização. A Figura 41a mostra um exemplo de trajetória de um objeto com o rastro suavizado. Na Figura 41b, o rastro deste objeto é cortado para destacar a velocidade atual do objeto. Por último, a Figura 41c mostra uma combinação dos dois aprimoramentos visuais anteriores, portanto, um rastro completo com a magnitude da velocidade realçada é exibido.

Em alguns casos em que a velocidade da animação em *fps* é alta e o tempo de visibilidade do rastro é pequeno, uma tolerância deve ser considerada na temperatura máxima individual dos objetos, isto é, $T_i^{max} - \varepsilon$, para que os segmentos destacados tenham tamanhos significativos. Quanto maior a tolerância ε , maior será o rastro. No Algoritmo 3, o ε deve ser considerado nas linhas 8 e 11, quando se compara a opacidade da cor *fg.a* com o máximo 1, que representa o limite de temperatura \bar{T} . Dessa forma, o valor máximo deve ser $1 - \varepsilon$.

Concluindo, a configuração dos parâmetros da abordagem proposta para visualização de ondas de tráfego é bem mais simples que a visualização de rotas quentes, pois o limiar de densidade é fixo e o parâmetro de tempo praticamente não é alterado durante a visualização. No entanto, esta visualização é muito dependente da densidade do volume de tráfego, ou seja, quanto mais objetos ativos houver, melhor será a percepção das ondas de tráfego. Por exemplo, se a Figura 39 mostrasse apenas 10% dos veículos que estão realmente circulando, não seria fácil

Figura 41 – O resultado da etapa de Processamento Visual aplicada ao rastro de um objeto. (a) O rastro suavizado. (b) O rastro cortado proporcionalmente à velocidade do objeto e com cor modificada. (c) a e b juntos.



Fonte: o autor.

perceber a onda de tráfego. Assim, essa visualização apresenta melhores resultados em aplicações nas quais a grande maioria dos veículos do fluxo de tráfego é rastreada. Por exemplo, dados de trajetórias de veículos capturados por câmeras (MAKRIS *et al.*, 2004) ou drones (AZEVEDO *et al.*, 2014) apresentariam bons resultados, mas dados de táxis não teriam bons resultados, pois a representação deles é pequena no fluxo de tráfego.

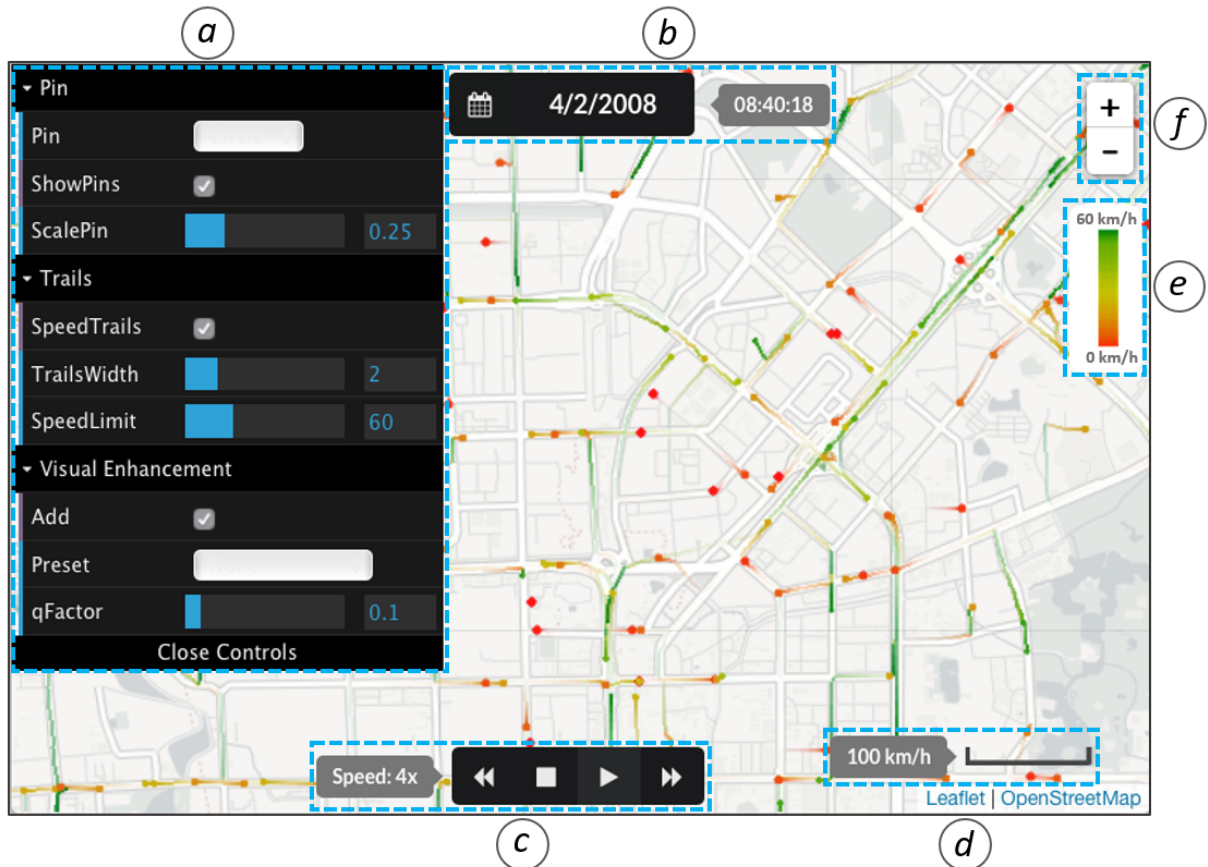
5.2.2 Avaliação por Especialista

Diferentemente da análise de rotas quentes que está relacionada com vários trabalhos incluindo os de clusterização de trajetórias, a visualização de ondas de tráfego não é difundida na literatura. Na realidade, a proposta de visualização desta tese é o primeiro trabalho neste tópico. Pelo fato de a configuração dos parâmetros da abordagem proposta para visualização de ondas de tráfego ser bem mais simples que a visualização de rotas quentes, uma avaliação experimental, como a realizada anteriormente, não seria adequada. Em seu lugar, uma ferramenta de visualização (Figura 42) foi desenvolvida para avaliar a proposta de visualização de ondas de tráfego. Para isso, foi adotada uma metodologia qualitativa de avaliação baseada na observação de campo (POLKINGHORNE, 2005). De acordo com essa metodologia, quatro cenários simulados foram apresentados a um especialista, e foi solicitado a ele que respondesse duas perguntas simples para cada cenário: 1) existe alguma anomalia de tráfego? e 2) qual foi a causa? Além disso, durante a avaliação, foi solicitado que o entrevistado descrevesse a dinâmica do fluxo de tráfego.

Como o foco era a avaliação da percepção das anomalias de tráfego, a ferramenta desenvolvida utilizou dados históricos, ou seja, foi executada de forma *offline*. Neste sentido, foi desenvolvido um controle de tempo (Figura 42c) com o qual era possível visualizar horas de dados de movimento em apenas alguns minutos com recursos de controle de tempo, como pausar, avançar e retroceder em diferentes velocidades. Assim, mesmo quando a visualização é acelerada em altas velocidades, os recursos visuais permitem que o especialista perceba a dinâmica do tráfego em uma visão geral.

Os dados de trajetória para esses cenários foram gerados por um simulador (TREIBER,) que usa três modelos de fluxo de tráfego: o Intelligent-Driver Model para simular as acelerações e desacelerações de frenagem dos motoristas (TREIBER *et al.*, 2000); o modelo MOBIL para permitir que os veículos mudem de faixa de acordo com os critérios de segurança e incentivo (TREIBER; HELBING, 2002); e o modelo Boundary Conditions para explorar

Figura 42 – Captura de tela de uma visualização do fluxo de tráfego de 1.800 veículos em movimento ao mesmo tempo. A imagem mostra as principais funcionalidades da ferramenta desenvolvida, como o (a) painel de personalização, onde o usuário pode definir o tamanho do marcador, o limite de velocidade, a largura do rastro, o filtro de processamento visual e outros parâmetros; (b) a exibição de data e hora; (c) o controle de tempo; (d) a escala da velocidade do rastro; (e) codificação da cor; e controle de zoom (f). Um vídeo mostrando a visualização animada está disponível em <<https://youtu.be/uycpCRSPyLo>>.



Fonte: o autor.

as condições de entrada e saída (HELBING *et al.*, 2002). Nos trabalhos citados, todos esses modelos foram validados com observações empíricas de tráfego.

Os cenários foram simulados para esta avaliação por duas razões: iniciar propositalmente distúrbios no fluxo de tráfego e ter um ambiente controlado sem ambiguidades. Assim, seria possível avaliar com certeza se o especialista percebe quando, onde e por que a anomalia ocorreu. Por exemplo, em um lugar e hora específicos, seria possível fazer um veículo frear abruptamente ou mudar de faixa, provocando uma reação em cadeia, ou seja, uma onda de tráfego. Essas simulações são amplamente utilizadas na Engenharia de Tráfego para investigar os efeitos de novas políticas antes de colocá-las em operação.

Para a avaliação, foi selecionado um especialista que possui mestrado na área de

Transporte e doutorado na área de Planejamento Urbano. Atualmente, é professor associado de uma universidade e ministra uma disciplina de Planejamento Urbano. Além de sua experiência acadêmica, ele também tem mais de dez anos de experiência em projetos de intervenção urbana, como túneis e viadutos, para melhorar o fluxo de tráfego de uma grande cidade.

No início da avaliação, o especialista usou livremente a ferramenta de visualização para se familiarizar com suas funcionalidades. Naquele momento, dados de trajetória real de duas grandes cidades foram utilizados. Depois disso, quatro cenários simulados foram apresentados a ele: via circular, via de acesso, bloqueio de pista e active.

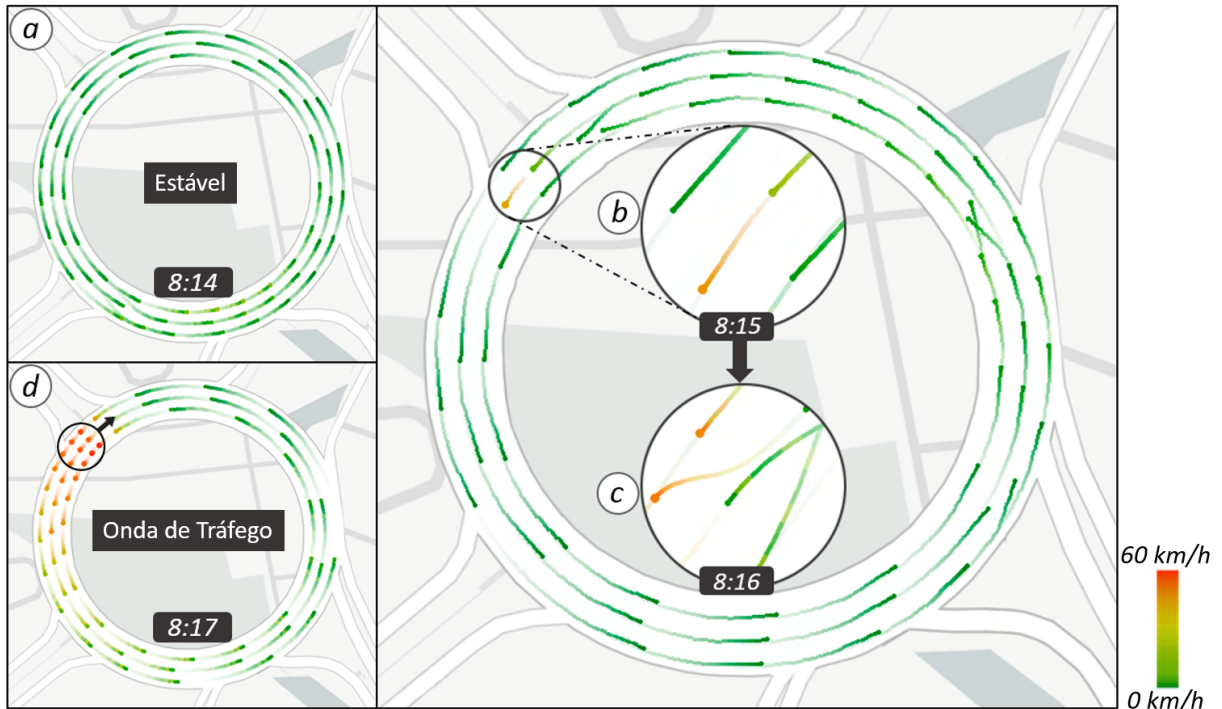
Via Circular

A via circular é uma simulação de tráfego com três faixas em um sistema fechado. Este é o cenário mais simples, pois depende principalmente do limite de velocidade e da densidade de veículos na via. Os gestores urbanos costumam usá-lo para testar o limite de velocidade ideal para uma determinada densidade de veículos (NI, 2015b). Além disso, este cenário é muito eficaz para analisar os efeitos e o comportamento das ondas de tráfego (STERN *et al.*, 2018; TREIBER; KESTING, 2013).

