

MODELAGEM E ANÁLISE DE DADOS DINÂMICOS DO TRÁFEGO URBANO: UMA REVISÃO CONCEITUAL E APLICADA

Hamifrancey Brito Meneses

Carlos Felipe Grangeiro Loureiro

Programa de Mestrado em Engenharia de Transportes - PETRAN
Universidade Federal do Ceará - UFC

RESUMO

A aplicação da tecnologia da informação à gestão do tráfego tem viabilizado a coleta de um grande volume de dados sobre os deslocamentos urbanos. O processamento adequado destes dados pode subsidiar múltiplas aplicações na operação e no planejamento da malha viária. Neste contexto, o presente trabalho tem como objetivo apresentar uma discussão conceitual sobre a modelagem e análise de dados dinâmicos do tráfego, de modo a orientar o uso de dados gerados por sistemas centralizados de controle semafórico. Inicialmente, são apresentados conceitos básicos de modelagem espaço-temporal de dados dinâmicos, com ênfase em modelos que incorporam a dimensão temporal aos SIG. Em seguida, são discutidos procedimentos de aferição de qualidade e modelos de mineração de dados dinâmicos. Depois, é efetuada uma revisão sobre análise espacial, incluindo ferramentas e recursos de visualização espacial. A discussão destes temas busca disponibilizar referências para o usuário de dados dinâmicos do tráfego urbano, visando auxiliar a reflexão e o uso prático destas tecnologias.

ABSTRACT

The application of the information technology to traffic management has made possible the collection of a great volume of data regarding urban movements. The appropriate data processing may help the development of multiple applications in urban traffic operations and planning. In such a context, this paper presents a theoretical review on dynamic traffic data modeling and analysis, in order to guide the application of UTC data. First, basic concepts about spatial and temporal modeling of dynamic data are presented, with emphasis on models that incorporate the time dimension to GIS. Second, this work discusses some concepts related to quality procedures and data mining models to dynamic data. Lastly, it is presented a theoretical and applied review on spatial analysis, including spatial visualization tools and resources. In discussing these issues, the main objective is to provide key references to urban traffic data analysts, helping to spread the usage of such technology.

1. INTRODUÇÃO

A população, as atividades econômicas e as facilidades de transportes mudam seu estado, suas configurações e seus valores de atributo ao longo do tempo, o que acarreta um aspecto dinâmico às atividades humanas (Ichikawa *et al.*, 2002). Por sua vez, o dinamismo das atividades humanas influencia diretamente na variabilidade espacial e temporal dos deslocamentos urbanos. O conhecimento contínuo desta dinâmica urbana tem papel fundamental na gestão eficiente e eficaz do tráfego urbano (Stathopoulos e Karlaftis, 2001).

Contudo, o diagnóstico do comportamento do tráfego urbano requer um esforço periódico de coleta e tratamento de dados, o que demanda consideráveis recursos financeiros e mão-de-obra especializada (Meneses, 2003). Uma alternativa eficiente para o monitoramento da dinâmica urbana consiste em operar sistemas em tempo real de Controle do Tráfego em Área (CTA). Como estes sistemas coletam dados continuamente, a sua base dinâmica de dados torna-se cada vez mais extensa, forçando alguns sistemas a limitarem o período de armazenagem dos dados. Porém, em seu trabalho diário, os engenheiros de tráfego necessitam de informações mínimas, representativas e acessíveis, sob pena de sobrecarga e conseqüente descarte de dados (Chen *et al.*, 2001). Assim, é preciso usar tecnologias específicas para aferir a qualidade dos dados, extrair informações específicas de grandes volumes de dados, assim

como apresentar estas informações num ambiente computacional capaz de representar o sistema de transporte urbano usando conceitos espaciais familiares aos técnicos.

Diante do uso potencial de dados de sistemas CTA para auxiliar a gestão do tráfego urbano, este trabalho tem como objetivo apresentar e discutir conceitos relativos à modelagem e análise de dados dinâmicos do tráfego urbano. Inicialmente, é abordada a modelagem espaço-temporal deste tipo de dado em ambiente SIG, apresentando-se o modelo espaço-temporal usado para importar e modelar dados de tráfego coletados pelo sistema CTA de Fortaleza (Meneses, 2003). Em seguida, são discutidos conceitos que abrangem a avaliação da qualidade e a mineração de dados de sistemas CTA. Depois, são apresentados conceitos básicos relativos à análise espacial de dados dinâmicos em ambiente SIG. Em cada um destes temas são apresentadas, para fins de ilustração, aplicações práticas destes conceitos na análise do sistema de tráfego urbano.

2. MODELAGEM DE DADOS DINÂMICOS DE TRÁFEGO EM AMBIENTE SIG

2.1. Modelagem de sistemas de transportes

A representação do mundo real implica na abstração de objetos e/ou fenômenos da realidade, à luz de um interesse particular, a qual deve ser armazenada numa base de dados específica. As características e os relacionamentos das entidades modeladas devem ser definidos de modo a viabilizar o funcionamento e a aplicação da modelagem desenvolvida (Davis, 1998). No caso de sistemas de transportes, a representação da realidade deve considerar a localização (espaço) e o dinamismo (tempo) de objetos e fenômenos que compõem estes sistemas, modelando atributos e relacionamentos das entidades geográficas. Esta modelagem deve usar os sistemas métricos padrões de aferição do espaço e do tempo, obedecendo as fundamentações teóricas que regem o funcionamento dos sistemas de transportes (Koncz e Adams, 2002).

