

NGHIÊN CỨU HOÀN THIỆN CẤU TRÚC CƠ SỞ DỮ LIỆU NỀN ĐỊA LÝ GẮN VỚI CÁC GIẢI PHÁP TỔNG QUÁT HÓA TỰ ĐỘNG

TS. ĐỒNG THỊ BÍCH PHƯƠNG

Viện Khoa học Đo đạc và Bản đồ

I. Đặt vấn đề

Sau khi trải qua giai đoạn số hóa bản đồ địa hình các tỷ lệ, các nước trên thế giới đều tiến tới xây dựng các cơ sở dữ liệu (CSDL) từ nguồn dữ liệu chính là bản đồ địa hình. Các CSDL thường có tỷ lệ, nói chính xác hơn là mức độ chi tiết tương đương bản đồ địa hình số cùng tỷ lệ. Các Cơ quan Đo đạc và Bản đồ quốc gia duy trì các CSDL đa tỷ lệ này song song với hệ thống bản đồ địa hình. Do được thành lập từ nguồn bản đồ địa hình dạng số, các CSDL không tránh khỏi kế thừa kết quả tổng quát hóa trong biên tập bản đồ tỷ lệ nhỏ từ tỷ lệ lớn hơn. Một nhược điểm khác của hệ thống bản đồ - CSDL này là khó khăn trong cập nhật, hiện chỉnh dữ liệu. Xu hướng hiện nay ở các quốc gia có công nghệ thông tin địa lý phát triển là xây dựng CSDL nền ở tỷ lệ lớn làm cơ sở để thành lập các CSDL và bản đồ địa hình tỷ lệ nhỏ hơn bằng các giải pháp tổng quát hóa tự động. Phương án này có ưu điểm là giảm kinh phí trong xây dựng và duy trì CSDL, đảm bảo tính đồng nhất của dữ liệu, tuy nhiên tốc độ và chất lượng thành lập các sản phẩm dẫn xuất phụ thuộc vào các tiến bộ của công nghệ tổng quát hóa tự động.

Ở Việt Nam, để phục vụ xây dựng CSDL nền địa lý tỷ lệ 1:10.000 đáp ứng các nhu cầu về CSDL địa lý của các ngành các cấp, Bộ Tài nguyên và Môi trường đã ban hành Quy định kỹ thuật tạm thời về mô hình cấu trúc và nội dung dữ liệu nền địa lý tỷ lệ 1/2.000, 1/5.000 và 1/10.000. Quy định kỹ thuật tạm thời đã hỗ trợ rất hiệu quả công

tác xây dựng CSDL nền địa lý tỷ lệ 1:10.000, đáp ứng về cơ bản nhu cầu chiết xuất và cập nhật dữ liệu ở các độ phân giải thấp hơn.

Tuy nhiên, để hướng tới mô hình trong đó CSDL nền địa lý là dữ liệu cơ bản để thành lập tất cả các bản đồ tỷ lệ nhỏ hơn và các datasets có mức độ chi tiết thấp hơn, đáp ứng các mục đích sử dụng khác nhau như thành lập bản đồ địa hình ở các tỷ lệ khác với dãy tỷ lệ cơ bản, bản đồ địa lý với các mức độ khái quát khác nhau, bản đồ nền với mức độ chi tiết khác nhau phục vụ công tác điều tra cơ bản và thành lập bản đồ chuyên đề, các datasets, các hệ thống thông tin địa lý phục vụ giải quyết các bài toán kinh tế xã hội, thì nội dung và cấu trúc CSDL nền địa lý cần phải tiếp tục hoàn thiện. Định hướng sử dụng các giải pháp tổng quát hóa, chiết xuất dữ liệu tự động cần phải quan tâm nghiên cứu ngay từ giai đoạn đầu của công tác xây dựng CSDL nền địa lý, tức là trong giai đoạn tổng quát hóa mô hình. CSDL được thiết kế hợp lý sẽ nâng cao mức độ tự động hóa trong thành lập bản đồ và dữ liệu dẫn xuất, tiết kiệm thời gian và kinh phí trong khai thác và bảo trì.

