

Metodologia para Representação da Estrutura de Dados Geoespacial Vetorial da Mapoteca Nacional Digital em Bancos de Dados Geográficos Relacionais

Rogério Luís Ribeiro Borba³, Guilherme L. A. Mota^{1,3},
Jorge Luís Nunes e Silva Brito^{2,3}

¹Departamento de Engenharia de Sistemas e Computação

²Departamento de Engenharia Cartográfica

³Programa de Pós-graduação em Engenharia de Computação - Geomática
Universidade do Estado do Rio de Janeiro
Rio de Janeiro – RJ – Brasil

rogerioluisribeiro@yahoo.com, {guimota, jsilvabr}@gmail.com

Abstract. *A single model of vector geospatial data structure that could be used by institutions responsible for systematic land mapping, has always been a major challenge in mapping national scene. The National Commission of Cartography, in 2007, launched the technical specifications for geospatial data vector that finally unite the form of mapping models and break with old paradigms. One of these specifications is a conceptual object oriented model. This paper presents a solution to map this conceptual model for a generic model of relational geographic database.*

Resumo. *Um modelo único de estrutura de dados geoespacial vetorial que pudesse ser utilizado pelas instituições responsáveis pelo mapeamento sistemático terrestre, sempre foi um dos grandes desafios no cenário cartográfico nacional. A Comissão Nacional de Cartografia, em 2007, lançou as especificações técnicas de dados geoespaciais vetoriais que finalmente unificam a forma de fazer mapeamento e quebram paradigmas com modelos antigos. Uma destas especificações é um modelo conceitual orientado a objeto. O presente trabalho descreve uma solução para mapear este modelo conceitual para um modelo genérico de banco de dados relacional geográfico.*

1. Introdução

No Brasil há duas instituições públicas responsáveis pelo mapeamento topográfico sistemático terrestre. O Exército Brasileiro, através da Diretoria de Serviço Geográfico (DSG), e o Instituto Brasileiro de Geografia e Estatística (IBGE). Até o início dos anos 90, a chamada cartografia convencional ou analógica era executada, pelas duas organizações, de forma padronizada, segundo especificações técnicas obrigatórias, de acordo com resoluções definidas pelos órgãos responsáveis, em obediência a legislação específica (Delou, 2006).

A partir da década de 90, a tecnologia digital passou a ser utilizada pelos órgãos encarregados da produção cartográfica. O crescimento da produção de bases cartográficas digitais por estes órgãos, sem normatização, causou a ruptura do processo padronizado de produção cartográfica. Cada instituição criou sua própria padronização para a estruturação dos elementos cartográficos contidos nos arquivos vetoriais por ela

gerados. Assim, apesar das facilidades e avanços proporcionados pela introdução da tecnologia digital, passa a haver dissonância no processo de produção cartográfica digital. Enquanto a DSG baseava a estruturação dos elementos cartográficos nas Tabelas da Base Cartográfica Digital – TBCD; o IBGE baseava a estruturação dos elementos cartográficos na Mapoteca Topográfica Digital – MTD. Em função desta separação de modelos, diversos problemas foram gerados, conforme observado por Delou (2006).

No ano de 2004, a Comissão Nacional de Cartografia (CONCAR), por intermédio da Subcomissão de Estruturação de Dados Espaciais (SDE) e do Comitê de Estruturação da Mapoteca Nacional Digital, deu início à definição de um novo modelo para unificar a estruturação de elementos cartográficos para a cartografia no Brasil, o que aconteceu em 2007, tendo sido lançadas as “especificações técnicas para estruturação de dados geoespaciais vetoriais (EDGV)” [CONCAR 2007]. A EDGV é parte componente da Mapoteca Nacional Digital (MND). A MND é componente da estruturação de dados cartográficos do Mapeamento Sistemático Terrestre, da Infra-estrutura Nacional de Dados Espaciais (INDE) [CONCAR 2007].

A EDGV foi modelada segundo o paradigma da orientação a objetos (OO), uma vez que esta abordagem é a mais adequada para criação de um modelo do mundo real. Além disto o uso da abordagem OO foi consenso entre os membros da CONCAR, dentre os quais se destacam a DSG e o IBGE.

A princípio, tanto as aplicações quanto as bases de dados deveriam seguir o paradigma da OO. Contudo, no que se refere ao armazenamento dos dados, o modelo relacional é adotado na grande maioria das empresas devido à sua maturidade e seu formalismo matemático. Além disso, os investimentos realizados para a aquisição de sistemas gerenciadores de bancos de dados relacionais (SGBDR) e treinamento de pessoal teriam que ser refeitos no caso da migração para bancos de dados OO. Pelas razões anteriormente listadas, faz-se necessária a definição de uma estratégia para a implementação da EDGV em SGBDR e, assim, possibilitar a utilização da infra-estrutura atualmente instalada. É preciso ainda destacar dois requisitos adicionais desta demanda. Primeiro, o fato deste mapeamento ser feito em bases de dados geoespaciais, portanto, considerando as componentes geométricas. Segundo, a estratégia de mapeamento deve ser genérica de forma a permitir sua implementação em quaisquer SGBDR que possuam extensão espacial.

