

DaFuS – ein automatisierter Geodatenfilter für die Erzeugung optimaler Geoinformationen

Hartmut ASCHE und Silvijja STANKUTE

Universität Potsdam, Potsdam · gislab@uni-potsdam.de

Dieser Beitrag wurde durch das Programmkomitee als „reviewed paper“ angenommen.

Zusammenfassung

Als wertvolle wirtschaftliche Ressource der Industrie gewinnt Geoinformation zunehmend an Bedeutung. Ein erheblicher Bedarf besteht an harmonisierten, einheitlichen Daten für eine bestimmte, oft interregionale räumliche Abdeckung. Eine effektive Nutzung und Pflege von Geodatenbeständen erfordert die Harmonisierung heterogener Geodaten nach den jeweiligen anwendungsspezifischen Qualitätsvorgaben. Der vorliegende Beitrag beschreibt ein neuartiges Data-Mining-Konzept. Dessen Ziel ist es, aus einer Vielfalt heterogener Datensätze einen optimalen Geodatenatz mit optimalen geometrischen und/oder semantischen Eigenschaften in einem automatisierten und benutzerdefinierten Fusionsprozess zu generieren. In einem laufenden Forschungsprojekt haben die Autoren ein automatisiertes Daten-Fusions-System *DaFuS* entwickelt, das neue raumbezogene Daten aus vorhandenen Vektor-Datensätzen generiert. Die erzeugten Datensätze besitzen eine höhere geometrische und/oder semantische Qualität als die unvollständigen Input-Datensätze.

1 Einführung

Als wertvolle wirtschaftliche Ressource der Industrie gewinnt Geoinformation zunehmend an Bedeutung. Das Ergebnis ist ein starker Anstieg der Nachfrage nach einer breiten Palette von Geodaten in verschiedenen Maßstabbereichen und für eine Vielzahl von Anwendungen. Ein erheblicher Bedarf besteht an harmonisierten, einheitlichen Daten für eine bestimmte, oft interregionale räumliche Abdeckung. Inzwischen übersteigt die verfügbare Menge an Geodaten bei Weitem die Fähigkeit, diese Datenmassen sinnvoll zu verarbeiten und zu nutzen. Aus diesem Grund können Geodaten einer bestimmten Region, eines Maßstabbereiches oder einer Qualitätsstufe zwar verfügbar, aber nicht zugänglich oder nutzbar sein. Dies hat verschiedene Ursachen, wovon politische Restriktionen und hohe Nutzungsgebühren die wichtigsten sind.

Geodaten von Objekten unserer physischen sowie der anthropogenen Umwelt werden turnusmäßig durch eine Reihe von Akteuren, darunter staatliche Institutionen, wie nationalen Einrichtungen der Landesvermessung und privater Unternehmen, erfasst. Die Erfassung räumlicher Daten erfolgt gemäß den spezifischen Anwendungsanforderungen dieser Akteure und der erforderlichen Flächendeckung. Es liegt auf der Hand, dass die Anforderungen an Daten und deren Nutzung sowie die Spezifikationen für die Datenerfassung unter den

Datenerfassern signifikant variieren. Im Ergebnis entsteht eine Vielzahl heterogener, häufig redundanter Datenbanken mit identischer räumlicher Abdeckung. Allerdings ist die geometrische und semantische Qualität der erfassten Daten oft unzureichend und ungenau (vgl. DEVILLERS & JEANSOULIN 2002, HUNTER & VIRRANTAS 2009, ONCHAGA & LAMBERT 2008). Abbildung 1 zeigt die Mittellinien eines Straßennetzes in einem Wohngebiet in Potsdam (Deutschland). Das Vektordatenbeispiel zeigt Straßendaten der Hersteller *TeleAtlas/TomTom*, *Navteq* sowie geotopographische Basisdaten des *ATKIS*-Geoinformationssystems der Brandenburgischen Landesvermessung (Deutschland).¹ Die Vektordatensätze sind einem hochauflösenden georeferenzierten Fernerkundungsbild überlagert, das mit dem *HRSC-A*-Sensor² aufgenommen wurde.

Es ist ersichtlich, dass die geometrische Qualität eines identischen Objektes der realen Welt (in diesem Fall: Straßen) in den entsprechenden Datensätzen stark variiert. Das Beispiel verdeutlicht auch, dass die Qualität der vorhandenen Geodatenätze häufig problematisch ist. Daten sind in ihren geometrischen und/oder semantischen Merkmalen oft unvollständig und damit ungeeignet für jede gezielte Verwendung der Daten.

