

Bài 1

SỰ HÌNH THÀNH VÀ ỨNG DỤNG CỦA NGÀNH KHOA HỌC HỆ THỐNG THÔNG TIN ĐỊA LÝ

Trong bài này giới thiệu về quá trình hình thành ngành khoa học này và những ứng dụng của nó.

1.1. Giới thiệu bản đồ

Thông tin địa lý thể hiện trước hết ở dạng bản đồ. Các bản đồ đầu tiên được phác thảo để mô tả vị trí, bản đồ địa hình thể hiện các nét chính về cảnh quan như sông ngòi, đường, làng bản, rừng cây... Chúng thường bao gồm địa hình với các ký hiệu điểm riêng biệt và đường contour. Các bản đồ này thích hợp cho mục đích chung hoặc cho quân đội. Các kiểu bản đồ cung cấp thông tin về từng vấn đề như địa chất, phân vùng lãnh thổ, tỉ lệ thất nghiệp được gọi là các bản đồ chuyên đề.

1.2. Những bất tiện của bản đồ in trên giấy

Trong quá trình sử dụng bản đồ giấy, những khuyết điểm được ghi nhận như sau:

- Bản đồ xây dựng với giá đắt và chi phí nhiều thời gian.

- Lượng thông tin hạn chế, nếu bản đồ chứa nhiều thông tin thì rất khó đọc.

- Không thể cập nhật thông tin theo thời gian.

- Bản đồ chỉ cho các tài liệu định tính, không thể phân tích định lượng các dữ liệu trên bản đồ.

- Không thể phân tích nhiều tập hợp dữ liệu không gian từ các bản đồ khác nhau (như đất, sườn dốc và lớp phủ thực vật để đánh giá mức độ xâm thực).

Hiện nay, nhu cầu các tài liệu sử dụng nhanh, có thể cung cấp thông tin cập nhật và chính xác cho các mục đích đặc biệt. Do vậy, bản đồ truyền thống không còn thuận tiện nữa.

1.3. Ra đời HTTTĐL

Vào những năm 1960, một số người đã có ý tưởng mô hình hóa không gian lưu trữ vào máy tính, đó là một bản đồ đơn giản có thể mã hóa, lưu trữ trong máy tính, sửa chữa khi cần thiết, có thể hiển thị trên màn hình và in ra giấy. Thời gian đầu, bản đồ điện toán (computer cartography) thể hiện những điểm, các đường thẳng (vector) và chữ (text). Các đồ thị phức tạp có thể được xây dựng từ những yếu tố này. Ví dụ; những đường không theo qui luật như sông, bờ biển sẽ được tạo ra liên tiếp từ các yếu tố vector nhỏ.

Các nhà nghiên cứu đã nhận thấy rằng nhiều vấn đề địa lý đòi hỏi thu thập và phân tích một khối lượng lớn thông tin không phải bản đồ. Ví dụ:

Điều tra dân số yêu cầu dữ liệu về người, hộ gia đình.
Ứng dụng địa chính yêu cầu thông tin về quyền sở hữu đất...

Vào lúc này thuật ngữ *Bản đồ máy tính* được thay thế bởi thuật ngữ HTTTĐL.

HTTTĐL đầu tiên xuất hiện vào năm 1964 thuộc dự án “Rehabilitation and Development Agency Program” của chính phủ Canada. Cơ quan “Hệ thống thông tin địa lý Canada-CGIS” đã thiết kế để phân tích, kiểm kê đất nhằm trợ giúp cho chính phủ trong việc sử dụng đất nông nghiệp. Dự án CGIS hoàn thiện vào năm 1971 và phần mềm vẫn sử dụng tới ngày nay. Dự án CGIS gồm nhiều ý tưởng sáng tạo mà đã được phát triển trong những phần mềm sau này.

Giữa những năm 60 và 70, HTTTĐL phát triển chủ yếu trong Chính phủ và các phòng thí nghiệm.

Năm 1964, Ông Howard Fisher thành lập “[Phòng thí nghiệm đồ họa máy tính Harvard](#)” phòng dẫn đầu về các công nghệ mới. Phòng thí nghiệm Harvard đã tạo ra một loạt các ứng dụng chính HTTTĐL bao gồm: SYMAP (Synagraphic Mapping System), CALFORM, SYMVU, GRID, POLYVRT, và ODYSSEY. ODYSSEY là mô hình đầu tiên vector HTTTĐL và nó trở thành chuẩn cho các phần mềm thương phẩm.

Hệ thống bản đồ tự động đã được phát triển bởi ([CIA](#)) trong cuối những năm 1960s. Dự án này tạo ra “[Ngân hàng dữ liệu Thế giới của CIA](#)”, thu thập thông tin đường bờ biển, con sông, ranh giới hành chính và phần mềm trọn gói [CAM](#) tạo ra những bản đồ những tỉ lệ khác nhau từ dữ liệu này. Đây là một hệ thống CSDL bản đồ đầu tiên trên Thế giới.

Hai công trình có giá trị khác là Hệ thống tin sử dụng đất New york (1967) và hệ thống tin quản lý đất Minnesota (1969).

Năm 1969, Jack Dangermond, một người trong nhóm nghiên cứu tại phòng thí nghiệm Harvard trong bộ phận đồ họa máy tính, đồng sáng lập ([ESRI](#)) cùng với vợ là Laura. [ESRI](#) trong ít năm vượt trội trong thị trường HTTTĐL và tạo ra các sản phẩm phần mềm ArcInfo và ArcView.

Hội nghị HTTTĐL đầu tiên vào 1970 tổ chức bởi Roger Tomlinson ([CGIS](#)) và Duane Marble (giáo sư tại Northwestern University).

Trong những năm 1980s và 1990s, nhiều ứng dụng được phát triển là những gói phần mềm phát triển bởi các công ty tư nhân như: ArcInfo, ArcView, MapInfo, SPANS GIS, PAMAP GIS, INTERGRAPH, và SMALLWORLD. Và rất nhiều ứng dụng đã chuyển từ hệ máy lớn vào sử dụng trong máy tính cá nhân (PC).

Ngày nay, HTTTĐL với phần cứng và phần mềm đồ họa hiện đại có sức mạnh trong hiển thị thế giới thực, các kỹ thuật 3D thể hiện cảnh quan, hình ảnh động thể hiện sự thay đổi theo thời gian.

1.4. Những ứng dụng của HTTTĐL

Môi trường

Trong lĩnh vực môi trường sử dụng HTTTĐL cho nhiều ứng dụng khác nhau từ kiểm kê đơn giản, chất vấn tới phân tích chồng lớp bản đồ, đưa ra quyết định. Các ứng dụng chính bao gồm:

- Mô hình hóa rừng
- Mô hình hóa khí/nước
- Quan trắc môi trường
- Thành lập bản đồ phân vùng nhạy cảm môi trường
- Phân tích về mối tương tác giữa sự thay đổi kinh tế, khí hậu, thủy văn địa chất.
- Phân tích tác động môi trường
- Chọn vị trí chôn lấp chất thải
- Giám sát sự thay đổi môi trường theo thời gian

Dữ liệu điển hình cho đầu vào những ứng dụng này bao gồm: độ cao địa hình, lớp phủ rừng, chất lượng lớp phủ đất, lớp phủ địa chất-thủy văn. Một số trường hợp ứng dụng HTTTĐL trong nghiên cứu môi trường là sự xem xét cân đối giữa phát triển kinh tế và những điều kiện về môi trường.

Cơ sở hạ tầng và những tiện ích

Những kỹ thuật HTTTĐL cũng được áp dụng rộng rãi trong việc thành lập các dự án và quản lý các tiện ích công cộng. Các cơ quan quản lý cơ sở hạ tầng và tiện ích công cộng tìm thấy ở HTTTĐL những công cụ mạnh mẽ để lập dự án, ra quyết định, phục vụ khách hàng, những yêu cầu cần điều chỉnh, và hiển thị máy tính. Những ứng dụng điển hình bao gồm những dịch vụ:

- Điện lực
- Khí đốt
- Nước
- Thoát nước
- Truyền thông
- Đường xá
- Hiệu quả truyền sóng TV/FM
- Những phân tích mối nguy hiểm, rủi ro
- Tình huống nguy kịch và dịch vụ khẩn cấp.

Những dữ liệu đầu vào cho những ứng dụng này bao gồm:

- Mạng đường phố,
- Dữ liệu địa hình,
- Dữ liệu về nhân khẩu,
- Ranh giới hành chính các cấp.

Kinh doanh và bán hàng

HTTTĐL sử dụng trong kinh doanh và bán hàng hiệu quả nhất trong một số lĩnh vực bao gồm:

- Vị trí có khả năng cạnh tranh.
- Cung cấp phân loại những mối nguy.
- Trợ giúp quản lý rủi ro trong công ty bảo hiểm.
- Tối ưu tuyến vận chuyển và phân phối.
- Gán địa chỉ và tìm kiếm vị trí.

Những dữ liệu đầu vào trong những ứng dụng này bao gồm:

- Mạng đường phố.
- Địa chỉ đường phố.
- Hồ sơ khách hàng.
- Những tài liệu kinh tế-xã hội.

Bản đồ máy tính

Sự phát triển máy tính trợ giúp bản đồ đã phát triển mạnh độc lập với phát triển vector-dựa trên HTTTĐL. Với trợ giúp HTTTĐL, quản lý những mảnh bản đồ theo tờ rất thuận

lợi, những kỹ thuật chồng lớp các chuyên đề thông tin bản đồ, những phép chiếu bản đồ vv... giúp cập nhật CSDL địa lý dễ dàng để tạo những bản đồ mới.

Thông tin đất

HTTTĐL trợ giúp cho quản lý thông tin sử dụng đất vì nó cho phép tạo và duy trì dữ liệu những thửa đất, những dự án đất, tình hình sử dụng. Nhiều nơi những chính quyền địa phương bắt đầu sử dụng HTTTĐL giúp quản lý thông tin đất của họ.

HTTTĐL cho phép dễ dàng nhập, thêm, phục hồi dữ liệu như thuế đất, dự án sử dụng đất, mã đất dễ dàng hơn rất nhiều so với thời đại bản đồ giấy.

Những ứng dụng tiêu biểu là quản lý thông tin đất là:

- Quản lý đăng ký đất sở hữu đất
- Chuẩn bị cho những dự án sử dụng đất và bản đồ phân vùng
- Bản đồ địa chính.

Nguồn vào dữ liệu bao gồm:

- Bản đồ quản lý ranh giới hành chính.
- Giao thông.
- Lớp phủ đất.

Các ngành liên quan

HTTTĐL là kết quả hội tụ kỹ thuật hiện đại của nhiều ngành:

Địa lý

Quan tâm đến hiểu biết thế giới và nơi loài người sinh sống. Các nhà địa lý có truyền thống làm việc lâu dài với các dữ liệu không gian và nhiều kỹ thuật được chuyển sang HTTTĐL.

Giao thông

- Quản lý mạng giao thông.
- Duy trì tín hiệu đèn giao thông.
- Phân tích điểm tai nạn, tìm các điểm nguy hiểm.
- Tuyến giao thông du lịch.
- Quản lý hệ thống ô tô, tìm vị trí, tuyến.

Lâm nghiệp

- Theo dõi thông tin những cây gỗ phát triển.
- Có thể lập dự án khai thác rừng.

