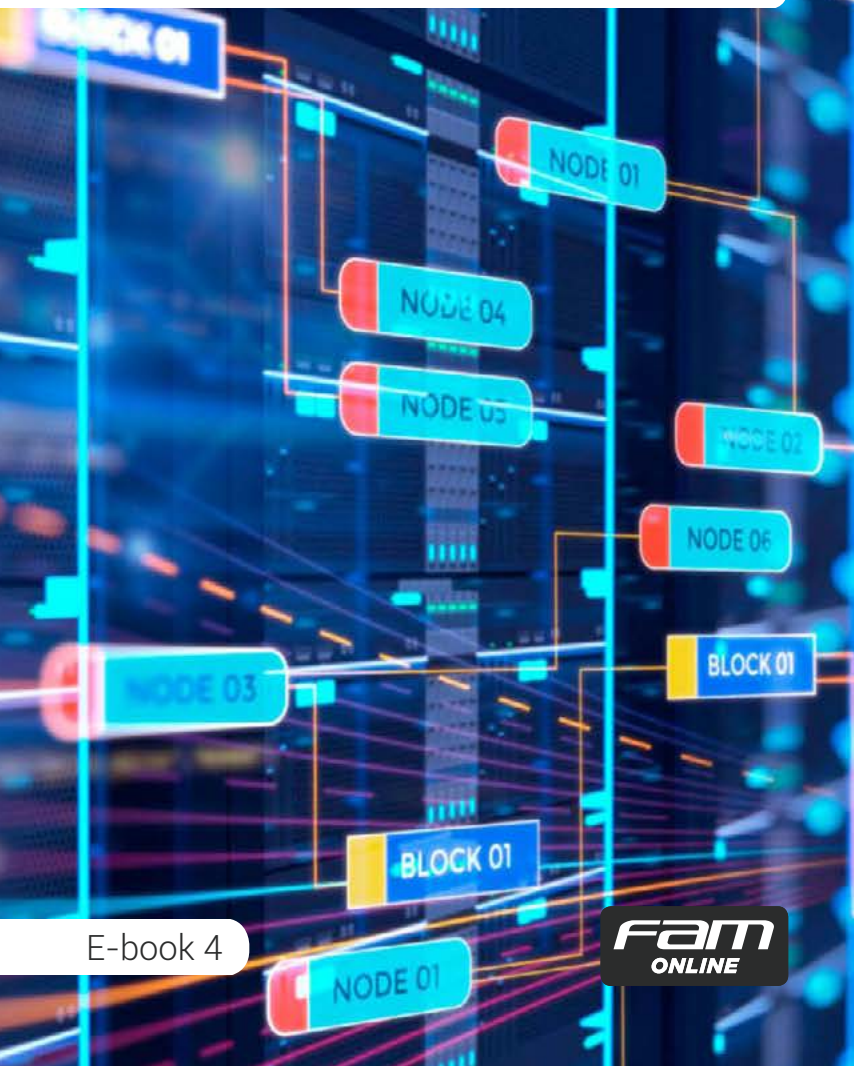


ESTRUTURA E MODELAGEM DE DADOS

César Torres Fernandes



E-book 4

FAM
ONLINE

Neste E-Book:

INTRODUÇÃO	3
SISTEMAS DE BANCO DE DADOS GEOGRÁFICOS	4
SISTEMAS DE BASE DE DADOS ESPACIAIS.....	11
SISTEMAS DE BANCO DE DADOS MULTIMÍDIA	19
VISÃO GERAL DA ARQUITETURA	26
CONSIDERAÇÕES FINAIS	30
SÍNTESE	31

INTRODUÇÃO

Neste módulo, nosso foco é aprender a respeito dos sistemas de bancos de dados geográficos, bem como os sistemas de bancos de dados multimídia.

SISTEMAS DE BANCO DE DADOS GEOGRÁFICOS

Conhecemos, neste momento, um pouco dos sistemas de banco de dados geográficos. Em tecnologia da informação, podemos afirmar que as mudanças e evoluções são constantes, conseqüentemente, esse processo tem ocorrido também no campo da tecnologia da informação espacial.

Desde os anos 1950, quando do início da ideia de usar o computador para cartografar e mapear, hardwares e softwares evoluíram ao ponto de permitir que os usuários de informações espaciais operassem mais facilmente, além de estender constantemente as fronteiras do conhecimento. Novos conceitos e técnicas de uso da informação espacial têm ocorrido em paralelo aos avanços em hardwares e softwares; assim, de maneira cumulativa, eles aceleraram a adoção de tecnologias de informação relacionadas.

