

Während in der Vergangenheit Geoinformationssysteme Geodaten eigenständig verwaltet haben, wird diese Aufgabe mehr und mehr separaten Geodatenbanksystemen übertragen. Damit erfolgt ein Übergang von proprietären Lösungen zu offenen Geodateninfrastrukturen, in denen Geodatenbanksysteme die Kernkomponente zur Verwaltung und Speicherung von Geodaten darstellen.

Offene Lösungen

Die Speicherung von Geodaten in Geoinformationssystemen war lange Zeit durch Eigenentwicklungen der GIS-Hersteller geprägt. So wurden bis Mitte der 1990er-Jahre die Daten vielfach ausschließlich in Dateien oder GIS-spezifischen Datenhaltungskomponenten gespeichert. Grund hierfür waren insbesondere die spezifischen funktionalen Anforderungen der Geoinformationssysteme. Erst mit der Entwicklung offener Geoinformationssysteme – insbesondere gemäß den Vorgaben des Open Geospatial Consortiums (OGC) – und mit der technischen Weiterentwicklung der Datenbanktechnologie fand dieser Zustand ein Ende.

Relationales Datenbankmodell

Einen entscheidenden Schritt für die heutige Datenbanktechnologie stellte die Entwicklung des relationalen Datenbankmodells dar. Anfang der 1980er-Jahre kamen relationale Datenbanksysteme (DBS) wie IBM DB2, Oracle und Ingres auf den Markt. Heute beträgt der Marktanteil von (objekt-) relationalen Datenbanksystemen rund 80 Prozent. Ebenso hat sich SQL als Datenbanksprache allgemein durchgesetzt.

Die Nutzung relationaler Datenbanksysteme zur Verwaltung räumlicher Daten ließ – wie bereits angedeutet – auf sich warten. Kernproblem ist, dass das relationale Datenbankmodell sich auf einfache Datentypen beschränkt, aber die Speicherung geometrischer Daten komplexe Datentypen wie Linienzüge und Multipolygone erfordert. Die Abbildung solcher Objekte in das relationale Datenbankmodell ist aufwändig und verhindert die Formulierung und Ausführung von räumlichen Anfragen.

Ein noch heute gängiger Ansatz, dieses Problem zu umgehen, stellt die Speicherung von Geometrien in Binary Large Objects (BLOBs) dar. Dies hat eine Reihe von Vorteilen: Es ist die integrierte Speicherung des gesamten Geoobjektes möglich, so dass DBS-Konzepte wie Zugriffskontrolle, gesicherter Mehrbenutzerbetrieb und Recovery auch auf die Geometrien angewendet werden können. Allerdings weist diese Lösung auch essenzielle Nachteile auf: Für das DBS ist ein BLOB eine Folge von Binärinformationen, die es nicht interpretieren kann. Da das DBS die Bedeutung der Binärdaten nicht kennt, können keine Mechanismen zur effizienten Anfragebearbeitung angewendet werden. Auch ist es nicht möglich, geometrische Operationen

zu verarbeiten oder durch die Anfragesprache zu unterstützen. Zudem sind BLOBs für externe Programme nicht ohne weiteres interpretierbar. Dies widerspricht aber der Idee offener Systeme.

Objektorientierung

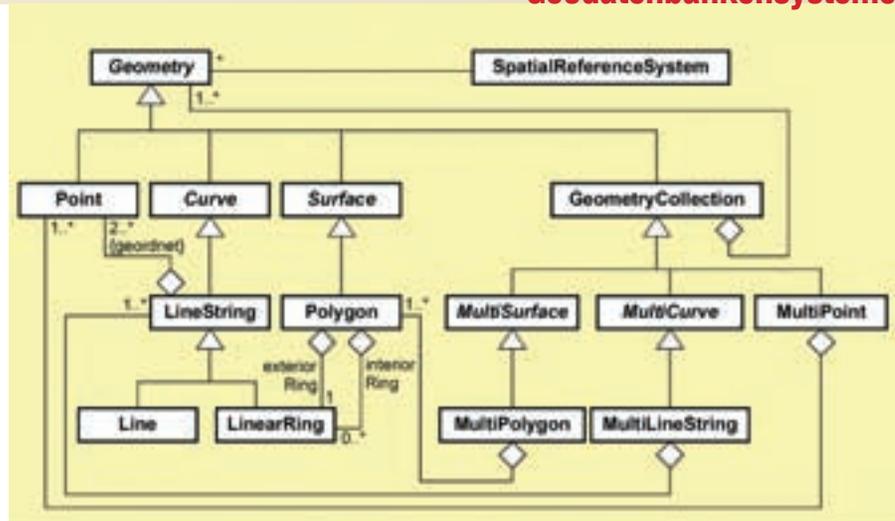
In den 1990er-Jahren wurde das Konzept der Objektorientierung in den Softwareentwurf und in Programmiersprachen eingeführt. Daher lag es nahe, das relationale Datenbankmodell um objektorientierte Konzepte zu erweitern. Dieses objektrelationale Datenbankmodell hat auch in neueren Versionen von SQL – SQL:1999 und SQL:2003 – seinen Niederschlag gefunden.