Eine effektive Nutzung und Pflege von Geodatenbeständen erfordert die Harmonisierung von heterogenen Geodaten nach den jeweiligen anwendungsspezifischen Qualitätsvorgaben. Eine neue Datenerfassung nach einheitlichen, anwendungsspezifischen Anforderungen ist komplex, zeitaufwendig und kostenträchtig. Insofern ist eine den individuellen Anforderungen einer speziellen Anwendung entsprechende Datenerfassung weder sinnvoll noch optimal. Vielmehr würde ein Verfahren benötigt, das einen optimalen Datensatz aus einer Reihe existierender suboptimaler Datensätze mit identischer räumlicher Abdeckung, Thematik und Maßstab erzeugt. Anders formuliert, wäre ein automatisierter Prozess zur Zusammenführung unvollkommener geometrischer und/oder semantischer Informationen aus zwei oder mehreren Quell-Datensätzen zu einem einzigen Zieldatensatz mit optimalen geometrischen und/oder semantischen Eigenschaften ein wichtiger Beitrag zur Reduktion wiederholter Datenerfassung, Minimierung der Datenredundanz und zur Verbesserung der vorhandenen Datenqualität. In einem laufenden Forschungsprojekt haben die Autoren ein automatisiertes Daten-Fusions-System *DaFuS* entwickelt, das neue raumbezogene Daten aus vorhandenen Vektor-Datensätzen generiert (STANKUTE & ASCHE 2009, 2010, 2011,

¹ Das niederländische Unternehmen *TeleAtlas/TomTom* (gegründet 1984, jetzt *TomTom*) und der US-Firma *Navteq* (gegründet 1985, jetzt Teil von *Nokia*) zählen zu den führenden Anbietern digitaler Vektordaten und Karten für die Fahrzeugnavigation. Die Geodatenbanken von *Tele Atlas/TomTom* und *Navteq* sind in Europa weitverbreitet. Die Daten werden nicht nur für die Fahrzeugnavigation, sondern auch für eine Vielzahl anderer Mapping-Anwendungen, z. B. Location Based Services, Marketing, Geomarketing usw. verwendet. Das deutsche Amtliche Topographisch-Kartographische Informationssystem (*ATKIS*) ist im Kern eine geotopographische Datenbank der Vermessungsbehörden der deutschen Bundesländer sowie des Bundes. *ATKIS* enthält verschiedene digitale Modelle der Erdoberfläche, sogenannte Digitale Landschaftsmodelle (*DLM*). Die in diesem Beitrag verwendeten *ATKIS*-Vektordatensätze entstammen den Basis-DLM der Landesvermessung Brandenburg.

² Die High Resolution Stereo Camera (*HRSC*) ist eine digitale hochauflösende multispektrale Stereokamera, die für die Mars-Missionen „Mars’96“, „Mars Express“ entwickelt wurde. Sie ermöglicht die multispektrale, dreidimensionale Kartierung des Mars durch Aufnahmen mit einer Bodenauflösung von 10 m bzw. 2m in bestimmten Bereichen. Eine modifizierte Version, die *HRSC-A*, wurde für die flugzeuggestützte Erfassung hochauflösender photographischer Bilder ausgelegt. Als Bodenauflösung kann aus 3.500 Metern Flughöhe bis zu 15 cm/Pixel erreicht werden. In diesem Beitrag wird als topographische Referenz mit freundlicher Genehmigung des DLR Berlin ein georeferenziertes *HRSC-A* Farbkomposit mit einer Bodenauflösung von 23 cm eingesetzt.

2012). Die erzeugten Datensätze besitzen eine höhere geometrische und/oder semantische Qualität als die unvollständigen Input-Datensätze.

Nach DIN ISO 8402 (1992) ist die Qualität von Geodaten (wie andere Informationsprodukte) nicht absolut, sondern anforderungsbezogen bestimmt. Dementsprechend wird die Qualität der in *DaFuS* verarbeiteten Geodaten als Eignungsoptimum für eine spezifische Anwendung verstanden. Die Qualität (das Eignungsoptimum) ergibt sich als Summe aus verschiedenen, gewichteten Merkmalen. Die hier relevanten Parameter sind: räumliche Vollständigkeit und geometrische Genauigkeit sowie inhaltliche Vollständigkeit und Attributgenauigkeit. Auf diese Qualitätsmerkmale wird in Kapitel 4 an Beispielen eingegangen.