Embora a simulação criada para este cenário tivesse uma duração de duas horas, o tempo de exibição poderia ser reduzido acelerando a simulação para ser visualizada em poucos minutos. A primeira hora retratava um tráfego denso e estável sem congestionamentos com 20 veículos/km/pista e um limite de velocidade de 60 km/h (Figura 43a). No início da segunda hora, um carro freou abruptamente (Figura 43b). Então, o veículo logo atrás teve que frear também para manter uma distância de segurança. Como consequência, o próximo veículo atrás teve que frear ainda mais e assim por diante. Esse distúrbio desestabilizou o fluxo de tráfego de todas as faixas causando uma onda de tráfego (Figura 43d).

No teste, o especialista identificou prontamente o local e o carro que causou a perturbação. Sua avaliação consistiu principalmente em observar as mudanças no comprimento do rastro, na cor e no tamanho do marcador (Figura 43b). Depois de identificar o local da perturbação, o especialista retrocedeu a visualização para iniciar uma exploração mais detalhada aumentando também o nível de zoom do mapa, reduzindo a velocidade da animação e alterando o filtro de Processamento Visual para estimar a velocidade do carro antes e depois da freada abrupta. Ele disse que a propagação da perturbação para as outras faixas ocorre por causa das mudanças de faixa dos veículos que seguem logo atrás (Figura 43c). Ele também percebeu

Figura 43 – O cenário de uma via circular. (a) Tráfego denso e estável sem congestionamentos. Os veículos estão se movendo a uma velocidade próxima ao limite. (b) Um carro freia abruptamente. (c) Veículos que seguem logo atrás desse carro mudam de faixa, propagando a perturbação para as outras faixas. (d) Então, uma onda de tráfego é formada e se move no sentido oposto ao fluxo.



Fonte: o autor.

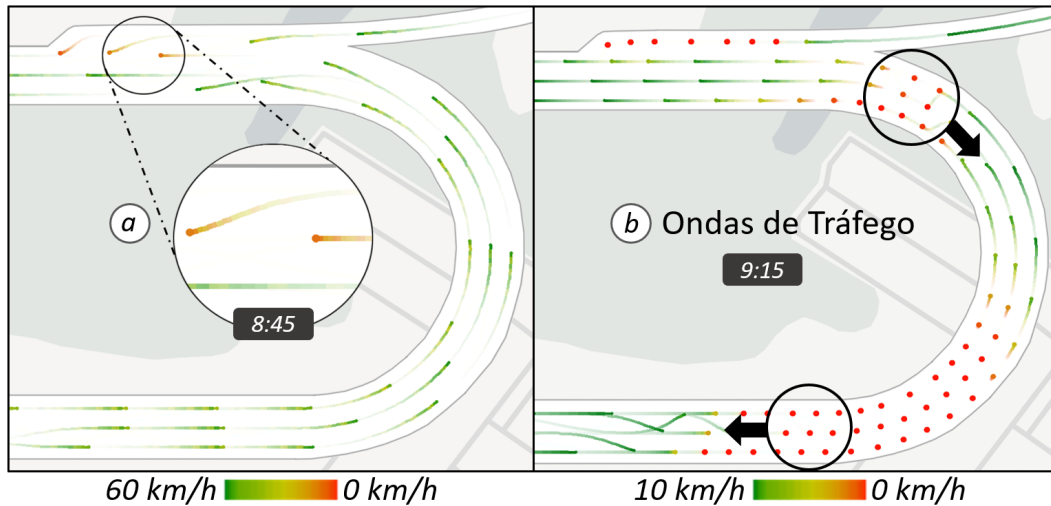
que o distúrbio estava desaparecendo nessa simulação, ou seja, o fluxo de tráfego retornaria ao equilíbrio inicial, embora ele ainda não tivesse visto a visualização até o fim.

Via de Acesso

Este cenário representa um sistema aberto que simula possíveis gargalos de tráfego causados por uma via de acesso (LEE *et al.*, 1999). Para isso, foram simuladas algumas situações em que veículos entram em uma via já ocupada, por exemplo, uma saída de estacionamento de um centro comercial para uma avenida movimentada. Este cenário é mais complicado que o anterior, pois é preciso considerar o fluxo de tráfego da via principal e o da via de acesso.

Este teste foi realizado a partir de uma simulação de duas horas de duração para avaliar os impactos de veículos acessando uma via em condições de tráfego intenso, mas com entrada e saída constantes de 3.600 veículos/h com limite de velocidade de 80 km/h. Após 30 minutos do início da simulação, foi adicionado um fluxo de um via de acesso com 20% do fluxo da via principal, ou seja, 720 veículos/h. Instantaneamente, ondas de tráfego surgiram na via principal com origem perto da região da via de acesso. Após uma hora de simulação, o fluxo da

Figura 44 – O cenário de uma via de acesso. (a) Os veículos acessam a via principal e provocam uma onda de tráfego. (b) A via de acesso gerou duas ondas de tráfego. Neste caso, a escala de cor foi configurada para representar um intervalo de velocidade menor de 0 a 10 km/h para realçar as diferenças de velocidade.



Fonte: o autor.

via principal foi reduzido para menos da metade do fluxo inicial (1.600 veículos/h).

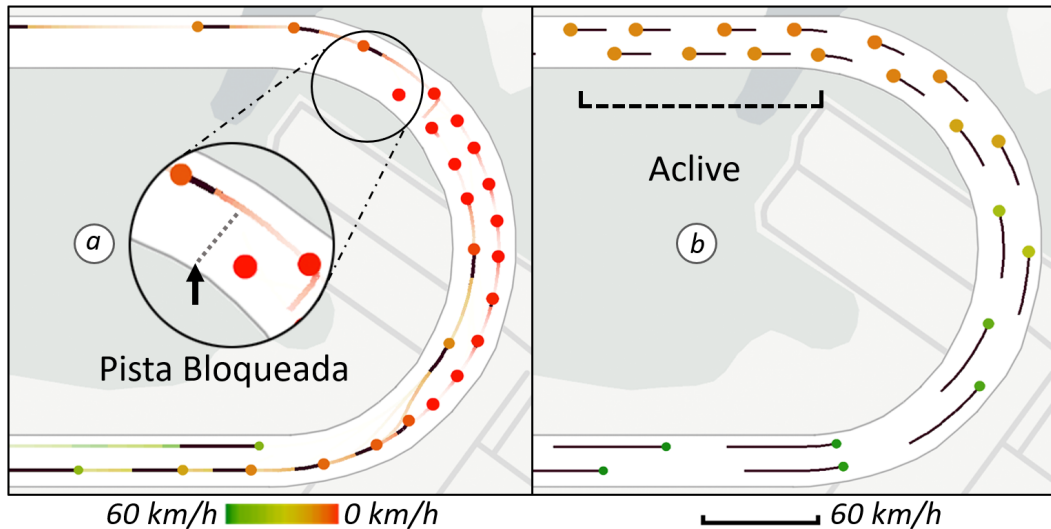
O especialista apontou o local, o momento em que surgiu a primeira onda de tráfego, e os veículos que a causaram (Figura 44a). Ele afirmou que a dinâmica da propagação das ondas de tráfego é semelhante ao cenário de via circular, ou seja, as mudanças de faixa são a principal causa da propagação. No entanto, o fluxo da via de acesso funciona como um gerador estacionário de ondas nesses casos. Por exemplo, a Figura 44b mostra duas ondas que se originaram próximo à via de acesso. Ele também notou a redução do fluxo da via principal após uma hora de simulação, mas disse que, nesses casos, o engarrafamento demora muito para desaparecer. De fato, após uma hora de simulação, o engarrafamento não se dissipou mesmo com um baixo fluxo. Por fim, ele citou os efeitos das fricções de tráfego, que acontecem quando dois fluxos disputam o mesmo espaço, e afirmou que um dos maiores desafios para os gestores urbanos é estimar os impactos desses efeitos.

Bloqueio de Pista

O cenário de bloqueio de pista reproduz os efeitos de eventos incomuns que ocorrem em uma via, como obras, acidentes ou condições ruins de conservação da via (TREIBER; KESTING, 2013). Neste sistema aberto, os gestores urbanos podem simular a propagação de ondas de tráfego causadas por esses eventos.

Para esta simulação, uma das duas faixas de uma via é bloqueada. Na primeira

Figura 45 – (a) O cenário de um bloqueio de pista. A faixa interna é obstruída sem nenhuma sinalização de bloqueio no mapa. (b) O cenário de um aclave. Os veículos pesados tiveram sua velocidade reduzida significativamente na seção do aclave prejudicando o fluxo de tráfego.



Fonte: o autor.

hora, o tráfego fluía normalmente com um fluxo de entrada de 1.500 veículos/h e um limite de velocidade de 60 km/h. Em seguida, o limite de velocidade é aumentado para 80 km/h. Depois disso, a simulação durou mais uma hora.

Mesmo sem nenhuma sinalização de bloqueio de pista no mapa, o especialista notou o problema naturalmente, devido ao comportamento dos veículos (Figure 45a). Ele disse que os rastros e cores ajudaram a caracterizar esse comportamento. Apesar de perceber o aumento de velocidade dos veículos, ele não o associou ao aumento no limite de velocidade. Porém, a mudança no limite de velocidade foi o principal motivo para a desestabilização do fluxo, pois quanto maior a velocidade, maior a dificuldade de um motorista mudar de faixa. Então, muitos veículos não conseguiram mudar de faixa, provocando uma onda de tráfego mesmo com um fluxo baixo de veículos.

Active

O cenário de aclave simula os impactos de veículos pesados, como ônibus e caminhões, no fluxo de tráfego. Esse cenário difere fundamentalmente dos anteriores por gargalos de conservação de fluxo (TREIBER *et al.*, 2000).

Uma estrada de duas faixas foi usada com um aclave relativamente íngreme para simular um tráfego muito leve com um fluxo de entrada de 1.900 veículos/h, um limite de velocidade de 80 km/h e 30% de veículos pesados. Nesta simulação não houve anomalias, ou

seja, o tráfego fluiu normalmente ao longo da simulação. No entanto, na subida, os veículos pesados tiveram sua velocidade reduzida, diminuindo significativamente a velocidade dos outros veículos (Figure 45b).

O especialista não descobriu o motivo da redução de velocidade. Ele supôs que poderia ser uma dificuldade devido à curva ou à má condição de conservação da via. Quando descobriu que 30% dos veículos eram pesados, ele sugeriu que esses veículos fossem representados com uma forma diferente.

Comentários do Especialista

O especialista reconheceu o potencial da abordagem proposta para visualizar padrões de oscilação de tráfego a partir de dados de trajetória. Ele considerou a visualização agradável e facilitou a percepção das ondas de tráfego. Além disso, ele afirmou que o rastro era um excelente indicador para visualizar as mudanças de faixa, que são uma das principais causas de perturbações em tráfego denso. O controle de tempo da visualização, em particular, a funcionalidade de retroceder recebeu comentários muito positivos.

Ele reconheceu que a proposta poderia ajudar significativamente em estudos que analisam os impactos de construção de novos edifícios em uma rede rodoviária. Por exemplo, a construção de um novo shopping em uma região populosa pode causar engarrafamentos em toda a vizinhança. A maioria dos gestores urbanos geralmente usa simuladores numéricos e ferramentas estatísticas. Assim, ele mencionou que o método proposto poderia ajudar esses gestores a conscientizar as autoridades públicas sobre esses impactos. Para aplicações *online*, ele considerou que o método poderia ser útil para prever o engarrafamento em tempo real.

Em relação às limitações da visualização proposta, o especialista percebeu a falta de gráficos estatísticos para descrever o fluxo de tráfego das vias. Na visualização inicial que utilizou dados de GPS de trajetórias reais, ele considerou que não era possível visualizar oscilações nessa base de dados, apenas uma visão geral do fluxo de tráfego. Ele concluiu que a abordagem proposta era mais apropriada para visualizar dados oriundos de sensores que capturam continuamente o tráfego, como etiquetas RFID, câmeras de vídeo e *scanners a laser*, que coletam dados de todo o fluxo de tráfego. No entanto, a representação contínua do fluxo total é muito difícil de acontecer. Talvez isso possa mudar, em um futuro próximo, quando todos os veículos forem rastreados por GPS. Além disso, ele considerou que a visualização proposta teria excelentes resultados em simulações.

5.3 Considerações Finais

Este capítulo apresentou duas aplicações da abordagem proposta para visualizar padrões de mobilidade em dinâmica de tráfego. A primeira foi a visualização de rotas quentes que visa monitorar vias com tráfego pesado. Esse é um problema relevante e desafiador no estudo de padrões de mobilidade global, no qual a análise visual pode desempenhar um papel importante na tomada de decisão, pois requer monitoramento e avaliação de diversas situações. Alguns casos de uso com conjuntos de dados reais descrevendo movimentos de táxis foram realizados para avaliar a eficácia do método proposto. Além disso, um estudo comparou a eficácia da abordagem proposta com uma solução tradicional da literatura. A segunda aplicação foi a visualização de ondas de tráfego para analisar padrões de mobilidade local. Nessa visualização, o foco é a análise visual de como o comportamento dos indivíduos afeta o fluxo de tráfego. Diferentemente da visualização de rotas quentes, a visualização de ondas de tráfego não tem grande dependência dos parâmetros de tempo e de limiar de densidade. Na realidade, esses parâmetros podem até ser configurados como constantes, independentemente do conjunto de dados utilizado. O que influencia o resultado dessa visualização é o quanto o fluxo de tráfego real foi representado no conjunto de dados – quanto maior, melhor. Por ter uma configuração de parâmetros mais simples, a eficácia da visualização de ondas de tráfego foi analisada através de uma avaliação com um especialista na percepção de anomalias em cenários simulados. Assim, foi possível identificar os benefícios e as limitações da visualização proposta.