II. Phương pháp

Việc thể hiện các đối tượng của bề mặt thực tế trên bản đồ địa hình và địa lý chung luôn tuân theo các nguyên tắc của bản đồ học truyền thống. Các nguyên tắc này được xây dựng với mục đích cân đối giữa khả năng có hạn của tải trọng bản đồ và yêu cầu vô hạn đối với thông tin về không gian và

thuộc tính của các đối tượng thực tế. Khác với bản đồ, CSDL địa lý không bị hạn chế bởi những yêu cầu về khả năng đọc được (readability). Tuy nhiên, là một mô hình thể giới thực, CSDL địa lý cũng có mục đích phản ánh chân thực đối tượng mô hình hóa là bề mặt Trái đất và các đối tượng liên quan. Do vậy, mức độ chi tiết mô tả mỗi đối tượng trong CSDL cần phải được xem xét trong quan hệ với các đối tượng khác và đặc biệt cần quan tâm đến mục đích sử dụng của CSDL.

Các yêu cầu về cấu trúc và nội dung CSDL nền địa lý tỷ lệ 1:10.000 gắn với các giải pháp tự động bao gồm:

1. Thiết kế cấu trúc CSDL trước hết phải tính đến các yêu cầu đặc thù cũng như qui trình xây dựng, quản lý, khai thác CSDL hiện đang sử dụng, tập quán, nhu cầu về sử dụng CSDL dẫn xuất và bản đồ địa lý, bản đồ chuyên đề ở Việt Nam. Cấu trúc CSDL phải thừa kế tối đa cấu trúc CSDL có sẵn, giảm tối thiểu khối lượng công tác chuẩn hóa, điều tra, thu thập dữ liệu bổ sung.

2. CSDL nền địa lý phải có mức độ chi tiết về hình học và ngữ nghĩa bằng hoặc cao hơn bản đồ địa hình cùng tỷ lệ. Hệ thống phân loại và danh mục đối tượng cần phải được xây dựng theo các tiêu chí tương tự hệ thống phân loại của bản đồ địa hình.

3. Các qui định về cấu trúc, nội dung và tiêu chí thu nhận của CSDL nền địa lý cần rõ ràng, đầy đủ, đảm bảo tính đơn nghĩa và tránh tối đa các trường hợp ngoại lệ.

4. CSDL nền địa lý cần phải cung cấp thông tin cho phép xác định mức độ quan trọng của đối tượng về các mặt hình học, quan hệ và ngữ nghĩa được mô hình hóa như là thuộc tính của từng đối tượng độc lập. Các thông tin này không chỉ phục vụ tổng quát hóa dữ liệu tự động mà còn hỗ trợ biên tập viên trong tổng quát hóa dữ liệu và bản đồ bán tự động, cũng như tham gia vào các bài toán phân tích địa lý.

III. Kết quả thử nghiệm

Trong khuôn khổ đề tài "Nghiên cứu hoàn thiện cấu trúc dữ liệu cho cơ sở dữ liệu nền địa lý gắn với các giải pháp tổng quát hóa dữ liệu tự động", dựa trên các yêu cầu về hoàn thiện cấu trúc CSDL nền địa lý được trình bày ở trên, nhóm tác giả đã đề xuất và thử nghiệm cấu trúc CSDL nền địa lý tỷ lệ 1:10.000 gắn với các giải pháp tổng quát hóa tự động. Thử nghiệm được thực hiện trong phạm vi mảnh bản đồ F-48-56-A (Thái Nguyên) và D-49-26-B (Quảng Ngãi).