Abb. 1: Suboptimale Vektordatensätze in einem Wohngebiet (Potsdam) (*TeleAtlas/TomTom* [weiß], *Naveq* [schwarz], *ATKIS* [grau]), überlagert in einem hochauflösenden *HRSC-A*-Bild

2 Konzept zur automatisierten Fusion suboptimaler Geodaten

Das nachfolgend vorgestellte *DaFuS*-Technologiekonzept zielt darauf ab, vorhandene räumliche Daten unterschiedlicher Herkunft zu kombinieren und in einen neuen, benutzerdefinierten Datensatz zu überführen. Hierfür werden die Attribute von zwei oder mehr unvollkommenen Datensätzen in einem automatisierten Prozess miteinander verschmolzen. Im Ergebnis entsteht ein optimierter Geodatensatz mit allen verfügbaren geometrischen und semantischen Informationen, die für eine bestimmte Anwendung erforderlich sind. Der so erzeugte Datensatz enthält alle ausgewählten geometrischen und thematischen Eigenschaften der Quelldatensätze. Attribute der Quelldaten, die für die Generierung des Zieldatensatzes für eine bestimmte Anwendung irrelevant sind, werden während des benutzerdefinierten Fusionsprozesses verworfen. Im Kern ermöglicht *DaFuS*, alle relevanten Informationen aus unterschiedlichen Quelldaten zu extrahieren, zu filtern, zu harmonisieren und zu kom-

binieren, um einen durch Anwender und Anwendung spezifizierten Datensatz optimaler Qualität zu erhalten.

Bei den prozessierten Daten handelt es sich um Vektordaten im ASCII-Format. Die Geometrie dieser Geodaten wird generell durch die geometrischen Primitive Punkt (0D), Linie (1D) und geschlossenes Polygon (2D) repräsentiert. Allerdings variiert der thematische Bezug stark. Er reicht bei einem Straßennetz von Attributen, z. B. Breite bzw. Anzahl der Fahrspuren oder Navigationsfähigkeit, bis Attributen der Flächennutzung, wie Größe, Nutzung oder Eigentümer von Flächen bzw. Parzellen. Von den Landesvermessungen oder größeren gewerblichen Unternehmen erfasste Geodaten entsprechen meistens den definierten und dokumentierten Qualitätsstandards. Derartige Geodaten können somit leicht validiert und für weitere Verarbeitung mit *DaFuS* genutzt werden. Geometrische wie thematische Genauigkeit und Vollständigkeit der wesentlichen Datenattribute ist ein wichtiger Qualitätsfaktor für die Datenfusion. Denn die Qualität des optimierten Zieldatensatzes ergibt sich direkt aus der Qualität der heterogenen Input-Datensätze.

Um unvollständige Geodaten aus unterschiedlichen Quellen zu bearbeiten, wurde das Datenfusionssystem als modulares System aus Komponenten und Modulen konzipiert. Abbildung 2 zeigt die Komponenten für die automatisierte Geodatenfusion mit optionaler Benutzerinteraktion. Das System ist generisch und skalierbar. Es kann dadurch leicht für verschiedene Maßstabbereiche, Anwendungen und Nutzeranforderungen angepasst werden. Einerseits kann der Fusionsprozess in einem automatisierten Arbeitsablauf durch die Eingabe der Inputdatensätze durchgeführt werden. Andererseits kann die Prozessausführung vom Nutzer beeinflusst werden. Dazu kann er in interaktiven Dialogfenstern die Einstellungen des Fusionsystems bestimmen und modifizieren. Insbesondere kann er ausgewählte Filterungs- und Harmonisierungseinstellungen für spezifische Anwendungen bzw. Bedürfnisse ändern.

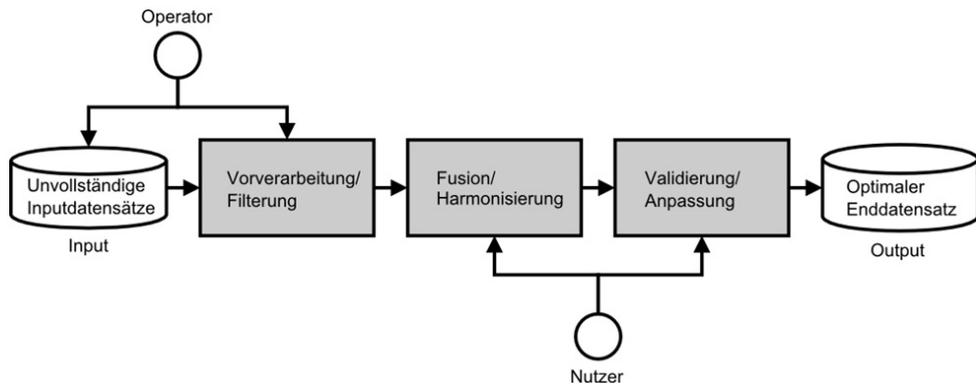


Abb. 2: Funktionsprinzip *DaFuS*: skalierbare Komponenten für die automatisierte Geodatenfusion mit optionaler Nutzer-Interaktion