No próximo capítulo, uma análise do desempenho da abordagem proposta é apresentada avaliando a sua eficiência no monitoramento de um grande volume de dados. Para isso, conjuntos de dados reais e sintéticos são utilizados.

6 ANÁLISE DE DESEMPENHO

Neste capítulo, o desempenho da abordagem proposta é analisado para avaliar a eficiência das visualizações produzidas. Essencialmente, pretende-se responder a seguinte pergunta: A abordagem proposta é adequada para visualizar o fluxo total de tráfego de uma grande cidade inteira com as tecnologias empregadas? Neste sentido, na Seção 6.1, as tecnologias empregadas são descritas. Na Seção 6.2, os conjuntos de dados utilizados são apresentados. A análise dos resultados de desempenho é discutida na Seção 6.3. Por último, as considerações finais são apresentadas na Seção 6.4.

6.1 Tecnologias Empregadas

Os protótipos foram implementados usando JavaScript e WebGL para renderização em GPU. Para a representação do mapa e a interação com o mesmo, foi utilizada a biblioteca Leaflet, que é baseada no OpenStreet Maps (AGAFONKIN,). A Leaflet é uma biblioteca JavaScript de código aberto para mapas interativos compatíveis com dispositivos móveis e máquinas *desktop*. Ela possui a maioria dos recursos de mapeamento presentes nas plataformas comerciais mais utilizadas, como o Google Maps. Além disso, suas funcionalidades podem ser facilmente estendidas através de *plug-ins*, o que facilitou muito o desenvolvimento deste trabalho. A implementação da abordagem proposta foi baseada na Web para aproveitar as vantagens da plataforma: disponibilidade, multiplataforma, fácil manutenção, adaptabilidade a aplicativos móveis, etc.

A ideia de visualizar em tempo real através de uma animação o fluxo total de tráfego de uma cidade inteira foi bastante desafiadora para as tecnologias disponíveis atualmente. Para representar os rastros dos objetos em movimento, esse fluxo pode demandar a renderização de centenas de milhares de segmentos de linha em um intervalo de milissegundos – por exemplo, cerca de 33ms para uma animação de 30fps.

Utilizando uma plataforma de baixo custo, que será descrita posteriormente, o primeiro protótipo foi desenvolvido usando primitivas geométricas como linhas e círculos. No entanto, mesmo utilizando GPU, os resultados não foram encorajadores. Apenas cerca de 20.000 elementos gráficos foram visualizados em uma animação a 20fps. O segundo protótipo foi desenvolvido utilizando texturas (*sprites*) ao invés de primitivas geométricas. Nesse protótipo, foi utilizada a biblioteca JavaScript de código aberto PixiJS, a qual foi criada para o desenvolvimento

de Jogos 2D e aplicações multimídia (GROVES; ENGLER,). Essa biblioteca é otimizada para animações de grandes quantidades de elementos gráficos com texturas usando WebGL e GPU. Ela também dá suporte a múltiplas plataformas e foi adaptada à Leaflet neste trabalho como um *plug-in*. Assim, foi possível visualizar 160.000 elementos gráficos a 20fps. Isso foi possível devido à capacidade das GPUs em processar paralelamente texturas (*buffers*) em altíssima velocidade. Dessa forma, esse protótipo foi o primeiro passo para viabilidade da abordagem proposta.

Contudo, a renderização desses elementos gráficos representava apenas uma atualização do rastro no intervalo de tempo de um quadro de animação. Era necessário renderizar os rastros por completo de todos os objetos em movimento. Dependendo do tamanho do rastro, isto é, do período de análise, o total de elementos gráficos poderia ser na casa de milhões. Por exemplo, em apenas um segundo poderia haver 4.800.000 (160.000 x 30fps) elementos. Além do alto poder computacional para renderizar essa quantidade de elementos, muita memória e armazenamento seriam demandados, inviabilizando a visualização. Assim, o segundo passo para viabilidade da abordagem proposta foi adotar um método incremental no qual as atualizações dos rastros são acumuladas de um quadro para outro, como discutido anteriormente. Dessa forma, não é necessário armazenar nenhum dado histórico, reduzindo significativamente os custos de armazenamento e memória, e tornando viável o emprego da abordagem proposta no monitoramento em tempo real. Isso foi possível devido ao fato da composição de *buffers* ser realizada por um processo conhecido em computação gráfica como *render baking*, que também é executado em paralelo na GPU.

As imagens e os testes de desempenho deste trabalho foram gerados em um *laptop* com processador Intel Core i5 de 2,5 GHz com 8 GB de RAM e uma NVIDIA GeForce GTX 1050 com 4 GB de memória de vídeo, ou seja, uma plataforma gráfica portátil de baixo custo com uma GPU com desempenho gráfico inferior à média. Como exemplo de comparação, a GPU comercialmente difundida com maior poder gráfico atualmente é a NVIDIA GeForce RTX 2080 TI, possuindo uma quantidade de núcleos de processamento paralelo quase seis vezes maior do que a utilizada neste trabalho. Sem mencionar que ela pertence a uma nova geração (série 20) com mais memória, mais processamento individual de cada núcleo, e melhorias na arquitetura. No entanto, o motivo da escolha dessa plataforma de baixo custo foi verificar a possibilidade da técnica ser utilizada por equipamentos comuns de usuários finais para acessar informações de tráfego, como acontece hoje com o Waze e Google Maps. É importante mencionar que tanto

o Waze como Google Maps apresentam informações de forma *online*, mas não em tempo real, pois eles agregam informações de tráfego em janelas de tempo. Além disso, as agências de trânsito poderiam utilizar a abordagem proposta em plataformas computacionais mais robustas para visualizar detalhadamente mais informações em múltiplas telas.

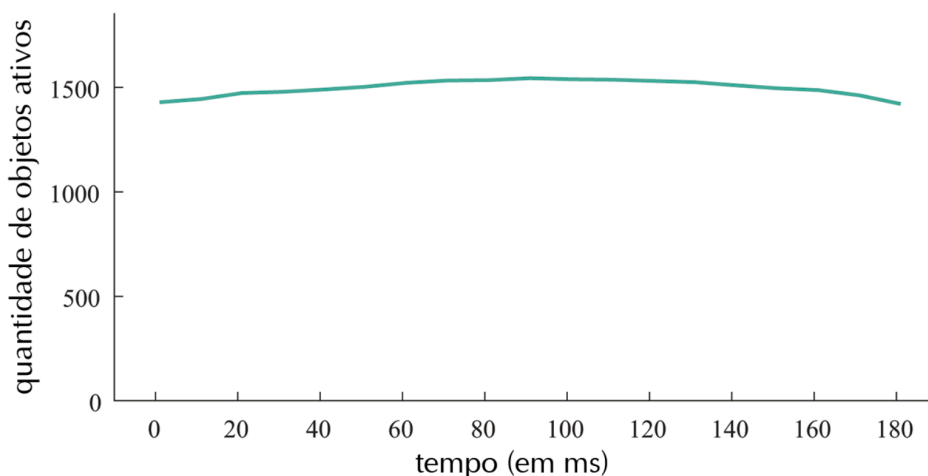
6.2 Conjuntos de Dados Analisados

Como mencionado anteriormente, a abordagem proposta apresenta rotas quentes com qualidade significativamente melhor recebendo como entrada um conjunto de dados com uma alta taxa de amostragem e com um grande número de objetos ativos. Para esse cenário, o desempenho da abordagem proposta foi testado no monitoramento de tráfego de uma cidade com uma grande quantidade de veículos monitorados. Para este propósito, conjuntos de dados reais e sintéticos foram utilizados. Um dos maiores conjuntos de dados reais disponibilizados publicamente com informações de trajetórias completas é o T-DRIVE. Esse conjunto de dados disponibiliza informações da segunda maior frota de táxis em uma cidade do mundo com mais de 70.000 táxis (GUI *et al.*, 2016), perdendo apenas para a cidade do México com cerca de 100.000.

Assim, um subconjunto de três minutos do T-Drive no horário de maior movimento foi utilizado. Em uma etapa de pré-processamento para convertê-lo em um subconjunto denso com alta taxa de amostragem, novos pontos de trajetória foram interpolados e mapeados para uma rede viária usando *map matching*. Em seguida, esse subconjunto foi usado como referência para gerar novos conjuntos de dados de tamanhos diferentes através de um processo de replicação. O principal objetivo do teste era verificar quantos objetos ativos a abordagem proposta suportaria. Nesse teste, a qualidade das rotas quentes não foi considerada. Assim, um período curto de três minutos foi selecionado para facilitar a replicação de dados.

Mesmo sendo um dos maiores conjuntos de dados públicos com trajetórias de mais de 70.000 táxis e o intervalo selecionado estar em um horário de maior movimento, o subconjunto de referência mantém um fluxo médio de apenas 1.500 veículos ativos durante esse intervalo, conforme mostrado na Figura 46. Apesar de apresentar um fluxo quase constante no intervalo selecionado, a Figura 47 mostra muitos objetos que entram e saem da visualização durante o intervalo de tempo – a primeira e a última medição foram descartadas neste gráfico para manter uma boa escala no eixo Y do gráfico, e a contagem de objetos ativos foi acumulada em intervalos de 10 segundos para melhorar a percepção da variação no eixo X. Para cada objeto que entra ou sai, a TOM é atualizada, afetando o desempenho. Além do número de objetos ativos, outro

Figura 46 – Distribuição do número de objetos ativos ao longo do tempo.



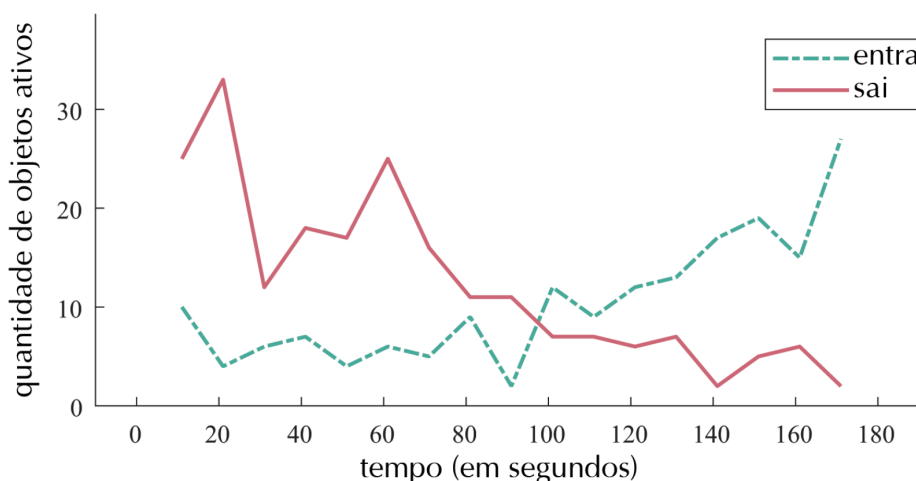
Fonte: o autor.

fator que afeta diretamente o desempenho é a resolução da visualização, a qual define o tamanho dos *buffers*. Assim, a partir do subconjunto de referência, 15 novos conjuntos de dados foram gerados. Cada um desses conjuntos foi testado com três resoluções diferentes: HD (1280 x 720), Full HD (1920 x 1080) e Ultra-Wide Full HD (2560 x 1080). Um conjunto de dados D_i foi gerado por i réplicas do subconjunto de referência, em que $i = 1, \dots, 16$, por exemplo, D_{16} tinha aproximadamente 24.000 (16 x 1.500) veículos ativos.

6.3 Análise dos Resultados de Desempenho

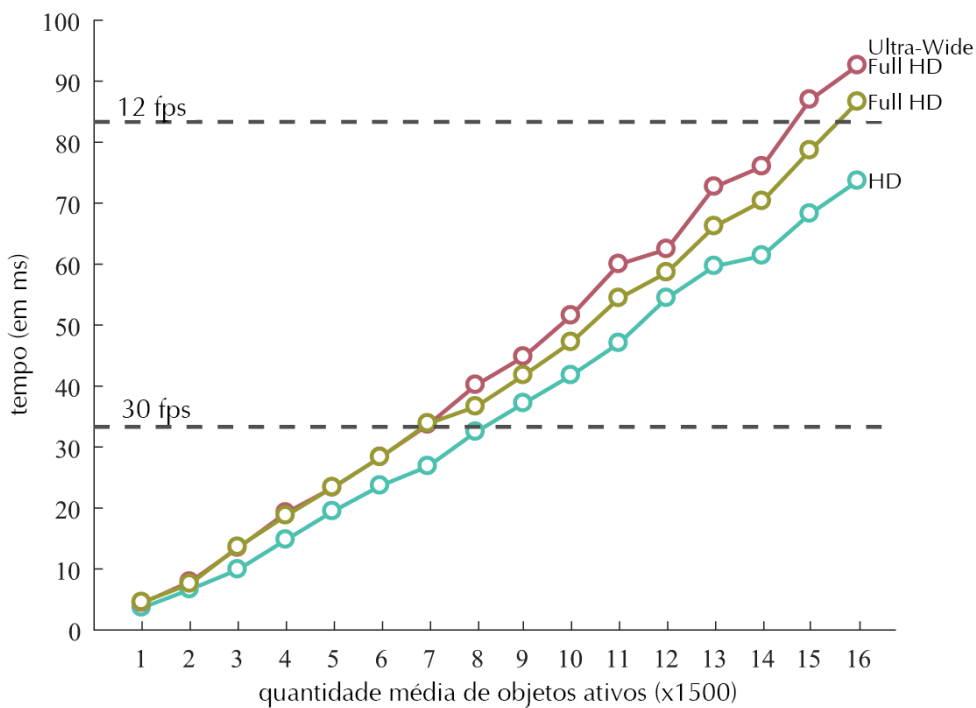
A Figura 48 apresenta os resultados de desempenho. Existem duas linhas de referência pontilhadas horizontais. A primeira linha marca o tempo 33,33ms, o que equivale a uma

Figura 47 – Evolução dos objetos ativos que entram e saem da visualização ao longo do tempo.