Làm sao cung cấp những dinh dưỡng cần thiết cho cây, duy trì bảo vệ tài nguyên rừng trong tương lai.

Lập kế hoạch thiết lập đường vận chuyển, phương pháp khai thác, di chuyển gỗ theo luật môi trường.

Quản lý rừng theo nhiều mục đích, bao gồm cả việc tái tạo lại.

Nông nghiệp – Trang trại

Tăng cường sử dụng các bản đồ chi tiết và những ảnh theo dõi mùa màng.

Phân tích sản lượng.

Có kế hoạch áp dụng hóa chất bảo vệ thực vật, hợp chất hóa học.

Những kỹ thuật dự báo nông nghiệp.

Những tiện ích

Bao gồm khí, điện thoại, điện tử, nước, truyền hình cáp.

Bản đồ học

Hiện thị các thông tin không gian dưới dạng các loại bản đồ. Bản đồ đang tồn tại là nguồn dữ liệu quan trọng cho hệ thống bản đồ điện toán.

Viễn thám

Có nghĩa là thu nhận thông tin từ tàu vũ trụ và vệ tinh. Theo truyền thống, các thông tin này gồm ảnh hàng không, hiện nay đó là các thông tin ảnh số thu nhận từ vệ tinh.

Trắc lượng ảnh

Sử dụng ảnh hàng không và kỹ thuật chiết xuất thông tin từ các ảnh này. Trước đây, trắc lượng ảnh sử dụng các nguồn dữ liệu địa hình (độ cao, đặc điểm nhìn thấy được như đường xá và mạng sông suối, sử dụng đất và lớp phủ đất...).

Khảo sát, cung cấp dữ liệu chính xác cao về vị trí ranh giới đất, công trình xây dựng, đặc điểm tự nhiên...

Số liệu quan sát tạo ra tại một điểm có rất nhiều nguồn dữ liệu cho HTTTĐL: bản đồ, biểu đồ khảo sát, ảnh hàng không, ảnh vệ tinh, bản câu hỏi, dữ liệu định vị toàn cầu.

Thống kê

Cung cấp nhiều phương pháp để xây dựng mô hình điện toán hoặc để phân tích dữ liệu. Kỹ thuật tối ưu hóa (như tìm đường ngắn nhất) là trọng tâm trong ứng dụng HTTTĐL.

Toán học

Cung cấp rất nhiều phương pháp, nhất là trắc địa và lý thuyết đồ họa.

Khoa học máy tính

Cung cấp nhiều phương pháp và công cụ phần mềm mà các nhà phân tích HTTTĐL có thể lựa chọn để giải quyết các vấn đề riêng biệt. Một số nhánh khoa học máy tính có thể khai thác gồm:

Trợ giúp thiết kế: cung cấp phần mềm dùng trong HTTTĐL các kỹ thuật nhập dữ liệu, trình bày, hiển thị.

Đồ họa máy tính: cung cấp phần cứng và phần mềm để thể hiện các đối tượng đồ thị.

Hệ thống quản trị CSDL (DBMS), hệ thống phần mềm để quản trị các bộ cơ sở dữ liệu lớn trong HTTTĐL như các ứng dụng về địa chính và điều tra dân số.

Trí tuệ nhân tạo: cung cấp nhiều kỹ thuật để trợ giúp ra quyết định.

Hành chính

Cơ sở dữ liệu về dân số và các bản đồ kết hợp.

Cơ sở dữ liệu địa chính và các bản đồ kết hợp.

Địa lý nhân khẩu học.

Bài 2

HỆ THỐNG THÔNG TIN ĐỊA LÝ VÀ CÁC THÀNH PHẦN CỦA HỆ THỐNG THÔNG TIN ĐỊA LÝ

Trong bài này sẽ giới thiệu các thành phần của HTTTĐL bao gồm phần cứng, phần mềm, con người, dữ liệu và các phương pháp. Các hợp phần này tạo nên 3 hệ thống con của HTTTĐL đó là:

*Hệ thống nhập dữ liệu.
Hệ thống quản trị dữ liệu.
Hệ thống xuất dữ liệu.*

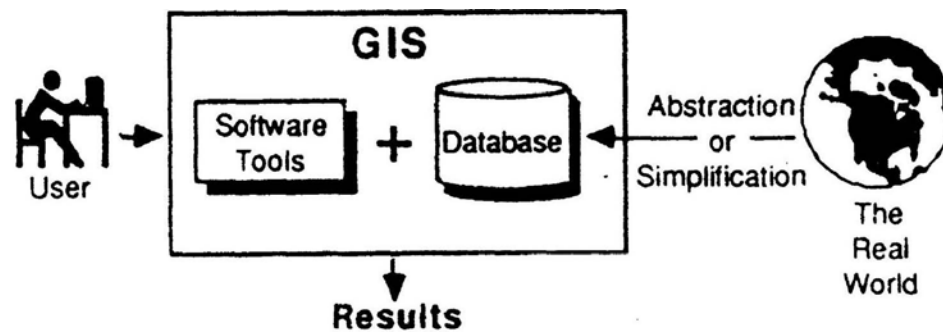


Hình 3.1 Các thành phần cơ bản của HTTTĐL

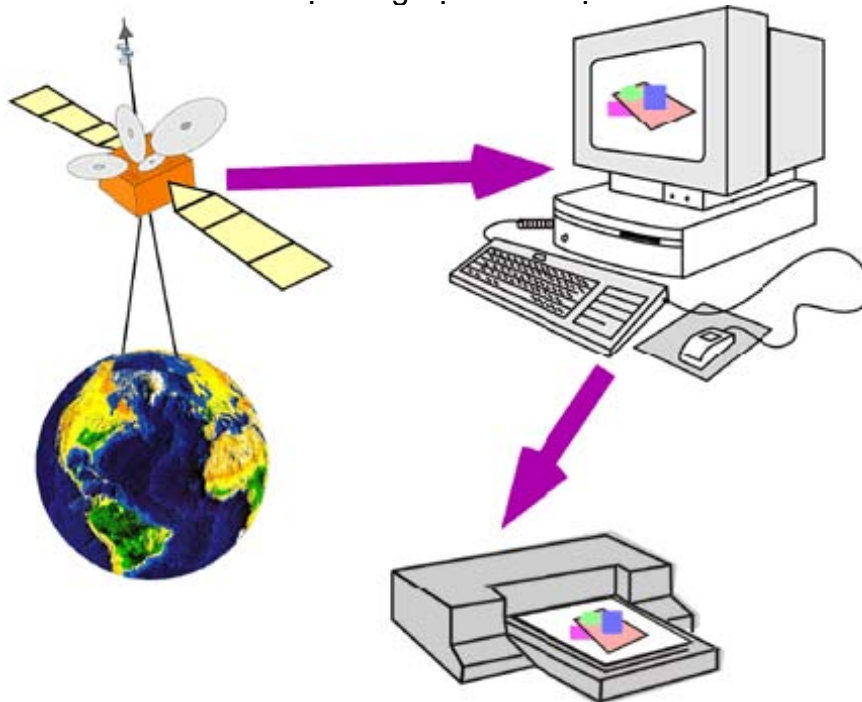
2.1. Hệ thống thông tin và hệ thống thông tin địa lý (HTTTĐL)

2.1. 1. Hệ thống vận hành HTTTĐL

Gồm 3 hệ thống con (Hình 3.3)



Hình 3.2 Hệ thống vận hành hệ HTTTĐL



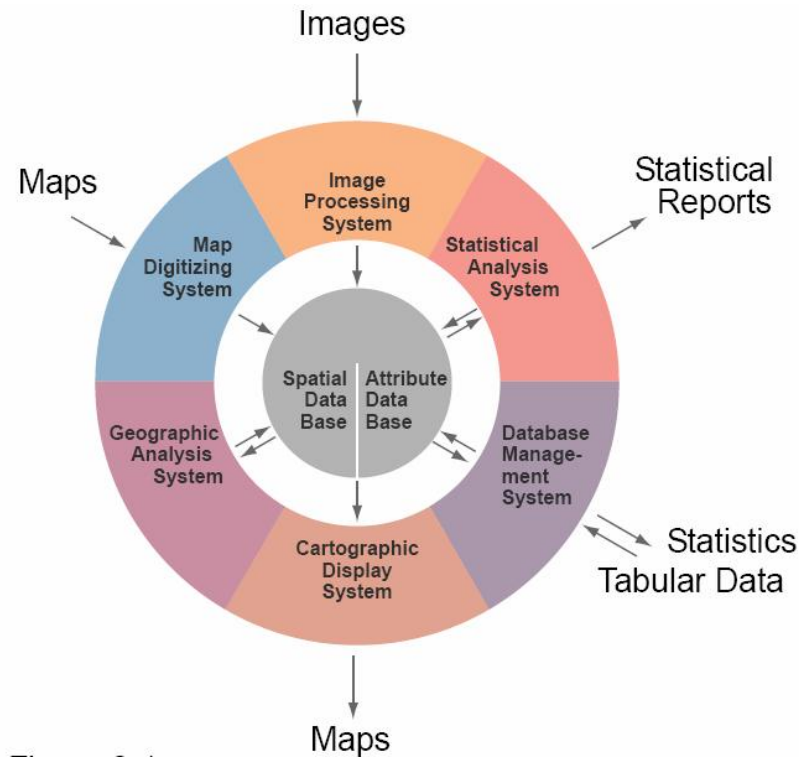
Hình 3.3 Hệ thống vận hành hệ HTTTĐL

Hệ thống đầu vào cho phép thu thập dữ liệu sử dụng cho phân tích theo mục đích.

Hệ thống phần mềm, phần cứng để lưu trữ, quản lý và phân tích dữ liệu hiển thị những thao tác trên màn hình máy tính.

Hệ thống đầu ra tạo ra các bản in (bản đồ, hình ảnh) và những kiểu kết quả khác.

Hệ thống của phần mềm HTTTĐL điển hình



Hình 3.4: các thành phần của một phần mềm HTTTĐL điển hình.

Hệ thống số hóa bản đồ

Hệ thống số hóa bản đồ sử dụng chuyển bản đồ giấy sang dạng bản đồ số để xây dựng CSDL. Một trong các phương pháp số hóa thông thường là dùng bàn số để nhập dữ liệu vector. Phương pháp thứ hai là sử dụng thiết bị scanner để quét bản đồ sau đó dùng các chức năng số hóa của phần mềm HTTTĐL chuyển raster sang vector.

Hệ thống thể hiện bản đồ

Hệ thống này cho phép thể hiện bản đồ trên màn hình máy tính, in bản đồ bằng máy in, máy vẽ. Các sản phẩm bản đồ trong các phần mềm HTTTĐL là rất lớn. HTTTĐL cung cấp các loại bản đồ với chất lượng cao, cho phép trình bày các thành phần bản đồ linh động và tương tác cao trên màn hình, bao gồm các chi tiết kỹ thuật của nhiều lớp dữ liệu phức tạp như chú giải, thước tỷ lệ, bản đồ nhiều màu sắc và những ký hiệu.

Hệ thống quản lý cơ sở dữ liệu

Theo truyền thống thuật ngữ này trình bày kiểu phần mềm dùng để nhập, quản lý và phân tích dữ liệu thuộc tính cũng như các dữ liệu không gian. HTTTĐL kết hợp chặt chẽ không chỉ QTCSDL truyền thống mà còn nhiều tiện ích để quản lý hợp phần không gian và thuộc tính các dữ liệu địa lý.