A atual disseminação de informações espaciais contrasta fortemente com a era anterior, que prevaleceu até o início dos anos 1990. Essa era caracterizou-se pelo termo genérico de informação geográfica (Sistemas de Informações Geográficas – SIG) como uma tecnologia altamente especializada, que interessa sobretudo aos usuários profissionais e pesquisadores para aplicações específicas, mas que alcança também a população em geral por meio de

diversos aplicativos com mapas e outras informações correlatas.

Uma das principais forças motrizes por trás da recente popularização da informação espacial é o aumento da disponibilidade de dados espaciais por parte de governos e instituições. Tais informações são distribuídas via internet por bases de dados espaciais, bibliotecas geográficas digitais e pela crescente conscientização da importância da informação espacial para todos os setores da sociedade moderna.

As mudanças no processo de coleta, gerenciamento e uso das informações espaciais provavelmente não teria acontecido, ao menos não nessa extensão, sem o poder dos bancos de dados espaciais. Por essa razão, vamos apresentar uma visão geral de conceitos relacionados a dados e informações espaciais e dos sistemas gerenciados de dados espaciais.

O projeto e a implementação da tecnologia da informação espacial dependem amplamente das tendências que caracterizaram a tecnologia da informação (TI). Hoje em dia, conceitos e técnicas utilizados em sistemas de banco de dados espaciais são retirados dos princípios e métodos gerais dos sistemas de banco de dados que têm evoluído cada vez mais.

Neste momento, são dados e usuários, não mais aplicativos ou tecnologia, que orientam o uso de informações espaciais em uma organização, levando à abordagem baseada em dados e centrada no usuário para informações que caracterizam as práticas

atuais na implementação de sistemas de banco de dados espaciais.

Tradicionalmente, a informação espacial era abordada a partir de três perspectivas: gerenciamento de dados espaciais, cartografia e análise espacial. Essas perspectivas podem ser descritas como a visão instrumental ou orientada à informação espacial. Com o tempo, surgiu uma nova conceituação de informações espaciais com quatro aspectos funcionais: sistemas de banco de dados, cartografia, comunicação e análise.

O componente de banco de dados espacial desempenha um papel central e crítico na nova visão da informação espacial, uma vez que os aplicativos fazem uso intenso de dados e recursos computacionais, bem como exigem o armazenamento de um grande volume tanto de dados tradicionais (alfanuméricos) quanto de dados não tradicionais (objetos geométricos, imagens, séries temporais).

Utilizando os recursos de indexação espacial e processamento de dados, um sistema de banco de dados pode reunir rapidamente dados relevantes para uma determinada aplicação e, juntamente com técnicas de análise de dados espaciais, desenvolver soluções para a tomada de decisão espacial.

A análise de informações espaciais para o apoio à decisão pode ser aprimorada por meio de técnicas em inteligência computacional que utilizam a inteligência artificial para tolerar imprecisões e incertezas em grandes bancos de dados, com vistas a

alcançar soluções robustas e aceitáveis. Pode-se, ainda, fazer uso de técnicas de mineração de dados, varrendo automaticamente os dados mantidos em um banco de dados espacial para descobrir possíveis relacionamentos entre os elementos de dados que, por sua vez, revelam os fenômenos difíceis ou impossíveis de se detectar.

Dessa forma, a abordagem do banco de dados mostra-se capaz de aprimorar as aplicações de suporte à tomada de decisão quando comparada a abordagens SIG tradicionais.

A mudança para usar uma abordagem de banco de dados para informações espaciais iniciada nos anos 1990 envolveu um esforço considerável e compromisso por parte dos fornecedores de SIG convencionais e dos bancos de dados. Atualmente, um sistema típico de banco de dados espaciais é um banco de dados com recursos e funções adicionais que seja capaz de lidar com dados espaciais. Esses recursos e funções incluem:

- Tipos de dados espaciais: são armazenados como tipos de dados especiais seguindo a definição do Open Geospatial Consortium (OGC, 1999) de "recursos simples" ou como um Binary Large Object (BLOB) em um banco de dados.
- Indexação espacial: mecanismo para facilitar o acesso a um espaço do banco de dados por meio de coordenadas armazenadas no espaço bidimensional. Há opções de indexação diferentes, como a árvore R, Quadtree e árvore B, cada qual tem seus

pontos fortes e seus pontos fracos, dependendo do formato de dados específico e da necessidade do aplicativo.

SAIBA MAIS

Caso você esteja familiarizado com a língua inglesa, vale a pena conferir R-Trees: A dynamic index structure for spatial searching, de Antomn Guttman, que trata em detalhes as Árvores R. Disponível em: <http://www-db.deis.unibo.it/courses/SI-LS/papers/Gut84.pdf>. Acesso em: 19 nov. 2019.