Danh sách các kiểu đối tượng và thuộc tính theo mô hình cấu trúc đề xuất được trình bày tại bảng 1. Những khác biệt cơ bản của mô hình cấu trúc được đề xuất so với mô hình cấu trúc của Qui định tạm thời bao gồm:

+ Một số đối tượng thuộc CSDL nền địa lý được đề xuất bổ sung thuộc tính `mucDoUuTien`. Đó là những đối tượng có số lượng lớn trên bản đồ địa hình, không đi kèm với các thuộc tính cho phép phân loại thứ bậc, có yêu cầu về phản ánh sự phân bố (đồ hình) của mạng lưới và ý nghĩa đặc biệt của đối tượng khi thu nhỏ tỷ lệ. Mức độ ưu tiên của một đối tượng trong chọn bỏ khi có nhu cầu về thu nhỏ tỷ lệ bản đồ, giảm độ phân giải không gian hoặc giảm dung lượng dữ liệu được xác định theo "mức độ quan trọng" của đối tượng ở các khía cạnh khác nhau, ví dụ mục đích sử dụng, giá trị kinh tế, thứ bậc trong phân tích đồ hình. Các thuộc tính hình học, quan hệ và ngữ nghĩa khác cũng có thể tham gia vào xác định mức độ quan trọng như chiều dài, diện tích, khoảng cách đến các đối tượng cùng loại và khác loại, phân cấp quản lý.... Cuối cùng, mức độ ưu tiên được kỹ thuật viên xem xét điều chỉnh tùy theo đặc điểm địa lý khu vực, yêu cầu về tổng quát hóa cũng như mục đích sử dụng của CSDL. Mức độ ưu tiên được định lượng hóa và nhận các giá trị từ 0 đến 8. Các đối tượng có Mức độ ưu tiên ≥ 8 là

những đối tượng quan trọng nhất và có mặt trong CSDL và được cân nhắc thể hiện trên bản đồ tất cả các tỷ lệ từ 1:2.000 đến 1:1.000.000. Các đối tượng có $0 \leq$ Mức độ ưu tiên < 1 là những đối tượng ít quan trọng hơn cả và chỉ có mặt trong CSDL tỷ lệ 1:2.000.

+ Bổ sung thuộc tính chieuRong cho các đối tượng TimDongChay, thuộc tính yNghiaDacBiet cho các loại đối tượng sau: Dao, BaiBoi, MatNuoc, TimDongChay, DoanTimDuongBo, DiemDanCu.

+ Gộp một số đối tượng có khái niệm và mô tả gần gũi như DiemDoCao và DiemDoSau, MocDiaGioi và MocBienGioi, DiemGocQuocGia, DiemCoSoQuocGia và DiemCoSoChuyenDung, Dao và BaiBoi.

+ Gộp các đối tượng KenhMuong,

SongSuoi và MangDanNuoc và tìm đường của các đối tượng KenhMuong, SongSuoi và MangDanNuoc được mô tả bằng kiểu G-Surface thành đối tượng TimDongChay.

+ Gộp các đối tượng KenhMuong, SongSuoi và MangDanNuoc được mô tả bằng kiểu G-Surface và đối tượng MatNuocTinh thành đối tượng MatNuoc.

+ Gộp các đối tượng DuongBoNuoc, DuongMepNuoc, BoKe và RanhGioi NuocMatQuiUoc vào thành đối tượng RanhGioiThuyHe.

+ Thống nhất xử lý với các đối tượng KhuChucNang, BenBai, TramDien: thuộc tính và vị trí tương đối mô tả bằng kiểu G-Point, ranh giới mô tả bằng đối tượng RanhGioi (thực tế và qui ước).