Während des Datenfusionsprozesses filtert *DaFuS* die mangelhaften geometrischen und thematischen Attribute heterogener Input-Datensätze. Anschließend werden die gewünschten Attribute extrahiert, um einen einzigen optimalen Datensatz mit den selektierten „Best-Fit“-Datenattributen zu erzeugen. Das System kann daher als automatisierter Geo-

datenfilter bezeichnet werden. Sein Datenharmonisierungs- und Fusionsprozess ermöglicht die Auswahl, Beseitigung und/oder Ersetzen der unerwünschter Quellattribute mit benutzerdefinierten geometrischen und/oder semantischen Attributen. Kurz gesagt bietet *DaFuS* einen benutzerdefinierten Filter für die Generierung der optimalen Geodatenätze in einem automatisiertem Fusions- und Harmonisierungsprozess.

3 Systemarchitektur

Das *DaFuS*-System basiert auf dem Prinzip generischer, modularer und skalierbarer Software. Die Systemkomponenten umfassen (a) räumliche Datenbanken oder Datenquellen, (b) Datenfusions-Back-End (Operator/Administrator), (c) Datenfusions-Front-End (Nutzer) und (d) Komponenten und Module. Letztere drei Module bilden das Systemframework (Abb. 3).

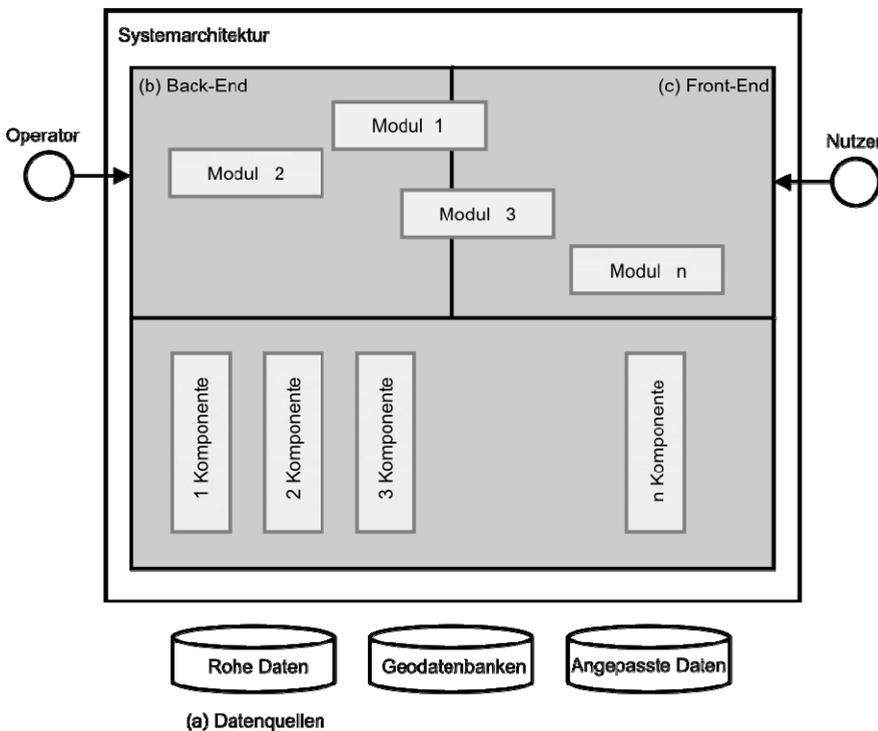


Abb. 3: *DaFuS* Systemarchitektur: generische, modulare Struktur (a) Inputdaten, (b) System Back-End (Operator), (c) System Front-End (Nutzer), Komponenten und Module

Zurzeit sind folgende Komponenten implementiert: Vorverarbeitung/Filterung, Fusion/Harmonisierung, Validierung/Anpassung und Output. Die Module beinhalten Geodaten-Input, Geodaten-Output, Geodaten-Filterung und Geodaten-Management. Zu den Kernmodulen

zählen: Daten-Viewer-Modul für die Datenausgabe einschließlich Standard-Datenschnittstellen, Geodaten-Filter mit kontextabhängigen Such- und Filterfunktionen und Geodaten-administration mit Systemfunktionalitäten für die Verwaltung heterogener räumlicher Daten und Geodatenbanken. Der modulare Systemaufbau ermöglicht eine einfache Anpassung an eine Reihe von einzelnen Anwendungs- und/oder benutzerspezifischen Anforderungen.