- Operadores espaciais: referem-se ao conjunto de funções de processamento de dados e processos que podem ser alcançados recorrendo-se a consultas estruturadas SQL, seja para consultar seja para recuperar o conteúdo selecionado do banco de dados, em tabelas de banco de dados segundo critérios espaciais e não espaciais específicos. Ademais, servem para gerar resultados do processamento em formatos específicos.
- Rotinas de aplicativos espaciais: inclui uma variedade de componentes de softwares para funções específicas de aplicativos de banco de dados, como carregamento de dados, controle de versão e transações longas, desempenho, *backup* e replicação de banco de dados.

O uso da abordagem de banco de dados para informações espaciais fornece alguns benefícios, como

centralizar dados espaciais e não espaciais em um único sistema para um processamento mais eficiente, aumentando assim o retorno sobre investimento (ROI), aproveitando as habilidades do pessoal de TI.

Podcast 1



Os sistemas de banco de dados espaciais são diferentes dos convencionais sistemas de banco de dados em dois aspectos importantes, a saber:

- a)** Requisito para armazenar tipos de dados complexos, como pontos, linhas e polígonos.
- b)** Funcionalidade necessária para processar tais tipos de dado complexos usando os operadores espaciais consideravelmente mais sofisticados do que os operadores convencionais de banco de dados.

É comum chamar dados espaciais de dados geograficamente referenciados ou de dados geoespaciais. Os espaciais são dados que podem ser exibidos, manipulados e analisados, pois representam um atributo espacial que indica uma localização na superfície da Terra. Esse atributo espacial é normalmente fornecido na forma de pares de coordenadas, permitindo tanto a determinação da posição quanto a forma de um determinado espaço.

Os dados espaciais têm duas propriedades importantes, a saber:

a) Referem-se a um espaço geográfico, o que significa que os dados são registrados em um sistema de coordenadas geográficas aceito.

b) Representam-se em uma variedade de escalas geográficas em escalas relativamente pequenas, representando grandes áreas da superfície terrestre, eles devem ser generalizados e simbolizados.

SAIBA MAIS

Saiba mais sobre as coordenadas geográficas, lendo os conceitos gerais tais quais definidos pelo IBGE. O texto encontra-se disponível em: <https://atlasescolar.ibge.gov.br/conceitos-gerais/o-que-e-cartografia/coordenadas-geograficas.html>. Acesso em: 19 nov. 2019.

SISTEMAS DE BASE DE DADOS ESPACIAIS

Os sistemas de banco de dados espaciais são um tipo especial de sistema de banco de dados específico para gerenciar e processar dados. Segundo Güting (1994), um sistema de banco de dados espacial tem as seguintes características:

- É um sistema de banco de dados.
- Oferece tipos de dados espaciais (Spatial Data Type – SDT) em seu modelo de dados e linguagem de consulta.
- Suporta tipos de dados espaciais em sua implementação, fornecendo pelo menos indexação espacial e algoritmos eficientes para junções espaciais.

Em sua definição de sistemas de banco de dados espaciais, Ralf H. Güting (1994) considerou os bancos de dados espaciais apenas como a tecnologia subjacente ao sistema de informação geográfica. Com isso, muitos bancos de dados espaciais são desenvolvidos como armazéns de dados espaciais que visam a fornecer informações oportunas e relevantes aos usuários do SIG.

Simultaneamente, os SIG são cada vez mais implementados com a utilização de recursos de dados externos obtidos das agências governamentais ou de fornecedores comerciais de dados espaciais, ao

invés da digitalização interna e criação de banco de dados.

Nos ambientes contemporâneos de processamento de dados espaciais, é bastante clara a divisão de trabalho entre sistemas de banco de dados espaciais e SIG. Banco de dados espaciais são implementados com a finalidade de armazenar e gerenciar grandes volumes de dados geograficamente referenciados ou baseados em localização. Esses sistemas indexam os dados para consulta e recuperação eficientes, além de fornecerem os mecanismos e procedimentos necessários para proteger os dados da destruição física, perda ou degradação da integridade lógica.

No entanto, sistemas de banco de dados espaciais são geralmente fracos na coleta e edição de dados, análise de dados espaciais e geração de mapas e outras formas de cartografia (produtos de informação). Esses processos e funções são mais bem tratados por SIG projetados principalmente para coletar e usar dados espaciais para esses propósitos.

Sendo assim, podemos verificar a divisão do trabalho entre sistemas de banco de dados espaciais e SIG da seguinte forma:

a) Tarefas principais - Sistemas de Informação Geográfica:

- Coleta e edição de dados.
- Análise de dados.

- Geração de mapas e produtos de informação cartográfica.

b) Sistemas de Banco de Dados Espaciais:

- Armazenamento e gerenciamento de dados.
- Indexação espacial.
- Segurança e integridade de dados.
- Consulta de dados espaciais.