Spaccapietra *et al.* (1998) e Borges (1997) recomendam que o processo de modelagem de dados geográficos em transportes deve ser orientado pelos seguintes critérios:

- a) Simplicidade: relativo a uma abstração simplificada da realidade, mas adequada às aplicações fins e compatível com a concepção cognitiva humana;
- b) Formalidade: relativo ao rigor formal na definição de conceitos usados no modelo, visando evitar inconsistências e ambigüidades;
- c) Expressividade: relativo à capacidade de reprodução consistente da realidade, por meio de diferentes classes e relacionamentos entre objetos, com diversos níveis de agregação;
- d) Ortogonalidade: relativo à capacidade de modelar o espaço e o tempo de forma independente, e à habilidade de combinar estas dimensões para aumento de expressividade;
- e) Manipulação: relativo ao poder de acesso (operações de consulta), descrição e manipulação de dados (operações analíticas), condicionado pela modelagem;
- f) Integridade: relativo à especificação de regras de integridade espacial e temporal das entidades e fenômenos modelados;

- g) Visualização: relativo à facilidade de visualizar e compreender as entidades e fenômenos modelados, por meio de múltiplos níveis de informação e variadas formas de visualização de entidades geográficas, em distintas versões temporais.

2.2. Modelagem de dados dinâmicos do tráfego urbano em ambiente SIG

As entidades do tráfego urbano (infra-estrutura, veículos, pessoas, cargas) estão localizadas e se deslocam no espaço, e seus fenômenos, atributos e relacionamentos são dinâmicos ao longo do tempo. Este caráter espacial e dinâmico faz dos Sistemas de Informações Geográficas (SIG) uma alternativa valiosa para a modelagem do tráfego urbano (Jia, 1999), embora os SIG estejam limitados à representação estática da realidade (Langran, 1993; Paret *et al.*, 1999).

No aspecto espacial, os SIG viabilizam a representação das entidades do sistema de tráfego urbano, bem como seus atributos e relacionamentos topológicos, por meio de uma estrutura de base de dados dual (características geográficas e descritivas) interativa e compatível com os conceitos espaciais familiares ao usuário. A capacidade de visualização e análise espacial, conferida por esta estrutura dual, permite que os SIG possibilitem a elucidação de relações e padrões não triviais do comportamento do tráfego urbano (Aranha, 2000). Já no âmbito temporal, embora tenha ocorrido uma difusão do uso do SIG em transportes, os sistemas convencionais ainda estão limitados à representação estática da realidade. Isto torna necessário uma adaptação ou remodelação da arquitetura de base de dados dos SIG, de modo a suportar o caráter dinâmico dos sistemas de transporte (Michinori, 2002).

Esta nova estrutura deve dispor das seguintes funções de manuseio de dados dinâmicos, conforme Langran (1993): inventário, análise, atualização, controle de qualidade, programação e apresentação de dados. O inventário consiste na armazenagem e descrição de áreas de estudo, incluindo modificações físicas e eventos de alteração computacional da base de dados. A análise envolve a investigação de alterações em padrões espaciais ao longo do tempo, bem como a determinação de causas ou efeitos destas mudanças. A atualização permite assegurar que os dados disponíveis são os mais recentes possíveis. O controle de qualidade visa minimizar a chance de erro, checando a validade e a consistência de novos dados em comparação a séries históricas. A programação permite efetuar operações padronizadas sobre a base de dados, como a eliminação de dados defasados, por exemplo. A visualização consiste na geração e apresentação da evolução temporal de um fenômeno real, devendo ser tanto qualitativa quanto quantitativa.

A implementação destas funcionalidades tem sido viabilizada por vários modelos que buscam representar o espaço e o tempo, em ambiente SIG, de forma integrada e complementar. Zhao *et al.* (1997) e Miller e Shaw (2001) descrevem seis modelos usados para incorporar a dimensão temporal ao SIG: modelo *snapshot*, modelo de atualização, modelo espaço-tempo, modelo multidimensional, modelo integrado e modelo de versões.

O primeiro, modelo *snapshot*, apresenta dados relativos a diferentes instantes de tempo, por meio de uma série sucessiva de camadas geográficas (*fotografias da realidade*). Cada camada descreve um mesmo fenômeno real, por meio de uma estrutura relacional, ao longo de seu espaço-tempo de abrangência. Este tipo de modelo pode implicar em elevada replicação de dados, no caso de fenômenos cujas entidades não sofrem alterações frequentes em sua forma, localização e atributos. Além disto, há uma carência de interligação entre os dados

armazenados em diferentes camadas. Contudo, as operações de consulta e manipulação dos dados são bastante simplificadas, devido à segregação de distintos estados temporais das entidades em diferentes arquivos.

O segundo, modelo de atualização, armazena uma representação totalizada da realidade em uma base de dados relacional única. Quando as entidades modeladas sofrem modificações, esta mudança é atualizada nesta base, por meio da criação de um novo registro. Apenas os atributos relativos às mudanças são atualizados neste registro. Este modelo minimiza a necessidade de arquivamento de dados, efetuando a ligação entre eventos em diferentes instantes. O terceiro, modelo espaço-tempo, consiste numa evolução do modelo de atualização. Este modelo cria novos objetos discretos, com sua própria história, a partir do objeto primitivo, quando ocorrem mudanças na realidade. Os objetos são armazenados numa estrutura relacional. Cada novo objeto contém dados temporais que indicam o período de tempo no qual os atributos associados são válidos. Como resultado, a topologia temporal é mantida.

O quarto, modelo multidimensional, trata o tempo como uma quarta dimensão, definindo cada entidade espacial por coordenadas (x,y,t) ou (x,y,z,t) . A inclusão do tempo como quarta dimensão torna a topologia mais complexa e não explicita as entidades espaciais, dificultando a modelagem conceitual e lógica. O quinto, modelo integrado, incorpora três domínios no processo de modelagem de fenômenos do mundo real: espacial, temporal e semântico.