Kern dieses modularen, skalierbaren Systems ist ein Satz von Datenfusions-Algorithmen, die die oben genannten Vorgänge durchführen. Beispiele für Algorithmen, die für das *DaFuS* System entwickelt wurden, finden sich in STANKUTE & ASCHE (2009, 2010, 2011). Spezifische Teilmengen der Algorithmen sind die Basis der Komponenten und Module, welche das *DaFuS* System bilden.

Die automatisierte Durchführung der Datenfusion benötigt eine minimale Anzahl an Algorithmen, welche sequentiell abgearbeitet werden. In der Systemstruktur ist diese vorgeschriebene Sequenz in einer Fusionspipeline (Abb. 4) gespiegelt, zu der die korrespondierenden Komponenten und Module gruppiert sind. Diese Prozesskette kann skaliert, d. h. durch zusätzliche Komponenten und Module erweitert werden, um sie an die Spezifikationen einer Anwendung anzupassen.

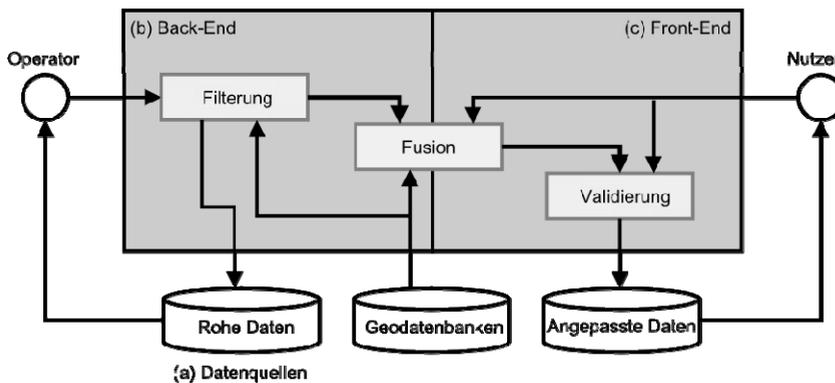


Abb. 4: *DaFuS* Fusions-Pipeline: die Prozesse und deren Verlaufsüberblick

4 *DaFuS* in der Anwendungspraxis: Fusion von Lineardaten

Wie das *DaFuS*-System im jetzigen Implementierungsstand arbeitet, sei an folgender typischer Datenfusions-Anwendung beispielhaft vorgestellt. In einem ersten Fall wird eine Anwendung betrachtet, die einen georeferenzierten, navigierbaren Straßennetz-Datensatz mit optimaler geometrischer und semantischer Information erfordert. *DaFuS* wird anstelle zeit- und kostenintensiver Neudatenerfassung eingesetzt, um den gewünschten Datensatz aus drei verfügbaren linearen Datensätzen, nämlich (a) *ATKIS*, (b) *TeleAtlas/TomTom*-, (c)

OpenStreetMap-(OSM)³-Straßen-Daten zu erzeugen. Bereits eine kursorische Bewertung der Quelldaten zeigt, dass sie sich in geometrischer Vollständigkeit, thematischem Umfang und Aktualität unterscheiden (Abb. 5). Deshalb werden die am besten geeigneten Datenattribute aus zwei unvollständigen Datensätzen extrahiert und auf den dritten Datensatz übertragen, der für die spezifische Anwendung am besten geeignet ist.

