

Distributed GIS / Netzbasierte GIS-Anwendungen

Stefan LEHMKÜHLER

(Dipl.-Ing. Stefan LEHMKÜHLER, Fakultät Raumplanung, Stadt- und Regionalplanung Universität Dortmund, D-44221 Dortmund;
e-mail: pinkpant@srpplus.Raumplanung.Uni-Dortmund.DE)

KURZFASSUNG

Der Vortrag führt in die Thematik „Distributed GIS“ ein.

Zu diesem Zweck werden zum einen Grundlagen möglicher kombinierter GIS / Netzanwendungen erläutert und zum anderen die Nutzung einer bereits realisierten ARC/INFO-WWW-Anwendung gezeigt. Diese Anwendung gibt beliebigen Nutzern über das WWW den Zugriff auf die volle Analysefunktionalität ARC/INFOs. Sie kann damit insbesondere für Nutzer von sogenannten "Desktop GIS"-Systemen relevant sein, die Daten in einem durch ARC/INFO nutzbaren Format (Shapefile, e00, alle durch ARC/INFO konvertierbaren Formate) verarbeiten wollen. Abschließend werden - im mündlichen Vortrag - Potentiale des "Distributed GIS"-Ansatzes, u.a. die Nutzung von VRML-Anwendungen diskutiert.

1. EINFÜHRUNG

Der Begriff „Distributed GIS“ wurde in den letzten Jahren auf unterschiedliche Weise interpretiert. Obwohl alle Interpretationen zumindest im wesentlichen der direkten Übersetzung „Verteilte GIS(-Anwendungen)“ entsprachen, repräsentiert dieser Begriff auch heute noch ein Spektrum, welches auch weiterhin in vielfältiger Weise gefüllt wird. Um dieses interessante Themenfeld, welches sich - pauschal gesprochen - aus der Kopplung von „Geographischen Informationssystemen“ und „Vernetzung“ ergibt¹, zu erschließen und gleichzeitig in Bezug auf die Erfüllung eigener Ansprüche oder auch prinzipiell neuer Möglichkeiten zu prüfen, bietet sich eine grundsätzliche Betrachtung an.

Dazu werden im folgenden die funktionalen Komponenten geographischer Informationssysteme und das grundlegende Prinzip des „Internets“ als dem erfolgreichsten Beispiel für „Vernetzung“ dargelegt und auf dieser Basis eine Typisierung möglicher verteilter oder netzbasierender GIS-Anwendungen erarbeitet.

2. GRUNDLAGEN

Die Typisierung der zur Zeit bekannten „Verteilten GIS-Anwendungen“ beruht zum einen auf dem weithin akzeptierten GIS-Komponentenmodell, welches die Funktionalität von geographischen Informationssystemen in folgende vier Gruppen faßt:

- Erfassung
- Verwaltung / Verarbeitung
- Analyse
- Ausgabe / Präsentation²

und zum anderen auf dem die Basis der Internetkommunikation bildenden

- „Client / Server-Prinzip“.

Während das GIS-Komponentenmodell weitgehend bekannt sein dürfte, ist das zuletzt genannte Prinzip eine weitergehende Betrachtung wert.

Der Grundgedanke des „Client / Server-Prinzips“ besteht in der Realisierung einer Aufgabenteilung. Während der Nachfrager oder „Client“ (Kunde) eine Leistung anfordert, besteht die Aufgabe des „Server's“ darin, diese Nachfrage zu befriedigen. Im Bereich von lokalen Netzen aber auch des globalen Internets sind vielfältige Ausprägungen von gewünschten und angebotenen Leistungen bekannt. Diese reichen von der Bereitstellung von Dateien (File-Server) oder Datenbanken (Data-Server) bis hin zur Bereitstellung von Druck- (Print-Server) und Rechenkapazität (Compute-Server). In neuerer Zeit wurden diese schon als

¹ Die Begriffe „Verteilte GIS-Anwendungen“ und „Netzbasierte GIS-Anwendungen“ werden hier synonym verwendet.