Hệ thống QTCSDL quản lý các dữ liệu thuộc tính như các thông tin dạng bảng, thống kê, chiết suất các thông tin đặc biệt để tạo các thông báo mới. Tuy vậy, quan trọng nhất, Hệ thống QTCSDL cung cấp khả năng phân tích dữ liệu thuộc tính.

Ví dụ: Thành lập bản đồ những ngôi nhà nơi mà chủ hộ gia đình có một hoặc nhiều con. Sản phẩm cuối cùng (một bản đồ) là dữ liệu không gian nhưng trong phân tích nó không có đặc tính không gian.

Hệ thống phân tích địa lý

Với hệ thống phân tích địa lý, chúng ta mở rộng khả năng chẩn vấn dữ liệu truyền thống bao gồm khả năng phân tích dữ liệu dựa vào vị trí của chúng.

Ví dụ: đơn giản nhất là khi chúng ta quan tâm sự kiện chung của các yếu tố địa lý khác nhau. Giả thiết chúng ta muốn tìm diện tích đất cư trú trên các loại đá gốc với mức khí radon cao. Đây là vấn đề mà HTQTCSDL đơn giản không thể giải quyết vì các kiểu đá gốc và phân chia sử dụng đất không chia sẻ cùng một dữ liệu địa lý. Chất vấn cơ sở dữ liệu truyền thống thực hiện tốt với điều kiện là chúng ta nói về các thuộc tính phụ thuộc vào cùng một yếu tố. Nhưng khi các yếu tố khác nhau, nó không thể thực hiện. Vì vậy chúng ta phải cần đến HTTTĐL. Thực ra, HTTTĐL có khả năng so sánh các yếu tố khác nhau dựa trên các sự kiện địa lý chung của chúng, đó là dấu hiệu xác nhận tiêu chuẩn của HTTTĐL. Sự phân tích này được thực hiện thông qua một quá trình gọi là "overlay".

Giống như HTQTCSDL hệ phân tích địa lý có hai cách tương tác với cơ sở dữ liệu: trong khi truy cập dữ liệu từ cơ sở dữ liệu, nó có thể đóng góp các kết quả mà nó phân tích như một phần thêm mới cho cơ sở dữ liệu.

Ví dụ, chúng ta có thể tìm mối liên quan giữa các bậc độ dốc và đất đai bị xâm thực do nông nghiệp và tạo ra bản đồ gọi là nguy cơ xâm thực đất. Bản đồ này không phải là bản đồ nguyên thủy mà nó xuất phát từ các dữ liệu hiện có và một tập hợp các dữ liệu xác định. Khả năng phân tích của hệ thống phân tích địa lý và HTQTCSDL giữ vai trò quan trọng trong việc mở rộng CSDL thông qua việc bổ sung các tri thức về mối quan hệ giữa các yếu tố.

Hệ thống xử lý ảnh

Một số phần mềm HTTTĐL còn có khả năng phân tích ảnh viễn thám và cung cấp các phân tích thống kê chuyên hóa. Phần mềm xử lý ảnh cho phép lấy ảnh viễn thám dạng thô và chuyển sang dạng dữ liệu bản đồ giải đoán (ảnh Landsat, SPOT) theo các thủ tục phân loại khác nhau.

Hệ thống phân tích thống kê

HTTTĐL cung cấp cả các thủ tục thống kê truyền thống cũng như một số thủ tục chuyên hóa để phân tích thống kê các dữ liệu không gian. Các nhà địa lý đã phát triển hàng loạt các thủ tục chuyên hóa để mô tả thống kê các dữ liệu không gian, một phần do tính chất đặc biệt của dữ liệu không gian, một phần do dữ liệu không gian đặt ra các vấn đề đặc biệt để suy luận bản đồ từ các thủ tục thống kê.

2.2. Các thành phần cơ bản HTTTĐL

2.2.1. Phần cứng (Hardware)

Phần cứng là các thiết bị sử dụng trong các thao tác HTTTĐL. Ngày nay phần mềm HTTTĐL chạy trên mọi kiểu phần cứng, Từ máy chủ trung tâm tới máy tính cá nhân, trên mạng hay máy đơn.

Máy tính. Máy tính sử dụng trong HTTTĐL có thể máy tính cá nhân, máy chủ và có thể làm việc trong môi trường mạng.

Vì có quá nhiều dữ liệu địa lý là hệ giao tiếp đồ họa, ví dụ: bao gồm các điểm, đường và các ảnh nên các nhà phân tích HTTTĐL muốn có các phần cứng đặc biệt để thực hiện, ví dụ: với công việc tương tác cần trạm làm việc giao tiếp đồ họa (workstation) với màn hình lớn, độ phân giải cao.

Thiết bị nhập dữ liệu. Bao gồm bàn số hóa (digitizer) và máy quét (scanner). Bàn số hóa dùng số hóa những yếu tố lựa chọn trên bản đồ giấy. Số hóa bàn số là một phương pháp phổ biến chuyển đổi bản đồ giấy và hình ảnh thành dạng số. Tuy vậy, đây là quá trình mệt mỏi, đặc biệt khi chuyển đổi những bản đồ mật độ cao. Máy quét ngày nay có thể thay thế bàn số bởi tự động chuyển đổi bản đồ giấy thành dạng số. Trong HTTTĐL, những ảnh raster có thể chuyển thành dạng vector thông qua quá trình chuyển đổi "raster-to-vector".

Máy in. Những thiết bị này dùng để in bản đồ. Gồm một số loại như: in kim, in phun, in laser. Kiểu máy vẽ gồm: bút vẽ (pen plotter), vẽ nhiệt thường đòi hỏi phần cứng chất lượng cao.

Hệ thống lưu trữ. Gồm: đĩa quang học, đĩa từ (ổ cứng máy tính), đĩa mềm, băng từ.

Theo quan điểm của các nhà địa lý, phần cứng đang được quan tâm hiện nay là hệ thống định vị toàn cầu.

2.2.2. Phần mềm

Phần mềm HTTTĐL cung cấp những chức năng và những công cụ cần thiết để nhập, lưu trữ, phân tích và hiển thị thông tin địa lý. Những chức năng chính là:

Những công cụ cho việc nhập và thao tác với thông tin địa lý.

Hệ thống lưu trữ và quản trị cơ sở dữ liệu.

Những công cụ cho phép chất vấn, phân tích, thể hiện, chuyển đổi dữ liệu.

Giao tiếp đồ họa với người sử dụng để dàng truy xuất, trình bày dữ liệu.

Phần mềm HTTTĐL bao gồm: Chương trình HTTTĐL. Là những gói ứng dụng chuyên dụng như mô hình hóa địa hình và phân tích mạng lưới. Những phần mềm HTTTĐL gồm:

- Modul GIS Environment của hãng Intergraph Corp.,
- Geo/SQL của Generation 5 Tech. Inc.,
- ARC/INFO của Environmental Systems Research Institute Inc., (ESRI)
- SPANS của Tydac Technologies,
- FMS/AC của Facility Mapping Systems Inc.

Những chương trình bản đồ máy tính cung cấp nhiều chức năng như HTTTĐL, nhưng bị giới hạn khả năng các phân tích không gian. Chúng được phát triển để thỏa mãn nhu cầu người sử dụng biểu diễn bản đồ. Một số chương trình loại này gồm:

- MapInfo phát triển bởi MapInfo Corp.,
- Atlas GIS phát triển bởi Strategic Mapping Inc.,
- MapGrafix phát triển bởi ComGrafix Inc.,
- QUIKMAP phát triển bởi AXYS Software Ltd. etc.

Phần mềm công cộng là những chương trình HTTTĐL phát triển bởi chính phủ hoặc các trường đại học, Cho phép miễn phí hoặc giá tượng trưng. Gồm các phần mềm như:

- IDRISI của trường Clark University.
- GRASS của GRASS Information Center.
- MOSS của Autometric Inc.

2.2.3. Dữ liệu

Thành phần quan trọng trong HTTTĐL là dữ liệu. Dữ liệu địa lý và những dữ liệu bảng biểu liên quan có thể thu thập hoặc mua từ những nhà cung cấp dữ liệu. HTTTĐL sẽ tích hợp dữ liệu trong HTQTDL nhằm tổ chức và duy trì dữ liệu không gian và thuộc tính. Khi tiến hành phân tích không gian, người dùng phải có các kỹ năng lựa chọn và sử dụng công cụ từ các hộp công cụ HTTTĐL và có những kiến thức sâu sắc về các dữ liệu sử dụng.

2.2.4. Con người (chuyên gia)

Con người quản lý hệ thống và phát triển các dự án nhằm ứng dụng HTTTĐL để nghiên cứu các vấn đề thực tế. Người sử dụng gồm các *chuyên gia kỹ thuật*, *người thiết kế và duy trì hệ thống*, *người sử dụng* nó để trợ giúp thực hiện những công việc hàng ngày.

Con người tham gia HTTTĐL gồm:

Những thành viên thực hiện, gồm:

Người vẽ bản đồ, theo dõi thiết kế hiển thị bản đồ, những chuẩn biểu tượng, ký hiệu bản đồ và những chuẩn loạt bản đồ.

Nhập dữ liệu, chuyển đổi bản đồ thành dạng số.
Những người sử dụng HTTTĐL.

Chuyên viên kỹ thuật:

Phân tích thông tin giải quyết các vấn đề, làm thỏa mãn những yêu cầu thông tin của người sử dụng.

Người quản trị hệ thống, luôn duy trì hệ thống hoạt động.

Lập trình viên, chuyển đổi những ứng dụng của người phân tích thành chương trình.

Người quản trị dữ liệu, trợ lý cho người phân tích, lập trình viên và người sử dụng nhằm tổ chức các yếu tố địa lý thành những lớp dữ liệu, xác định nguồn dữ liệu, phát triển cấu trúc mã cho các dữ liệu thuộc tính, và những tài liệu thông tin về nội dung CSDL.

Tổ chức:

Người quản lý, theo dõi thực hiện dự án HTTTĐL.

Người quản lý chất lượng.

2.2.5. Những phương pháp

Những phương pháp thực hiện sẽ quyết định sự thành công một dự án HTTTĐL, tùy thuộc vào những kế hoạch thiết kế, luật lệ chuyển giao vv...