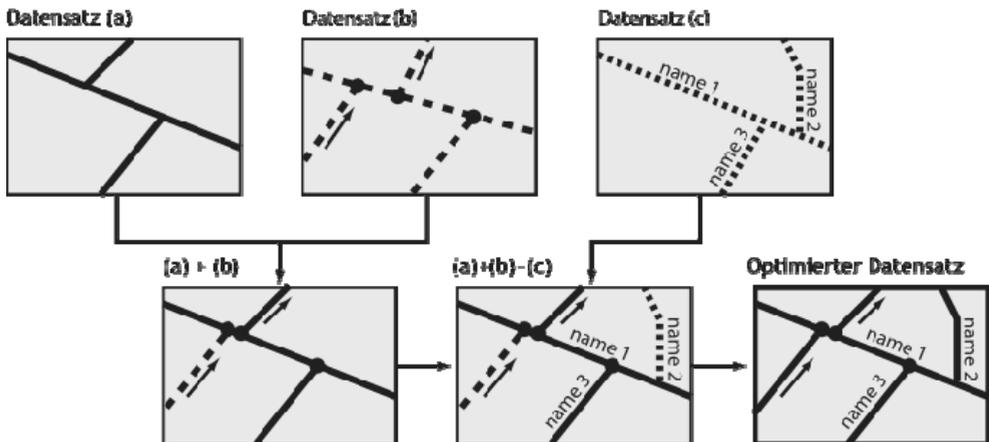


Abb. 5: *DaFuS*-Anwendungsbeispiel: Fusion linearer Datensätze. (a+b) Fusion geometrischer Attribute, ((a+b)+c) Fusion thematischer Attribute

Im Fusionsprozess wird der Ziel-Datensatz so modifiziert, dass er der einzige optimale lineare Datensatz ist, der für die Anwendung benötigt wird. Hierfür wird zunächst die geometrische und semantische Qualität der Quelldaten in der Filterungskomponente analysiert. Dabei wird festgestellt, dass der Quelldatensatz (a) die umfassendste geometrische Information bietet. Datensatz (b) enthält im Gegensatz dazu unvollständige geometrische Attribute, aber alle notwendigen Routing-Informationen. Datensatz (c) beinhaltet relevante thematische Attribute, die in Datensätzen (a, b) nicht vorhanden sind. Durch Eingabe von der Inputdatensätze in die *DaFuS*-Vorverarbeitungs-/Filterungs-Komponente wird ein einziger, geometrisch und thematisch optimaler, navigierbarer Ziel-Datensatz generiert. Diese Komponente validiert die Datenqualität im Vergleich zu den vorherigen Teilergebnissen.

Entsprechend der Gewichtung der anwendungsbezogenen Anforderungen – ein navigierbarer, geometrisch korrekter, thematisch vollständiger Straßennetz-Datensatz – wird der

³ *OpenStreetMap* oder *OSM* ist ein von Wikipedia inspiriertes Kooperationsprojekt, das 2004 begonnen wurde. Ziel des Projektes ist es, frei editierbare räumliche Daten zu erzeugen und jedermann frei nutzbar zu machen. Geodaten der rasch anwachsenden *OSM*-Datenbank werden in erster Linie durch eine Vielzahl von Akteuren mit tragbaren GPS-Geräten und Notepads gesammelt. Entsprechend den Aktivitäten der Datenerfasser variieren räumliche Abdeckung und Detaillierungsgrad sowohl global als auch regional. Deutschland ist mit *OSM*-Daten sehr gut abgedeckt, die derzeit von rund 20 Millionen Nutzern verwendet werden. Wesentliche Treiber für die Gründung und das Wachstum von *OSM* sind einerseits verschiedene Beschränkungen in der Verfügbarkeit von Geodaten, restriktive Nutzungsrechte oder hohe Anschaffungs- oder Nutzungskosten der nationalen oder kommerziellen raumbezogenen Daten, wie *ATKIS*, *TeleAtlas/TomTom* oder *Navteq*, andererseits die Verfügbarkeit preiswerter GPS-Geräte.

Datensatz (a) als Zieldatensatz gewählt und mit den geometrischen Attributen des Datensatzes (b) optimiert. Aus den Inputdatensätzen (a, b) erzeugt die Fusionskomponente den fusionierten Datensatz (ab). Aufgrund der bei der Datenerfassung verwendeten verschiedenen Datenerfassungsregeln variieren die geometrischen Attribute der Datensätze (a) und (b). Das Ergebnis sind topologische Inkonsistenzen zwischen den beiden Quellgeometrien. Um dieses Problem zu beheben, werden die entsprechenden geometrischen Eigenschaften des Datensatzes (b) auf die Geometrie des Datensatzes (a) übertragen. Damit enthält der resultierende Datensatz (ab) eine begrenzte Anzahl neuer (übertragener) Geometrieattribute, denen allerdings die erforderlichen Routinginformationen fehlen. Der nunmehr geometrisch vollständige Zieldatensatz wird in einem Folgeschritt um die fehlenden thematischen Informationen aus Quelldatensatz (c) ergänzt. Hierfür werden die Liniengeometrien des Datensatzes (ab) aufgespalten. Die benötigten thematischen Attribute werden sodann auf die relevanten Liniensegmente übertragen. Damit ist der Zieldatensatz (abc) erzeugt, der alle geforderten geometrischen und semantischen Attribute enthält.