A partir dessa visão sistêmica, os sistemas de banco de dados espaciais são configurados como parte de uma infraestrutura destinada a atender às necessidades gerais de informação de uma organização, com ou sem o uso de um SIG específico.

Os sistemas espaciais de banco de dados permitem o gerenciamento e processamento de dados geograficamente referenciados. Para entendermos melhor, conceituaremos brevemente os tipos de dados espaciais, a indexação de banco de dados e as restrições de integridade de dados espaciais.

Convencionou-se que os sistemas de banco de dados foram projetados para gerenciar e processar dados alfanuméricos representados por cadeias de caracteres, valores numéricos, datas e expressões booleanas ou lógicas. Não houve provisão em sistemas de bancos de dados relacionais convencionais para suportar o armazenamento e o processamento de dados espaciais representados por pontos, linhas, polígonos e superfícies.

Os bancos de dados orientados a objetos e objeto-relacionais permitem que os usuários definam tipos abstratos de dados que descrevam estruturas complexas em colunas em um banco de dados. Vários fornecedores de software de banco de dados se utilizam de tais recursos a fim de definir tipos de dados abstratos que permitam que seus produtos gerenciem e processem dados espaciais.

Além de dados espaciais vetoriais de pontos (linhas, polígonos e superfícies), é possível definir dados espaciais rasterizados (georreferenciados, fotografias aéreas digitais, mapas digitalizados, imagens de sensoriamento remoto e superfícies contínuas, como um modelo de elevação digital) enquanto tipos abstratos de dados, podendo ser gerenciados e processados em um único sistema de banco de dados espacial.

A indexação espacial serve ao mesmo propósito que a indexação de tabela em um banco de dados relacional, ou seja, para acelerar o acesso e o retorno de dados a um usuário de um banco de dados. Entretanto, a indexação espacial é muito mais complicada do que a indexação de tabela, pois lida com um espaço bidimensional em vez de uma matriz linear direta de dados baseados em texto em tabelas.

Assim, um conceito fundamental de indexação espacial é o uso de aproximação, segundo o qual o processo de acesso espacial restringe gradualmente sua área de busca até que os objetos de banco de dados necessários sejam encontrados.

São diversos os métodos de indexação espacial usados para sistemas de banco de dados espaciais. Um dos mais adotados é a Árvore R, em que o "R" significa "região". Trata-se de uma árvore de vários níveis que armazena um conjunto de retângulos em cada nó.

Em um índice espacial da árvore R, objetos de dados individuais são representados pelo mínimo retângulo envolvente (Minimum Bounding Rectangle – MBR). O índice da árvore R armazena números de referência e coordenadas de seus quatro cantos em um arquivo de índice. Quando o banco de dados recebe uma solicitação de pesquisa espacial, o sistema varre as coordenadas armazenadas do MBR para determinar qual MBR se enquadra no nível da ramificação, na janela de pesquisa.

Depois de identificar o MBR, o sistema verifica os retângulos visando a identificar o MBR de objetos para selecionar aqueles que se enquadram dentro dos limites da janela de pesquisa. Quando esses MBR são identificados, o sistema utiliza as informações de referência entre MBR e número de identificação do objeto para acessar a tabela de atributos com informações pertinentes ao contexto da pesquisa.

As restrições de integridade são mais bem descritas como regras de negócios do que são aplicadas em sistemas de banco de dados para proteger o "valor" intangível dos dados, garantindo com isso sua precisão, correção e validade. Por esse motivo, foram introduzidas três classes de restrições de

integridade de modelagem de dados e operações de banco de dados, a saber:

- I) Restrições de domínio.
- II) Restrições de chave e de relação.
- III) Restrições de integridade semântica.

Cockcroft (1997) estendeu essa classificação para abarcar inclusive aspectos de requisitos de dados espaciais. Para isso, baseou-se na distinção entre regras topológicas, semânticas e de usuário, como podemos observar:

- Restrições de integridade topológica: preocupação com as propriedades geométricas de relações espaciais entre objetos no banco de dados.
- Restrições de integridade semântica: regras do banco de dados que regem o comportamento espacial de objetos no banco de dados.
- Restrições definidas pelo usuário: regras de negócios semelhantes às identificadas na modelagem de dados não espaciais.

Em trabalho posterior a esse, Cockcroft (1998) propôs ainda que cada uma das três classes de integridade seja aplicada aos dados no estado consistente do banco de dados e aos dados no processamento de transações, resultando nas seguintes classes de restrições de integridade de dados espaciais:

- Restrições estáticas de integridade topológica, por exemplo, os polígonos devem ser fechados.