2.3. Cấu trúc dữ liệu trong gis

Dữ liệu của một hệ thống thông tin địa lý có thể chia thành hai dạng:

-Hình ảnh (không gian)

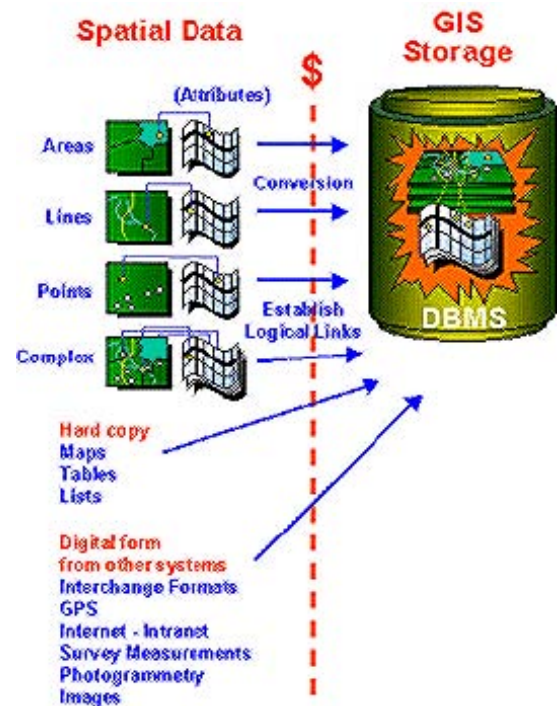
-Phi hình ảnh (thuộc tính)

2.3.1. Dữ liệu không gian

Số liệu hình ảnh hay còn gọi là dữ liệu không gian (graphic) là sự mô tả bằng kỹ thuật số các phần tử bản đồ. GIS sử dụng dữ liệu hình ảnh để thể hiện bản đồ ra màn hình hay ra giấy.

Trong máy tính, dữ liệu không gian thường được thể hiện dưới các dạng sau: -Điểm: được thể hiện bằng những biểu tượng dạng điểm -Đường gấp khúc hay đoạn cong -Vùng hay đa giác -**Các điểm ảnh**

Các thành phần đồ họa trong cơ sở dữ liệu GIS thường được mô tả bằng nhiều lớp (layer), mỗi lớp chứa một nhóm đối tượng thuần nhất với vị trí của chúng theo hệ tọa độ chung của tất cả các lớp.

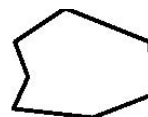
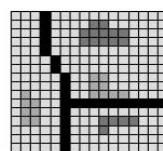
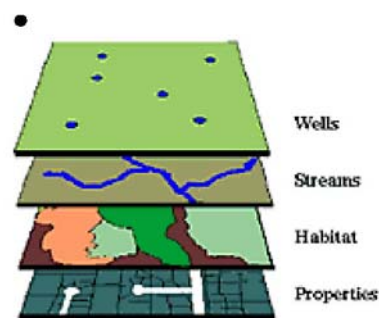


2.3.2. Dữ liệu phi không gian

Số liệu thuộc tính thể hiện các tính chất, số lượng, chất lượng hay mối quan hệ của các phần tử bản đồ và các vị trí địa lý. Chúng được lưu trữ dưới dạng số hay ký tự. Thông thường dữ liệu được quản lý dưới dạng bảng (table) bao gồm cột (column) hay còn được gọi là trường (field), hàng (row) hay còn gọi là mẫu tin (record).

Để định nghĩa một trường phải có tên trường (field name) và kiểu dữ liệu của trường (type), kiểu dữ liệu có thể là: kiểu ký tự (character), kiểu số nguyên (integer), kiểu số thực (real), kiểu logic,...

Ví dụ: ta có bảng dữ liệu về thế giới như sau



Tân trảng

	Tân quốc gia	Tân thu ảả	Dán sã -94	TL GT DS
Mẫu tin	☐ Afghanistan	Kabul	15,513,267	5.2
	☐ Albania	Tirane	1,626,315	1.8
	☐ Algeria	Algiers	22,600,957	2.5
	☒ Andorra	Andorra La Vella	61,599	2.4

Mẫu tin thể hiện tổng hợp các tính chất của đối tượng mà nó miêu tả, ví dụ như ở bảng trên, các mẫu tin thể hiện các tính chất, số liệu về các quốc gia như tên quốc gia, tên thủ đô ỉ, dân số, tỉ lệ gia tăng dân số,...

2.4. Khái niệm về bản đồ

Bản đồ là một mô hình của các thực thể và các hiện tượng trên trái đất, trong đó các thực thể được thu nhỏ, đơn giản hóa và các hiện tượng được khái quát hóa để có thể thể hiện được trên mặt phẳng bản vẽ. Bản đồ chứa các thông tin về vị trí và các tính chất của vật thể và các hiện tượng mà nó trình bày.

Thế giới thực rất rộng lớn và phức tạp, ngoài ra thế giới thực có quá nhiều kích thước để chúng ta có thể thấy bao quát được. Nếu một phần không gian được chọn để trình bày dưới một tỉ lệ nhỏ hơn thực tế thì chúng ta có thể thấy được cấu trúc và dạng của phần không gian đó dễ hơn nhiều và từ đó có thể hiểu thấu đáo được khu vực nghiên cứu và có thể đưa ra được quyết định đúng đắn (như việc tìm đường đi, việc qui hoạch một tuyến đường, việc tìm kiếm một vị trí thích hợp để xây dựng khu công nghiệp,...)

Thông thường bản đồ là một mô hình theo tỉ lệ. Có nghĩa là tỉ lệ của khoảng cách trên bản đồ và khoảng cách trên thực tế sẽ bằng nhau ở mọi vị trí trên bản đồ, mặc dù có một vài sai số không thể tránh khỏi nếu một phần của mặt cầu được thể hiện trên mặt phẳng. Chúng ta thường gặp vấn đề này trong bản đồ có tỉ lệ nhỏ trình bày một khu vực rộng lớn.

Thực chất bản đồ là một hệ thống thông tin về không gian vì nó có khả năng truyền đổi thông tin như báo chí, sách hay vô tuyến truyền hình. Chúng ta có thể xem bản đồ và tìm thấy các thông tin mà người vẽ bản đồ muốn truyền tải, ví dụ như bản đồ bản đồ địa hình, bản đồ dân số, bản đồ quy hoạch sử dụng đất, bản đồ địa chất thủy văn, bản đồ địa chất môi trường,...

Bài 3

CƠ SỞ DỮ LIỆU CỦA HỆ THỐNG THÔNG TIN ĐỊA LÝ

HTTTĐL thuộc loại ứng dụng máy tính để xây dựng một CSDL lớn. Không giống như các ứng dụng máy tính khác, người sử dụng có thể dùng ngay sau khi mua phần cứng và phần mềm. Để sử dụng HTTTĐL yêu cầu một CSDL không gian được xây dựng thích hợp với phần cứng, phần mềm và những ứng dụng đã phát triển, những thành phần đã thiết đặt, tích hợp và kiểm tra trước khi có thể sử dụng CSDL HTTTĐL.

Trong BÀI này sẽ trình bày những vấn đề liên quan tới xây dựng Cơ Sở Dữ Liệu (CSDL HTTTĐL)

3.1. Dữ liệu và cơ sở dữ liệu

Cơ sở dữ liệu là sự chọn lọc các dữ liệu cần thiết nhất (không có số liệu thừa: redundant data) và các dữ liệu này có thể chia sẻ giữa nhiều hệ thống ứng dụng khác nhau. CSDL có thể được xem như là giao diện giữa số liệu và các chương trình ứng dụng.



Hình 3.1. CSDL và ứng dụng Để cơ sở dữ liệu (CSDL) có thể làm việc tốt, thì trong CSDL không có dữ liệu thừa.

3.2. Liên kết dữ liệu trong CSDL

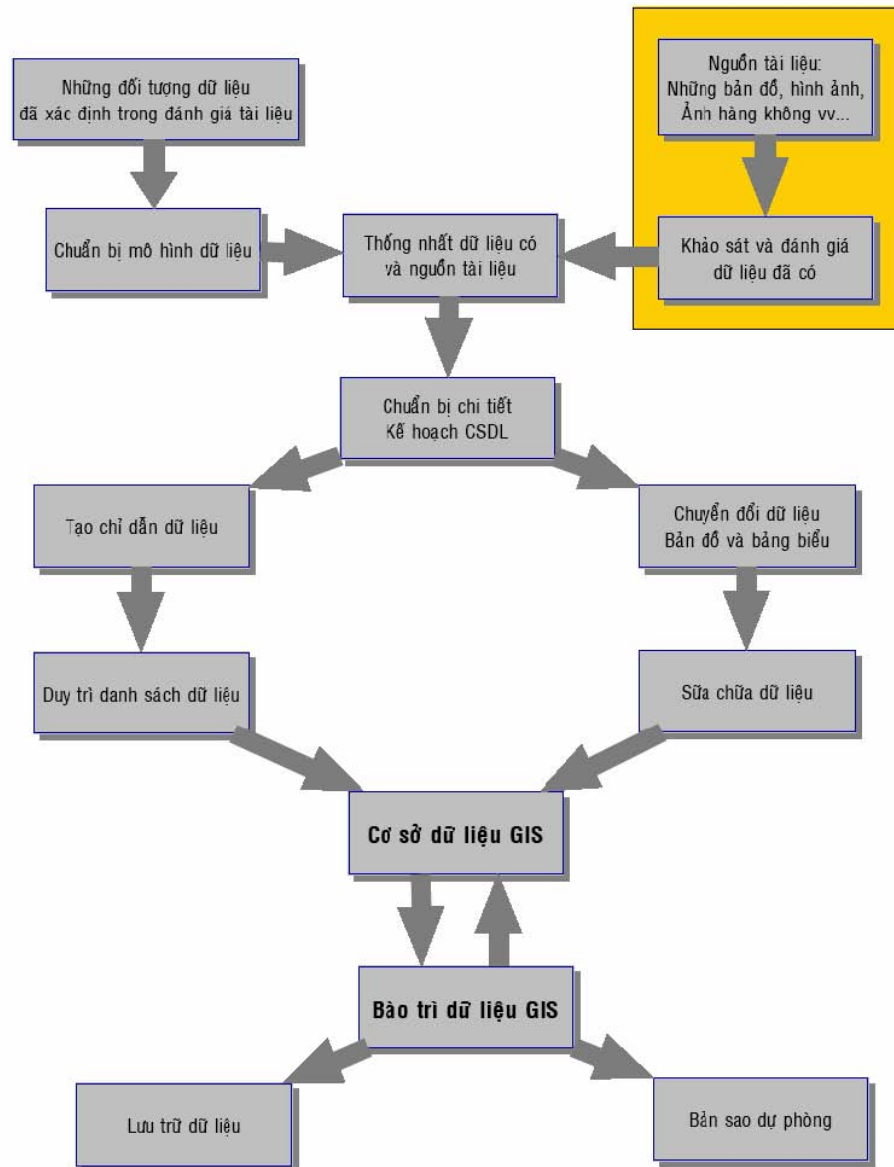
CSDL bao gồm nhiều tập tin (hay bảng) dữ liệu, các tập tin dữ liệu này nếu mô tả cho cùng một loại đối tượng sẽ được liên kết với nhau bởi các trường khóa (identifier), ví dụ: có hai bảng dữ liệu về sinh viên như sau:

MSSV	Họ tên	Ngày sinh	Nơi sinh
------	--------	-----------	----------

25412562	Nguyễn Văn Tuấn	12/8/76	Cần Thơ
25412563	Trần Văn Hoàng	4/10/76	An Giang
25412564	Lê Hoàng Anh	7/04/76	Sóc Trăng
MSSV	Nền móng	Kết cấu bê tông thép gỗ	Kiến trúc
25412562	8	9	9
25412563	7	8	5
25412564	5	6	7

Trường khóa trong trường hợp này là trường MSSV. Để biết điểm của sinh viên, ta phải truy số liệu của bảng 1 để biết tên và bảng 2 để biết điểm. Chương trình máy tính sẽ dựa vào trường khóa MSSV để lấy số liệu theo yêu cầu của người truy cập số liệu.