Der erläuterte Anwendungsfall zeigt exemplarisch die Datenfusion linearer Geodaten in einem automatisierten Verarbeitungsprozess. Das Ergebnis, ein einziger optimaler Datensatz, enthält die entsprechend den Vorgaben am besten geeigneten Attribute jedes Quelldatensatzes. Die lineare Datenfusion kann als Standardanwendung des *DaFuS*-Systems angesehen werden. Dabei kann die Datenfusion nicht nur auf einfache Liniennetze, sondern auch auf komplexere Linearstrukturen von Geobjekten, z. B. Kreisverkehre, angewendet werden, wie in STANKUTE & ASCHE (2011) gezeigt wurde.

Eine weitere häufige Anwendung von *DaFuS* betrifft die Verbesserung unvollständiger Polygondaten. Dabei werden die heterogenen polygonalen Ausgangsdaten in einem mehrschrittigen Prozess, der dem der linearen Datenfusion ähnelt, zu einem Zieldatensatz fusioniert: Filtern und Vorverarbeitung der Quelldaten, Referenzierung der Objekte und Fusion der Polygone. Im Unterschied zu Linien handelt es sich bei Polygonen jedoch um komplexere Strukturen als bei Linien. Dementsprechend unterscheidet sich die Komplexität der Algorithmen, die für die Erzeugung eines einzigen optimalen Datensatzes benötigt werden, signifikant von denen für lineare Daten.

5 Fazit

Der vorliegende Beitrag beschreibt ein neuartiges Data-Mining-Konzept. Dessen Ziel ist es, aus einer Vielfalt heterogener Datensätze einen optimalen Geodatenatz mit optimalen geometrischen und/oder semantischen Eigenschaften in einem automatisierten und benutzerdefinierten Fusionsprozess zu generieren. Das innovative Fusionskonzept wurde in dem generischen *DaFuS* Softwaresystem umgesetzt. Aufgrund des modularen und skalierbaren Aufbaus kann das System an die Anforderungen der zu verarbeitenden Daten und/oder die jeweiligen Anwendungsspezifikationen angepasst werden. Durch Verarbeitung verbreiteter Vektor-Geodatenätze (z. B. *Tele Atlas/TomTom*- und *Navigo*-Daten oder *ATKIS*-Daten), die als Inputdatensätze dienten, wurde in umfangreichen Tests nachgewiesen, dass ein einzelner optimaler Geodatenatz aus bestehenden suboptimalen Datensätzen erzeugt werden kann. Abbildung 6 zeigt einen fusionierten Straßen-Datensatz mit optimaler geometrischer Qualität, die aus den Inputdatensätzen, dargestellten in Abbildung 1, gewonnen wurde. Vergleicht man den fusionierten Datensatz mit dem im Hintergrund liegenden georefe-

renzierten *HRSC*-Bild, ist leicht erkennbar, dass die optimierte geometrische Qualität deutlich über der Qualität der einzelnen Quelldatensätze liegt. Auf die gleiche Weise wird auch die semantische Qualität der Inputdatensätze in einem Zieldatensatz verbessert.



Abb. 6: Ein fusionierter optimaler Straßenvektordatensatz mit hochauflösendem *HRSC-A*-Bild im Hintergrund (Potsdam)

Vorerst bietet *DaFuS* Fusionsmodule weitverbreitete GI-Vektordaten an.⁴ Deren Qualitätsmerkmale werden mithilfe der jeweiligen Dokumentation oder Metadaten bewertet. Derzeit werden zusätzliche Fusions- und Harmonisierungsmodule für andere nationale Geodatenbanken als die deutsche *ATKIS*-Geodatenbank entwickelt. Ein der Ziele in dieser Entwicklungsphase ist die Konzeption und Implementierung einer Komponente, die Geodaten aus benachbarten Regionen oder Ländern harmonisiert, und damit die *INSPIRE*-Direktive der Europäischen Union befördert. Mit dieser Entwicklung wird *DaFuS* seinen Funktionsumfang jenseits der klassischen Datenfusion in den entstehenden Bereichen der Datengenerierung aus teilweise inkompatiblen Geodatenmodellen erweitern. Es wird derzeit auch ein weiteres Entwicklungsziel verfolgt: Die Erweiterung der Fusionskomponente durch ein Fusionsmodul für Geodaten mit mehr oder weniger undeutlichen geometrischen als auch semantischen Eigenschaften, z. B. nicht von Fachleuten generierte *OpenStreetMap*-Datenbank.