Fonte: o autor.

Figura 48 – Os resultados de desempenho. Um total de 16 conjuntos de dados foram testados com três resoluções diferentes: HD (1280 x 720), Full HD (1920 x 1080) e Ultra-Wide Full HD (2560 x 1080). As linhas de referência pontilhadas representam a velocidade de animação de 30fps e de 12fps.



Fonte: o autor.

velocidade de animação de 30fps. A segunda linha marca o tempo 83,33ms, indicando 12fps, que é a taxa mínima de quadros por segundo para se perceber a visualização como movimento. No limite de 30fps, as resoluções Full HD e Ultra-Wide Full HD suportam no máximo 10.500 (7 x 1500) objetos ativos, enquanto a HD suporta 12.000 (8 x 1500) objetos ativos. Considerando 12fps como limite, as três resoluções mantêm mais de 21.000 (14 x 1500) objetos ativos. Observe que o tempo cresce linearmente com o tamanho do conjunto de dados e que a área total das resoluções não afeta significativamente o desempenho. Por exemplo, mesmo a resolução Ultra-Wide Full HD, com uma área 33% maior que a Full HD, tem um desempenho similar à Full HD para até 30fps. Isso se deve à complexidade linear dos Algoritmos 1 e 2 para acessar todos os objetos ativos. Os processos executados em paralelo na GPU – suavização e filtragem no Algoritmo 3, além da composição e renderização de *buffers* – não consideram o tamanho do conjunto de dados, já que sua função principal é lidar com *buffers*. Portanto, o tamanho do *buffer* para esses processos é o fator mais predominante no desempenho, ou seja, o tamanho do conjunto de dados não tem um impacto significativo no desempenho do sistema. No entanto, esse fato não diminui a importância do processamento paralelo em GPU, uma vez que não é

viável obter uma renderização de alto desempenho com visualização interativa sem GPU. Além disso, as GPUs de baixo desempenho não suportam visualizações efetivas usando *viewports* de alta resolução.

6.4 Considerações Finais

Concluindo, a análise mostra que o desempenho da abordagem proposta é afetado basicamente pela quantidade de veículos ativos e pelo tamanho da *viewport*. Primeiro, a quantidade de veículos ativos afeta o desempenho pela necessidade de acessar e atualizar a TOM. Como mostrado na Figura 48, o tempo para renderizar um quadro de animação cresce linearmente com essa quantidade de veículos. Isso ocorre devido aos Algoritmos 1 e 2, que acessam e atualizam a TOM, possuem complexidade linear e utilizarem estruturas indexadas. Segundo, o tamanho da *viewport* afeta o desempenho devido à necessidade de atualizar, compor e renderizar os *buffers*. Nessa lógica, quanto maior a resolução da *viewport*, maior o tempo para processar a quantidade de suas células. Logo, a quantidade de objetos ativos não tem um impacto significativo no desempenho do processamento de *buffers*. Assim, o comportamento linear apresentado na Figura 48 é devido basicamente às operações de leitura e escrita da TOM. Mesmo assim, a GPU tem um papel fundamental na visualização processando os *buffers* em paralelo. Sem ela, não seria possível visualizar o movimento com animação.

É importante mencionar que o desempenho da abordagem proposta não depende da taxa de amostragem, ao contrário da qualidade das visualizações, que demanda uma alta taxa de amostragem. Isso acontece devido à utilização do método incremental em que dados históricos não são armazenados. Por exemplo, a abordagem proposta tem o mesmo desempenho para processar um conjunto de dados com uma taxa de amostragem de um minuto e um outro com taxa de um segundo. No entanto, a qualidade da visualização do primeiro será bem pior que a do segundo.

Em relação a pergunta feita no começo deste capítulo, a resposta deve considerar os dois padrões de mobilidade analisados pelas visualizações propostas. Para a visualização de ondas de tráfego, a qual demanda uma representação considerável do fluxo total de tráfego, a resposta é positiva, pois a quantidade de objetos ativos alcançada é suficiente para retratar possivelmente todo o fluxo de tráfego de uma área representada nos níveis de zoom mais alto, considerando apenas os objetos que estão dentro da *viewport*. Para a visualização de rotas quentes que visa analisar padrões de mobilidade globais, a resposta depende da plataforma de

hardware utilizada e de quanto o fluxo total foi representado. Para uma cidade como Pequim, por exemplo, com a plataforma utilizada nos testes, não seria possível visualizar uma representação considerável do fluxo total. No entanto, a visualização de rotas quentes não é tão dependente da representação do fluxo total como na visualização de ondas de tráfego, como já discutido antes. É possível ter bons resultados analisando apenas trajetórias de táxis. Como de fato aconteceu com os trabalhos que utilizaram o T-Drive para analisar padrões de mobilidade globais, já que táxis são excelentes sensores para analisar fluxo de tráfego. Neste caso, a proposta para visualização de rotas quentes seria adequada para analisar dados de táxi, mesmo utilizando a plataforma de baixo custo utilizada nos testes. Por fim, se houver necessidade de uma representação maior do fluxo total, utilizar uma plataforma mais robusta possibilita aumentar consideravelmente o limite de objetos ativos.

Outro aspecto que deve ser considerado na visualização de rotas quentes é a área monitorada em quilômetros quadrados. Por exemplo, nos testes realizados, foram visualizados os veículos ativos de uma grande cidade como Pequim, com 16.808km^2 . Presumivelmente, isso não é prático. Reduzindo a área de monitoramento para visualizar uma parte menor da cidade, o limite de veículos ativos aumenta. Por exemplo, se a área for reduzida para visualizar metade da cidade de Pequim, será possível visualizar em Full HD um total de até 42.000 veículos distribuídos pela cidade.

O próximo capítulo apresenta as conclusões acerca deste trabalho, destacando os resultados alcançados, e apresentando os trabalhos futuros oriundos desta tese de doutorado.

7 CONCLUSÕES E TRABALHOS FUTUROS

Este capítulo finaliza o trabalho com as considerações finais e as etapas futuras para continuidade da abordagem proposta nesta tese. A Seção 7.1 apresenta os resultados alcançados destacando as principais contribuições e limitações deste trabalho. Em seguida, na Seção 7.2, as possibilidades de trabalhos futuros são elencadas.

7.1 Resultados Alcançados

Esta tese investigou o problema de descobrir e monitorar padrões de mobilidade local e global em tempo real através de visualização interativa a partir de dados de trajetórias. Como visto anteriormente, agregar movimento empregando discretização do tempo em janelas pode apresentar inconsistências no tempo e no espaço com o fluxo de tráfego real. Embora essa agregação possibilite extrair conhecimento potencialmente útil, nesses casos, não é possível visualizar as dinâmicas de tráfego em tempo real, pois o resultado é exibido apenas no final de uma janela de tempo. Nesse sentido, a abordagem proposta representa o movimento de objetos sem usar as janelas de tempo empregando algoritmos otimizados que usam GPU para a análise de toda a dinâmica do movimento, o que é vital para os sistemas de monitoramento de tráfego. As principais contribuições desta tese são:

- **Uma arquitetura** para monitorar em tempo real objetos ativos em movimento através de dados coletados por dispositivos GPS. Nela, apenas as posições dos objetos ativos (TOM) e um *buffer* são temporariamente armazenados para serem visualizados, ou seja, os dados históricos não são mantidos. Portanto, diferentemente das abordagens encontradas na literatura, o modelo proposto é compacto e adequado para uma aplicação de monitoramento de dinâmicas de tráfego;
- **Uma solução incremental baseada em grade** que usa uma estrutura de dados indexada para rastrear os objetos ativos e suas posições atuais. Essa solução pode monitorar cenários dinâmicos com mais eficiência e requer um pequeno número de parâmetros: um período de análise e um limiar de densidade. Ambos os parâmetros podem ser ajustados em tempo real e as alterações são visualizadas instantaneamente. Diferentemente de outras abordagens, não é necessário que os dados sejam mapeados com *map-matching* ou submetidos a uma análise estatística prévia para a definição dos parâmetros;
- **Uma metodologia sistemática** que engloba as principais etapas da visualização de dados

de trajetória. A partir de dados brutos, ela pode ser utilizada para remover inconsistências nos dados, pré-processar atributos espaço-temporais, sincronizar as posições dos objetos ativos para a mesma unidade de tempo discreta, renderizar elementos gráficos para a visualização, e realçar os atributos das imagens geradas para facilitar a análise visual;

- **Uma solução para visualização de padrões de mobilidade global** através do monitoramento das vias com tráfego pesado (rotas quentes). Uma validação experimental com dados reais e sintéticos demonstrou a eficácia dessa solução. Além disso, um estudo equiparou a eficácia da abordagem proposta com uma solução tradicional da literatura;
- **Uma proposta para analisar padrões de mobilidade local** através da visualização de ondas de tráfego. A eficácia dessa proposta foi avaliada por um especialista na percepção de anomalias em cenários simulados, identificando os benefícios e as limitações da visualização proposta;
- **Algoritmos otimizados de renderização e filtragem** que são executados em paralelo em GPU para manter taxas interativas de exibição de quadros, mesmo com grandes conjuntos de dados. Esses algoritmos exploram a capacidade das GPUs em processar paralelamente texturas (*buffers*) em altíssima velocidade. Os resultados demonstram que a abordagem proposta é escalonada linearmente com o tamanho do conjunto de dados. Assim, ela pode ser utilizada para visualizar o fluxo total de tráfego de uma grande cidade inteira, dependendo da plataforma de *hardware* utilizada.

Apesar dos bons resultados obtidos, a abordagem proposta neste trabalho apresenta algumas limitações discutidas a seguir:

- **Dependência de um fluxo de alta densidade.** Para visualização de rotas quentes, quando maior a quantidade de veículos circulando, menor será a contribuição de cada um e, assim, os erros GPS serão minimizados. Além disso, um fluxo de baixa densidade não gera rotas quentes significativas, pois, dependendo do limiar, ou quase todas as rotas serão quentes ou quase nenhuma será. Para a visualização de ondas de tráfego, é necessário uma representação significativa do fluxo do tráfego total. Caso contrário, com uma pequena amostra do fluxo, não é possível perceber as ondas de tráfego. Essa limitação afeta a eficácia da abordagem;
- **Dependência de uma alta taxa de amostragem.** As representações visuais de trajetórias terão melhor qualidade com altas taxas de amostragem. De fato, essa dependência existe para a maioria das abordagens de visualização de trajetórias que não usam *map-matching*.

As taxas de amostragem com frequência de coleta de posições entre um e cinco segundos são as ideais para este trabalho. No entanto, como dados históricos não são armazenados pelo método proposto, a taxa de amostragem não afeta o desempenho. Essa limitação afeta a qualidade visual dos resultados;

- **Dependência da escala de representação do mapa**, que é definida pelo nível de zoom. Como a abordagem utiliza uma solução baseada em grade, a qual agrega os dados em células, utilizar um nível de zoom errado poderá afetar a eficácia e a qualidade da visualização. Para o monitoramento de rotas quentes de cidades, regiões e bairros, os níveis de 12 a 15 são geralmente os mais relevantes, enquanto que, para visualizar ondas de tráfego, são indicados os níveis 16, 17 e 18 em que as faixas são distinguíveis;
- **Dependência de valores significativos para o limiar de densidade e o período de análise** para obter bons resultados na visualização de rotas quentes. De fato, essa dificuldade existe para a maioria das abordagens, incluindo a clusterização de trajetórias que geralmente usa análise estatística prévia para definir esses valores. Na abordagem proposta, o ajuste desses valores é realizado em tempo real, e o resultado é visualizado instantaneamente. Na visualização de ondas de tráfego, essa limitação não existe, já que esses valores podem ser definidos como constantes.

7.2 Trabalhos Futuros

Como trabalho futuro, as visualizações propostas podem ser integradas em uma ferramenta completa de análise visual, incluindo suporte a consultas espaço-temporais e gráficos estatísticos. Em seguida, o sistema integrado pode ser testado em um estudo de campo com dados oriundos de sensores que captura continuamente o tráfego, como etiquetas RFID, câmeras de vídeo e *scanners a laser*, visando uma representação completa do fluxo de tráfego real. Depois, os benefícios e as limitações dessa ferramenta de análise visual podem ser avaliados sistematicamente por especialistas do domínio. Além disso, a percepção gráfica dos padrões de mobilidade pode ser avaliada por usuários comuns.