² Vgl. Bill / Fritsch 1991, S.5.

traditionell zu bezeichnenden Dienste um die Bereitstellung von multimedialen Daten durch sogenannte „WWW-Server“ im World-Wide-Web (WWW) ergänzt.³ Bei diesen durch einen Web-Server angebotenen „Multimedia-Daten“ kann es sich um Grafiken, Geräusche, bewegte Bilder und um Text handeln. Diese Daten werden durch ein spezielles Anzeigeprogramm, einen sogenannten „WWW-Client“ (Browser) angefordert und nach Lieferung durch den Server auf dem Bildschirm des Nachfragers dargestellt. Im Kontext dieses Themas ist - neben der Möglichkeit Grafiken darzustellen - vor allem die Tatsache relevant, daß sich neben der Lieferung von Daten auch die Erfüllung von anderen, weitergehenden Aufgaben durch einen Server realisieren läßt.

Generell ist die Verteilung von Aufgaben (z.B. in Form von Unteraufträgen) auch bei GIS-Anwendern nicht unbekannt und zum Teil jahrelange Praxis. Digitalisier- oder Analysetätigkeiten, aber auch Plot- bzw. Ausgabebetätigkeiten können bei Bedarf „ausgelagert“ und durch Dritte gegen Rechnung wahrgenommen werden. Auf diese Art und Weise ergeben sich insbesondere für kleinere Planungsbüros Möglichkeiten, mit einem preiswerten „Desktop GIS“ komplexe Anforderungen eines Auftraggebers bei vertretbaren Investitionskosten zu erfüllen. Je nach personellen oder softwaretechnischen Ansprüchen⁴ können jedoch erhebliche Kosten durch die angesprochene „Auslagerung“ entstehen. Diese Kosten resultieren in der Regel aus der partiellen Weitergabe der hohen Anschaffungskosten höher qualifizierter GIS-Software, eben den Kosten, die eine Anschaffung dieser Software durch eines kleineres Planungsbüro unmöglich machen. Neben dieser Weitergabe von Investitionskosten werden ebenso die durch Koordination und Aufgabendurchführung entstehenden Personalkosten an den (Unter-)Auftraggeber zurückgegeben.

Die Realisierung dieser aus der GIS-Praxis bekannten Aufgabenteilung auf Basis z.B. des Internets bzw. des WWW bietet - einen auftraggeber- und auftragnehmerseitig vorhandenen Netzzugang vorausgesetzt - die Möglichkeit, beliebige Funktionalität schnell und kostengünstig anzubieten bzw. zu nutzen. Kostengünstig deshalb, da zum einen der Kreis von Nachfragern nicht mehr räumlich beschränkt ist und zum anderen Personalkosten entfallen können. Die Lockerung der räumlichen Beschränkung ergibt sich aus der Tatsache, daß der Zugang zum WWW in der Regel Fernsprechgebühren in Höhe des Ortstarifs erfordert, die Nutzung geographisch weit entfernter Ressourcen jedoch, genauso wie die Nutzung von Angeboten der Nachbarstadt, nicht kostenwirksam ist.⁵ Die Reduktion der Personalkosten wiederum resultiert einerseits aus der vollständig digitalen Abwicklung von Transaktionen, die eine manuelle Bearbeitung in der Vielzahl der Fälle überflüssig macht und andererseits aus der vom Nachfrager selbständig vorzunehmenden Auftragspezifikation. Unter Nutzung dieser beiden Charakteristika hat sich in bezug auf verteilte oder netzbasierte GIS-Anwendungen vor allem ein Anwendungstyp herausgebildet.

Bei diesem Typ verteilter GIS-Anwendungen handelt es sich um eine sogenannte „Atlas-Anwendung“. Analog zu einer Recherche in einem Schulatlas wird ein gegebener Datenbestand abgefragt und orientiert an den Vorgaben des Kunden präsentiert. Basierend auf dem oben angeführten GIS-Komponentenmodell und unter Integration des Client / Server-Prinzips ergibt sich ein Typ mit folgender Charakteristik:

Client	Server
	Erfassung
	Verwaltung / Verarbeitung
	Analyse
	Präsentation

Abb. 1: Atlas-Typ, Quelle: Eigene Darstellung

Als Beispiel dieses Typs ist eine Bildschirmphotographie einer „MapObjects Internet Map Server“-Recherche der Firma ESRI Inc. in Abb. 2 dargestellt.⁶