O sexto, modelo de versões, produz uma nova versão da realidade quando ocorrem mudanças nas características da área de estudo. A incorporação destas alterações é feita numa base de dados relacional, por meio de três tipos de estruturas: tabelas, registros e atributos (Huiguo *et al.*, 2002). A estrutura de tabelas consiste em criar uma nova versão da base de dados, por meio de uma nova tabela, sempre que ocorrem modificações nas entidades modeladas. Este tipo de modelo ocasiona alta duplicação de dados no caso de fenômenos mais estáticos, bem como carece de interligação entre as várias versões. Contudo, esta estrutura facilita as operações de recuperação de dados.

Já a estrutura de registros consiste em acrescentar uma linha de registro numa base de dados única (tabela), cada vez que uma entidade sofre alterações. Esta estrutura implica em menor replicação de dados que a primeira. Entretanto, parte dos dados do novo registro é replicada e as operações analíticas sobre a base de dados requerem maior capacidade de processamento que a estrutura de tabelas, devido aos vários estados da realidade representados numa mesma tabela. Por sua vez, a estrutura de atributos cria uma nova coluna para registrar as modificações nas características de entidades modeladas. Esta estrutura é a mais compacta em termos de espaço computacional, sendo a mais extensa em termos de número de atributos, em comparação às duas anteriores. Contudo, a manipulação dos atributos é bastante complexa, devido ao crescimento do número de atributos gerados ao longo do tempo.

Cabe destacar que qualquer uma das estruturas relacionais do modelo espaço-temporal de *versões* não busca reformular a arquitetura da base de dados do ambiente SIG. Na verdade, este modelo adapta a estrutura relacional dual do SIG de modo a incorporar a dimensão temporal.

2.3. Aplicação prática de modelo espaço-temporal de versões

O sistema CTA de Fortaleza - CTAFOR dispõe de uma base de dados dinâmica do tráfego na área controlada. Esta base agrega 18 variáveis de tráfego (ex. fluxo veicular, atraso) para sete categorias de elementos viários (região, área, rota, semáforo, logradouro, detector, estágio), conforme modelado pelo sistema *Split Cycle Optimization Offset Time* (SCOOT) de controle semafórico em tempo real. O acesso a esta base de dados é feito por meio do *software Automatic SCOOT Traffic Information Database* (ASTRID), sub-sistema do SCOOT. O ASTRID permite consultar dados desagregados até o nível de 15 minutos, para períodos de até doze meses, individualmente, por mês e elemento viário consultados (TRL, 2000).

Meneses (2003) desenvolveu uma interface lógica, denominada de TRANSCOOT, para a importação de dados da base dinâmica do CTAFOR, para o SIG TransCAD (Caliper, 1996), com base numa modelagem que combina uma estrutura flexível de representação de entidades viárias e um modelo espaço-temporal de versões para representar a dinâmica do tráfego urbano. No âmbito espacial, a modelagem em questão usa a estrutura genérica apresentada na Tabela 1 para representar as sete categorias de entidades viárias modeladas pelo SCOOT. Contudo, o usuário pode variar a forma de representação espacial destas entidades, tendo em vista que a elaboração da base geográfica no TransCAD fica a critério do usuário.

Tabela 1: Estrutura de representação geográfica de entidades modeladas pelo SCOOT

Tipo de entidade viária modelada pelo SCOOT	Correspondente entidade geográfica no SIG
Semáforo, Estágio	Ponto
Detector, Logradouro (<i>link</i>), Rota	Linha
Região, Área	Polígono

Fonte: Meneses (2003)

Quanto à dimensão temporal, a modelagem em estudo usa um modelo de versões para representar dados dinâmicos de tráfego em ambiente SIG. Este modelo combina estruturas relacionais de tabelas e de atributos para consolidar dados do sistema SCOOT no SIG TransCAD. Estas estruturas geram um conjunto de tabelas distintas, na qual cada tabela representa uma versão agregada da realidade modelada pelo SCOOT, em um dado escopo estatístico-temporal. O escopo estatístico é definido pelo usuário dentre as opções: soma, mínimo, máximo, média e desvio padrão. O escopo temporal de cálculo destas estatísticas é delimitado pelos níveis de agregação dos dias e dos horários pesquisados. A agregação do período de dias requisitado deve ser definida entre as opções: diária, semanal, mensal, semestral, anual e total. Já a agregação dos horários consultados é feita segundo as opções 15, 30 ou 60 minutos ou horário total. A Figura 1 apresenta a estrutura do modelo de versões desenvolvido para consolidar dados dinâmicos de tráfego modelados pelo sistema SCOOT em ambiente SIG TransCAD.

Conforme mostrado na Figura 1, a execução do modelo de versões em estudo é realizada em três etapas. Na primeira etapa, arquivos temporários (*.TMP) compilados pelo SCOOT são convertidos para o formato DBASE, compatível com o TransCAD. Estes arquivos são gerados pela interface lógica a partir dos parâmetros de consulta definidos pelo usuário, por meio de linhas padronizadas de extração de dados enviadas ao servidor do sistema SCOOT, via conexão remota *Internet Protocol/Transference Control Protocol* (IP/TCP).

Na segunda etapa, os arquivos DBASE de cada elemento viário são submetidos a uma agregação temporal, de acordo com o escopo estatístico-temporal definido pelo usuário. Nesta etapa, cada elemento viário, representado por um código *SCOOT Control Number* (SCN), é agregado no tempo (dias e horas) e segundo uma ou mais estatísticas especificadas. Esta agregação é feita de forma individual para cada SCN, e seus resultados são acumulados em registros temporários, agrupados por tipo de estatística. O número de registros temporários “T” depende dos níveis de agregação temporal definidos pelo usuário.