A presente iniciativa faz parte do projeto E-FOTO que desenvolve uma estação fotogramétrica digital educacional livre (licença GNU/GPL). A intenção é criar condições para que a restituição fotogramétrica digital do E-FOTO possa ser armazenada em SGBDR livres ou proprietários.

A seguir, na seção 2, é apresentada brevemente a EDGV. Na seção 3, é abordada a componente geométrica; na seção 4, é apresentada a metodologia proposta e o mapeamento em SGBDR; na seção 5, são apresentados os resultados obtidos. Por fim, a seção 6 apresenta as conclusões da pesquisa.

2. Estruturação dos dados geoespaciais vetoriais

As especificações técnicas da EDGV definem um padrão único a ser seguido pelas instituições responsáveis pelo mapeamento topográfico sistemático terrestre. No contexto da presente proposta, dentre as especificações da EDGV, o modelo conceitual OO é a mais importante. Esse modelo é organizado em treze categorias (pacotes-conforme UML): (1) Abastecimento de Água e Saneamento Básico; (2) Administração

Pública; (3) Educação e Cultura; (4) Energia e Comunicações; (5) Estrutura Econômica; (6) Hidrografia; (7) Limites; (8) Localidades; (9) Pontos de Referência; (10) Relevo; (11) Saúde e Serviço Social; (12) Sistema de Transportes; (13) Vegetação.

Cada categoria/pacote na EDGV contempla um diagrama de classes. Existem basicamente três tipos de classes:

- Classes de Feição - possuem geometria associada (ponto, linha ou polígono);
- Classes Agregadoras - não possuem geometria, somente agregam classes que possuem geometria;
- Classes de Domínio - não possuem geometria e nem agregam outras classes, simplesmente descrevem conjuntos de valores comuns. Podem ser referenciadas por classes do tipo Feição e Agregadora.

O modelo conceitual OO da EDGV após traduzido para o esquema relacional, poderá ser utilizado, por exemplo, no armazenamento de dados geográficos 3D adquiridos através do processo de restituição digital. Portanto, é necessário que o armazenamento de informações alfanuméricas e espaciais seja uniforme. A seção a seguir trata deste suporte por parte dos SGBDR.

3. O suporte a informação geográfica em SGBDR

Muitos SGBDR, como o Postgresql e o Oracle, disponibilizam mecanismos para que, além de informações alfanuméricas, sejam armazenados dados espaciais. Assim, a informação geográfica num registro de uma tabela pode coexistir com os respectivos dados alfanuméricos. Além disto, é importante mencionar que os bancos de dados supracitados, por seguirem os padrões definidos pelo Open Geospatial Consortium (OGC), possibilitam a interoperabilidade no tratamento da informação geográfica.

O modelo geométrico desenvolvido pela OGC tem como classe ancestral a *Geometry* que é abstrata e especializada em outras classes. Assim, é possível tratar adequadamente classes cuja forma geométrica não seja conhecida a priori. Este é o caso da classe *Elemento_Fisiografico_Natural* da categoria Relevo da EDGV, que pode ser representada como: ponto, linha e polígono. Assim ao mapear esta classe para uma tabela, uma coluna do tipo *Geometry* deve ser criada para tratar sua representação.

Por outro lado, alguns softwares não possuem a capacidade de lidar com mais de uma forma geométrica para uma dada feição. Este é o caso, por exemplo, do ArcGis, onde ao criar um elemento geográfico é preciso definir concretamente qual o tipo esperado de geometria (ponto, linha, polígono entre outros). Nesse caso, não é possível definir um tipo abstrato como por exemplo *Geometry*. Uma solução alternativa seria criar para cada forma geométrica uma tabela e repetir os atributos alfanuméricos.

4. Implementando o mapeamento do modelo conceitual orientado a objeto para o modelo relacional

Uma diferença importante entre os modelos OO e relacional reside nos tipos de relacionamentos possíveis. O modelo relacional possui apenas um tipo de relacionamento, enquanto o modelo OO possui mais de um tipo de relacionamento, cada um com um significado particular. Dentre os relacionamentos da OO, o modelo conceitual da EDGV emprega os mencionados a seguir: (1) associação; (2) agregação; (3) composição; (4) generalização/especialização. Além destes, a EDGV define um novo tipo de relacionamento para tratar a espacialidade. As subseções a seguir tratam do

mapeamento em modelo relacional dos relacionamentos citados acima.