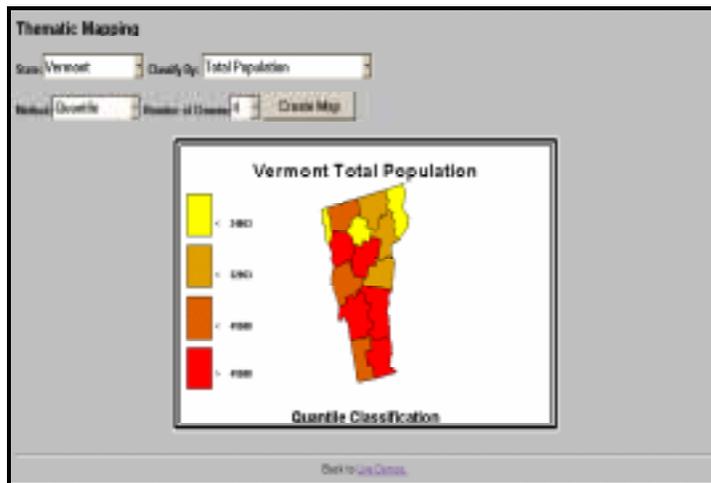
³ Ohne in unzulässiger Weise zu verkürzen, kann das „World-Wide-Web“ als Multimediaeteil des Internets bezeichnet werden.

⁴ Hier ist vorrangig an Ansprüche in bezug auf die Analysefunktionalität von GIS-Software gedacht.

⁵ Angemerkt sei an dieser Stelle, daß somit auch kleineren Planungsbüros prinzipiell die Möglichkeit der Abwicklung globaler Geschäfte gegeben wird.

⁶ Die Inhalte basieren auf dem von ESRI Inc. vertriebenen „Maps & Data Volume 1“ von 1996. Hier wurde ein „Census Data 1990“-Datensatz dargestellt.

Abb. 2: „MapObjects Internet Map Server“-Recherche



Quelle: <http://maps.esri.com/scripts/esrimap.dll?name=ThemeMap&Cmd=Map&State=Vermont&Attribute=Total+Population&Method=Qua>

konstatieren, daß dieser Typ die Verwendung „eigener“ untersuchungsspezifischer Daten nicht zuläßt. Diese „Distributed GIS“-Anwendung - die Abfrage einer bestehenden Datenbank mit anschließender grafischer „Onscreen-Präsentation“ des Rechercheergebnisses - berücksichtigt demnach die Belange kleinerer Planungsbüros oder allgemein die Belange der Anwender von sog. „Desktop GIS“-Systemen nicht hinreichend.

3. POTENTIALE

Nutzer sog. „Desktop GIS“-Systeme zeichnen sich aber nicht nur durch besondere Belange bzw. Wünsche aus, sondern der Einsatz von Software dieses spezifischen Typs bietet auch spezifische Potentiale. So ist es zum Beispiel nicht erforderlich, die komplette Funktionalität an einen Dienstleister oder Server zu übertragen. Zur Zeit verfügbare „Desktop GIS“-Systeme (z.B. ArcView 3.0 aus dem Hause ESRI) bieten neben der Möglichkeit Daten zu verwalten und zu präsentieren auch Möglichkeiten, projektbezogenen Daten zu erfassen. Auch die Analysekomponente wurde in letzter Zeit weiter entwickelt, sodaß diese „kleinen“ GIS-Systeme durchaus bei einer Vielzahl von Anwendungsfällen ausreichend sind. Lediglich komplexe Aufgabenstellungen, die eine umfassende Analysefunktionalität erfordern, zwingen den Anwender, Leistungen in Form von Unteraufträgen bei Dritten „einzukaufen“.

- Da diese Externalisierung vielfach mit erheblichem Kosten- und Koordinationsaufwand verbunden ist⁸, stellt die Prüfung der durch eine „Distributed GIS“-Anwendung entstehenden Möglichkeiten einen durchaus praxisrelevanten Ansatz dar.

Wie festgestellt, ist die Funktionalität in bezug auf die Komponenten „Erfassung“, „Verwaltung / Verarbeitung“ sowie „Präsentation“ in der Regel bei den auf dem Markt verfügbaren „Desktop GIS“-Systemen ausreichend.