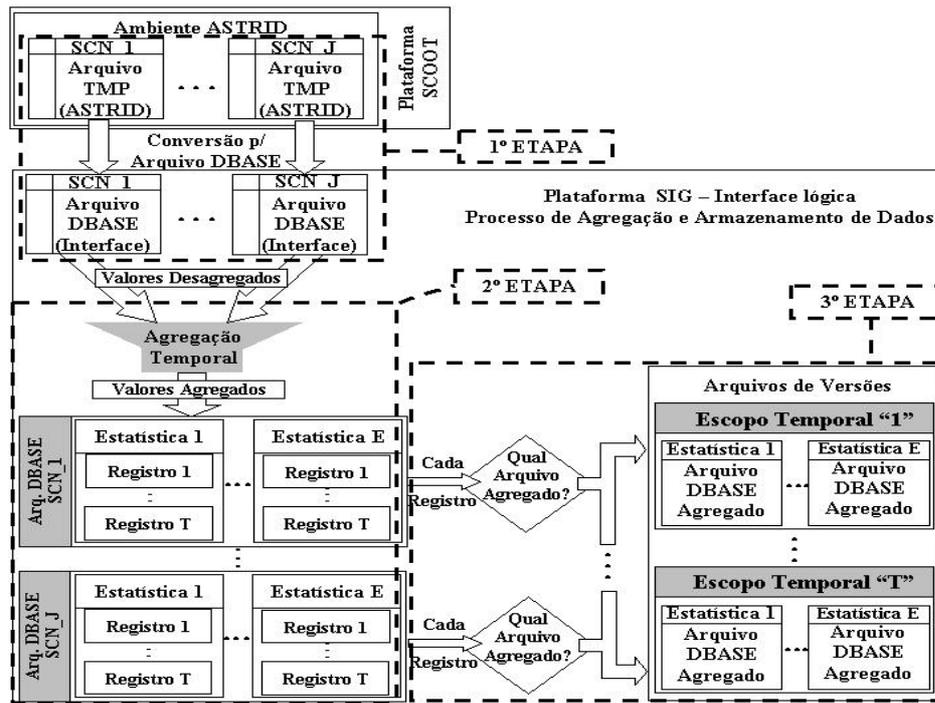


Figura 1: Estrutura de execução do modelo espaço-temporal de versões usados pela interface lógica de importação e modelagem de dados dinâmicos de tráfego (Meneses, 2003).

Por sua vez, a terceira etapa consolida os registros temporários em arquivos de *versões* distintos para cada escopo estatístico-temporal. Para tanto, são efetuados testes lógicos que agrupam versões agregadas de dados de tráfego de diferentes SCN num mesmo arquivo DBASE, relativo a um dado escopo temporal e uma dada estatística de agregação. Cada arquivo possui uma estrutura de atributos com codificação padronizada, cuja quantidade e forma é definida pelos parâmetros temporais (dias da semana, datas) e pelo nível de agregação horária definido pelo usuário.

Com esta estrutura final, os arquivos de *versões* podem ser georeferenciados à sua respectiva base de dados geográfica, o que permite a realização de análises espaciais sobre estes dados. Contudo, duas atividades fundamentais para a consistência das análises realizadas não são contempladas pela modelagem apresentada nesta seção, são elas: qualidade e mineração de dados dinâmicos de tráfego urbano. A primeira atividade consiste em aferir a representatividade dos dados coletados por sistemas CTA. No caso do modelo de *versões* apresentado, não são executados procedimentos abrangentes de aferição da qualidade dos dados importados, limitando-se apenas a eliminação de dias atípicos e valores nulos na base

de dados do SCOOT. A segunda atividade tem a função de converter grandes volumes de dados em informações úteis para o cotidiano do engenheiro de tráfego. Na seção seguinte, são discutidos estes dois temas com maior nível de detalhe.

3. QUALIDADE E MINERAÇÃO DE DADOS DINÂMICOS DO TRÁFEGO URBANO

Inicialmente, a seção 3.1 apresenta uma discussão sobre o conceito e a medição da qualidade de dados dinâmicos de tráfego, sendo apresentados dois procedimentos de correção de inconsistência deste tipo de dado. Em seguida, a seção 3.2 conceitua a mineração de dados (*data mining*) e descreve os tipos de modelos que regem esta tecnologia. Além disto, é apresentado um exemplo prático de utilização de *data mining* na detecção de incidentes do tráfego urbano, reproduzido do trabalho de Lee *et al.* (2002).

3.1. Qualidade de dados

No caso específico de sistemas CTA, os dados coletados em campo podem apresentar inconsistências que devem ser filtradas e eliminadas para aumento da sua confiabilidade. Estas falhas podem ser caracterizadas por: ausência de dados, dados nulos e/ou por valores máximos e mínimos extremos, bem distantes da tendência central. Normalmente, a inconsistência de dados coletados por sistemas CTA está associada a imperfeições dos equipamentos de mensuração, tanto na instalação como na coleta e transcrição dos dados (Shaladover, 2002).

Desta forma, é necessário aferir a qualidade de dados coletados e modelados por sistemas CTA, com base em critérios e procedimentos específicos, que considerem as restrições e necessidades específicas de cada aplicação. Neste contexto, Wilson *et al.* (1999) propõem uma estrutura genérica de medição da qualidade de dados de sistemas CTA, amparada em cinco categorias: acurácia, localização e alcance geográfico, abrangência temporal, padrões de confiabilidade, e prioridade de aferição da qualidade. Cada uma destas categorias determina um conjunto de parâmetros para aferir o nível de qualidade de um dado atributo. A definição dos parâmetros de qualidade depende do processo de coleta dos dados e das aplicações fins.