4.1. Associação, agregação e composição

São formas de relacionamentos entre classes, parecidas com os relacionamentos entre tabelas no banco de dados relacional, que utilizam os mecanismos de restrição de integridade. No caso mais específico da composição, onde as partes só existem enquanto o todo existir, no modelo OO basta excluir o pai (todo) para que os filhos (partes) também sejam excluídos. No modelo relacional este mecanismo pode ser efetuado colocando na chave estrangeira restrição de integridade como: **on delete cascade**.

4.2. Generalização, Especialização

É um forma de relacionamento natural no modelo OO, porém no modelo relacional não existe tal mecanismo. Para simular o relacionamento de Generalização/Especialização no modelo relacional, há basicamente três estratégias: (1) Tabela por hierarquia; (2) Tabela por classe concreta; (3) Uma tabela por classe.

A presente proposta adotou a estratégia de uma tabela por hierarquia, que prioriza a simplicidade e a performance. Para este caso, todos os atributos são colocados em uma única tabela (Ambler, 2008). Dentre suas vantagens podem ser mencionadas: (1) simples implementação; (2) fácil inclusão de classes; (3) bom suporte a polimorfismo e (4) acesso rápido. Dentre as desvantagens podem ser enumeradas: (1) a modificação de uma classe pode afetar outras classes na hierarquia; (2) desperdício de espaço; (3) pode gerar tabelas muito grandes; (4) colunas não podem ter restrições do tipo *not-null* em alguns casos; (5) necessidade de uma coluna extra, para indicar a classe. Existem alguns casos de herança múltiplo no modelo conceitual da EDGV. Para mapeá-los também pode ser seguida a estratégia de tabela única. A figura 1 abaixo, exemplifica o uso da estratégia uma tabela por hierarquia.

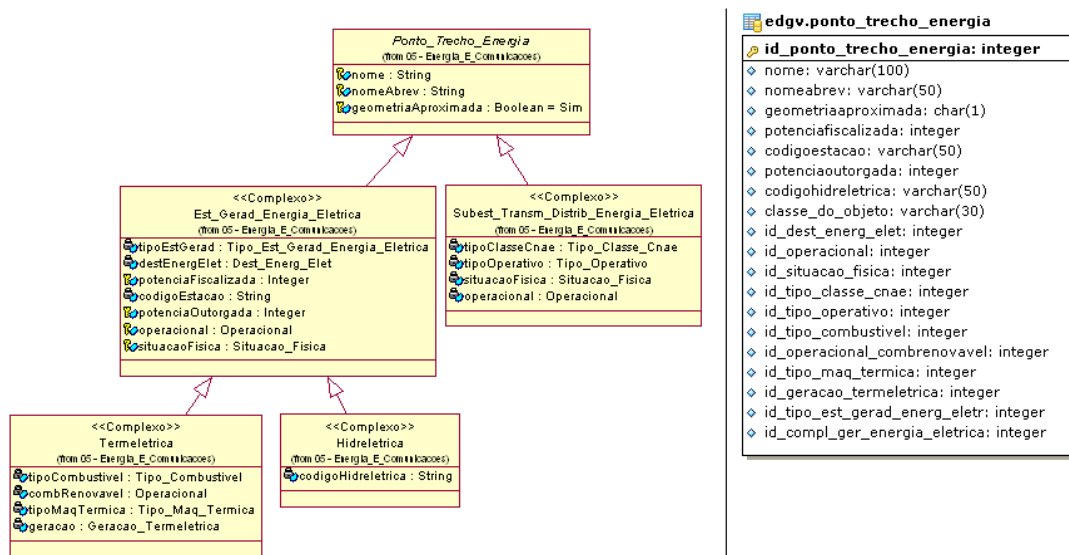


Figura 1. Exemplo da estratégia uma tabela por hierarquia

4.3. Relacionamentos espaciais

No modelo da EDGV além dos relacionamento padrões da orientação a objeto há os relacionamentos espaciais (estereotipo <<espacial>>). Este tipo de relacionamento pode

ser implementado diretamente através das operações espaciais dos SGBDR, sem necessitar da definição de relacionamento entre as respectivas tabelas. Um exemplo é apresentado na figura 2.

Neste exemplo, uma área político administrativa contém uma ou mais localidades. Observe que este relacionamento é espacial, ou seja, através de uma operação espacial é possível saber quais localidades pertencem a determinada área político administrativa. Assim, no banco de dados não seria necessário criar relacionamento entre as tabelas.