Somit bietet sich unter Nutzung einer Client / Server-Konfiguration folgende Aufgabenverteilung an:

Client	Server
Erfassung	
Verwaltung / Verarbeitung	
	Analyse
Präsentation	

Abb. 3: Mehrwert-Typ
Quelle: Eigene Darstellung

Obwohl dieser Typ zur Erschließung z.B. durch die öffentliche Hand vorgehaltener Datenbestände prädestiniert ist und dazu auch - vor allem in den Vereinigten Staaten von Amerika - genutzt wird, macht die Verlagerung der kompletten GIS-Funktionalität auf einen WWW-Server nur sehr eingeschränkt von den Möglichkeiten des „Distributed GIS“ Gebrauch. Darüber hinaus schränkt die Zahl der unterstützten Grafikformate die Qualität der präsentierten Karte erheblich ein. Die resultierende Grafik ist in der Regel auf eine Darstellung von 500 * 300 Bildpunkten beschränkt.⁷

Neben diesen Mängeln hinsichtlich der Qualität der präsentierten Karte ist zu

⁷ Kritiker bezeichnen diesen Typ einer „Distributed GIS“-Anwendung daher auch als „Fax-Lösung“, da sie nur Bilder von Karten (und nicht Kartenoriginale) liefert.

⁸ Vgl. die unter 2. GRUNDLAGEN angeführten Rahmenbedingungen.

Diese Aufteilung bietet, neben der Nutzung von projektspezifischen - durch den „Client“ - erzeugten Daten, die Möglichkeit, die Analyseergebnisse in Form von Daten und nicht als „Bilder von Daten“ bereitzustellen. Der Kunde - in diesem Szenario das kleine Planungsbüro - behält die Freiheit, die resultierenden Daten weiterzuverarbeiten und nach eigenen Wünschen zu präsentieren.⁹

4. DIE ANWENDUNG: PINKNET

Eine Applikation dieses „Mehrwert-Typs“ wurde im Winter 1995 an der Fakultät Raumplanung der Universität Dortmund entwickelt.¹⁰ Dieser Entwicklung lagen folgende Ziele zugrunde:

- Erschließung der Analysefunktionalität ARC/INFO's für jedermann
- weltweite Nutzbarkeit
- integrierte Abrechnungsfunktionalität
- Mehrbenutzerbetrieb
- geringer Schulungsaufwand.

Ausgehend von einem auftraggeberseitig (Client) vorhandenen „Desktop GIS“-System und der Nutzung des Internets als weltweit etablierter Kommunikationsinfrastruktur, wurde die Kopplung des GIS ARC/INFO mit einem WWW-Server als technische Basis der Applikation definiert. Das WWW-Serverprogramm (Cern-httpd) wird auf einer SUN 2 (IPX) mit 32 MB Hauptspeicher und einer 1 GB Festplatte unter Solaris 2.4 (UNIX) betrieben und stellt gleichzeitig die Präsentation der Fakultät Raumplanung im WWW bereit¹¹. Bei der auf dieser Hardware ebenfalls installierten GIS-Software handelt es sich um ARC/INFO in der Version 6.1.

Neben der einfachen Lieferung von statischen WWW-Seiten durch einen WWW-Server auf Anfrage eines WWW-Clients (Browsers) bot die zugrunde liegende Software¹² schon früh die Möglichkeit, dynamisch - auf Anforderung des Clients - Informationsseiten zu erstellen und dem Nachfrager zuzusenden. Hierzu werden durch den Server Möglichkeiten geboten, serverseitig vorhandene Programme durch den Nutzer des WWW-Browsers aufzurufen. Die Definition der bei diesem Aufruf anzuwendenden Konventionen (z.B. die Nutzung bestimmter Variablen) ist Bestandteil des zugrundeliegenden Kommunikationsprotokolls und somit weltweit eindeutig. Im Herbst 1995 wurde zudem durch den Hersteller des führenden (und besten!!) WWW-Browsers eine Möglichkeit geschaffen, den Transfer von Dateien nicht nur vom Server auf den Client (download), sondern auch in Gegenrichtung (http-upload) durchzuführen. Die Möglichkeit, Dateien beliebig zwischen Client und Server auszutauschen - die Grundlage für den Transfer der „Rohdaten“¹³ zum Dienstleister - war somit gegeben.

Der wesentliche Bestandteil der Applikationsentwicklung lag somit, neben der Erstellung einer intuitiven Benutzeroberfläche in HTML, in der Entwicklung serverseitiger Programme, welche:

- die Benutzeroberfläche dynamisch generieren
- die Benutzereingaben und die transferierten „Rohdaten“ entgegennehmen
- die Durchführung der gewünschten Analyse durch ARC/INFO steuern
- die resultierenden Analyseergebnisse zur Verfügung stellen
- die Abrechnung vornehmen und
- den Auftraggeber zum Abruf der Analyseergebnisse auffordern.