Outra possibilidade de continuidade deste trabalho é estudar meios de proteção de privacidade dos dados dos usuários para dar suporte à visualização das dinâmicas de tráfego por meio de um aplicativo para dispositivos móveis, assim como acontece hoje nos aplicativos Google Maps e Waze, mas sem usar agregação de dados em janela de tempo.

Outra oportunidade de trabalho futuro é desenvolver uma solução de alertas por

regiões pré-definidas, como bairros ou vias inteiras. Assim, quando uma região possui uma certa quantidade de células quentes, o sistema acionaria um aviso. Esse processo pode ser automatizado por técnicas de aprendizagem automática de máquinas para identificar eventos anormais comparando com *buffers* de referência calculados por dados históricos, que representariam o comportamento normal das rotas observadas em um horário específico.

Por fim, outra possibilidade seria visualizar informações não espaço-temporais sobre os objetos em movimento, por exemplo, tipo de veículo: carro, caminhão, transporte público etc. Esses estudos podem ajudar a criar outros meios para melhorar a percepção da dinâmica do movimento de objetos.

REFERÊNCIAS

- AGAFONKIN, V. **Leaflet: an open-source JavaScript library for mobile-friendly interactive maps**. Disponível em: <https://leafletjs.com>. Acesso em: 27 ago. 2018.
- AGUILERA, M. K.; CHEN, W.; TOUEG, S. Heartbeat: A timeout-free failure detector for quiescent reliable communication. **Distributed algorithms**, Springer, p. 126–140, 1997.
- AHEARN, S. C.; DODGE, S. Recursive multi-frequency segmentation of movement trajectories (ReMuS). **Methods in Ecology and Evolution**, Wiley, v. 9, n. 4, p. 1075–1087, jan 2018.
- AIGNER, W.; MIKSCH, S.; SCHUMANN, H.; TOMINSKI, C. **Visualization of time-oriented data**. [S.l.]: Springer Science & Business Media, Berlin, Alemanha, 2011.
- ALEMI, F.; CIRCELLA, G.; HANDY, S.; MOKHTARIAN, P. What influences travelers to use uber? exploring the factors affecting the adoption of on-demand ride services in california. **Travel Behaviour and Society**, Elsevier BV, v. 13, p. 88–104, oct 2018.
- ALGIZAWY, E.; OGAWA, T.; EL-MAHDY, A. Real-time large-scale map matching using mobile phone data. **ACM Transactions on Knowledge Discovery from Data**, Association for Computing Machinery (ACM), v. 11, n. 4, p. 1–38, jul 2017.
- ALMEIDA, A. M.; LEITE, J. L.; MACEDO, J. A.; MACHADO, J. C. Gps2gr: Optimized urban green routes based on gps trajectories. In: **ACM. Proc. of the 8th ACM SIGSPATIAL Workshop on GeoStreaming**. [S.l.], 2017. p. 39–48.
- AMINI, F.; RUFIANGE, S.; HOSSAIN, Z.; VENTURA, Q.; IRANI, P.; MCGUFFIN, M. J. The impact of interactivity on comprehending 2d and 3d visualizations of movement data. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 21, n. 1, p. 122–135, jan 2015.
- ANDRIENKO, G.; ANDRIENKO, N.; CHEN, W.; MACIEJEWSKI, R.; ZHAO, Y. Visual analytics of mobility and transportation: State of the art and further research directions. **IEEE Transactions on Intelligent Transportation Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 18, n. 8, p. 2232–2249, aug 2017.
- ANDRIENKO, G.; ANDRIENKO, N.; DEMSAR, U.; DRANSCH, D.; DYKES, J.; FABRIKANT, S. I.; JERN, M.; KRAAK, M.; SCHUMANN, H.; TOMINSKI, C. Space, time and visual analytics. **IJGIS**, Taylor & Francis, v. 24, n. 10, p. 1577–1600, 2010.
- ANDRIENKO, G.; ANDRIENKO, N.; FUCHS, G.; GARCIA, J. M. C. Clustering trajectories by relevant parts for air traffic analysis. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 24, n. 1, p. 34–44, jan 2018.
- ANDRIENKO, G.; ANDRIENKO, N.; HURTER, C.; RINZIVILLO, S.; WROBEL, S. From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In: **2011 IEEE Conference on Visual Analytics Science and Technology (VAST)**. [S.l.]: Institute of Electrical and Electronics Engineers (IEEE), 2011.
- ANDRIENKO, G.; ANDRIENKO, N.; HURTER, C.; RINZIVILLO, S.; WROBEL, S. Scalable analysis of movement data for extracting and exploring significant places. **IEEE Transactions**

on **Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 19, n. 7, p. 1078–1094, jul 2013.

ANDRIENKO, G.; ANDRIENKO, N.; JANKOWSKI, P.; KEIM, D.; KRAAK, M.-J.; M., A.; WROBEL, S. Geovisual analytics for spatial decision support: Setting the research agenda. **IJGIS**, Informa UK Limited, v. 21, n. 8, p. 839–857, 2007.

ANDRIENKO, G.; ANDRIENKO, N.; WROBEL, S. Visual analytics tools for analysis of movement data. **ACM SIGKDD Explorations Newsletter**, ACM, v. 9, n. 2, p. 38–46, 2007.

ANDRIENKO, N.; ANDRIENKO, G. **Exploratory analysis of spatial and temporal data: a systematic approach**. [S.l.]: Springer Science & Business Media, 2006.

ANDRIENKO, N.; ANDRIENKO, G. Visual analytics of movement: An overview of methods, tools and procedures. **Information Visualization**, SAGE Publications, v. 12, n. 1, p. 3–24, 2012.

ANDRIENKO, N.; ANDRIENKO, G. Visual analytics of movement: An overview of methods, tools and procedures. **Information Visualization**, Sage Publications Sage UK: London, England, v. 12, n. 1, p. 3–24, 2013.

ANWAR, A.; NAGEL, T.; RATTI, C. Traffic origins: A simple visualization technique to support traffic incident analysis. In: IEEE. **Visualization Symposium (PacificVis), 2014 IEEE Pacific**. [S.l.], 2014. p. 316–319.

ASSAF, M. H.; PETRIU, E. M.; GROZA, V. Ship track estimation using GPS data and kalman filter. In: **2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)**. [S.l.]: IEEE, 2018.

AZEVEDO, C. L.; CARDOSO, J. L.; BEN-AKIVA, M.; COSTEIRA, J. P.; MARQUES, M. Automatic vehicle trajectory extraction by aerial remote sensing. **Procedia - Social and Behavioral Sciences**, Elsevier BV, v. 111, p. 849–858, feb 2014.

BACH, B.; SHI, C.; HEULOT, N.; MADHYASTHA, T.; GRABOWSKI, T.; DRAGICEVIC, P. Time curves: Folding time to visualize patterns of temporal evolution in data. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 22, n. 1, p. 559–568, jan 2016.

BAGLIONI, M.; MACÊDO, J. A. F. de; RENSO, C.; TRASARTI, R.; WACHOWICZ, M. Towards semantic interpretation of movement behavior. In: **Advances in GIScience**. [S.l.]: Springer Berlin Heidelberg, 2009. p. 271–288.

BAMBA, B.; LIU, L.; PESTI, P.; WANG, T. Supporting anonymous location queries in mobile environments with privacygrid. In: **Proceeding of the 17th international conference on World Wide Web - WWW'08**. [S.l.]: Association for Computing Machinery (ACM), 2008.

BARBOSA, R.; CARDOSO, D. O.; CARVALHO, D.; FRANÇA, F. M. Weightless neuro-symbolic GPS trajectory classification. **Neurocomputing**, Elsevier BV, v. 298, p. 100–108, jul 2018.

BECCARIO, C. **Earth Wind Map**. Disponível em: <http://earth.nullschool.net>. Acesso em: 11 ago. 2018.

BELCASTRO, L.; MAROZZO, F.; TALIA, D.; TRUNFIO, P. G-RoI. **ACM Transactions on Knowledge Discovery from Data**, Association for Computing Machinery (ACM), v. 20, n. 3, p. 1–22, jan 2018.

BERMINGHAM, L.; LEE, I. A probabilistic stop and move classifier for noisy gps trajectories. **Data Mining and Knowledge Discovery**, Springer, p. 1–29, 2018.

BERTIN, J. **Semiology of graphics**. [S.l.]: University of Wisconsin press Madison, Wisconsin, US, 1983. v. 3.

BIAN, J.; TIAN, D.; TANG, Y.; TAO, D. A survey on trajectory clustering analysis. **arXiv preprint arXiv:1802.06971**, 2018.

BOGORNY, V.; RENSO, C.; AQUINO, A. R. de; SIQUEIRA, F. de L.; ALVARES, L. O. CONSTanT - a conceptual data model for semantic trajectories of moving objects. **Transactions in GIS**, Wiley, v. 18, n. 1, p. 66–88, feb 2013.

BORRUSO, G. Network density estimation: A GIS approach for analysing point patterns in a network space. **Transactions in GIS**, Wiley, v. 12, n. 3, p. 377–402, jun 2008.

BOSS, D.; NELSON, T.; WINTERS, M.; FERSTER, C. J. Using crowdsourced data to monitor change in spatial patterns of bicycle ridership. **Journal of Transport & Health**, Elsevier BV, v. 9, p. 226–233, jun 2018.

BOUKHERS, Z.; WANG, Y.; SHIRAHAMA, K.; UEHARA, K.; GRZEGORZEK, M. Convoy detection in crowded surveillance videos. In: **Human Behavior Understanding**. [S.l.]: Springer International Publishing, 2016. p. 137–147.

BUCHIN, M.; DRIEMEL, A.; KREVELD, M. van; SACRISTÁN, V. An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In: **Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS'10**. [S.l.]: Association for Computing Machinery (ACM), 2010.

CAI, L.; LI, S.; WANG, S.; LIANG, Y. GPS trajectory clustering and visualization analysis. **Annals of Data Science**, Springer Nature, v. 5, n. 1, p. 29–42, nov 2017.

CAI, L.; ZHOU, Y.; LIANG, Y.; HE, J. Research and application of GPS trajectory data visualization. **Annals of Data Science**, Springer Nature, v. 5, n. 1, p. 43–57, mar 2018.

CAO, F.; ESTERT, M.; QIAN, W.; ZHOU, A. Density-based clustering over an evolving data stream with noise. In: SIAM. **Proceedings of the 2006 SIAM international conference on data mining**. [S.l.], 2006. p. 328–339.

CAO, Y.; XIONG, L.; YOSHIKAWA, M.; XIAO, Y.; ZHANG, S. Contpl: Controlling temporal privacy leakage in differentially private continuous data release. **Proceedings of the VLDB Endowment**, v. 11, n. 12, 2018.

CHAMBERS, E.; FASY, B. T.; WANG, Y.; WENK, C. Map-matching using shortest paths. In: **Proceedings of the 3rd International Workshop on Interactive and Spatial Computing - IWISC18**. [S.l.]: ACM Press, 2018.

CHEN, C.; DING, Y.; XIE, X.; ZHANG, S. A three-stage online map-matching algorithm by fully using vehicle heading direction. **Journal of Ambient Intelligence and Humanized Computing**, Springer Nature, mar 2018.

CHEN, W.; GUO, F.; WANG, F.-Y. A survey of traffic data visualization. **IEEE TITS**, IEEE, v. 16, n. 6, p. 2970–2984, 2015.

CHEN, Y.; TU, L. Density-based clustering for real-time stream data. In: **ACM. Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining**. [S.l.], 2007. p. 133–142.

CHEN, Z.; SHEN, H. T.; ZHOU, X.; ZHENG, Y.; XIE, X. Searching trajectories by locations. In: **Proceedings of the 2010 International Conference on Management of Data - SIGMOD10**. [S.l.]: ACM Press, 2010.

CHEVALIER, F.; DRAGICEVIC, P.; FRANCONERI, S. The not-so-staggering effect of staggered animated transitions on visual tracking. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 20, n. 12, p. 2241–2250, dec 2014.

CHOONG, M. Y.; ANGELINE, L.; CHIN, R. K. Y.; YEO, K. B.; TEO, K. T. K. Modeling of vehicle trajectory clustering based on LCSS for traffic pattern extraction. In: **IEEE 2nd International Conference on Automatic Control and Intelligent Systems -I2CACIS**. [S.l.]: IEEE, 2017.

COLLBERG, C.; KOBOUROV, S.; NAGRA, J.; PITTS, J.; WAMPLER, K. A system for graph-based visualization of the evolution of software. In: **Proceedings of the 2003 ACM symposium on Software visualization - SoftVis03**. [S.l.]: ACM Press, 2003.

CONROW, L.; WENTZ, E.; NELSON, T.; PETTIT, C. Comparing spatial patterns of crowdsourced and conventional bicycling datasets. **Applied Geography**, Elsevier BV, v. 92, p. 21–30, mar 2018.

COOK, K. A.; THOMAS, J. J. **Illuminating the path: The research and development agenda for visual analytics**. [S.l.], 2005.