Das objektrelationale Datenbankmodell ist zur Speicherung von Geodaten gut geeignet: Komplexe geometrische Datentypen können durch entsprechende Klassen repräsentiert werden. Auch ist es möglich, die zur Analyse und Anfragebearbeitung benötigten geometrischen Funktionen dem DBS hinzuzufügen. Ein objektrelationales Geodatenbanksystem stellt bereits systemseitig einen Satz von vordefinierten Geometrieklassen und -methoden zur Verfügung. Der Zugriff auf die Geometrieobjekte ist von allen Anwendungsprogrammen aus mit Hilfe von SQL möglich. Damit wird einer wichtigen Anforderung für offene GI-Systeme Rechnung getragen. Da einem objektrelationalen Geodatenbanksystem die Struktur und die Bedeutung der geometrischen Datentypen bekannt sind, kann dieses Wissen auch bei der Anfragebearbeitung berücksichtigt werden.

Geometriedatenmodelle

Einen wichtigen Aspekt offener interoperabler Geoinformationssysteme stellt die Standardisierung des verwendeten Geometrieklassenmodells dar. Hier haben das OGC und die International Organization for Standardization (ISO) eine Reihe von Spezifikationen entwickelt.

In diesem Zusammenhang ist insbesondere das Simple Feature Model zu nennen, das durch das OGC erarbeitet wurde und seit zwei Jahren auch als ISO 19125:2004 genormt ist. Simple Features sind Geometrien im zweidimensionalen Raum, deren Stützpunkte geradlinig miteinander verbunden sind. Beispiele hierfür sind Punkte, Streckzüge und Polygone. Der erste Teil der Norm beschreibt das Klassenmodell mit geometrischen Datentypen



Geometrieklassen des Simple Feature Models.

und zugehörigen Operationen. Hierunter fallen Methoden zur Prüfung topologischer Beziehungen, zur Approximation von Geometrien, zur Berechnung geometrischer Eigenschaften wie Länge und Flächeninhalt und zur Verschneidung von Geometrien. Der zweite Teil der Norm behandelt die Umsetzung des Klassenmodells in ein Datenbankschema.

Der ISO-Standard 13249-3:2003 „SQL/MM Spatial“ spezifiziert ebenfalls ein Geometrieklassenmodell. Es erweitert das Simple Feature Model unter anderem um die Unterstützung von Kreisbögen, um Koordinatentransformationen und um Methoden zur Bereitstellung von GML (Geography Markup Language).

Um in einem Geodatenbanksystem aus den Koordinaten korrekte Längen und Flächen berechnen zu können, ist die Zuordnung und Interpretation eines räumlichen Bezugssystems zu einer Geometrie erforderlich. Für diesen Zweck haben sich die EPSG-Bezugssysteme durchgesetzt. Lineare Bezugssysteme erlauben die Identifizierung von Punkten auf einer Linie durch Abstandsangaben zu einem ausgezeichneten Anfangspunkt. Gerade in Netzwerken – zum Beispiel in Straßennetzen oder in Ver- und Entsorgungsnetzen wie Wasser- und Kanalnetzen, wie sie von Netzinformationssystemen verwaltet werden – haben lineare Bezugssysteme eine große Bedeutung und werden daher von gängigen Geodatenbanksystemen unterstützt.

Topologie- und Rasterdatenmodelle werden bislang von den erwähnten Standards nicht datenbankbezogen abgedeckt, sodass derzeitige Systeme hier entweder keine oder nur eine proprietäre Unterstützung anbieten. Als Beispiel seien hier die Erweiterungen in Oracle Spatial 10 ge-

nannt. Diese umfassen ein Netzwerkdatenmodell, das unter anderem Operationen zur Berechnung kürzester Wege anbietet, und ein Topologiedatenmodell, das die Speicherung topologischer Primitive und die Ableitung topologischer Beziehungen aus dem Datenmodell erlaubt. Das Rasterdatenmodell von Oracle unterstützt die Speicherung und Abfrage georeferenzierter Rasterkarten mit der Möglichkeit der Kachelung, dem Aufbau von Bildpyramiden und der Kompression der Daten.