- Restrições de integridade topológica de transição, por exemplo, se o limite de um polígono é modificado, tanto o próprio polígono quanto todo o conjugado de polígonos devem ser atualizados simultaneamente.
- Restrições estáticas de integridade semântica, por exemplo, a área de uma terra não deve ser negativa.
- Restrições de integridade semântica de transição, por exemplo, após a subdivisão de uma parcela de terreno, a soma das áreas das unidades subdivididas deve ser igual à área da parcela original.
- Restrições estáticas de integridade definidas pelo usuário, por exemplo, rios e córregos com mais de 2 metros devem ser armazenados como recursos de polígono.
- Transição de restrições de integridade definidas pelo usuário, por exemplo, após um pedido de zoneamento de uma parcela de terreno tenha sido aprovado, o *status* de uso da parcela em questão deve ser atualizado dentro de dois dias.

As restrições de integridade são capturadas durante a modelagem conceitual de dados e devem ser aplicadas durante todo o ciclo de vida do desenvolvimento de sistemas. Também devem ser devidamente documentadas e armazenadas como um componente integral dos metadados do banco de dados.

Nessa parte, introduzimos o conceito de sistemas de banco de dados espacial como uma classe es-

pecial de sistemas de banco de dados e dos tipos de dados espaciais. Introduzimos igualmente a definição de Güting (1994) de sistemas de banco de dados espaciais como uma classe específica de sistemas de banco de dados capaz de executar todas as funções de um sistema de banco de dados típico, como indexação e junções, mas que possui funcionalidades adicionais para manipular dados com atributos espaciais.

Nesse contexto, características e requisitos das bases de dados espaciais foram discutidos com referência aos conceitos de restrições de integridade de dados espaciais, métodos de acesso espacial por indexação em árvore R e dados espaciais.

SISTEMAS DE BANCO DE DADOS MULTIMÍDIA

Em um passado recente, o foco principal em pesquisa e desenvolvimento de multimídia estava nas comunicações, na autoria e apresentação de multimídia. Posto que cada vez mais dados multimídias digitais em formato de imagem, vídeo e áudio estão sendo capturados e armazenados, o foco mudou nos últimos anos para armazenamento e recuperação de informações multimídia eficientes e eficazes.

Situação semelhante aconteceu quando mais e mais dados alfanuméricos foram armazenados em formato legível por computador. Isso levou ao desenvolvimento de sistemas de gerenciamento de banco de dados (SGBD), um dos aplicativos de computador mais populares, que são usados em quase todas as organizações.

Dadas as diferenças entre características e requisitos de dados alfanuméricos e dados multimídia, um SGBD tradicional não é capaz de manipular efetivamente dados multimídias. Assim, novas técnicas de indexação e recuperação de multimídia se colocam como necessárias.

Neste tópico, estudaremos alguns conceitos importantes a respeito das principais características e dos requisitos dos dados multimídia. Em seguida, passaremos à discussão sobre os problemas gerais de design dos sistemas de gerenciamento de banco

de dados multimídia para atender a tais características e requisitos, além disso, sobre as técnicas para indexação e recuperação de documentos de texto, áudio, imagens e vídeo.

Os bancos de dados multimídia geralmente são acessados remotamente por uma rede. Objetos multimídia identificados como relevantes para a consulta devem ser recuperados do servidor e transmitidos ao cliente para apresentação. Um conjunto de requisitos envolvendo fatores como atraso e instabilidade de retardo (chamado qualidade de serviço), devem ser atendidos para obter uma apresentação multimídia de maneira adequada.

Nos sistemas de gerenciamento de banco de dados tradicionais, a principal preocupação de desempenho é a eficiência (quanto tempo leva para responder a uma consulta). Nos sistemas de gerenciamento de banco de dados multimídia (SGBDMM ou Multimedia Database Management System – MMDBMS), a eficiência também é importante, mas considera-se igualmente a eficácia da recuperação (capacidade de recuperar itens importantes, bem como a capacidade de rejeitar itens irrelevantes). Isso ocorre porque os SGBDMM recuperam itens com base na similaridade medida, ao utilizar uma métrica de similaridade em vez da sua correspondência exata.

O problema, porém, está em **como** extrair todos os recursos de itens multimídia e projetar uma métrica de similaridade que esteja exatamente de acordo com o julgamento humano. É provável que alguns

itens tidos como relevantes pelo sistema sejam realmente considerados irrelevantes pelo usuário, não sendo, dessa forma, recuperados.

Neste momento, algumas definições mostram-se importantes em relação aos tipos de mídias e multimídia. Sendo assim, **mídia** refere-se aos tipos de informações ou tipos de representação de informações, como dados alfanuméricos, imagens, áudio e vídeo. Há várias maneiras de classificar mídia, e as mais comuns são baseadas em formatos físicos e relacionamentos de mídia com o tempo. Deste modo, classificamos mídia com base na existência ou não de dimensões de tempo. Essa convenção nos leva a duas classes de mídia: a estática e a dinâmica (ou contínua no tempo):

- **Mídia estática:** sem uma dimensão temporal, seu conteúdo e significado não dependem da hora da apresentação. A mídia estática inclui dados alfanuméricos, gráficos e imagens estáticas.