Não obstante os critérios de aferição da qualidade de dados de sistemas CTA, faz-se necessário definir métodos práticos para a correção de inconsistências na base de dados destes sistemas. Dois procedimentos podem ser destacados: a otimização *fuzzy* e a filtragem digital. O primeiro usa a lógica *fuzzy* para estimar valores de variáveis do tráfego urbano, por meio de valores coletados em campo, considerando as restrições teóricas do fluxo do tráfego. O método *fuzzy* supõe que o valor real de uma variável, num dado instante e local, está próximo do valor de campo. Assim, o processo de otimização *fuzzy* busca ajustar um valor próximo do valor real, por meio de uma combinação dos valores coletados em campo em um dado escopo espaço-temporal, respeitando o princípio da conservação do fluxo veicular na rede viária (Kikuchi *et al.*, 2000).

Por sua vez, o segundo procedimento usa princípios de filtragem digital personalizada de dados para suavizar valores extremos numa base de dados. O processo pode ser resumido nas seguintes etapas sucessivas: escolha de uma frequência de resposta para filtragem de dados, definição de um filtro para encontrar características específicas, conversão de dados filtrados no período de domínio, e cálculo acumulado da soma ponderada dos dados convertidos para gerar o valor agregado no período de domínio (Coifman, 1998).

3.2. Mineração de dados

Conforme Francisco (2002), a mineração de dados consiste num processo de busca de padrões não triviais e estatisticamente representativos em grandes volumes de dados. No caso de base de dados dinâmicos de sistemas CTA, esta tecnologia permite extrair informações de grandes volumes de dados, gerando subsídios que auxiliam a tomada de decisão de técnicos gestores do tráfego urbano, bem como viabilizam a disseminação de informações ao usuário.

De acordo com Klein *et al.* (2002), a maioria das técnicas de *data mining* derivam de três tipos básicos de modelos: físicos, característicos e cognitivos. O primeiro, com base em objetos pré-estabelecidos, replica objetos discriminatórios que podem ser observados de forma precisa numa base de dados. O segundo modelo classifica e/ou identifica eventos por meio de dados geográficos numa base de dados. O terceiro extrai informações de uma base de dados por meio de simulação e automatização do processo de tomada de decisão humano. Dentre as técnicas de *data mining*, é possível destacar as seguintes: lógica *fuzzy*, redes neurais, algoritmos genéticos, análise de agrupamento, espectrograma e procura relacional.

Lee *et al.* (2002) apresentam um exemplo de aplicação da técnica de mineração de dados denominada procura relacional, para identificar o padrão espaço-temporal de ocorrência de incidentes do tráfego, num corredor de tráfego em Singapura. Esta técnica identifica um conjunto de variáveis fortemente relacionadas numa dada base de dados. No exemplo em questão, a procura relacional foi aplicada a um conjunto de dados de tráfego coletados durante 3 h, num dia típico, em um corredor de tráfego de Singapura, com 30 laços detectores. Dentre oito variáveis pesquisadas, a técnica de procura relacional identificou forte relacionamento entre as variáveis: velocidade (km/h), atraso veicular (s) e ocupação veicular (%).

O resultado da interação entre estas variáveis pode ser visualizado por meio de um diagrama espaço-tempo de retângulos, conforme ilustra a Figura 2. A cor, a espessura e a altura dos retângulos de cada detector representam, respectivamente, as variáveis: velocidade (km/h), atraso veicular (s) e ocupação veicular (%), num intervalo de tempo. A cor escura indica baixa velocidade. Alturas e espessuras maiores indicam maior atraso e ocupação, respectivamente. A análise da Figura 2 deixa clara a ocorrência de dois incidentes nas proximidades dos detectores 24 e 26, nos instantes T1 e T3, respectivamente. O primeiro incidente ocasiona aumento gradativo de atraso e redução de velocidade, formando uma onda de choque de congestionamento que se estende até o período T4 e afeta os detectores a jusante do detector 24. Já o incidente no detector 26 se propaga aos detectores anteriores na forma de uma onda de choque com menores impactos que a primeira, a julgar pelo menor período de duração e menor extensão de detectores afetados. Além disto, este incidente causa menores atrasos e menor redução de velocidade.