Bảng 1: Danh sách các kiểu đối tượng và thuộc tính

STT	Loại đối tượng CSDL thử nghiệm	Kiểu đối tượng	Thuộc tính	Loại đối tượng CSDL 1:10.000 (hiện hành)
I	Gói Cơ sở			
1	DiemCoSoDoDac	P	maDoiTuong, soHieuDiem, toaDoX, toaDoY, docaoH, loaiCapHang, loaiMoc, loaiDiemCSDD	DiemGocQuocGia, DiemCoSoQuocGia, DiemCoSoChuyenDung
II	Gói Biên giới Địa giới			
2	MocBienGioiDiaGioi	P	maDoiTuong, soHieuMoc, toaDoX, toaDoY, loaiMocBGDG	MocDiaGioi, MocBienGioi
3	DuongBienGioi	C	maDoiTuong, loaiHienTrangPhapLy, quocGiaLienKe, chieuDai,	DuongBienGioi
4	DuongDiaGioi	C	maDoiTuong, loaiHienTrangPhapLy, chieuDai,	DuongDiaGioi
5	DiaPhan	S	maDoiTuong, maDonViHanhChinh, dienTich, ten,	DiaPhan
6	DiemCoSoLanhHai	P	maDoiTuong, soHieuDiem, KinhDo, ViDo, doCaoH,	DiemCoSoLanhHai
7	DuongCoSoLanhHai	C	maDoiTuong, chieuDai,	DuongCoSoLanhHai
8	VungBien	S	maDoiTuong, dienTich,	VungBien

III	Gói Địa hình			
9	DiaHinhDacBiet	P,C,S	maDoiTuong, ten, tyCaoTySau, mucDoUuTien	DiaDanhSonVan DiaHinhDacBiet
10	DuongBinhDo	C	maDoiTuong, loaiDuongBinhDo, loaiKhoangCaoDeu, doCaoH,	DuongBinhDo, DuongBinhDoSau
11	DiemDoCaoDoSau	P	maDoiTuong, loaiDiemDoCao, doCaoH, mucDoUuTien	DiemDoCao, DiemDoSau
12	TaLuy	C	maDoiTuong, loaiTaLuy, loaiHinhThaiTaLuy, loaiThanhPhanTaLuy, tyCaoTySau	TaLuyCongTrinhThuyLoi, TaLuyGiaoThong
IV	Gói Thủy hệ			
13	Bien	P,S	maDoiTuong,ten,	Bien
14	DaoBai	P,S	maDoiTuong,ten,loaiDaoBai,mucDoUuTien, yNghiaDacBiet	Dao, BaiBoi
15	BienDoiDongChay	P,C,S	maDoiTuong, loaiTrangThaiXuatLo, tyCaoTySau, chieuDai,ten,mucDoUuTien	BienDoiDongChay
16	RanhGioiThuyHe	C	maDoiTuong, loaiRanhGioiNuocMat,loaiTrangThaiDuong BoNuoc,loaiRanhGioiThuyHe	DuongBoNuoc, DuongMepNuoc, RanhGioiNuocMatQuyUoc
17	TimDongChay	C	maDoiTuong,ten, hienTrangSuDung, loaiTrangThaiNuocMat,loaiDongChay, mucDoUuTien, chieuRong, yNghiaDacBiet	SongSuoi, KenhMuong, MangDanNuoc
18	MatNuoc	S	maDoiTuong,ten,kieuMatNuoc, mucDoUuTien,yNghiaDacBiet	MatNuocTinh, SongSuoi, KenhMuong, ĐamLay
19	NguồnNuoc	P	maDoiTuong, ten,loaiNgonNuoc,mucDoUuTien, yNghiaDacBiet	NgonNuoc