⁹ Da bei Nutzung dieses „Distributed GIS“-Ansatzes jeweils „echte“ Daten übertragen werden und die Präsentation auf herkömmlichem Wege beim Planungsbüro erfolgt, entfallen somit auch alle Restriktionen hinsichtlich der Auflösung grafischer Darstellungen.

¹⁰ Obwohl die Entwicklung inkl. intensiver Testläufe 1996 erfolgreich abgeschlossen wurde, ist die Nutzung der Anwendung aus lizenzrechtlichen Gründen nicht möglich, da die Fakultät Raumplanung eine Hochschullizenz an ARC/INFO besitzt, welche eine kommerzielle Verwertung der Software ausschließt.

¹¹ Diese Präsentation ist unter „<http://www.raumplanung.uni-dortmund.de>“ abzurufen.

¹² Hierbei handelt es sich beim zugrundeliegenden Kommunikationsprotokoll um das „HyperText Transfer Protocol“ (HTTP) und bei der zur Darstellung genutzten Programmiersprache um die „HyperText Markup Language“ (HTML).

¹³ Mit „Rohdaten“ werden im folgenden die durch den Server zu analysierenden Datenbestände bezeichnet.

Zur Realisierung dieser Anforderungen wurden drei Programme in der Programmiersprache PERL entwickelt, die aufeinander aufbauen.

Das erste Programm (arc_customer.cgi)

- identifiziert den Kunden aufgrund einer zwingend erforderlichen Benutzer / Passwort-Angabe anhand der Umgebungsvariable „REMOTE_USER“ und
- erzeugt unter Angabe des vorhandenen Guthabens und der ggf. bereits auf dem Server vorhandenen Coverages und Dateien des Kunden die Benutzeroberfläche (s. Abb. 4).

Das zweite Programm (arcnow.pl)

- prüft das angezeigte Guthaben des Kunden auf einen Wert größer 50 (DM)
- nimmt - bei positiv verlauteter Prüfung - die Benutzereingaben und Rohdaten auf
- speichert sie im Verzeichnis des Kunden auf dem Server ab
- erzeugt eine „todo“-Anweisungsdatei für das dritte Programm und
- gibt dem Kunden ein unmittelbares Feedback.

Das dritte Programm (arc.pl)

- wertet die vom zweiten Programm erzeugte „todo“-Anweisungsdatei aus
- nimmt die erforderlichen Dekomprimierungs- und Extraktionsschritte vor¹⁴
- führt die rohdatentyp-spezifischen Vorarbeiten durch (shapearc, import, clean, usw.)
- veranlaßt ARC/INFO, das durch den Kunden angegebene Kommando abzuarbeiten
- konvertiert das Analyseergebnis ins E00-Format und komprimiert dieses
- erzeugt eine „log“-Datei
- schickt dem Auftraggeber eine Benachrichtigung per email (s. Abb. 5) und
- löscht die „todo“-Anweisungsdatei.

Während die beiden ersten Programme auf Anforderung eines Kunden gestartet werden, ist das dritte Programm permanent aktiv und prüft in einem vorgegebenen Intervall die Existenz einer „todo“-Anweisungsdatei. Aufgrund der Entkopplung von Dateitransfer, Analyse durch ARC/INFO und Abruf des Analyseergebnisses durch den Kunden nach Benachrichtigung besteht kein Erfordernis, eine permanente Verbindung zum Server aufrecht zu erhalten. Dieser Aspekt ist insbesondere bei zeitintensiven

Abb. 4: Benutzeroberfläche „Pinknet“; Quelle: Eigene Darstellung



Analysen relevant. Diese zeitintensiven Analysen müssen nicht nur aus einem ARC/INFO-Kommando bestehen, da - wie auf Abbildung 4 gezeigt - durchaus die Möglichkeit besteht, mehrere Dateien gleichzeitig auf den ARC/INFO-WWW-Server zu transferieren und es sich dabei auch um AML-Dateien handeln kann, welche z.B. die Abarbeitung einer ganzen Folge von Kommandos ermöglichen.

Abschließend bleibt somit festzuhalten, daß dieser Ansatz des „Distributed GIS“ die Durchführung selbst komplexer Simulationsmodelle erlaubt und dem

¹⁴ Zur Zeit wird die Annahme von mit „compress“ (*.Z) und „gzip“ (*.gz) komprimierten und von mit „tar“ archivierten Rohdaten unterstützt.