- **Mídia dinâmica:** com dimensões temporais, seus significados e exatidão dependem da taxa em que são apresentados. A mídia dinâmica inclui animação, áudio e vídeo, além disso, têm seus intervalos ou taxas intrínsecas de unidade.

Em uma velocidade mais lenta, a reprodução distorce o significado ou a qualidade do som. Como essas mídias devem ser reproduzidas continuamente a uma taxa fixa, no geral são chamadas de **mídia contínua**, ou **mídia isócrona**, devido ao relacionamento fixo entre cada unidade de mídia e o tempo.

Por sua vez, **multimídia** refere-se a uma coleção de tipos de mídia usados em conjuntos em que está implícito que pelo menos um tipo de mídia não é um dado alfanumérico, ou seja, pelo menos um tipo de mídia é imagem, áudio ou vídeo.

Os **dados multimídia** referem-se à representação legível por computador de vários tipos de mídia. **Informações multimídias** referem-se às informações transmitidas por vários tipos de mídia. Às vezes, utilizamos itens e objetos de mídia ou multimídia para nos referirmos a qualquer entidade autônoma em um sistema de recuperação de informações multimídia (SRIM) que pode ser consultada, recuperada e apresentada.

Em relação aos bancos de dados e SGBDs, às vezes são usados de forma intercambiável, mas a definição de banco de dados refere-se a uma coleção ou a um repositório de dados ou itens de mídia; o termo SGBD é utilizado para nos referirmos a todo o sistema que gerencia o banco de dados.

Desde a década de 1940, foram desenvolvidos sistemas automatizados de recuperação de informações (RI) com a finalidade de auxiliar a armazenagem e gerenciar um grande número de documentos de texto de modo que os documentos relevantes para as consultas do usuário possam ser recuperados rapidamente.

Os sistemas de gerenciamento de banco de dados recuperam itens com base em dados estruturados usando a correspondência exata. O sistema de recu-

peração de informação (Information Retrieval – IR) também é chamado de recuperação baseada em texto. A recuperação baseada em conteúdo refere-se à recuperação baseada em recursos reais de mídia, como cor e forma, em vez de anotação de texto do item de mídia.

Podcast 2



Normalmente, a recuperação baseada em conteúdo ancora-se na similaridade em vez de uma correspondência exata entre a consulta e um conjunto de itens do banco de dados. Um sistema de recuperação de informações multimídia refere-se a um sistema básico fornecedor de recuperação de informações multimídias, usando uma combinação de SGBD, recuperação de informação e técnicas de recuperação baseadas em conteúdo. Dessa forma, um SRIM completo é chamado de SGBD Multimídia.

Nos sistemas de recuperação de informação multimídia, uma das questões mais importantes é a extração de recursos ou a representação de conteúdo, por exemplo, quais são os principais recursos ou conteúdo de um item de multimídia. A extração de recursos pode ser um processo automático ou semiautomático.

Em algumas literaturas de recuperação baseadas em conteúdo, a extração de recursos também é chamada de indexação, sendo que o termo índice se refere a uma estrutura de dados ou à organização

de recursos extraídos para busca e recuperação eficientes.

Podemos explicar a necessidade de sistemas de recuperação de informação multimídia por meio de três fatos:

1. Cada vez mais dados multimídias estão sendo capturados e armazenados; para usar as informações contidas nesses dados, é necessário um sistema de indexação e recuperação eficiente e eficaz.
2. Os dados multimídia têm características e requisitos especiais significativamente diferentes dos dados alfanuméricos. Desse modo, um SGBD tradicional não é adequado para manipular dados multimídia.
3. Embora as técnicas de recuperação de informação possam ajudar na recuperação de multimídia, elas sozinhas não são adequadas para lidar com dados multimídia eficazmente.

Na atualidade, estamos diante de uma explosão de informações multimídias. Por exemplo, uma grande quantidade de imagens e vídeos está sendo criada e armazenada na internet, conversões de pinturas e figuras no formato impresso ou digital, bem como imagens de jornais e TV para facilitar o processamento, distribuição e preservação. Muitas imagens médicas estão sendo capturadas todos os dias e os satélites estão produzindo muito mais.

Essa tendência continua com o avanço das tecnologias digitais e de armazenamento, por isso, é pouca

utilidade criar um mero repositório para essa quantidade cada vez maior de informações multimídia.