3.3. Đặc điểm CSDL HTTTĐL



Hình 4.1 Chu trình CSDL HTTTĐL

CSDL rất quan trọng vì để tạo ra nó thường tới 3/4 thời gian để phát triển HTTTĐL. Mỗi lần tổ chức thông tin, CSDL xây dựng từ 10 tới 15 năm.

CSDL tóm lược rành mạch, rõ ràng loại thông tin về thế giới thực và tổ chức nó theo phương thức chứng tỏ sự hiệu quả (hữu ích). CSDL được xem như biểu diễn hay mô hình của thực tế (world) được phát triển cho ứng dụng cụ thể.

Một trong những lý do có rất nhiều hệ thống phần mềm và phần cứng sử dụng cho GIS vì mỗi hệ thống cho phép người sử dụng biểu diễn hay mô hình những kiểu nào

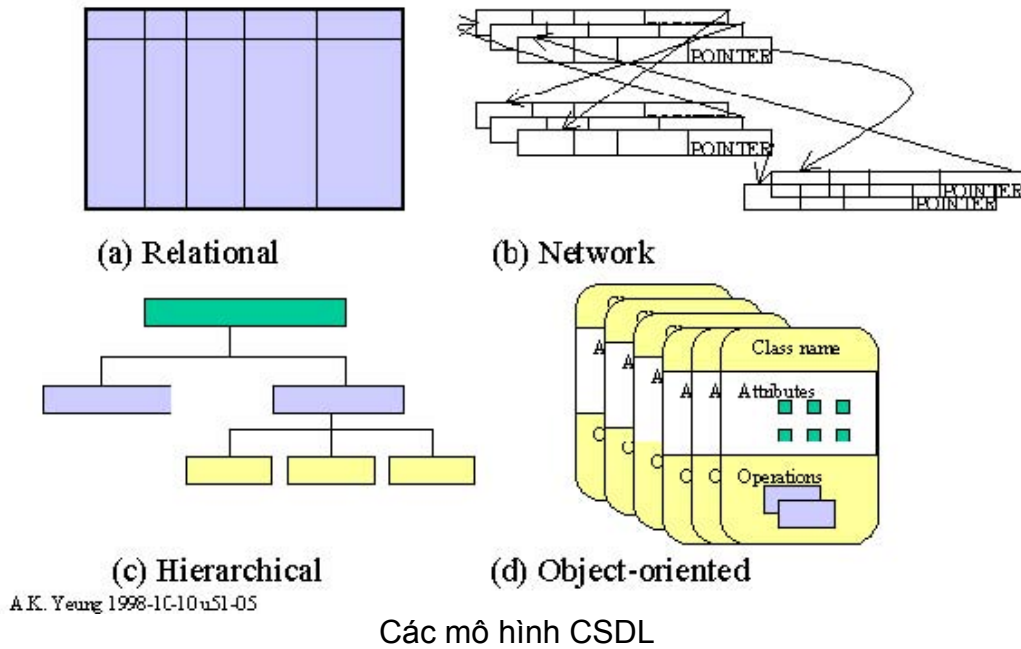
đó của tự nhiên.

3.4. Tổ chức CSDL HTTTĐL hay những mô hình CSDL

Có 4 mô hình CSDL cơ bản là

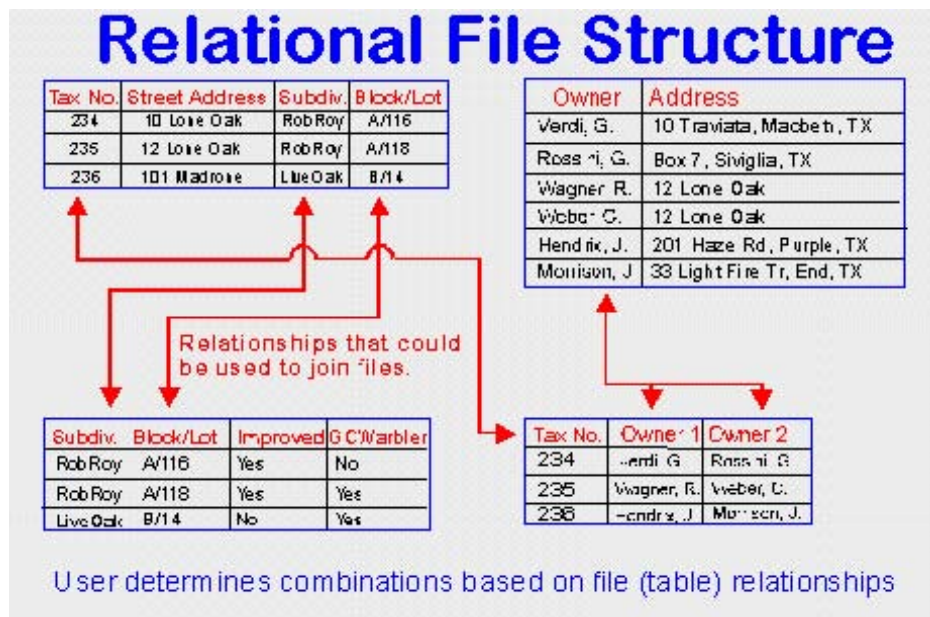
Mô hình quan hệ, Mô hình mạng, Mô hình phân nhánh, Mô hình hướng đối tượng.

Figure 5: Database models



3.4.1. Mô hình quan hệ

Dữ liệu được tổ chức bởi những bản ghi giống như bảng (xem chi tiết hơn trong Chương 5)



Hình 4.4 Mô hình dữ liệu quan hệ

- Trong mô hình dữ liệu quan hệ (hình 4.4) không có cấp bậc của trường dữ liệu, mỗi trường dữ liệu được dùng như là key.
- Mỗi bảng hai chiều thường được lưu trữ như một tập tin riêng.
- Bảng thể hiện mối quan hệ giữa tất cả thuộc tính được chứa trong bảng.
- Việc tìm kiếm những thuộc tính quan hệ được lưu trữ trong những bảng khác nhau có thể được làm bằng cách nối hai hoặc nhiều bảng dùng thuộc tính giống nhau.
- Đây là hệ thống linh động nhất và thích hợp cho việc sử dụng SQL (structured query language). Vì tính linh động hệ thống này nên phần lớn mô hình dữ liệu quan hệ được sử dụng nhiều nhất để lưu trữ thông tin thuộc tính trong hệ HTTTDL.

Điều quan trọng là những dữ liệu phi không gian được phân chia ra một vài dạng tùy theo sự cần thiết truy xuất chúng như: file ngang hàng, file phân nhánh và file quan hệ. Phương pháp đơn giản nhất là file ngang hàng trong đó mỗi yếu tố địa lý tương xứng một hàng dữ liệu.

3.4.2. Mô hình mạng

Dữ liệu được tổ chức bởi những bản ghi với con trỏ liên kết.

3.4.3. Mô hình phân nhánh

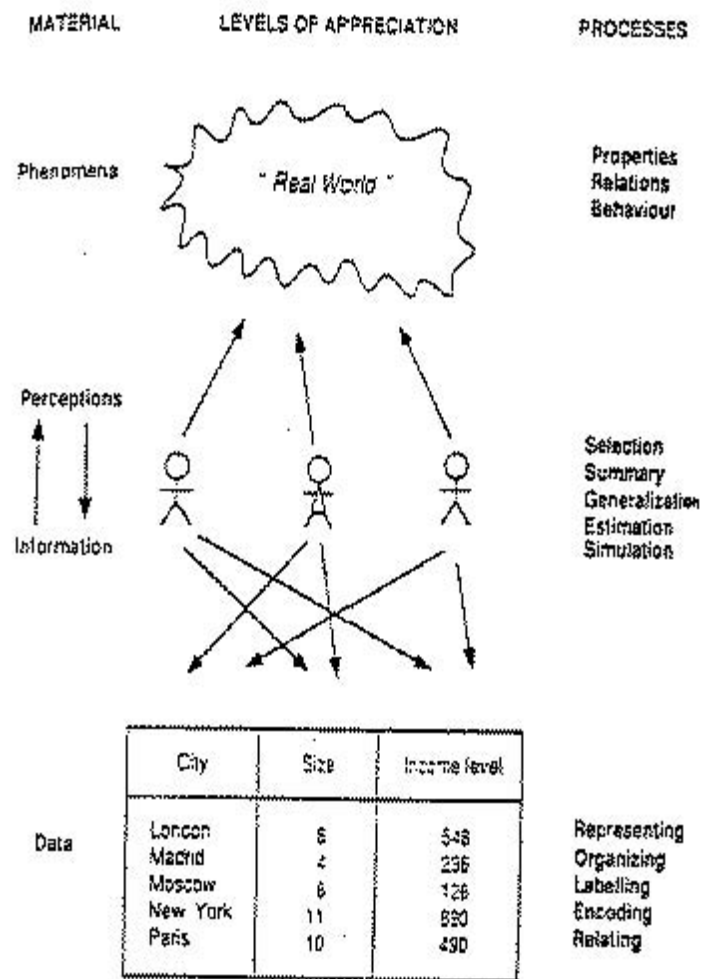
- Dữ liệu được tổ chức bởi những bản ghi trong tổ chức cha-con một-tới-nhiều mối quan hệ.
- Dữ liệu được tổ chức theo cấu trúc hình cây (Hình 4.3). Cấp bậc cao nhất

- Đây là loại mô hình có lớp trên, lớp dưới. Mỗi thành phần chỉ có một cha nhưng có nhiều con.
- Trong mô hình phân cấp, mọi quan hệ là quan hệ nhiều– một hoặc quan hệ một– một.
- Việc truy tìm dữ liệu sẽ hiệu quả nếu không có nhiều cấp trung gian trong CSDL.

3.4.4. Hướng đối tượng

Mô hình mới nổi lên, dữ liệu duy nhất xác định như những đối tượng riêng biệt phân loại thành những kiểu đối tượng hay lớp tùy thuộc vào đặc điểm (những thuộc tính và những phép toán)

3.5. Thiết kế một CSDL HTTTĐL



Hình 4.2 Các bước phát triển CSDL HTTTĐL.

Thiết kế một CSDL HTTTĐL bao gồm các bước:

Thiết kế khái niệm

Ở mức thiết kế này là cơ sở hình thành CSDL cần xây dựng, được xây dựng trên cơ sở khảo sát nhu cầu, thông tin, nguồn dữ liệu. Xây dựng một sơ đồ tổng quát cho các yêu cầu cho CSDL HTTTĐL. Mức thiết kế này không phụ thuộc vào phần cứng hoặc phần mềm. Chỉ quan tâm đến các mục tiêu ứng dụng mà người dùng đòi hỏi. (Bước 1, 2, 3, 4 trên sơ đồ Hình 4.2).