CRAMER, J.; KRUEGER, A. B. Disruptive change in the taxi business: The case of uber. **American Economic Review**, American Economic Association, v. 106, n. 5, p. 177–182, may 2016.

CUI, W.; ZHOU, H.; QU, H.; WONG, P. C.; LI, X. Geometry-based edge clustering for graph visualization. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, n. 6, p. 1277–1284, nov 2008.

DABIRI, S.; HEASLIP, K. Inferring transportation modes from GPS trajectories using a convolutional neural network. **Transportation Research Part C: Emerging Technologies**, Elsevier BV, v. 86, p. 360–371, jan 2018.

DAMIANI, M. L.; HACHEM, F.; ISSA, H.; RANC, N.; MOORCROFT, P.; CAGNACCI, F. Cluster-based trajectory segmentation with local noise. **Data Mining and Knowledge Discovery**, Springer Nature America, Inc, v. 32, n. 4, p. 1017–1055, mar 2018.

DAVIDSON, A.; PETERS, J.; BRAKEWOOD, C. **Interactive Travel Modes: Uber, Transit, and Mobility in New York City**. [S.l.], 2017.

DEAR, M.; SCOTT, A. J. **Urbanization and urban planning in capitalist society**. [S.l.]: Routledge, 2018. v. 7.

- DEVOGELE, T.; ETIENNE, L.; ESNAULT, M.; LARDY, F. Optimized discrete fréchet distance between trajectories. In: **Proceedings of the 6th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data - BigSpatial17**. [S.l.]: ACM Press, 2017.
- DING, X.; CHEN, L.; GAO, Y.; JENSEN, C. S.; BAO, H. UITraMan: a unified platform for big trajectory data management and analytics. **Proceedings of the VLDB Endowment**, VLDB Endowment, v. 11, n. 7, p. 787–799, mar 2018.
- DING, Y.; CHEN, C.; XIE, X.; LIU, K.; FENG, L. An online trajectory compression system applied to resource-constrained GPS devices in vehicles. In: **2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)**. [S.l.]: IEEE, 2018.
- DORAISWAMY, H.; ZACHARATOU, E. T.; MIRANDA, F.; LAGE, M.; AILAMAKI, A.; SILVA, C. T.; FREIRE, J. Interactive visual exploration of spatio-temporal urban data sets using urbane. In: **Proceedings of the 2018 International Conference on Management of Data - SIGMOD18**. [S.l.]: ACM Press, 2018.
- DORFMAN, D. M.; LAPLANTE, C. D.; POZDNYAKOVA, O.; LI, B. Flock cluster analysis of mast cell event clustering by high-sensitivity flow cytometry predicts systemic mastocytosis. **American Journal of Clinical Pathology**, Oxford University Press, v. 144, n. 5, p. 764–770, nov 2015.
- DOUGLAS, D. H.; PEUCKER, T. K. ALGORITHMS FOR THE REDUCTION OF THE NUMBER OF POINTS REQUIRED TO REPRESENT a DIGITIZED LINE OR ITS CARICATURE. **Cartographica: The International Journal for Geographic Information and Geovisualization**, University of Toronto Press Inc. (UTPress), v. 10, n. 2, p. 112–122, dec 1973.
- DU, J.; WU, F.; LI, J.; HE, H. A progressive method of simplifying polylines with multi-bends simplification templates. In: **Communications in Computer and Information Science**. [S.l.]: Springer Singapore, 2017. p. 515–528.
- DUAN, X.; JIN, C.; WANG, X.; ZHOU, A.; YUE, K. Real-time personalized taxi-sharing. In: SPRINGER. **International Conference on Database Systems for Advanced Applications**. [S.l.], 2016. p. 451–465.
- DYKES, J.; MOUNTAIN, D. Seeking structure in records of spatio-temporal behaviour: visualization issues, efforts and applications. **Computational Statistics & Data Analysis**, Elsevier BV, v. 43, n. 4, p. 581–603, aug 2003.
- ELLIS, G.; DIX, A. A taxonomy of clutter reduction for information visualisation. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 13, n. 6, p. 1216–1223, nov 2007.
- ERSOY, O.; HURTER, C.; PAULOVICH, F. V.; CANTAREIRO, G.; TELEA, A. Skeleton-based edge bundling for graph visualization. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 17, n. 12, p. 2364–2373, dec 2011.
- FURTADO, A. S.; ALVARES, L. O. C.; PELEKIS, N.; THEODORIDIS, Y.; BOGORNY, V. Unveiling movement uncertainty for robust trajectory similarity analysis. **International Journal of Geographical Information Science**, Informa UK Limited, v. 32, n. 1, p. 140–168, sep 2017.

- GILBERT, A.; GUGLER, J. **Cities poverty and development: Urbanization in the third world**. [S.l.]: Oxford University Press, 1992. ISBN 9780198741602.
- GOMES, G. A.; SANTOS, E.; VIDAL, C. A. Interactive visualization of traffic dynamics based on trajectory data. In: IEEE. **Graphics, Patterns and Images, 2017 30th SIBGRAPI Conference on**. [S.l.], 2017. p. 111–118.
- GOMES, G. A.; SANTOS, E.; VIDAL, C. A.; SILVA, T. L. C. da; MACEDO, J. A. F. Real-time discovery of hot routes on trajectory data streams using interactive visualization based on GPU. **Computers & Graphics**, Elsevier BV, v. 76, p. 129–141, nov 2018.
- GONZALEZ, M. C.; HIDALGO, C. A.; BARABASI, A.-L. Understanding individual human mobility patterns. **nature**, Nature publishing group, v. 453, n. 7196, p. 779, 2008.
- GROVES, M.; ENGLER, C. **PixiJS: The HTML5 Creation Engine**. Disponível em: <http://www.pixijs.com/>. Acesso em: 22 ago. 2018.
- GUDMUNDSSON, J.; KREVELD, M. van; SPECKMANN, B. Efficient detection of motion patterns in spatio-temporal data sets. In: **Proceedings of the 12th annual ACM International Workshop on Geographic Information Systems**. [S.l.]: ACM Press, 2004.
- GUI, Z.; YU, H.; TANG, Y. Locating traffic hot routes from massive taxi tracks in clusters. **J. Inf. Sci. Eng.**, v. 32, n. 1, p. 113–131, 2016.
- GUO, H.; WANG, Z.; YU, B.; Z., H.; Y., X. Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: IEEE. **PacificVis**. [S.l.], 2011. p. 163–170.
- GUO, S.; LIU, C.; GUO, Z.; FENG, Y.; HONG, F.; HUANG, H. Trajectory prediction for ocean vessels base on k-order multivariate markov chain. In: SPRINGER. **International Conference on Wireless Algorithms, Systems, and Applications**. [S.l.], 2018. p. 140–150.
- GÜTING, R. H.; ALMEIDA, V. T. de; DING, Z. Modeling and querying moving objects in networks. **The VLDB Journal**, Springer Nature, v. 15, n. 2, p. 165–190, dec 2005.
- HACHEM, F.; DAMIANI, M. L.; ISSA, H. Discovering gatherings based on individual mobility patterns: Challenges and direction. In: **2017 IEEE International Conference on Data Mining Workshops (ICDMW)**. [S.l.]: IEEE, 2017.
- HAGAN, R. D.; PHILLIPS, C. A.; LANGSTON, M. A.; RHODES, B. J. Classification and anomaly detection in traffic patterns of new york city taxis: A case study in compound analytics. In: **2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)**. [S.l.]: IEEE, 2018.
- HALL, J. V.; KRUEGER, A. B. An analysis of the labor market for uber’s driver-partners in the united states. **ILR Review**, SAGE Publications, v. 71, n. 3, p. 705–732, jun 2017.
- HAN, B.; LIU, L.; OMIECINSKI, E. Neat: Road network aware trajectory clustering. In: **ICDCS**. [S.l.: s.n.], 2012. p. 142–151.
- HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. [S.l.]: Elsevier, 2011.

- HAN, W.; DENG, Z.; CHU, J.; ZHU, J.; GAO, P.; SHAH, T. A parallel online trajectory compression approach for supporting big data workflow. **Computing**, Springer, v. 100, n. 1, p. 3–20, 2018.
- HANCOCK, M.; NACENTA, M.; GUTWIN, C.; CARPENDALE, S. The effects of changing projection geometry on the interpretation of 3d orientation on tabletops. In: **Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS09**. [S.l.]: ACM Press, 2009.
- HAUSMANN, A.; TOIVONEN, T.; SLOTOW, R.; TENKANEN, H.; MOILANEN, A.; HEIKINHEIMO, V.; MININ, E. D. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. **Conservation Letters**, Wiley, v. 11, n. 1, p. e12343, feb 2017.
- HE, S.; SHIN, K. G. (re) configuring bike station network via crowdsourced information fusion and joint optimization. In: ACM. **Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing**. [S.l.], 2018. p. 1–10.
- HE, X.; MONTILLET, J.-P.; FERNANDES, R.; BOS, M.; YU, K.; HUA, X.; JIANG, W. Review of current GPS methodologies for producing accurate time series and their error sources. **Journal of Geodynamics**, Elsevier BV, v. 106, p. 12–29, may 2017.
- HELBING, D.; HENNECKE, A.; SHVETSOV, V.; TREIBER, M. Micro-and macro-simulation of freeway traffic. **Mathematical and computer modelling**, Elsevier, v. 35, n. 5-6, p. 517–547, 2002.
- HINT.FM. **Wind Map**. Disponível em: <http://hint.fm/wind>. Acesso em: 10 ago. 2018.
- HOLTEN, D.; WIJK, J. J. van. Force-directed edge bundling for graph visualization. **Computer Graphics Forum**, Wiley, v. 28, n. 3, p. 983–990, jun 2009.
- HOREMUŽ, M.; ANDERSSON, J. V. Polynomial interpolation of GPS satellite coordinates. **GPS Solutions**, Springer Nature, v. 10, n. 1, p. 67–72, jan 2006.
- HUANG, X.; ZHAO, Y.; MA, C.; YANG, J.; YE, X.; ZHANG, C. Trajgraph: A graph-based visual analytics approach to studying urban network centralities using taxi trajectory data. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 22, n. 1, p. 160–169, 2016.
- HUGHES, R.; MACKENZIE, D. Transportation network company wait times in greater seattle, and relationship to socioeconomic indicators. **Journal of Transport Geography**, Elsevier BV, v. 56, p. 36–44, oct 2016.
- HWANG, S.; EVANS, C.; HANKE, T. Detecting stop episodes from GPS trajectories with gaps. In: **Springer Geography**. [S.l.]: Springer International Publishing, 2017. p. 427–439.
- HWANG, S.; VANDEMARK, C.; DHATT, N.; YALLA, S. V.; CREWS, R. T. Segmenting human trajectory data by movement states while addressing signal loss and signal noise. **International Journal of Geographical Information Science**, Informa UK Limited, v. 32, n. 7, p. 1391–1412, jan 2018.
- ISAKSSON, C.; DUNHAM, M.; HAHSLER, M. Sostream: Self organizing density-based clustering over data stream. **Machine Learning and Data Mining in Pattern Recognition**, Springer, p. 264–278, 2012.

JEONG, M.-H.; YIN, J.; WANG, S. Outlier Detection and Comparison of Origin-Destination Flows Using Data Depth. In: WINTER, S.; GRIFFIN, A.; SESTER, M. (Ed.). **10th International Conference on Geographic Information Science (GIScience 2018)**. [S.l.]: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018. (Leibniz International Proceedings in Informatics (LIPIcs), v. 114), p. 6:1–6:14. ISBN 978-3-95977-083-5.

JEUNG, H.; YIU, M. L.; ZHOU, X.; JENSEN, C. S.; SHEN, H. T. Discovery of convoys in trajectory databases. **Proceedings of the VLDB Endowment**, VLDB Endowment, v. 1, n. 1, p. 1068–1080, aug 2008.

JIANG, N.; GRUENWALD, L. Research issues in data stream association rule mining. **ACM Sigmod Record**, ACM, v. 35, n. 1, p. 14–19, 2006.

JIANG, R.; JIN, C.-J.; ZHANG, H.; HUANG, Y.-X.; TIAN, J.-F.; WANG, W.; HU, M.-B.; WANG, H.; JIA, B. Experimental and empirical investigations of traffic flow instability. **Transportation Research Procedia**, Elsevier BV, v. 23, p. 157–173, 2017.

JIANG, S.; CHEN, L.; MISLOVE, A.; WILSON, C. On ridesharing competition and accessibility: Evidence from uber, lyft, and taxi. In: INTERNATIONAL WORLD WIDE WEB CONFERENCES STEERING COMMITTEE. **Proceedings of the 2018 World Wide Web Conference on World Wide Web**. [S.l.], 2018. p. 863–872.

JIANG, W.; ZHANG, L. The impact of the transportation network companies on the taxi industry: Evidence from beijing's GPS taxi trajectory data. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 6, p. 12438–12450, 2018.

KALIRAJAN, K.; SUDHA, M. Moving Object Detection for Video Surveillance. **The Scientific World Journal**, Hindawi, v. 2015, p. 10, 2015.

KANCHARLA, S. R.; RAMADURAI, G. Incorporating driving cycle based fuel consumption estimation in green vehicle routing problems. **Sustainable Cities and Society**, Elsevier BV, v. 40, p. 214–221, jul 2018.