Será impossível usar totalmente essas informações multimídias, a menos que sejam organizadas para recuperação rápida sob demanda. Não apenas uma quantidade crescente de dados está sendo armazenada, mas também os tipos de dados e suas características são diferentes dos dados alfanuméricos. As principais características dos dados multimídia são discutidas a seguir:

- Dados multimídias, especialmente áudio e vídeo, são muito intensivos em dados.
- Áudio e vídeo têm uma dimensão temporal e devem ser reproduzidos a uma taxa fixa para alcançar o efeito desejado.
- Áudio, imagem e vídeo digitais são representados em uma série de valores de amostra individuais e carecem de estrutura semântica óbvia para os computadores reconhecerem automaticamente o conteúdo.
- Muitos aplicativos multimídias exigem a apresentação simultânea de vários tipos de mídia de maneira coordenada espacial e temporalmente.
- O significado dos dados multimídias às vezes é confuso e subjetivo.
- Os dados multimídias são ricos em informações. Muitos parâmetros são necessários para representar seu conteúdo adequadamente.

VISÃO GERAL DA ARQUITETURA

Os principais módulos de um sistema de recuperação de informações multimídias são interface do usuário, extrator de recursos, gerenciador de comunicações, mecanismo de indexação e, por fim, pesquisa e gerenciador de armazenamento. As funções desses módulos são explicadas pela descrição de dois tipos principais de operações do sistema de recuperação de informação multimídia: inserção e recuperação de itens de banco de dados.

O modelo de dados afeta todos os aspectos do design e operação do sistema de recuperação da informação multimídia, determinando de que maneira as informações se organizam e armazenam, e que tipos de consultas suportam.

A interface do usuário fornece ferramentas para o usuário inserir itens de banco de dados, emitir consultas e apresentar os resultados da consulta ao usuário. O design da interface do usuário é complicado pela diversidade de mídias, recursos e imprecisões dos significados dos dados multimídia.

Nos sistemas de recuperação de informação multimídia, recursos e atributos dos itens de dados multimídia são extraídos e armazenados com os próprios itens de dados. Logo, esses recursos são organizados com certas estruturas de índice para recuperação eficiente.

A recuperação é baseada em medidas de similaridade entre os recursos de consulta e os itens de dados. A recuperação e apresentação de multimídia devem atender a restrições espaciais e temporais. Essas restrições são especificadas como requisitos de qualidade de serviço (QoS) dos objetos ou aplicativos multimídia, sendo que todos os subsistemas, incluindo os sistemas operacionais, sistemas de comunicação e gerenciador de armazenamento, devem trabalhar juntos para alcançar a QoS desejada.

SAIBA MAIS

Saiba mais sobre Qualidade de Serviço na Internet, ao ler o artigo Qualidade de serviço na internet, de Ana Paula Silva dos Santos (1999), que se encontra disponível em: <https://memoria.rnp.br/newsgen/9911/qos.html>. Acesso em: 21 nov. 2019.

Assim, a arquitetura de um sistema de recuperação de informações multimídia deve ser flexível e extensível para suportar diversos aplicativos, tipos de consulta e conteúdo (recursos). A fim de atender a esses requisitos, os sistemas de recuperação de informações multimídias comuns consistem em vários módulos funcionais (biblioteca de funções).

Outra característica desses sistemas é que são normalmente distribuídos, consistindo em vários servidores e clientes. Isso resulta do grande tamanho dos dados multimídia (a duplicação de bancos

de dados para cada usuário não pode ser fornecida) e da maneira como as informações multimídia são usadas (geralmente são acessadas por muitos usuários, como em uma biblioteca digital ou em um vídeo no sistema de demanda).

O papel de um modelo de dados nos DBMS é fornecer uma estrutura (ou linguagem) para expressar as propriedades dos itens de dados que devem ser armazenados e recuperados usando o sistema. A estrutura deve permitir que designers e usuários definam, insiram, excluam, modifiquem e pesquisem itens e propriedades do banco de dados.

Nos sistemas de recuperação da informação multimídia, bem como nos sistemas gerenciadores de banco de dados multimídia, o modelo de dados assume a função adicional de especificar e calcular diferentes níveis de abstração dos dados de multimídia, pois modelos de dados multimídia capturam as propriedades estáticas e dinâmicas do banco de dados. Portanto, eles fornecem uma base formal para o desenvolvimento das ferramentas apropriadas ao uso dos dados multimídia.

As propriedades estáticas podem incluir objetos que compõem os dados de multimídia, os relacionamentos entre objetos e atributos do objeto. As propriedades dinâmicas incluem aquelas relacionadas à interação entre objetos, operações em objetos, interação do usuário e assim por diante.

A riqueza do modelo de dados desempenha um papel fundamental na usabilidade do sistema de recu-

peração de informações multimídia. Embora os tipos básicos de dados multimídia devam ser suportados, eles fornecem apenas a base sobre a qual criar recursos adicionais. Os espaços de recurso multidimensionais são uma característica da indexação multimídia. Um modelo de dados deve suportar a representação desses espaços multidimensionais, especialmente as medidas de distância.