Thiết kế logic

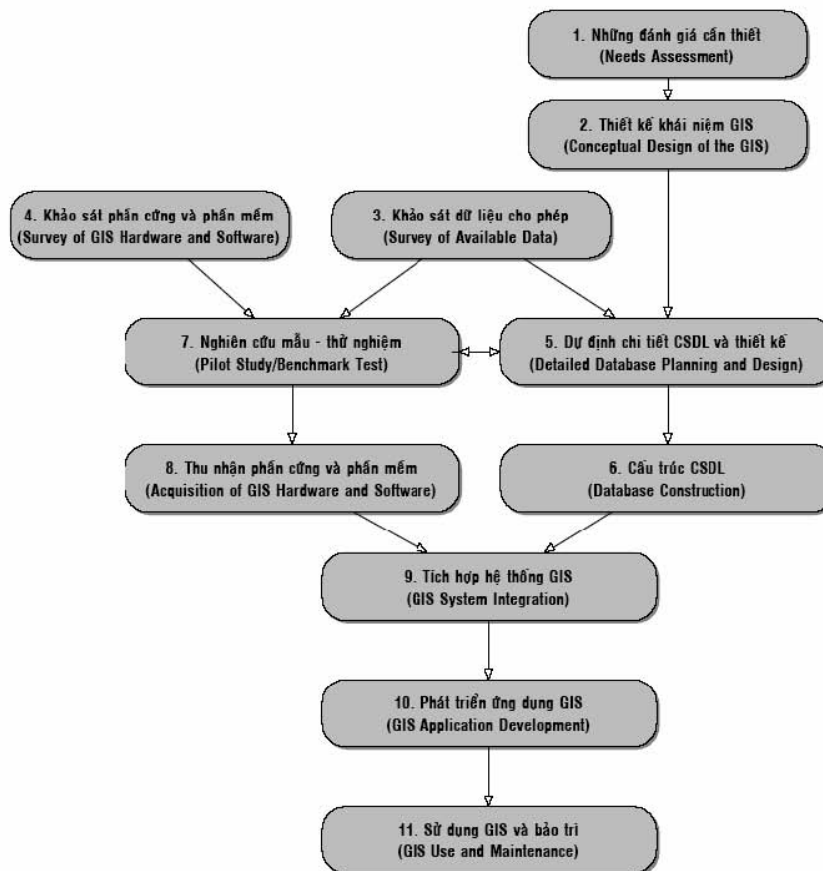
Trong mức thiết kế này, CSDL được mô tả chi tiết, bao gồm các hạng mục tin, các mối quan hệ dữ liệu, đặt mức độ chính xác, các thủ tục đảm bảo toàn vẹn dữ liệu. Thiết kế logic cũng đưa ra cấu trúc của các thành phần trong CSDL (còn gọi là cấu trúc CSDL). Mức thiết kế này là khởi điểm của các công việc tin học. Người thiết kế phải hiểu rõ tính năng của một hệ thống phần mềm quản trị CSDL. Trong mức này, người thiết kế đưa ra các phương án để lựa chọn các thành phần của CSDL HTTTĐL. (Bước 5, 6)

Thiết kế vật lý

Mức này là sự triển khai và điều chỉnh thành quả của mức logic trên các phần cứng, phần mềm cụ thể của HTTTĐL.

Phát triển CSDL HTTTĐL không chỉ đơn thuần mua phần cứng hay phần mềm. Phần đòi hỏi khắt khe nhất quá trình phát triển HTTTĐL là xây dựng CSDL. Đòi hỏi nhiều thời gian nhất, chi phí nhiều tiền nhất, và yêu cầu nỗ lực trong lập kế hoạch và quản lý. Mặc dù chu trình phát triển HTTTĐL hiện nay phần lớn tập trung vào xây dựng CSDL, nhưng một số địa phương vẫn tập trung vào mua phần cứng và phần mềm. Việc chọn lựa đúng cho xây dựng CSDL đúng đắn phải dựa trên sự hiểu biết của những cơ quan có kinh nghiệm.

3.6. Các bước phát triển CSDL HTTTĐL



1. Những đánh giá cần thiết (Needs Assessment)
2. Thiết kế khái niệm GIS (Conceptual Design of the GIS)
3. Khảo sát những dữ liệu cho phép (Survey of Available Data)
4. Khảo sát phần cứng và phần mềm (Survey of GIS Hardware and Software)
5. Dự định chi tiết CSDL và thiết kế (Detailed Database Planning and Design)
6. Cấu trúc CSDL (Database Construction)
7. Nghiên cứu mẫu - thử nghiệm (Pilot Study/Benchmark Test)
8. Thu nhận phần cứng và phần mềm (Acquisition of GIS Hardware and Software)
9. Tích hợp hệ thống GIS (GIS System Integration)
10. Phát triển ứng dụng GIS (GIS Application Development)
11. Sử dụng GIS và bảo trì (GIS Use and Maintenance)

Hình 4.2 Các bước phát triển CSDL HTTTĐL.

Gồm 11 bước bắt đầu bằng việc đánh giá và kết thúc với việc sử dụng và duy trì hệ thống HTTTĐL. Những bước này có mối liên quan mật thiết với nhau, mỗi bước hoàn thiện sẽ khởi đầu cho những bước tiếp theo. Trong khi thực tế một số hoạt động này có khi xảy ra đồng thời.

3.6.1. Mô hình hóa dữ liệu

Mô hình hóa dữ liệu là một quá trình định nghĩa các hiện tượng hay các yếu tố địa lý

mà đặc điểm và những mối quan hệ chúng được quan tâm. Liên quan tới thực hiện tổ chức thông tin và cấu trúc dữ liệu.

Có ba mức trong quá trình mô hình hóa dữ liệu, những mô hình dữ liệu tăng dần nghi thức định nghĩa chính xác hơn trong CSDL HTTTĐL.

Mô hình hóa khái niệm-định nghĩa rộng và tổng quát phạm vi và yêu cầu của CSDL.

Mô hình hóa Logic – xác định yêu cầu người sử dụng của CSDL với những định nghĩa rõ ràng những thuộc tính và những mối quan hệ.

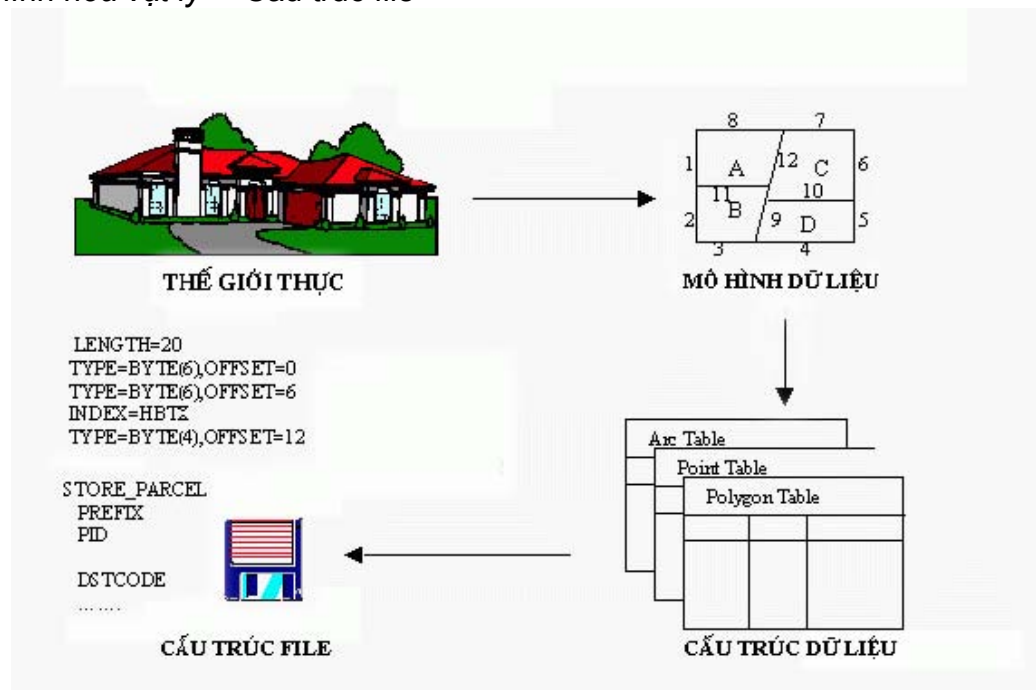
Mô hình hóa vật lý – xác định cấu trúc lưu trữ bên trong và tổ chức các file dữ liệu trong CSDL.

MHHDL liên quan tới 3 mức mô hình dữ liệu trong thiết kế CSDL:

Mô hình hóa khái niệm Mô hình dữ liệu

Mô hình hóa Logic Cấu trúc dữ liệu

Mô hình hóa vật lý Cấu trúc file



H 4.1 Những mức độ rút gọn dữ liệu trong tổ chức thông tin

3.6.2. Mô hình hóa khái niệm

Mô hình hóa dữ liệu khái niệm: *nhận biết, nhận diện nội dung dữ liệu và mô tả nó*

trong dạng tóm tắt, hay khái niệm, mức độ của nó. Bước này xác định mục tiêu CSDL HTTTĐL cần làm gì, làm sao sẽ thực hiện được. Trên cơ sở đó, xác định tất cả các dạng nhu cầu về dữ liệu của người dùng nhằm thực hiện mục tiêu nói trên.

Thiết kế khái niệm hệ thống HTTTĐL là bước tiền thân của thiết kế CSDL HTTTĐL, bao gồm các mô hình nghi thức, thủ tục (chuẩn bị cho mô hình dữ liệu) cho CSDL HTTTĐL và khởi đầu cho chiến lược xây dựng CSDL. Thiết kế khái niệm CSDL là hoạt động quan trọng nhất trong phát triển HTTTĐL.

Chuẩn bị mô hình dữ liệu HTTTĐL

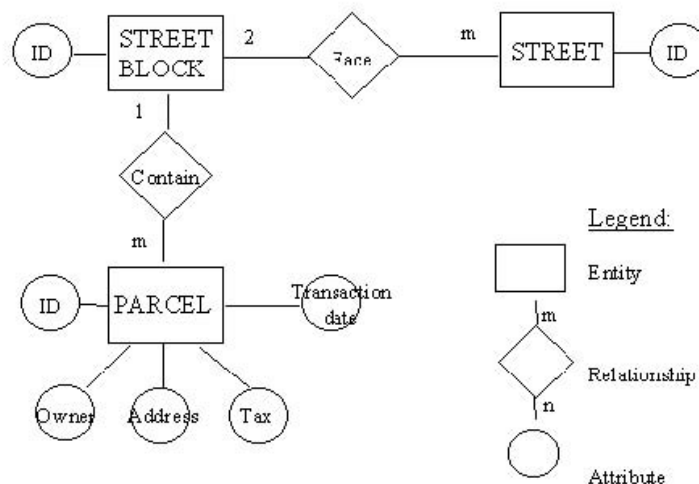
Mô hình dữ liệu là hình thức xác định yêu cầu dữ liệu trong HTTTĐL. Mô hình dữ liệu có thể trong một vài dạng:

Biểu đồ Thực thể-mối quan hệ.

Mục đích của mô hình dữ liệu và quá trình xác định mô hình là đảm bảo rằng dữ liệu được nhận dạng và mô tả trong tính toàn vẹn chính xác và kiểu cách hình thức được người sử dụng và người phân tích dữ liệu chấp nhận.

Mô hình dữ liệu lúc này có hình thức đặc biệt của thực thể những thuộc tính và tất cả mối quan hệ thực thể trong HTTTĐL (Dạng xác định của thực thể, đặc tính của chúng và tất cả các mối quan hệ giữa những thực thể trong HTTTĐL). Việc xây dựng mô hình dữ liệu không cần thiết cho dự án qui mô nhỏ.