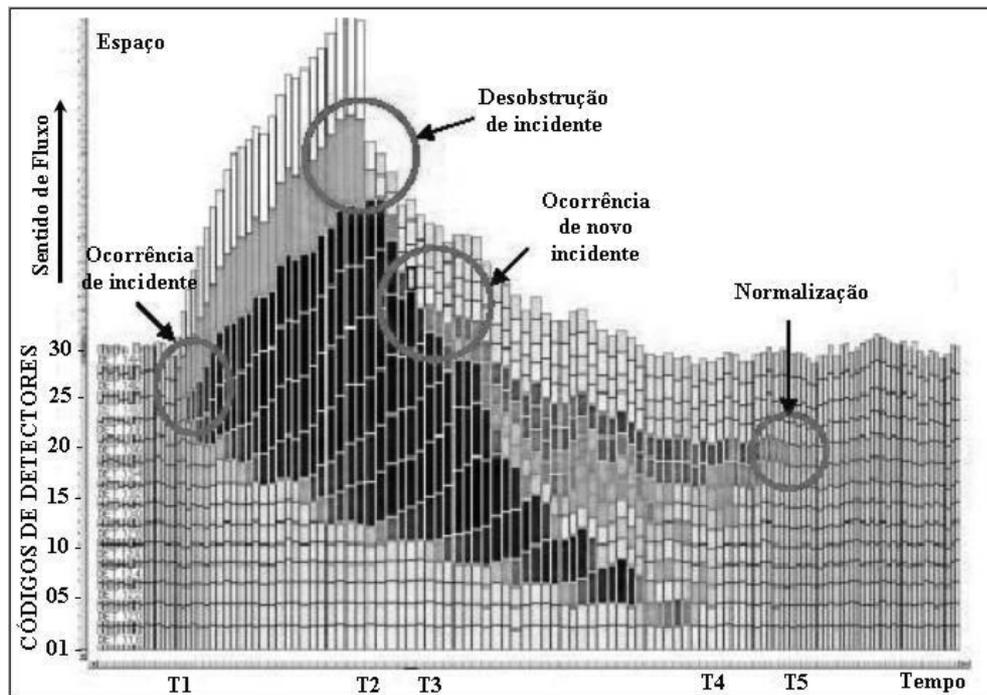


Figura 2: Visualização da técnica de *data mining* denominada procura relacional a dados dinâmicos de tráfego num corredor de tráfego em Singapura (Lee *et al.*, 2002).

Os procedimentos de aferição de qualidade de dados e as técnicas de mineração de dados são fundamentais para melhor explorar um grande volume de dados de tráfego e obter informações consistentes sobre a dinâmica urbana. Entretanto, a inclusão da dimensão espacial ao processo de análise de sistemas urbanos pode viabilizar a elucidação de padrões não triviais do tráfego urbano. A seguir, a seção 4 apresenta uma discussão sobre análise espacial, incluindo conceitos básicos, bem como um exemplo de utilização da análise espacial para gestão do tráfego urbano.

4. ANÁLISE ESPACIAL DE DADOS DINÂMICOS EM AMBIENTE SIG

4.1. Conceitos básicos

Conforme Miller e Shaw (2001), a análise espacial é um campo da geografia que estuda propriedades que variam geograficamente, enquadrando questões relativas à extensão, padrões, associações, interações e mudanças destas propriedades no espaço. O princípio da análise espacial consiste em medir propriedades e relacionamentos de entidades, considerando a localização geográfica no processo analítico descritivo e/ou inferencial do espaço.

A análise espacial envolve três tipos básicos de operações em bases de dados dinâmicos: consultas sobre atributos, operações espaciais e geração de dados. A primeira envolve o processamento de atributos alfanuméricos durante a operação requisitada. Por exemplo, a identificação de vias com volume de tráfego superior a 100 veic/h. Já as operações espaciais envolvem o uso da informação espacial no processo de análise. Por exemplo, a determinação

do volume de tráfego que passa pela área central de uma cidade, a partir de dados de volume de tráfego nos logradouros. Neste caso, é preciso usar a informação espacial de contingência ou não de um logradouro na área em questão. Estas duas operações espaciais podem gerar novos dados ou não. Em caso afirmativo, a terceira operação gera um novo conjunto de dados (atributos relacionais e dados geográficos) a partir dos dados iniciais, conforme a operação de análise espacial especificada. Por exemplo, a identificação de interseções críticas com base no volume de tráfego dos logradouros. Esta operação ocasiona a geração de novas entidades geográficas (interseções críticas) com um conjunto de atributos específicos (Chou, 1997).

Além das operações básicas, Chou (1997) destaca outras ferramentas de análise espacial: operações em uma ou múltiplas camadas geográficas, modelagem espacial, análise de padrões pontuais, análises de superfície, análises em áreas e em redes. A primeira envolve a execução das operações básicas de análise espacial em uma ou entre múltiplas camadas geográficas. A segunda trata da construção de modelos explicativos e de previsão da realidade, com base em estatística espacial. A terceira lida com o exame e a avaliação de padrões espaciais de fenômenos pontuais (ex. acidentes de trânsito). A quarta trata da distribuição espacial de informações em uma estrutura tri-dimensional. A quinta envolve o processamento de dados em áreas de abrangência regulares (ex. grade retangular) ou não (ex. setores censitários). A sexta envolve a modelagem de redes, tipicamente aplicadas a problemas de transportes.

A visualização dos resultados de operações de análise espacial sobre dados dinâmicos é fundamental para o processo de tomada de decisão, especialmente em sistemas de transportes. Miller e Shaw (2001) destacam dois tipos básicos de visualização espacial: a cartográfica convencional e a visualização realista. A primeira inclui mapas bi e tridimensionais, além de modelos de representação 3D, relativos a um ou mais temas. A segunda usa técnicas de animação, simulação e realidade virtual para representar objetos e fenômenos do mundo real, explorando as percepções sensoriais humanas e reduzindo a necessidade da abstração. Por meio de recursos de visualização espacial, é possível desenvolver e testar hipóteses, explorar e identificar padrões não triviais sobre dados dinâmicos derivados de *Intelligent Transportation Systems* (ITS), como ocorre com dados coletados por sistemas CTA.

A implementação de todas estas funcionalidades da análise espacial requer uma plataforma computacional capaz de reproduzir e reconhecer os relacionamentos espaciais entre as entidades modeladas. Conforme Câmara (2000), os SIG permitem a operacionalização de múltiplas ferramentas de análise espacial, tendo em vista a sua capacidade de modelar as relações topológicas entre entidades do mundo real. Este atributo faz com que os SIG se diferenciem de sistemas de gerenciamento de dados relacionais e sistemas cartográficos de geração de mapas. Além disto, os SIG permitem múltiplas formas de visualização espacial (ex. mapas e superfícies) ou não de dados (ex. tabelas e matrizes). Contudo, estes sistemas estão limitados à visualização cartográfica e estática de resultados de análise espacial.