Um modelo de dados do sistema de recuperação de informação multimídia deve atender aos seguintes requisitos:

- O modelo de dados deve ser extensível para que novos tipos de dados possam ser adicionados.
- O modelo de dados deve ser capaz de representar tipos básicos de mídia e objetos compostos com relacionamentos espaciais e temporais.
- O modelo de dados deve ser flexível para que os itens possam ser especificados, consultados e pesquisados em diferentes níveis de abstração.
- O modelo de dados deve permitir armazenamento e pesquisa eficientes.

CONSIDERAÇÕES FINAIS

Neste módulo, apresentamos uma visão geral dos sistemas de banco de dados geográficos, bem como uma breve apresentação dos conceitos deste tipo de banco de dados. Na sequência, estudamos conceitos básicos dos sistemas de banco de dados multimídias e suas principais características.

SÍNTESE



ESTRUTURA E MODELAGEM DE DADOS

MODELAGEM DE DADOS 4

Neste módulo, apresentamos uma breve visão sobre os sistemas de banco de dados geográficos, bem como as condições e demandas que culminaram com o seu surgimento.

Em seguida, discutimos alguns conceitos importantes, como dados, informação, formas de indexação, sua relação com os sistemas de informações geográficas, tipos de dados espaciais, indexação e restrições de integridade.

– Por fim, estudamos os **sistemas de banco de dados multimídia**, a partir dos quais discutimos os principais conceitos relacionados a esse tipo de sistema, bem como sua relação com os **sistemas de recuperação da informação** (SRI). Encerramos o módulo com uma visão geral da **arquitetura básica**.

Referências Bibliográficas & Consultadas

ALVES, W. P. **Banco de Dados: teoria e desenvolvimento**. São Paulo: Érica, 2014.

ASCENCIO, A. F. G.; ARAÚJO, G. S. **Estruturas de dados: algoritmos, análise da complexidade e implementações em JAVA e C/C++**. São Paulo: Pearson Prentice Hall, 2010 [Biblioteca Virtual].

COCKCROFT, S. A Taxonomy of Spatial Constraints. **GeoInformatica**, v. 1, n. 4, p. 327-349, 1997.

COCKCROFT, S. User defined spatial business rules: Storage, management and implementation – A Pipe Network Case Study. *In: Information Science Discussion papers Series*, v. 98, n. 01. University of Otago, New Zealand, 1998.

ELMASRI, R.; NAVATHE, S. B. **Sistema de banco de dados**. 6. ed. São Paulo: Pearson Addison-Wesley, 2010 [Biblioteca Virtual].

GÜTING, R. H. An introduction to spatial database systems. **VLDB Journal**, v. 03, n. 01, p. 357-399, 1994. Disponível em: <http://www.vldb.org/journal/VLDBJ3/P357.pdf>. Acesso em: 21 nov. 2019.

GUOJUN, L. **Multimedia Database Management Systems**. Artech House, Inc. 1999.

HALL, G.B; YEUNG, A.K.W. **Spatial Database Systems: design, implementation and Project Management**. Springer. 2007.

HEUSER, C. A. **Projeto de banco de dados**. 6. ed. Porto Alegre: Bookman, 2008 [Minha Biblioteca].

MEDEIROS, L. F.; **Banco de dados: princípios e prática**. Curitiba: InterSaberes, 2013 [Biblioteca Virtual].

OGC. OPEN GIS CONSORTIUM. **OpenGIS simple feature specification for SQL (OpenGIS Project Document 99-049)**. Revisão 1.1, Wayland, MA: Open GIS Consortium, Inc. 1999.

PUGA, S.; FRANÇA, E.; GOYA, M. **Banco de dados: implementação em SQL, PL/SQL e Oracle 11g**. São Paulo: Pearson Education do Brasil, 2013 [Biblioteca Virtual].

RAMAKRISHNAN, R.; GEHRKE, J. **Sistemas de gerenciamento de banco de dados**. 3. ed. Porto Alegre: AMGH, 2008 [Minha Biblioteca].

ROB, P.; CORONEL, C. **Sistemas de banco de dados: projeto, implementação e gerenciamento**. 8. ed. São Paulo: Cengage Learning, 2011.

SANTOS, A. P. S. Qualidade de serviço na internet. **Boletim bimestral sobre tecnologia de redes**, v. 3, n. 6, nov.-1999. Disponível em: <https://memoria.rnp.br/newsgen/9911/qos.html>. Acesso em: 21 nov. 2019.

SARACEVIC, T. Information Science. **Journal of the American Society for Information Science**, v. 50, n. 12, p. 1051-1063, 1999.

FaTM
ONLINE