Figure 21: A portion of an entity-relationship (E-R) diagram



A.K. Yeung 1998-10-10 u51-21

Hình Mô hình thực thể-mối quan hệ

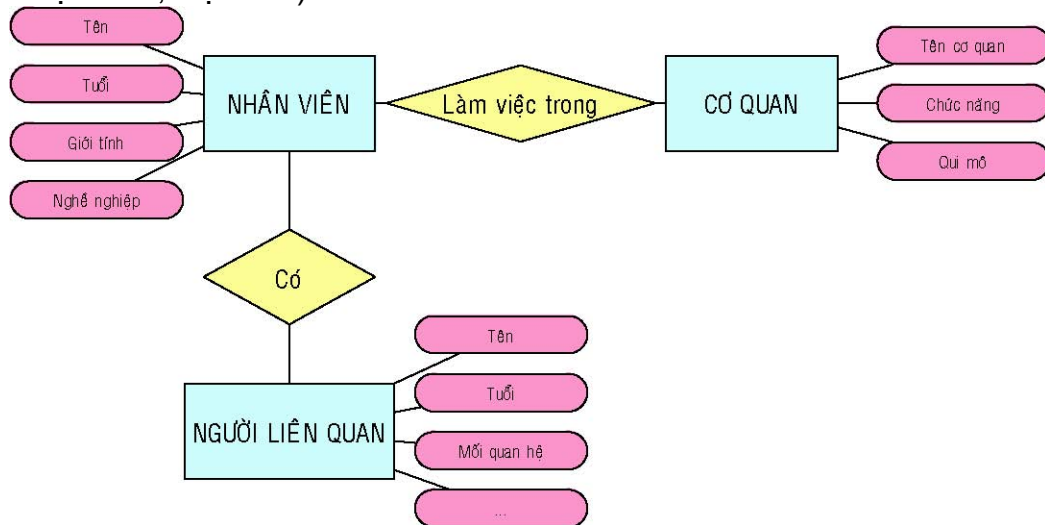
Mô hình hóa dữ liệu thực thể-mối quan hệ (E-R)

Mô hình thực thể-mối quan hệ (entity-relationship), hai ví dụ sẽ xem xét. Những CSDL bình thường (thực tế)

Ví dụ CSDL về công chức trong một tổ chức gồm thông tin (nhân viên, người phụ thuộc, văn phòng, vv...). Những mối quan hệ giữa những thực thể NHÂN VIÊN làm việc trong CƠ QUAN và NHÂN VIÊN có NGƯỜI LIÊN QUAN. Một số thuộc tính của mỗi thực thể như sau: (Hình 4.4).

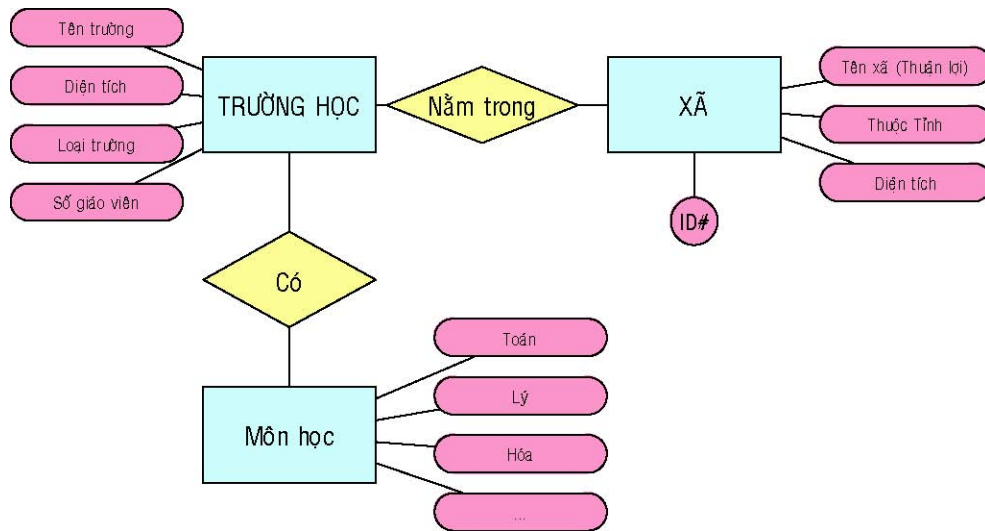
NHÂN VIÊN (tên, tuổi, giới tính, nghề) NGƯỜI LIÊN QUAN (tên, tuổi, quan hệ với NHÂN VIÊN) CƠ QUAN (tên cơ quan, chức năng, qui mô). CSDL không gian đơn giản sẽ như sau Ví dụ CSDL không gian đơn giản gồm thông tin (TRƯỜNG HỌC, MÔN HỌC, THUỘC TỈNH, vv...). Hình 4.5

Những mối quan hệ giữa những thực thể TRƯỜNG HỌC thuộc TỈNH và TRƯỜNG HỌC có MÔN HỌC. Một số thuộc tính của mỗi thực thể như sau: TRƯỜNG HỌC (tên trường, diện tích, loại trường, số giáo viên,...) MÔN HỌC (toán, lý, hóa,...) XÃ (Mã xã, tên xã, thuộc Tỉnh, diện tích).



Hình 4.4 CSDL đơn giản một cơ quan

Dạng sơ đồ CSDL không gian như sau:



Hình 4.5 CSDL không gian đơn giản.

Những sơ đồ trên là ví dụ sử dụng hai chuẩn ký hiệu thiết kế CSDL khái niệm:

- Thực thể: tên thực thể và danh sách thuộc tính
- Biểu đồ quan hệ thực thể: trình bày những thực thể, những thuộc tính của chúng và mối quan hệ giữa các thực thể.

Trên hình 4.5, cần hai điều cần chú ý:

- Biểu đồ tiêu chuẩn thực thể-mối quan hệ không biểu diễn hay miêu tả thực thể không gian (point, line, polygon) của dữ liệu.
- Những miêu tả các thuộc tính (biểu diễn bằng hình ellipse) có thể bất tiện vì chiều dài tên khác nhau và số lượng các thuộc tính sẽ trình bày.

Mô hình cơ bản thực thể-mối quan hệ

Khi xây dựng mô hình cơ bản thực thể-mối quan hệ chú ý ba nội dung (Chen 1976):

1. Những thực thể.
2. Quan hệ giữa những thực thể.
3. Những thuộc tính thực thể hay mối quan hệ.

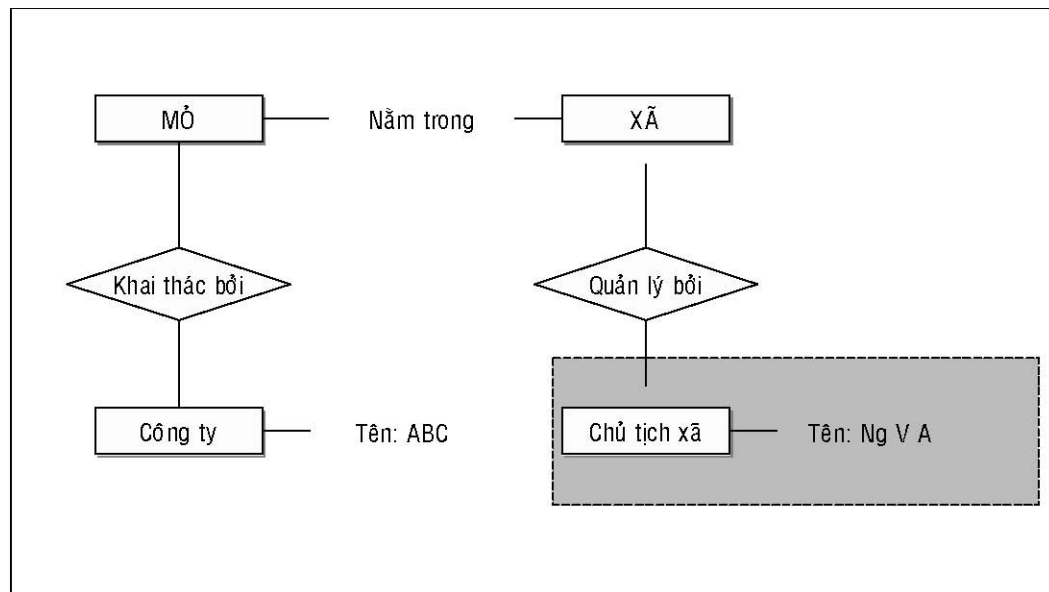
Mỗi thành phần có một biểu tượng và chúng tồn tại theo một tập hợp qui luật xây dựng biểu đồ (ví dụ, mô hình E-R) CSDL sử dụng 3 biểu tượng cơ bản. Hình 4.6.

Những thực thể biểu diễn như những hình chữ nhật. Những mối quan hệ biểu diễn như

hình thoi.

Những thuộc tính biểu diễn như ellipse. Những mối quan hệ thông thường trong mô hình E-R cơ bản là:

1. Những mối liên quan.
2. Tập hợp và nhóm mối quan hệ.
3. Mối quan hệ cha-con.
4. Những mảng thành phần mỗi đối tượng.



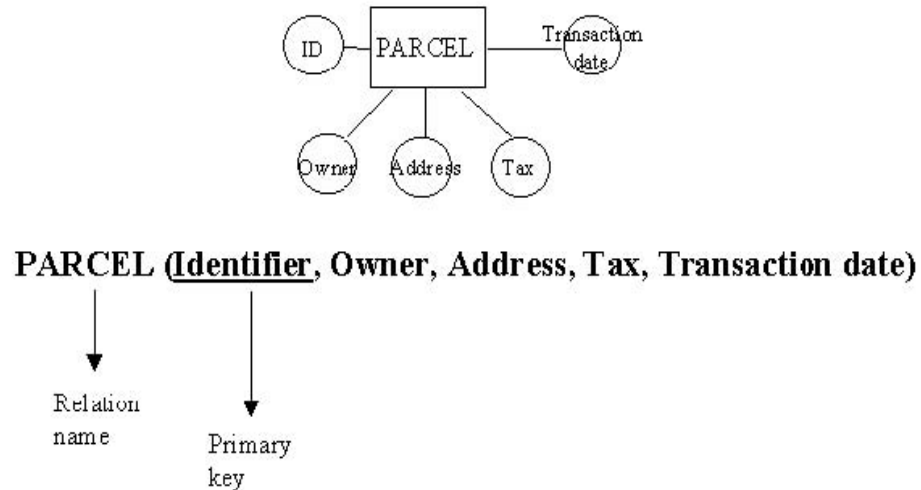
Hình 4.6 Ví dụ biểu đồ thực thể-mối quan hệ đơn giản

Có thể xây dựng nhiều dạng biểu đồ E-R cho dữ liệu. Trong xây dựng biểu đồ E-R (mô hình khái niệm) của CSDL cần xác định khi nào, cái gì đó biểu diễn là thực thể hay như là thuộc tính của thực thể khác.

Trong quá trình xây dựng biểu đồ E-R sẽ xuất hiện bất hợp lý trong định nghĩa những thực thể, những mối quan hệ và những thuộc tính. Kết quả biểu đồ E-R nên loại bỏ những bất hợp lý, khi có một biểu đồ rõ ràng có thể trực tiếp chuyển thành những giản đồ thiết kế logic và vật lý.