4.2. Aplicação prática da análise espacial à gestão do tráfego urbano

Como exemplo de aplicação do ferramental de análise espacial na gestão do tráfego urbano, Paiva Neto (2002) apresenta uma metodologia de análise espacial desenvolvida para avaliar duas estratégias (cenários) de controle semaforico centralizado (tempo fixo e tempo real), com base em dados dinâmicos do tráfego. Esta aplicação consistiu em determinar a diferença percentual de atraso veicular entre as estratégias de tempo fixo e tempo real, relativamente ao

cenário de tempo fixo, para sub-áreas de controle semafórico. O atraso veicular de cada cenário foi determinado pelo somatório do atraso médio dos logradouros que compõe uma dada sub-área. A determinação da contingência dos logradouros em suas respectivas sub-áreas foi feita por meio de operações espaciais entre múltiplas camadas geográficas em ambiente SIG. Este sistema também foi usado para permitir a visualização espacial dos resultados.

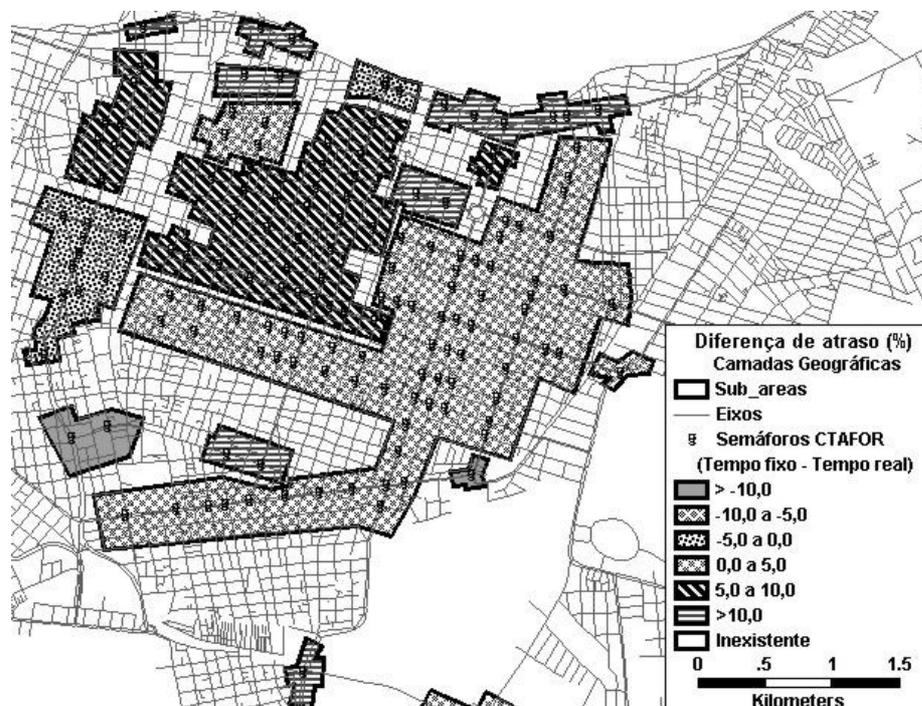


Figura 3: Resultado da análise espacial de avaliação de desempenho das estratégias de controle semafórico de tempo fixo coordenado e tempo real, em Fortaleza (Paiva, 2002).

A metodologia de análise foi aplicada para as sub-áreas controladas pelo CTAFOR com base em dados de atraso para os cenários de tempo fixo e de tempo real, implementados em campo no horário de 7:00 às 8:00 horas, dos dias 7 e 21 de novembro de 2002. Os resultados obtidos são apresentados na Figura 3.

A análise da Figura 3 permite concluir que a estratégia de controle em tempo real proporciona melhores taxas de redução de atraso veicular médio em sub-áreas menores do que naquelas de maior proporção. Isto ocorre porque o sistema usado para otimizar o controle semaforizado trabalha de forma mais eficiente em sub-áreas menores com um corredor principal de deslocamento.

5. CONSIDERAÇÕES FINAIS

A disseminação da tecnologia da informação para a gestão do tráfego tem gerado grandes bases de dados dinâmicos, com potencial para fornecer variadas e importantes informações sobre o comportamento do tráfego urbano, para diferentes grupos de usuários. Contudo, a conversão de um grande volume de dados em informações úteis para motoristas, técnicos e

gestores públicos requer a modelagem e o tratamento prévio destes dados, aliados a um conjunto de ferramentas de análise que permitam uma visualização dos resultados de uma forma mais perceptível ao usuário.