Räumliche Basisanfragen

Eine Hauptaufgabe eines Geodatenbanksystems ist die Bestimmung von Geobjekten, deren Geometrie vorgegebene geometrische oder topologische Bedingungen erfüllt. Für diesen Zweck sollte das Geodatenbanksystem einen Satz von räumlichen Basisanfragen möglichst effizient bearbeiten können:

- Punktanfragen bestimmen zu einem Anfragepunkt alle Geobjekte, die den Punkt enthalten. Ein typischer Anwendungsfall ist die Selektion von Geobjekten über eine Koordinate oder eine interaktive Auswahl.
- Rechteckanfragen berechnen zu einem Anfragerechteck alle Geobjekte, die das Rechteck schneiden. Die Darstellung von (rechteckigen) Kartenausschnitten in einem GIS beruht auf der Rechteckanfrage.
- Eine Regionsanfrage bestimmt zu einem gegebenen Anfragepolygon alle Geobjekte, die das Polygon schneiden. Eine typische Anwendung ist die Berechnung aller Objekte innerhalb einer Pufferzone (etwa aller Spielplätze in einem 150 Meter breiten Streifen um alle Hauptstraßen einer Stadt).

■ Der räumliche Verbund findet in zwei Tabellen die Paare von Geoobjekten, die eine räumliche Verbundbedingung erfüllen, zum Beispiel sich schneiden oder eine Abstandsbedingung einhalten. Der räumliche Verbund ist Grundvoraussetzung für eine effiziente Verschneidung von Kartenebenen in einem GIS.

■ Nächste-Nachbarn-Anfragen berechnen in Hinblick auf einen Anfragepunkt die nächstgelegenen Geoobjekte. Gerade ortsbezogenen Dienste benötigen diese Anfrage für die Suche nach den nächstgelegenen Points of Interest.

Die effiziente Bearbeitung räumlicher Basisanfragen wird durch einige Charakteristika von Geodaten erschwert: So ist die Anzahl der Objekte in einer Geodatenbank unter Umständen sehr hoch. Außerdem ist verglichen mit Zeichenketten oder Zahlen die Verarbeitung geometrischer Datentypen weitaus aufwändiger. Daher werden räumliche Anfragen in mehreren Filterschritten bearbeitet. In frühen Filterschritten versucht man mit möglichst geringem Aufwand, die Geometrien zu bestimmen, die die Anfragebedingung erfüllen oder nicht erfüllen. Geometrien, bei denen man dies nicht entscheiden kann, werden danach in aufwändigeren Filterschritten geprüft.

Zwei wichtige Formen von Filterschritten stellen die Betrachtung von Geometrieapproximationen und der Einsatz von räumlichen Indizes dar. In Datenbanksystemen dienen Indizes dazu, die Daten so zu organisieren, dass Anfragen effizient die gesuchten Daten finden. Solche Indizes benötigen eine lineare Ordnung für die Daten, um die Datensätze sortiert anordnen zu können. Eine solche lineare Ordnung ist geometrischen Datentypen nicht offenkundig. Daher können herkömmliche Indizes nicht (ohne weiteres) für Geodaten eingesetzt werden; stattdessen werden spezifische räumliche Indizes wie der R-Baum benötigt. Ein R-Baum zerlegt den Datenraum schrittweise in Rechtecke, die die in dem jeweiligen Teilbereich gespeicherten Daten umfassen. Diese Rechtecke dürfen sich überlappen. Dadurch kann ein willkürliches Zerschneiden von Geometrien vermieden werden. Allerdings wird durch die Überlappungen die Effizienz räumlicher Anfragen negativ beeinflusst, da die Verarbeitung dort verzweigen muss. Daher versuchen verschiedene R-Baum-Varianten, solche Überlappungen möglichst klein zu halten.



Räumliche Basisanfragen

Aktuelle Geodatenbanksysteme

Derzeit sind eine Reihe objektrelationaler Geodatenbanksysteme auf dem Markt:

■ Für den IBM Informix Dynamic Server werden zur Verwaltung von Geodaten zwei unterschiedliche Erweiterungen angeboten: Das „Spatial DataBlade“ stellt geometrische Datentypen und Methoden gemäß dem Simple Feature Model zur Verfügung. Daneben gibt es zusätzliche Funktionen, die zum Beispiel den Geodaten-Server ArcSDE von Esri unterstützen. Alternativ wird das „Geodetic DataBlade“ angeboten, das geodätische Koordinaten verwaltet, sodass Höhenangaben unterstützt werden. Allerdings ist dieses DataBlade nicht OGC-konform. Zur Unterstützung räumlicher Basisanfragen setzen beide DataBlades räumliche Indizes ein.