V	Gói Giao thông			
20	CauGiaoThong	P	maDoiTuong, loaiCauGiaoThong, loaiChucNangCau, taiTrong, chieuDai, chieuRong,ten,mucDoUuTien	CauGiaoThong
21	HamGiaoThong	P	maDoiTuong, loaiHamGiaoThong,chieuCao, chieuRong,chieuDai,ten,mucDoUuTien	HamGiaoThong
22	BenBai	P	maDoiTuong,ten,mucDoUuTien	BenBai
23	DoiTuongGTKhac	P	maDoiTuong,ten,loaiDoiTuongGT, mucDoUuTien	BaoHieuGiaoThong, CongGiaoThong, DoanVuotSongSuoi,CauDiBo, HamDiBo, Deo
24	DoanTimDuongBo	C	maDoiTuong, loaiDuongBo, loaiChatLieuTraiMat, loaiHienTrangSuDung, loaiKetCau, doRong, ten, tenTuyen1, tenTuyen2, tenTuyen3,loaiDoanTimDuong, mucDoUuTien, yNghiaDacBiet	DoanTimDuongBo, CauGiaoThong, HamGiaoThong, DoanVuotSongSuoi
25	MatDuongBo	S	maDoiTuong,	MatDuongBo, MatBoKenhMuong, CauGiaoThong, HamGiaoThong,DuongNoiBo, De, Dap
26	NutMangGiaoThong	P	maDoiTuong,	NutMangDuongBo, NutDuongSat
27	RanhGioiDuongBo	C	maDoiTuong,loaiRanhGioiDuongBo	RanhGioiDuongBo
28	DuongCapTreo	C	maDoiTuong, ten, chieuDai,	DuongCapTreo
29	DuongNoiBo	C	maDoiTuong, loaiChatLieuTraiMat,	DuongNoiBo
30	DoanDuongSat	C	maDoiTuong, loaiDuongSat, loaiChucNangDuongSat, loaiHienTrangSuDung, loaiKhoDuongSat, loaiKetCauDuongSat,mucDoUuTien, loaiDoanDuongSat	DoanDuongSat, CauGiaoThong

VI	Gói dân cư			
31	DiemDanCu	P,S	maDoiTuong, loaiDiemDanCu,ten,mucDoUuTien, yNghiaDacBiet,soDan,soNha, capHanhChinh	DiemDanCu, Nha
32	DuongDayTaiDien	C	maDoiTuong, dienAp,	DuongDayTaiDien
33	TramDien	P	maDoiTuong, loaiViTriTramDien,	TramDien
34	DoiTuongDinhVi	P	maDoiTuong, ten,loaiDoiTuongDinhVi, mucDoUuTien	ThapNuocBeNuoc, TramQuanTrac, TramThuPhatSong, CongTrinhKienTrucDacBiet
35	Nha	S	maDoiTuong, Sotang,chieuCao, mucDoUuTien	Nha
36	HaTangKinhTeXaHoi	P,S	maDoiTuong, diaChi,ten,capHaTangKinhTeXaHoi, mucDoUuTien	HaTangKinhTeXaHoi, BenBai, TramDien
37	RanhGioi	C	maDoiTuong,loaiRanhGioi, kieuRanhGioi	RanhGioi, RanhGioiPhuBeMat
38	DeDap	C	maDoiTuong, dungLamGiaoThong,ten, loaiDeDap	De, Dap
39	CongTrinhThuyLoi	P	maDoiTuong, ten,loaiCongTrinhThuyLoi	CongThuyLoi, TramBom, CongTrinhTrenDe
VII	Gói Thực vật			
40	ThucVat	S	maDoiTuong, ten, loaiPhuBeMat,	PhuBeMat

+ Bổ sung kiểu G-Surface cho đối tượng DiemDanCu để thu nhận các khu dân cư phân bố không tập trung (mã IA02)

+ Gộp một số đối tượng gói Giao thông có cùng kiểu đối tượng vào RanhGioiDuongBo, MatDuongBo và DoanTimDuongBo

Với phương án mô hình cấu trúc được đề xuất, 273 loại đối tượng thế giới thực được định nghĩa trong chuẩn thông tin địa lý cơ sở có thể được mô hình hóa bằng 48 kiểu đối tượng thay vì 96 kiểu đối tượng được đề xuất trong Dự thảo qui chuẩn quốc gia về CSDL nền địa lý tỷ lệ 1:10.000.