Diante deste contexto, este trabalho buscou apresentar uma revisão conceitual dos principais temas que envolvem a aplicação prática de dados dinâmicos à gestão do tráfego, buscando pontuar os temas tratados por meio de exemplos práticos. Assim, espera-se disponibilizar uma referência para os usuários de bases de dados dinâmicos, capaz de orientar e promover reflexões.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aranha, F. (2000) Nomes e Tempo Circular. *Revista InfoGEO*, Ano 2, Nº 11, Janeiro/Fevereiro-2000, p. 30-31.
- Borges, K. A. de V. (1997) *Modelagem de Dados Geográficos: uma Extensão do Modelo OMT para Aplicações Geográficas*. Dissertação Mestrado, Escola de Governo Fundação João Pinheiro, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Caliper (1996) *TransCAD – User’s Guide, V. 3.1 for Windows*. Caliper Corporation, Newton, Massachusetts, USA.
- Câmara, G. (2000) As Roupas Novas do GIS (Parte II). *Revista InfoGEO*, Ano 3, Nº 13, Maio/Junho-2000, p. 21-22.
- Chen, S.; K.F. Petty; A. Skabardonis; P. Varaiya; Z. Jia (2001) Freeway Performance Measurement System: Mining Loop Detector Data. *Transportation Research Board*, The 80th Annual Meeting, Washington, D.C., USA.
- Chou, Y. (1997) *Exploring Spatial Analysis in Geographic Information Systems*. Onword Press, Santa Fe, USA.
- Coifman, B. A. (1998) New Methodology for Smoothing Freeway Loop Detector Data: Introduction to Digital Filtering. *Transportation Research Board*, The 77th Annual Meeting, Washington, D.C., USA.
- Davis, C. Jr (1998) Modelagem de Dados Geográficos. *Revista InfoGEO*, Ano 2, Nº 1, Julho/Agosto-1998, p. 38-40.
- Francisco, E. R. (2002) GIS+DW=Spacial Data Warehouse? *InfoGEO*, Ano 4, Nº 24, Maio/Junho-2002, p. 52-53.
- Huiguo, L.; F. Kezhong; W. Qing (2002) Time Geography and Theoretical Issues of Spatial-Temporal Data Model. *Geographical Survey Institute*, Technical Report 1, N.275-2, Beijing, China.
- Ichikawa, S. M.; C. S. Pitombo e E. Kawamoto (2002) Aplicação de Minerador de Dados na Obtenção de Relações entre Padrões de Viagens Encadeadas e Características Sócio-Econômicas. *Anais do XVI Congresso de Pesquisa e Ensino em Transportes*, ANPET, Natal, RN, vol. 2, p. 175-186.
- Jia, X. (1999) A Method for Representing and Reasoning Spatial Knowledge in Transportation. *Transportation Research Board*, The 78 th Annual Meeting, Washington, D.C., USA.
- Kikuchi, S.; D. Miljkovic e H. J. V. Zuylen (2000) Examination of Methods that Adjust Observed Traffic Volumes on a Network. *Transportation Research Board*, The 79 th Annual Meeting, Washington, D.C., USA.
- Klein, L.; P. Yi e H. Teng (2002) Decision Support System for Advanced Traffic Management Through Data Fusion and Mining. *Transportation Research Board*, The 81 th Annual Meeting, Washington, D.C., USA.
- Koncz, N. e T. M. Adams (2002) Temporal Data Constructs for Multidimensional Transportation GIS Applications. *Transportation Research Board*, The 81 th Annual Meeting, Washington, D.C., USA.
- Langran, G. (1993) *Time in Geographic Information Systems*. Taylor & Francis, Washington, D.C., USA.
- Lee, H. N.; S. Jeng e P. Chandrasekar (2002) Investigating an Incident Situation Using Data Mining. *Transportation Research Board*, The 81 th Annual Meeting, Washington, D.C., USA.
- Meneses, H. B. (2003) *Interface lógica em ambiente SIG para bases de dados de sistemas centralizados de controle do tráfego urbano em tempo real*. Dissertação de Mestrado, Universidade Federal do Ceará, Fortaleza, Brasil.
- Michinori, H. (2002) Development of Spatial Temporal GIS Basic System. *Geographical Survey Institute*, Technical Report 1, Nº 275-2, Beijing, China.
- Miller, H. J. e S. Shaw (2001) GIS-Based Spatial Analysis and Modeling. *Ins: Oxford University Press (eds) Geographic Informaion Systems for Transportation: Principles and Applications*, London, UK.
- Paret, C.; S. Spaccapietra e E. Zimanyi (1999) Spatial-Temporal Conceptual Models: Data Structures + Space + time”. *7 th ACM Symposium on Advances in GIS*, November 5-6, Kansas City, Kansas, USA.

- Shaladover, S. E. (2002) Pathfinders. *Traffic Technology International*, April/May-2002, p. 53-58.
- Spaccapietra, S; C. Paret e E. Zimanyi (1998) Modeling Time from a Conceptual Perspective. *International Conference on Information and Knowledge Management, CIK98'*, November 3-7, Washington D. C., USA.
- Stathopoulos, A. e M. Karlaftis (2001) Temporal and Spatial Variations of Real-Time Traffic Data In Urban Areas. *Transportation Research Board*, The 80th Annual Meeting, Washington, D.C., USA.
- Paiva Neto, P.M. (2003) *Análise Espacializada dos Benefícios Operacionais do Controle Semafórico Centralizado em Tempo Real do CTAFOR*. Projeto de Graduação, Engenharia Civil, Universidade Federal do Ceará.
- TRL (2000) SCOOT 0414 – *Executive Summary – Description of SCOOT*. SCOOT Traffic Handbook. Transportation Research Laboratory.
- Wilson, C.; D. Bowers; P. Welsh; D. Wright; R. Pearce e M. Patel (1999) Fit for a Purpose? Defining the Quality of Traffic Management Data. *Traffic Technology International*, December-1998/January-1999, p. 75-78.
- Zhao, F.; L. Wang; H. Elbadrawi e L. D. Shen (1997) Temporal Geographic Information System and Its Application to Transportation. *Transportation Research Board*, The 76 th Annual Meeting, Washington, D.C., USA.

Universidade Federal do Ceará
Departamento de Engenharia de Transportes
Campus do Pici – Bloco 703 - Fortaleza - CE - 60.455-760

Fone/Fax: (0xx85) 288-9571/9488
E-mail: hamifrancy@zipmail.com
felipe@det.ufc.br