KANRAR, S.; MANDAL, N. K. Video traffic analytics for large scale surveillance. **Multimedia Tools and Applications**, Springer, v. 76, n. 11, p. 13315–13342, 2017.

KEIM, D. A.; MANSMANN, F.; SCHNEIDEWIND, J.; THOMAS, J.; ZIEGLER, H. Visual analytics: Scope and challenges. In: **Lecture Notes in Computer Science**. [S.l.]: Springer Berlin Heidelberg, 2008. p. 76–90.

KHARRAT, A.; POPA, I. S.; ZEITOUNI, K.; FAIZ, S. Clustering algorithm for network constraint trajectories. In: **Headway in Spatial Data Handling**. [S.l.]: Springer Berlin Heidelberg, 2008. p. 631–647.

KIM, S.; JEONG, S.; WOO, I.; JANG, Y.; MACIEJEWSKI, R.; EBERT, D. S. Data flow analysis and visualization for spatiotemporal statistical data without trajectory information. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 24, n. 3, p. 1287–1300, mar 2018.

KONG, X.; SONG, X.; XIA, F.; GUO, H.; WANG, J.; TOLBA, A. Lotad: long-term traffic anomaly detection based on crowdsourced bus trajectory data. **World Wide Web**, Springer, v. 21, n. 3, p. 825–847, 2018.

- KVELADZE, I.; SPEK, S. van der; KRAAK, M.-J. The recognition of temporal patterns in pedestrian behaviour using visual exploration tools. In: **7th International Conference, GIScience 2012, Columbus, OH, USA, September, Proceedings (Paperback)**. [S.l.: s.n.], 2012.
- KWON, N.; KIM, J.; LIM, S.; CHAE, J.; PARK, H. T-reduce: Route-aware mobile trajectory data reduction in transportation management systems. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–1, 2018.
- KYAW, T.; OO, N. N.; ZAW, W. Building travel speed estimation model for yangon city from public transport trajectory data. In: SPRINGER. **International Conference on Big Data Analysis and Deep Learning Applications**. [S.l.], 2018. p. 250–257.
- LAMBERT, A.; BOURQUI, R.; AUBER, D. Winding roads: Routing edges into bundles. **Computer Graphics Forum**, Wiley, v. 29, n. 3, p. 853–862, aug 2010.
- LAMPE, O. D.; HAUSER, H. Interactive visualization of streaming data with kernel density estimation. In: **2011 IEEE Pacific Visualization Symposium**. [S.l.]: IEEE, 2011.
- LANDESBERGER, T. V.; BRODKORB, F.; ROSKOSCH, P.; ANDRIENKO, N.; ANDRIENKO, G.; KERREN, A. Mobilitygraphs: Visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. **IEEE TVCG**, IEEE, v. 22, n. 1, p. 11–20, 2016.
- LANGE, R.; DÜRR, F.; ROTHERMEL, K. Efficient real-time trajectory tracking. **The VLDB Journal—The International Journal on Very Large Data Bases**, Springer-Verlag New York, Inc., v. 20, n. 5, p. 671–694, 2011.
- LEAL, E.; GRUENWALD, L.; ZHANG, J. A survey of parallel indexing techniques for large-scale moving object databases. In: **Advances in Business Information Systems and Analytics**. [S.l.]: IGI Global, 2018. p. 72–105.
- LEE, H.; LEE, H.-W.; KIM, D. Dynamic states of a continuum traffic equation with on-ramp. **Physical Review E**, APS, v. 59, n. 5, p. 5101, 1999.
- LEE, J.-G.; HAN, J.; WHANG, K.-Y. Trajectory clustering: a partition-and-group framework. In: **Proceedings of the 2007 ACM SIGMOD international conference on Management of data - SIGMOD07**. [S.l.]: ACM Press, 2007.
- LEVIS, D.; LIEBCHEN, B. Micro-flock patterns and macro-clusters in chiral active brownian disks. **Journal of Physics: Condensed Matter**, IOP Publishing, v. 30, n. 8, p. 084001, 2018.
- LI, M. Z.; SUH, D. Y.; RYERSON, M. S. Visualizing aviation impacts: Modeling current and future flight trajectories with publicly available flight data. **Transportation Research Part D: Transport and Environment**, Elsevier BV, v. 63, p. 769–785, aug 2018.
- LI, X.; CUI, J.; AN, S.; PARSAFARD, M. Stop-and-go traffic analysis: Theoretical properties, environmental impacts and oscillation mitigation. **Transportation Research Part B**, Elsevier, v. 70, p. 319–339, 2014.
- LI, X.; HAN, J.; LEE, J.-G.; GONZALEZ, H. Traffic density-based discovery of hot routes in road networks. **Advances in Spatial and Temporal Databases**, Springer, p. 441–459, 2007.

- LI, Y.; HAN, J.; YANG, J. Clustering moving objects. In: **Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD04**. [S.l.]: ACM Press, 2004.
- LI, Z.; DING, B.; HAN, J.; KAYS, R. Swarm: Mining relaxed temporal moving object clusters. **Proceedings of the VLDB Endowment**, VLDB Endowment, v. 3, n. 1-2, p. 723–734, sep 2010.
- LI, Z.; LEE, J.-G.; LI, X.; HAN, J. Incremental clustering for trajectories. In: **Database Systems for Advanced Applications**. [S.l.]: Springer Berlin Heidelberg, 2010. p. 32–46.
- LIENHARD, J. H. **A heat transfer textbook**. [S.l.]: Courier Corporation, 2013.
- LIN, C.-Y.; WANG, Y.-C.; FU, W.-T.; CHEN, Y.-S.; CHIEN, K.-C.; LIN, B.-Y. Efficiently preserving privacy on large trajectory datasets. In: **2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)**. [S.l.]: IEEE, 2018.
- LIN, L.; LI, J.; CHEN, F.; YE, J.; HUAI, J. Road traffic speed prediction: A probabilistic model fusing multi-source data. **IEEE Transactions on Knowledge and Data Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 30, n. 7, p. 1310–1323, jul 2018.
- LIN, X.; MA, S.; ZHANG, H.; WO, T.; HUAI, J. One-pass error bounded trajectory simplification. **Proceedings of the VLDB Endowment**, VLDB Endowment, v. 10, n. 7, p. 841–852, mar 2017.
- LIU, S.; PU, J.; LUO, Q.; QU, H.; NI, L. M.; KRISHNAN, R. VAIT: A visual analytics system for metropolitan transportation. **IEEE Transactions on Intelligent Transportation Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, n. 4, p. 1586–1596, dec 2013.
- LIU, X. C.; TAYLOR, J.; PORTER, R. J.; WEI, R. Using trajectory data to explore roadway characterization for bikeshare network. **Journal of Intelligent Transportation Systems**, Informa UK Limited, p. 1–17, feb 2018.
- LORENZO, G. D.; SBODIO, M.; CALABRESE, F.; B., M.; PINELLI, F.; NAIR, R. AllAboard: visual exploration of cellphone mobility data to optimise public transport. **IEEE TVCG**, v. 22, n. 2, p. 1036–1050, 2016.
- LUKMAN, R. K.; CERINŠEK, M.; VIRTIČ, P.; HORVAT, B. Improving efficient resource usage and reducing carbon dioxide emissions by optimizing fleet management for winter services. **Journal of Cleaner Production**, Elsevier BV, v. 177, p. 1–11, mar 2018.
- MAKRIS, D.; ELLIS, T.; BLACK, J. Bridging the gaps between cameras. In: IEEE. **Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on**. [S.l.], 2004. v. 2, p. II–II.
- MAO, J.; SONG, Q.; JIN, C.; ZHANG, Z.; ZHOU, A. Tscluwin: Trajectory stream clustering over sliding window. In: SPRINGER. **International Conf. on DSAA**. [S.l.], 2016. p. 133–148.
- MARKOVIC, N.; SEKULA, P.; LAAN, Z. V.; ANDRIENKO, G.; ANDRIENKO, N. Applications of trajectory data from the perspective of a road transportation agency: Literature review and maryland case study. **IEEE Transactions on Intelligent Transportation Systems**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–12, 2018.

MATASCI, G.; HERMOSILLA, T.; WULDER, M. A.; WHITE, J. C.; HOBART, G. W.; ZALD, H. S. J.; COOPS, N. C. A space-time data cube: Multi-temporal forest structure maps from landsat and lidar. In: **2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**. [S.l.]: IEEE, 2017.

NASERIAN, E.; WANG, X.; XU, X.; DONG yu-ning. A framework of loose travelling companion discovery from human trajectories. **IEEE Transactions on Mobile Computing**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–1, 2018.

NI, D. **Traffic flow theory: Characteristics, experimental methods, and numerical techniques**. [S.l.]: Butterworth-Heinemann, 2015.

NI, D. **Traffic Flow Theory: Characteristics, Experimental Methods, and Numerical Techniques**. [S.l.]: Butterworth-Heinemann, 2015. ISBN 9780128041475.

NOGUEIRA, T. P.; BRAGA, R. B.; OLIVEIRA, C. T. de; MARTIN, H. FrameSTEP: A framework for annotating semantic trajectories based on episodes. **Expert Systems with Applications**, Elsevier BV, v. 92, p. 533–545, feb 2018.

ORAKZAI, F.; CALDERS, T.; PEDERSEN, T. B. Distributed convoy pattern mining. In: **2016 17th IEEE International Conference on Mobile Data Management (MDM)**. [S.l.]: IEEE, 2016.

PACK, M. L.; WONGSUPHASAWAT, K.; VANDANIKER, M.; FILIPPOVA, D. Ice–visual analytics for transportation incident datasets. In: IEEE. **Information Reuse & Integration, 2009. IRI'09. IEEE International Conference on**. [S.l.], 2009. p. 200–205.

PAHINS, C. A. L.; STEPHENS, S. A.; SCHEIDEGGER, C.; COMBA, J. L. D. Hashedcubes: Simple, low memory, real-time visual exploration of big data. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 23, n. 1, p. 671–680, jan 2017.

PARK, H.; OH, C.; MOON, J.; KIM, S. Development of a lane change risk index using vehicle trajectory data. **Accident Analysis & Prevention**, Elsevier BV, v. 110, p. 1–8, jan 2018.

PATIL, V.; SINGH, P.; PARIKH, S.; ATREY, P. K. GeoSClean: Secure cleaning of GPS trajectory data using anomaly detection. In: **2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)**. [S.l.]: IEEE, 2018.

PERIN, C.; WUN, T.; PUSCH, R.; CARPENDALE, S. Assessing the graphical perception of time and speed on 2d+time trajectories. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 24, n. 1, p. 698–708, jan 2018.

POCO, J.; DORAISWAMY, H.; VO, H.; COMBA, J. L.; FREIRE, J.; SILVA, C. *et al.* Exploring traffic dynamics in urban environments using vector-valued functions. In: WILEY O. LIBRARY. **CGF**. [S.l.], 2015. v. 34, n. 3, p. 161–170.

POLKINGHORNE, D. E. Language and meaning: Data collection in qualitative research. **Journal of counseling psychology**, American Psychological Association, v. 52, n. 2, p. 137, 2005.

PU, J.; LIU, S.; DING, Y.; QU, H.; NI, L. T-watcher: A new visual analytic system for effective traffic surveillance. In: IEEE. **Mobile Data Management (MDM), 2013 IEEE 14th International Conference on**. [S.l.], 2013. v. 1, p. 127–136.

PUNTHEERANURAK, S.; SHEIN, T. T.; IMAMURA, M. Efficient discovery of traveling companion from evolving trajectory data stream. In: IEEE. **2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)**. [S.l.], 2018. p. 448–453.

QIAN, H.; LU, Y. Simplifying GPS trajectory data with enhanced spatial-temporal constraints. **ISPRS International Journal of Geo-Information**, MDPI AG, v. 6, n. 11, p. 329, oct 2017.

RAJALAKSHMI, K.; GOYAL, M. Location-based services: Current state of the art and future prospects. In: **Optical and Wireless Technologies**. [S.l.]: Springer, 2018. p. 625–632.

ROGERS, B. The social costs of uber. **U. Chi. L. Rev. Dialogue**, HeinOnline, v. 82, p. 85, 2015.

ROY, P.; PATRA, N.; MUKHERJEE, A.; ASHOUR, A. S.; DEY, N.; BISWAS, S. P. Intelligent Traffic Monitoring System Through Auto and Manual Controlling using PC and Android Application. In: **Applied Video Processing in Surveillance and Monitoring Systems**. [S.l.]: IGI Global, 2017. p. 244–262.

SACHARIDIS, D.; PATROUMPAS, K.; TERROVITIS, M.; KANTERE, V.; POTAMIAS, M.; MOURATIDIS, K.; SELLIS, T. On-line discovery of hot motion paths. In: ACM. **Advances in database technology**. [S.l.], 2008. p. 392–403.

SANTIPANTAKIS, G. M.; VOUIROS, G. A.; GLENIS, A.; DOULKERIDIS, C.; VLACHOU, A. The datAcron ontology for semantic trajectories. In: **Lecture Notes in Computer Science**. [S.l.]: Springer International Publishing, 2017. p. 26–30.

SCHEEPENS, R.; HURTER, C.; WETERING, H. V. D.; WIJK, J. J. V. Visualization, selection, and analysis of traffic flows. **IEEE TVCG**, v. 22, n. 1, p. 379–388, 2016.