Basierend auf einer Qualitätsanalyse der fusionierten optimalen Datensätze – inklusive des Vergleichs mit neu erworbenen räumlichen Daten gleicher Art und räumlicher Abdeckung – kann abschließend festgestellt werden, dass die wiederholte Datenerfassung von identischen oder ähnlichen Geodaten für jede neue Anwendung in den meisten Fällen nicht er-

⁴ *DaFuS*-Dienstleistungen werden von map/x/tek (www.mapxtek.de) angeboten.

forderlich ist. Insofern trägt *DaFuS* dazu bei, die Heterogenität und Redundanz von Geodaten in massiven GI-Datenbanken zu reduzieren. Gleichzeitig unterstützt das System durch seinen modularen Aufbau und Skalierbarkeit die effiziente und sinnvolle Nutzung der raumbezogener Massendaten. Damit fördert *DaFuS* nicht zuletzt ein kostengünstiges Geodatenmanagement, indem es die mehrfache Wiederverwendung von bestehenden GI-Datensätzen ermöglicht, die auf einzelne Benutzer und/oder Anwendungsanforderungen angepasst werden können.

Literatur

- DELVAR, M. & DEVILLERS, R. (2010), Spatial data quality: From process to decisions. *Transactions in GIS*, 14 (4), 379-386.
- DEVILLERS, R., GERVAIS, M., BEDARD, Y. & JEANSOULIN, R. (2002), Spatial Data Quality: From Metadata to Quality Indicators and Contextual End-User Manual. In: *Proceedings of the OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management*, 45-55.
- DEVILLERS, R. & JEANSOULIN, R. (2006), Spatial Data Quality: Concepts. In: DEVILLERS, R. & JEANSOULIN, R. (Eds.), *Fundamentals of Spatial Data Quality*. Geographical Information Series. ISTE, London, 31-42.
- HUNTER, G. J., BREGT, A. K., HEUVELINK, G. B. M., DE BRUIN, S. & VIRRANTAU, K. (2009), Spatial Data Quality: Problems and Prospects. In: NAVRATIL, G. (Eds.) *Research Trends in Geographic Informations Science, Lecture Notes in Geoinformation and Cartography*. Springer, Heidelberg, 101-121.
- JEANSOULIN, R., PAPINI, O., PRADE, H. & SCHOCKAERT, S. (Eds.) (2002), *Methods for Handling Imperfect Spatial Information*. Studies in Fuzziness and Soft Computing, 256. Springer-Verlag, Berlin/Heidelberg.
- ONCHAGA, R., MORALES, J. WIDYA, I. & LAMBERT, J. M. (2008), An Ontology Framework for Quality of Geographical Information Services. In: *Proceedings of the 16th International Symposium on Advances in Geographic Information Systems, ACM GIS2008*.
- STANKUTE, S. & ASCHE, H. (2009), An Integrative Approach to Geospatial Data Fusion. In: GERVASI, O., TANIAR, D. & MURGANTE, B. (Eds.), *Computational Science and Its Applications – ICCSA 2009*. Springer-Verlag, Heidelberg, LNCS, 5592, 490-504.
- STANKUTE, S. & ASCHE, H. (2010), Geometrical DCC-Algorithm for Merging Polygonal Geospatial Data. In: GERVASI, O., TANIAR, D. & MURGANTE, B. (Eds.), *Computational Science and Its Applications – ICCSA 2010*. Springer-Verlag, Heidelberg, LNCS, 6016, 531-543.
- STANKUTE, S. & ASCHE, H. (2011), Improvement of Spatial Data Quality Using the Data Conflation. In: GERVASI, O., TANIAR, D. & MURGANTE, B. (Eds.) *Computational Science and Its Applications – ICCSA 2011*. Springer-Verlag, Heidelberg, LNCS, 6782, 492-500.
- STANKUTE, S. & ASCHE, H. (2012), A Data Fusion System for Spatial Data Mining, Analysis and Improvement. In: GERVASI, O., TANIAR, D. & MURGANTE, B. (Eds.), *Computational Science and Its Applications – ICCSA 2012*. Springer-Verlag, Heidelberg, LNCS, 7334, 439-449.