■ Die Erweiterungskomponente für Geodaten im Datenbanksystem IBM DB2 ist der „DB2 Spatial Extender“. Auch der „DB2 Spatial Extender“ bietet OGC-konform geometrische Datentypen und Methoden an. Räumliche Anfragen werden durch eine hierarchische Gitterstruktur unterstützt.

■ Oracle ist derzeit Marktführer im Bereich (objekt)relationaler Datenbanksysteme. Bereits Mitte der 1990er-Jahre entstand „Oracle Multidimension“ zur Speicherung mehrdimensionaler Daten. Die spätere Umbenennung in „Spatial Data Option“ zeigt eine geänderte Zielrichtung: die Unterstützung zweidimensionaler Geodaten. Mit dem Release 8.1.5 erfolgte die Umbenennung in Oracle Spatial. Seitdem ist es in Oracle möglich, Geometrien objektrelational zu speichern, wobei auch die Vorgaben von SQL/MM Spatial berücksichtigt werden. Die Unterstützung geographischer Koordinatensysteme und von R-Bäumen erfolgte mit der Version 9. Oracle 10g

unterstützt topologische Datenmodelle und die Speicherung von georeferenzierten Rasterdaten und erlaubt den Einsatz von EPSG-Koordinatensystemen.

■ Im Bereich der Open-Source-Bewegung ist das objektrelationale Datenbanksystem PostgreSQL zu erwähnen. PostgreSQL unterstützt den R-Baum als räumlichen Index und stellt einige geometrische Datentypen bereit, die allerdings nicht OGC-konform sind. Dies wird durch die Erweiterung PostGIS erreicht, die auf PostgreSQL aufsetzt und ein voll funktionsfähiges, OGC-konformes Geodatenbanksystem bereitstellt.

■ MySQL als weiteres Open-Source-System ist ein relationales Datenbanksystem ohne objektrelationale Erweiterungen. Nichtsdestotrotz hat MySQL seit der Version 4.1 geometrische Datentypen „eingebaut“. Als räumlicher Index wird eine einfache Variante des R-Baums verwendet.

Moderne Geoinformationssysteme ermöglichen die direkte Speicherung ihrer Daten in Geodatenbanksystemen. Auch beruhen die derzeit verbreiteten Geodaten-Server, die Geodaten als WMS oder WFS anbieten, oftmals auf Geodatenbanksystemen.

Da die genannten Datenbanksysteme häufig auch als betriebliche Datenbanksysteme genutzt werden, ist eine integrierte Speicherung aller Geo- und Sachdaten einer Organisation möglich. Dies entlastet den Betriebsaufwand erheblich und erlaubt eine interoperable Nutzung. Durch die Bereitstellung der Daten über SQL-Anfragen und herkömmliche Datenbankschnittstellen können Geodaten ohne größere Probleme für Anwendungen aller Art verwendet werden. Auch ist ein direkter Zugriff auf räumliche Daten mit Programmiersprachen wie Java oder C++ möglich.

Die Standardisierung von SQL hat (bislang leider) nicht zu syntaktisch ein-

heitlichen SQL-Dialekten geführt. Dies gilt auch für die Geodatenbankerweiterungen. Somit ist es nicht möglich, mit einer syntaktisch identischen SQL-Anweisung Geodaten beispielsweise in IBM DB2, Oracle und PostgreSQL zu speichern oder anzufragen.

Zukünftige Entwicklungen

In den letzten Jahren haben 3D-Geoinformationssysteme zum Beispiel für die Erzeugung von 3D-Stadtmodellen deutlich an Bedeutung gewonnen. Dies legt die Verwendung von 3D-Geodatenbanksystemen nahe. Allerdings benötigt man für die Speicherung von 3D-Geodaten nicht immer ein 3D-Geodatenbanksystem. Oftmals reicht es, die Höheninformation als zusätzliches thematisches Attribut zu speichern (2,5D-Datenmodell). Der Bedarf, ein 3D-Geodatenbanksystem einzusetzen, ist davon abhängig, ob die dritte Dimension gleichberechtigt mit den ersten beiden Dimensionen genutzt wird. Auch ist zu betrachten, ob dreidimensionale Funktionen (zum Beispiel die Berechnung des Schnitts zwischen zwei Körpern) in den Anfragen benötigt werden. Ein 3D-Geodatenbanksystem muss dem Anwender ein geeignetes Datenbankmodell bereitstellen und die für 3D-Anwendungen erforderlichen räumlichen Basisanfragen durch entsprechende Indizes und geometrische Algorithmen unterstützen. Die datenbankbezogenen ISO- und OGC-Spezifikationen berücksichtigen 3D-Geometrien bislang nur sehr rudimentär. Gleiches gilt für die derzeit auf dem Markt befindlichen Geodatenbanksysteme.