Figura 2 - Relacionamento espacial

5. Experimentos

Os experimentos foram realizados em ambiente de software livre, utilizando o SGBD PostgreSQL com a extensão espacial PostGIS. Portanto, foi necessário o desenvolvimento de scripts SQL. Os seguintes passos foram executados: (1) desenvolvimento e execução de um script para a criação das tabelas de feição; (2) adição de uma coluna, através da função *AddGeometryColumn*, para armazenar a geometria das feições nas tabelas criadas no passo 1 (esta função adiciona informações na tabela de metadados); (3) desenvolvimento e execução de um script para a criação das tabelas derivadas das classes agregadoras e de domínio; (4) desenvolvimento e execução de um script de restrição de integridade para gerar os relacionamentos entre as tabelas. Note que este mesmo procedimento poderia ser adotado para qualquer outro SGBD.

Na figura 3 é mostrada a visualização da inserção de um ponto e um polígono na tabela *Elemento_Fisiografico*, cuja tipo da coluna geométrica é *Geometry*, permitindo o armazenamento de um ponto, linha ou polígono.

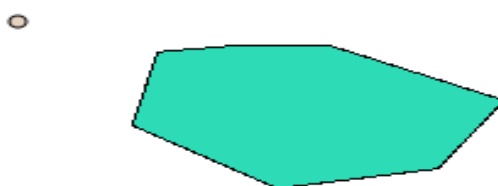


Figura 3. Resultado da inserção de duas feições na tabela Elemento_Fisiográfico visualizado no Software livre QGIS

O resultado alcançado é um esquema relacional/objeto-relacional genérico derivado do modelo conceitual da EDGV, que pode ser utilizado e especializado para PostgreSQL, Oracle Spatial entre outros. A portabilidade deste repositório de dados geoespaciais baseado nos padrões do OGC permite seu compartilhamento com aplicações e softwares geoespaciais.

6. Conclusão

Este trabalho apresentou uma solução para mapear a Estrutura de Dados Geoespacial Vetorial (EDGV), parte componente da MND para bancos de dados geográficos. Neste contexto, os relacionamentos do tipo composição, agregação e

associação simples no modelo orientado a objeto da EDGV tornaram-se, no modelo relacional, relacionamentos normais entre tabelas com suas respectivas restrições de integridade. Cada conjunto de classes do relacionamento de generalização/especialização condensou-se em uma tabela com uma coluna especial (*classe_do_objeto*) para identificar a classe correspondente. Ainda na generalização/especialização os relacionamentos das subclasses foram condensados nesta tabela. Os relacionamentos estereotipados de espacial da EDGV não produziram nenhum efeito no modelo relacional, uma vez que são derivados de funções espaciais. O produto obtido foi um esquema de banco de dados geográfico que será utilizado no software E-FOTO para armazenamento das feições obtida por restituição fotogramétrica digital.

Como consequência, este mapeamento disponibiliza um padrão de esquema em banco de dados relacional/objeto-relacional com extensão espacial, permitindo assim uma interoperabilidade em nível de base de dados. Acredita-se que esta solução de geração de esquema possa ser utilizada por qualquer instituição que pretenda realizar seu processos de mapeamento cartográfico em conformidade com a EDGV definida pela Comissão Nacional de Cartografia.

Referências bibliográficas

- Ambler, W. Scott. (2008) Mapping Objects to Relational Databases: O/R Mapping In Detail <http://www.agiledata.org/essays/mappingObjects.htm> [Acessado em agosto].
- Booch et. al. (2000)UML Guia do Usuário. Rio de Janeiro: Campus.
- Brito, J. L. N. et. al. (2006) Projeto E-FOTO: Disponibilizando uma Ferramenta educacional livre para a Fotogrametria Digital COBRAC 2006 Congresso Brasileiro de Cadastro Técnico Multifinalitário UFSC Florianópolis 15 a 19 de Outubro.
- Coelho L. e Brito J. L. N. (2007) Fotogrametria Digital. Rio de Janeiro: UERJ.
- CONCAR (2007)Especificações Técnicas para Estruturação de Dados Geoespaciais Digitais Vetoriais. Brasília.
- CONCAR (2007)- RESOLUCAO_CONCAR_001-2007. Brasília.
- Delou, A. L. (2006.)A. Estruturação de Bases Cartográficas Para SIG: Um estudo comparativo dos modelos do IBGE e da DSG. Rio de Janeiro: UERJ.
- Fowler, M. (2002)Analysis Patterns Reusable Object Models. MA.: Addison-Welsey
- Lunardi, O. A. & Augusto M. J. C. (2006)Infra-Estrutura dos Dados Espaciais Brasileira – Mapoteca Nacional Digital. COBRAC 2006 · Congresso Brasileiro de Cadastro Técnico Multifinalitário · UFSC Florianópolis · 15 a 19 de Outubro .
- Mapoteca Topográfica Digital(1996): Documento Geral – Versão 3.0. IBGE, Diretoria de Geociências, Departamento de Cartografia. Rio de Janeiro.