Thử nghiệm chiết xuất CSDL nền địa lý và bản đồ địa hình tỷ lệ 1:50.000 từ CSDL tỷ lệ 1:10.000 với cấu trúc được đề xuất và các thuộc tính bổ sung cho thấy mức độ tự động hóa đạt trên 90% cho giai đoạn tổng quát hóa dữ liệu và trên 80% cho giai đoạn tổng quát hóa bản đồ đối với đa số các kiểu đối tượng. Phần mềm ứng dụng là FME 2011 (Safe Software).

IV. Kết luận

CSDL nền địa lý với cấu trúc được đề xuất có ưu điểm giảm tính chủ quan trong tổng quát hóa chọn bỏ ở mức đối tượng,

đảm bảo tính thống nhất trong thu nhận và thể hiện đối tượng theo dãy tỷ lệ, tăng tốc độ xử lý trong tổng quát hóa tự động. Ngoài ra, kết quả làm giàu đối tượng không chỉ hỗ trợ quá trình tổng quát hóa dữ liệu, tổng quát hóa bản đồ tự động mà còn được tham khảo trong quá trình sửa lỗi, giải quyết xung đột cũng như tham gia vào các bài toán phân tích địa lý. Việc bổ sung thuộc tính *mucDoUuTien* cũng giải quyết triệt để vấn đề dẫn xuất dữ liệu với các mục đích sử dụng khác nhau mà không phải gắn vào dãy tỷ lệ và các qui định của bản đồ địa hình như phương án bổ sung thuộc tính LOD (Mức độ chi tiết) do Cục Bản đồ Bộ Tổng tham mưu sử dụng trong xây dựng bản đồ địa hình quân sự tỷ lệ 1:100.000 từ CSDL nền địa hình - địa lý quân sự tỷ lệ 1:50.000. Nhược điểm của phương án này là đầu tư ban đầu lớn cho xây dựng cơ sở dữ liệu và các qui định, qui chuẩn kỹ thuật, tăng dung lượng CSDL nền, cũng như yêu cầu có những nghiên cứu sâu về xác định mức độ quan trọng của từng loại đối tượng làm cơ

sở để xây dựng qui trình làm giàu dữ liệu và các qui định, qui chuẩn kỹ thuật.

CSDL nền địa lý được thiết kế với cấu trúc và nội dung hợp lý có thể đảm bảo tự động hóa giai đoạn tổng quát hóa dữ liệu và hỗ trợ hiệu quả giai đoạn tổng quát hóa bản đồ. Vấn đề chiết xuất dữ liệu tự động xây dựng các bản đồ, dataset dẫn xuất có thể giải quyết tương đối triệt để bằng các phần mềm thương mại thông dụng như ArcGIS, FME.○

TÀI LIỆU THAM KHẢO

[1]. Bộ Tài nguyên và Môi trường. Quy định kỹ thuật (tạm thời) về mô hình cấu trúc và nội dung dữ liệu nền địa lý tỷ lệ 1:2.000, 1:5.000 và 1:10.000 (2009).

[2]. Bộ Tài nguyên và Môi trường. Qui chuẩn kỹ thuật quốc gia về chuẩn thông tin địa lý cơ sở (2012).

[3]. Bộ Tài nguyên và Môi trường. Dự thảo Quy chuẩn kỹ thuật quốc gia cơ sở dữ liệu nền địa lý tỷ lệ 1:10.000 (2011).○

Summary

COMPLETE RESEARCH DATABASE STRUCTURE ASSOCIATED WITH THE GEOGRAPHICAL BACKGROUND AUTOMATIC GENERALIZATION SOLUTION

Dr. Dong Thi Bich Phuong

Vietnam Institute of Geodesy and Cartography

Recently, automated generalization has become the prerequisite for the maintenance of multi-representation spatial database. Well-organized, purpose-oriented and enriched with semantic and relation attributes core database may act as solution to the problem of on-demand map and dataset automated derivation. The paper proposed and tested new approach to restructuring and enriching core database toward automated map extraction.○