Hier sind aber wohl demnächst entsprechend erweiterte Klassenmodelle zu erwarten.

Verwaltung von Messdaten

Mit dem Aufkommen moderner Messverfahren wie dem Laserscanning steigen die Datenvolumina georeferenzierter Messdaten massiv an. Dadurch entstehen sehr große Geodatenbanken im Terabyte-Bereich. Ein einzelnes Messdatum hat dabei eine recht einfache Struktur: typischerweise eine (3D- oder 4D-) Punktkoordinate mit einem oder mehreren Messwerten. Messsysteme verwalten Messdaten meist dateigestützt und überlassen die Auswertung speziellen Programmen oder CAD-Erweiterungen. Aufgrund der Datenvolumina stößt dieser Ansatz aber an seine Grenzen. Daher muss es ein Ziel sein, Anfragen, die zur Analyse von Messdaten erforderlich sind, durch Geodatenbanksysteme zu unterstützen.

Spatio-temporale Datenbanken

Bei Geodaten spielt insbesondere das Miteinander von Raum und Zeit eine entscheidende Rolle. So möchte man zum Beispiel

- wissen, wann das Ozonloch seine größte Ausdehnung in einem Jahr besitzt,
- alle Fahrzeuge abfragen, die schneller als 125 km/h fahren oder
- alle Personen bestimmen, die in den nächsten zehn Minuten eine Filiale der Firma X passieren werden.

Bei allen drei Beispielen liegt eine Kombination zwischen räumlichen und zeitlichen Anfragekriterien vor. Dafür bieten aber derzeitige Geodatenbanksysteme keine geeigneten Operationen und Indices an. Dies hat in den letzten Jahren zur Entwicklung von spatio-temporalen Datenbankmodellen und Indices geführt, die zurzeit aber noch Gegenstand der Forschung sind. ■

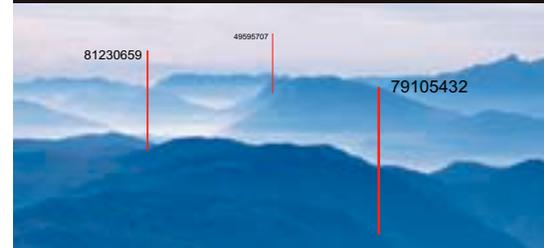
AUTOR

Prof. Dr. Thomas Brinhoff
Institut für Angewandte Photogrammetrie und Geoinformatik (IAPG)
FH Oldenburg/Ostfriesland/Wilhelmshaven
E-Mail: thomas.brinkhoff@fh-oldenburg.de

 www.fh-ooow.de



GIS nach Maß



an über 100 bayerischen Behörden im Einsatz

- ▶ ALB, DFK
- ▶ GPS Unterstützung
- ▶ ROK Signaturenkatalog
- ▶ dynamische Layer
- ▶ webbasierte Auskunft
- ▶ stufenweiser Ausbau
- ▶ zentraler Datenkatalog
- ▶ Schnittstellen
- ▶ OGC Dienste

SIRIS Produkte:

- ▶ **RISView**
Java Webauskunft
- ▶ **FINView**
GIS Client auf MO Basis
- ▶ **RISEdit**
ArcGIS Lösung mit Datenmanagement

Gesellschaft für Umweltplanung und Computergrafik mbH

Rosenkavalierplatz 10
81925 München

fon 089 / 92 14 33 81
fax 089 / 92 14 33 71

info@guc.de, www.guc.de



Links

- Geodatenbanksysteme in Theorie und Praxis: <http://www.geodbs.de>
- IBM DB2 Spatial Extender: <http://www-306.ibm.com/software/data/spatial/>
- IBM Informix Geodetic DataBlade: <http://www-306.ibm.com/software/data/informix/blades/geodetic/>
- IBM Informix Spatial DataBlade: <http://www-306.ibm.com/software/data/informix/blades/spatial/>
- Oracle Spatial: <http://www.oracle.com/technology/products/spatial/index.html>
- PostGIS: <http://www.postgis.org/>
- Rasdaman Rasterserver: <http://www.rasdaman.com/>