SCHEEPENS, R.; WILLEMS, N.; WETERING, H. van de; ANDRIENKO, G.; ANDRIENKO, N.; WIJK, J. J. van. Composite density maps for multivariate trajectories. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 17, n. 12, p. 2518–2527, dec 2011.

SCHEEPENS, R.; WILLEMS, N.; WETERING, H. van de; WIJK, J. J. van. Interactive visualization of multivariate trajectory data with density maps. In: **2011 IEEE Pacific Visualization Symposium**. [S.l.]: IEEE, 2011.

SILVA, J. A.; FARIA, E. R.; BARROS, R. C.; HRUSCHKA, E. R.; CARVALHO, A. C. de; GAMA, J. Data stream clustering: A survey. **ACM Computing Surveys (CSUR)**, ACM, v. 46, n. 1, p. 13, 2013.

SILVA, T. Coelho da; ZEITOUNI, K.; MACÊDO, J. de. Online clustering of trajectory data streams. In: **MDM**. [S.l.: s.n.], 2016.

SOARES JUNIOR, A.; TIMES, V. C.; RENSO, C.; MATWIN, S.; CABRAL, L. A. A semi-supervised approach for the semantic segmentation of trajectories. In: **2018 19th IEEE International Conference on Mobile Data Management (MDM)**. [S.l.]: IEEE, 2018.

STERN, R. E.; CUI, S.; MONACHE, M. L. D.; BHADANI, R.; BUNTING, M.; CHURCHILL, M.; HAMILTON, N.; HAULCY, R.; POHLMANN, H.; WU, F.; PICCOLI, B.; SEIBOLD, B.; SPRINKLE, J.; WORK, D. B. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. **Transportation Research Part C: Emerging Technologies**, Elsevier BV, v. 89, p. 205–221, apr 2018.

SUI, P.; YANG, X. A privacy-preserving compression storage method for large trajectory data in road network. **Journal of Grid Computing**, Springer Nature, v. 16, n. 2, p. 229–245, feb 2018.

SUN, H.; HU, S.; MCINTOSH, S.; CAO, Y. Big data trip classification on the new york city taxi and uber sensor network. **Journal of Internet Technology**, v. 19, n. 2, p. 591–598, 2018.

TANG, L.-A.; ZHENG, Y.; YUAN, J.; HAN, J.; LEUNG, A.; HUNG, C.-C.; PENG, W.-C. On discovery of traveling companions from streaming trajectories. In: **2012 IEEE 28th International Conference on Data Engineering**. [S.l.]: IEEE, 2012.

TERROVITIS, M.; POULIS, G.; MAMOULIS, N.; SKIADOPOULOS, S. Local suppression and splitting techniques for privacy preserving publication of trajectories. **IEEE Transactions on Knowledge and Data Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 29, n. 7, p. 1466–1479, jul 2017.

THOMAS, F.; JOHNSTON, O.; THOMAS, F. **The illusion of life: Disney animation**. [S.l.]: Hyperion New York, 1995.

THUDT, A.; BAUR, D.; CARPENDALE, S. Visits: A spatiotemporal visualization of location histories. In: **Proceedings of the eurographics conference on visualization**. [S.l.: s.n.], 2013. p. 79–83.

TOMINSKI, C.; SCHUMANN, H.; ANDRIENKO, G.; ANDRIENKO, N. Stacking-based visualization of trajectory attribute data. **IEEE TVCG**, IEEE, v. 18, n. 12, p. 2565–2574, 2012.

TREIBER, M. **Microsimulation of Traffic Flow**. Disponível em: <http://www.traffic-simulation.de>. Acesso em: 17 maio 2018.

TREIBER, M.; HELBING, D. Realistische mikrosimulation von strassenverkehr mit einem einfachen modell. In: **16th Symposium Simulationstechnik ASIM**. [S.l.: s.n.], 2002. v. 2002, p. 80.

TREIBER, M.; HENNECKE, A.; HELBING, D. Congested traffic states in empirical observations and microscopic simulations. **Physical review E**, APS, v. 62, n. 2, p. 1805, 2000.

TREIBER, M.; KESTING, A. **Traffic Flow Dynamics: Data, Models and Simulation**. [S.l.]: Springer, 2013. ISBN 978-3-642-32460-4.

TSENG, C.-M.; CHAU, S. C.-K.; LIU, X. Improving viability of electric taxis by taxi service strategy optimization: A big data study of new york city. **IEEE Transactions on Intelligent Transportation Systems**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–13, 2018.

TU, W.; CAO, R.; YUE, Y.; ZHOU, B.; LI, Q.; LI, Q. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. **Journal of Transport Geography**, Elsevier BV, v. 69, p. 45–57, may 2018.

- TVERSKY, B.; MORRISON, J. B.; BETRANCOURT, M. Animation: can it facilitate? **International Journal of Human-Computer Studies**, Elsevier BV, v. 57, n. 4, p. 247–262, oct 2002.
- UBER. **Deck.gl**. Disponível em: <http://uber.github.io/deck.gl>. Acesso em: 20 ago. 2018.
- VISHNU, V. M.; RAJALAKSHMI, M.; NEDUNCHEZHIAN, R. Intelligent traffic video surveillance and accident detection system with dynamic traffic signal control. **Cluster Computing**, Springer, p. 1–13, 2017.
- WANG, S. J.; MORIARTY, P. Big data for sustainable urban transport. In: **Big Data for Urban Sustainability**. [S.l.]: Springer, 2018. p. 81–103.
- WANG, X.; ZHAO, S.; DONG, L. Research and application of traffic visualization based on vehicle gps big data. In: SPRINGER. **International Conference on Intelligent Transportation**. [S.l.], 2016. p. 293–302.
- WANG, Y.; ARCHAMBAULT, D.; SCHEIDEGGER, C. E.; QU, H. A vector field design approach to animated transitions. **IEEE Transactions on Visualization and Computer Graphics**, Institute of Electrical and Electronics Engineers (IEEE), v. 24, n. 9, p. 2487–2500, sep 2018.
- WANG, Y.; LEE, K.; LEE, I. DNA-chart visual tool for topological higher order information from spatio-temporal trajectory dataset. **Expert Systems with Applications**, Elsevier BV, v. 108, p. 28–35, oct 2018.
- WANG, Z.; LU, M.; YUAN, X.; ZHANG, J.; WETERING, H. V. D. Visual traffic jam analysis based on trajectory data. **IEEE TVCG**, v. 19, n. 12, p. 2159–2168, 2013.
- WARE, C. **Information visualization: perception for design**. [S.l.]: Elsevier, Amsterdam, Netherlands, 2012.
- WEINTRIT, A.; KOPACZ, P. A novel approach to loxodrome (rhumb-line), orthodrome (great circle) and geodesic line in ecdis and navigation in general. **Methods and Algorithms in Navigation, Marine Navigation and Safety of Sea Transportation. Leiden: A Balkema Book, CRC Press, Taylor & Francis Grup**, p. 123–132, 2011.
- WEISKOPF, D. Vector field visualization. In: **GPU-Based Interactive Visualization Techniques**. [S.l.]: Springer Berlin Heidelberg, 2014. p. 81–159.
- WOLSHON, B. One-Way-Out: Contraflow Freeway Operation for Hurricane Evacuation. **Natural Hazards Review**, American Society of Civil Engineers, v. 2, n. 3, p. 105–112, 2001.
- WU, G. P.; CHAN, K. C. Privacy-preserving trajectory classification of driving trip data based on pattern discovery techniques. In: **2017 IEEE International Conference on Big Data (Big Data)**. [S.l.]: IEEE, 2017.
- WU, J.; ZHU, Y.; CHEN, H. Road network construction and hot routes inferring with vehicular trajectories. In: SPRINGER. **International Conference on Intelligent Computing**. [S.l.], 2014. p. 351–362.
- WU, R.; LUO, G.; SHAO, J.; TIAN, L.; PENG, C. Location prediction on trajectory data: A review. **Big Data Mining and Analytics**, Tsinghua University Press, v. 1, n. 2, p. 108–127, jun 2018.

- XIA, F.; WANG, J.; KONG, X.; WANG, Z.; LI, J.; LIU, C. Exploring human mobility patterns in urban scenarios: A trajectory data perspective. **IEEE Communications Magazine**, Institute of Electrical and Electronics Engineers (IEEE), v. 56, n. 3, p. 142–149, mar 2018.
- XIE, Z.; YAN, J. Kernel density estimation of traffic accidents in a network space. **Computers, Environment and Urban Systems**, Elsevier BV, v. 32, n. 5, p. 396–406, sep 2008.
- XU, Y.; BELYI, A.; BOJIC, I.; RATTI, C. Human mobility and socioeconomic status: Analysis of singapore and boston. **Computers, Environment and Urban Systems**, Elsevier BV, apr 2018.
- XU, Y.; SHAW, S.-L.; LU, F.; CHEN, J.; LI, Q. Uncovering the relationships between phone communication activities and spatiotemporal distribution of mobile phone users. In: **Human Dynamics Research in Smart and Connected Communities**. [S.l.]: Springer, 2018. p. 41–65.
- YANG, H.-L.; LIN, S.-L. User adoption of location-based service. In: **2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)**. [S.l.]: IEEE, 2018.
- YANG, L.; KWAN, M.-P.; PAN, X.; WAN, B.; ZHOU, S. Scalable space-time trajectory cube for path-finding: A study using big taxi trajectory data. **Transportation Research Part B: Methodological**, Elsevier BV, v. 101, p. 1–27, jul 2017.
- YANG, X.; TANG, L.; ZHANG, X.; LI, Q. A data cleaning method for big trace data using movement consistency. **Sensors**, MDPI AG, v. 18, n. 3, p. 824, mar 2018.
- YAO, D.; ZHANG, C.; ZHU, Z.; HUANG, J.; BI, J. Trajectory clustering via deep representation learning. In: **2017 International Joint Conference on Neural Networks (IJCNN)**. [S.l.]: IEEE, 2017.
- YU, Y.; WANG, Q.; WANG, X.; WANG, H.; HE, J. Online clustering for trajectory data stream of moving objects. **Computer science and information systems**, v. 10, n. 3, p. 1293–1317, 2013.
- YUAN, G.; SUN, P.; ZHAO, J.; LI, D.; WANG, C. A review of moving object trajectory clustering algorithms. **Artificial Intelligence Review**, Springer, v. 47, n. 1, p. 123–144, 2017.
- YUAN, J.; ZHENG, Y.; XIE, X.; SUN, G. Driving with knowledge from the physical world. In: **Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD11**. [S.l.]: ACM Press, 2011.
- YUAN, J.; ZHENG, Y.; ZHANG, C.; XIE, W.; XIE, X.; SUN, G.; HUANG, Y. T-drive: driving directions based on taxi trajectories. In: **ACM. Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems**. [S.l.], 2010. p. 99–108.
- ZENG, W.; FU, C.-W.; ARISONA, S. M.; QU, H. Visualizing interchange patterns in massive movement data. **Computer Graphics Forum**, Wiley, v. 32, n. 3pt3, p. 271–280, jun 2013.
- ZHANG, A.; SONG, S.; WANG, J.; YU, P. S. Time series data cleaning: from anomaly detection to anomaly repairing. **Proceedings of the VLDB Endowment**, VLDB Endowment, v. 10, n. 10, p. 1046–1057, jun 2017.
- ZHANG, D.; LEE, K.; LEE, I. Hierarchical trajectory clustering for spatio-temporal periodic pattern mining. **Expert Systems with Applications**, Elsevier BV, v. 92, p. 1–11, feb 2018.

- ZHANG, G.-Y.; ZHANG, J. Trajectory clustering based on trajectory structure and longest common subsequence. **Transactions on Computer Science and Engineering**, DEStech Publications, aug 2018.
- ZHANG, Y.; MI, Z. Environmental benefits of bike sharing: A big data-based analysis. **Applied Energy**, Elsevier BV, v. 220, p. 296–301, jun 2018.
- ZHENG, K.; ZHENG, Y.; YUAN, N. J.; SHANG, S. On discovery of gathering patterns from trajectories. In: **2013 IEEE 29th International Conference on Data Engineering (ICDE)**. [S.l.]: IEEE, 2013.
- ZHENG, K.; ZHENG, Y.; YUAN, N. J.; SHANG, S.; ZHOU, X. Online discovery of gathering patterns over trajectories. **IEEE Transactions on Knowledge and Data Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 26, n. 8, p. 1974–1988, aug 2014.
- ZHENG, Y. Trajectory data mining: An overview. **ACM Transactions on Intelligent Systems and Technology**, Association for Computing Machinery (ACM), v. 6, n. 3, p. 1–41, may 2015.
- ZHENG, Y.; CAPRA, L.; WOLFSON, O.; YANG, H. Urban computing: concepts, methodologies, and applications. **ACM Transactions on Intelligent Systems and Technology (TIST)**, ACM, v. 5, n. 3, p. 38, 2014.
- ZHENG, Y.; ZHOU, X. (Ed.). **Computing with Spatial Trajectories**. [S.l.]: Springer Nature, 2011.
- ZHENG, Y.; ZHOU, X. **Computing with spatial trajectories**. [S.l.]: Springer Science & Business Media, 2011.