3.6.3. Mô hình hóa dữ liệu logic

Figure 22: Logical schema of entity PARCEL in Figure 21

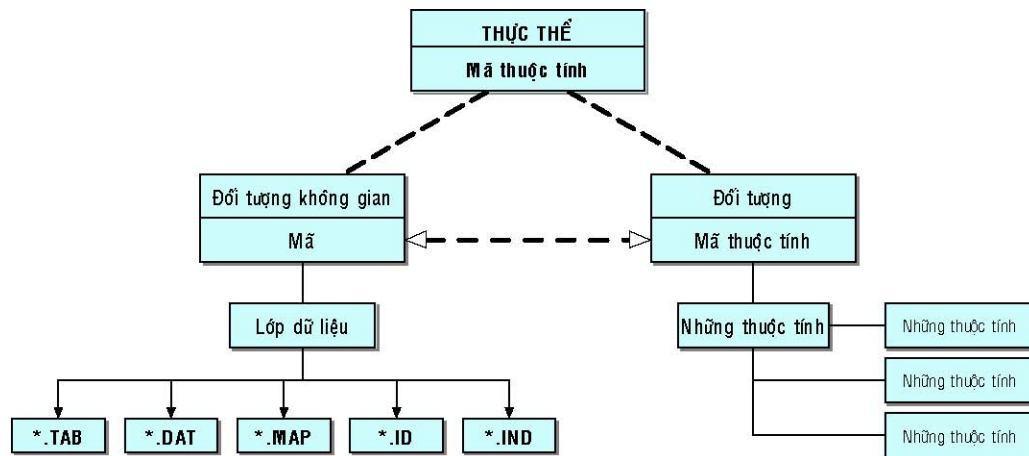


A.K. Yeung 1998-10-10 u5l-22

Trong mức thiết kế này, CSDL được đặc tả chi tiết, bao gồm các hạng mục tin, các mối quan hệ dữ liệu, đặt độ chính xác, các thủ tục đảm bảo sự toàn vẹn dữ liệu. Thiết kế logic đưa ra cấu trúc của các thành phần trong CSDL (còn gọi là cấu trúc CSDL). Mức thiết kế này là khởi điểm của các công việc tin học. Người thiết kế phải hiểu rõ tính năng của một hệ thống phần mềm quản trị CSDL. Trong mức này, người thiết kế đưa ra các phương án để lựa chọn các thành phần của CSDL HTTTĐL.

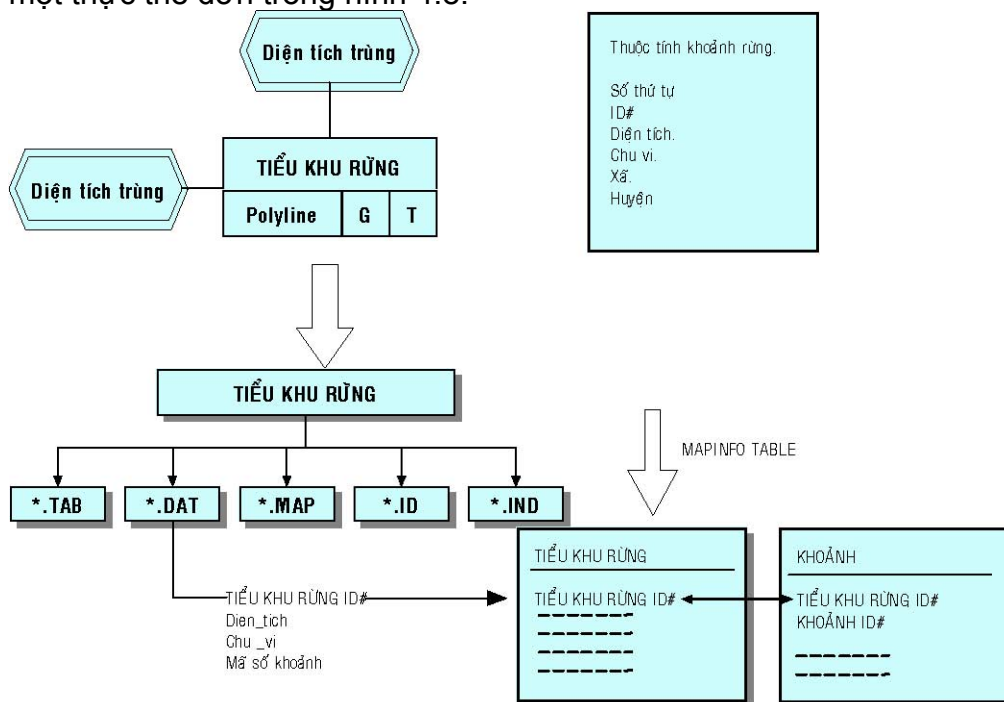
Đây là công việc nhằm chuyển đổi thiết kế khái niệm thành thiết kế (logic) CSDL HTTTĐL. Trên cơ sở các thành phần đã được liệt kê ra trong phần thiết kế mức quan niệm, trong phần thiết kế này sẽ tập trung vào thiết kế chi tiết dữ liệu.

Chuyển đổi từ biểu diễn thực thể biểu đồ E-R thành thiết kế logic CSDL cho thực thể đơn. Hình 4.8 minh họa tách riêng giữa thuộc tính thực thể và thông tin không gian trong MapInfo.



Hình 4.8 HTTTĐL biểu diễn đối tượng và đối tượng không gian

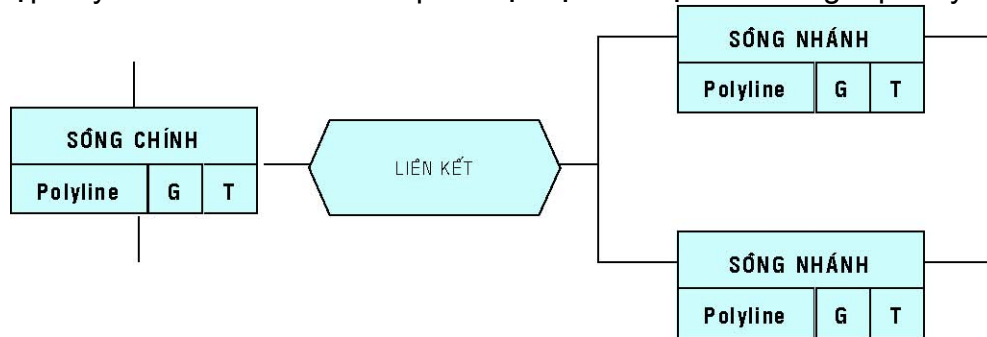
Sự chuyển đổi từ biểu diễn thực thể trong mô hình E-R thành thiết kế vật lý CSDL cho một thực thể đơn trong hình 4.8.



Hình 4.8 Ví dụ minh họa bản đồ hóa E-R và thuộc tính trong MapInfo.

Mỗi thực thể trong biểu đồ E-R sẽ biểu diễn trong một lớp dữ liệu (table trong phần mềm MapInfo). Thông thường không chỉ một thực thể đơn được chuyển thành một lớp dữ liệu mà nhiều thực thể trong một lớp dữ liệu và chúng có mối quan hệ với nhau. Do vậy có thể có một vài bảng dữ liệu được xây dựng để mô tả những mối quan

hệ phức tạp này. Hình 4.9 mô tả mối quan hệ một vài thực thể trong lớp thủy văn.



Hình 4.9 Ví dụ minh họa bản đồ hóa E-R và thuộc tính trong MapInfo.

Trong hình 4.10 đối tượng sông chính, sông phụ (nhánh) trong cùng một lớp dữ liệu như những đối tượng đường, mối quan hệ dữ liệu sẽ tạo ra do phần mềm MapInfo.

Mỗi thực thể thể hiện trong sơ đồ E-R được chuyển đổi tới lớp dữ liệu HTTTĐL trong đó gồm cả đối tượng không gian và mối liên hệ. Hơn nữa, mối liên hệ biểu diễn trong CSDL (hình lục giác đơn) cần chuyển đổi thành mã nguyên thủy và mã thứ sinh trong bảng dữ liệu của những thực thể biểu diễn. Trong hình 4.11 thực thể “XÃ” có “CHÚA” thực thể “KHOẢNH RỪNG”, trong bảng dữ liệu thuộc tính của mỗi lớp dữ liệu có chứa mã nguyên thủy cho mỗi đối tượng (XÃ ID#). Nhưng trong dữ liệu khoảnh rừng ngoài mã “KHOẢNH RỪNG ID#” cần có mã thứ sinh “XÃ ID#”.

Để hoàn thành thiết kế logic CSDL cần kiểm toán tất cả các thực thể và thuộc tính của chúng như một đối tượng không gian với tọa độ và topology với tất cả mối quan hệ chứa trong CSDL. Sao cho những thông tin này có thể sử dụng trong phần mềm HTTTĐL.

3.6.4. Mô hình vật lý (physical data modeling)

Figure 23: Example of a physical schema

Item Definition Table:

COLUMN	ITEM NAME	WIDTH	OUTPUT	TYPE	N.DEC	ALT NAME
1	PARCEL-ID	10	15	B	-	-
11	AREA	10	15	F	2	-
21	LJ-CLASS	3	5	C	-	-
24	OWNER-LN	15	20	C	-	-
29	OWNER-FN	15	20	C	-	-
44	ADDRESS-1	30	35	C	-	-
74	ADDRESS-2	30	35	C	-	-
104	TRNS-DATE	8	10	B	-	-
112	ASS-VALUE	10	15	I	-	-
122	TAX-RATE	5	10	F	3	-
127	TAX	8	10	F	2	-

Explanatory notes:

PARAMETER	DESCRIPTION
Item name	Any name to 16 characters
Width	No. of space used to store item values
Output	No. of spaces used to display item values
Type	Data item type:
C	Character
I	Integer
B	Binary
N	Number
D	Date mm/dd/yr
F	Floating point
N.DEC	No. of decimal points

A.K. Yeung 1998-10-10 u51-23

Mức này là sự triển khai và điều chỉnh thành quả của mức logic trên các phần cứng, phần mềm cụ thể của HTTTĐL. Đồng thời nó sẽ yêu cầu các cấu hình của phần cứng cũng như phần mềm thích hợp để đảm bảo sự thành công của việc triển khai này.

Phần mềm HTTTĐL tạo ra phần lớn thiết kế vật lý CSDL. Cấu trúc hay dạng của dữ liệu trong HTTTĐL, như ARC/INFO™, Intergraph™, System 9™, MapInfo™ vv... thực sự được xác định trong từng phần mềm. Có thể thấy rằng thiết kế vật lý những thực thể không gian hoàn toàn xác định bởi các phần mềm và người thiết kế HTTTĐL không cần làm bất kỳ gì thêm. Những thuộc tính những thực thể được khống chế trong hệ thống quản lý mối quan hệ liên kết trong HTTTĐL. Trong trường hợp này, người làm HTTTĐL cần thiết kế những bảng quan hệ cho thông tin thuộc tính.

4.3.1. Thiết lập CSDL

Công việc cài đặt này tuân thủ các đòi hỏi của phần mềm quản trị CSDL. Thành lập hệ quản trị CSDL không gian có khả năng đọc CSDL thuộc tính và kết nối nó với dữ liệu không gian.

Bên cạnh đó có thể sử dụng hệ quản trị CSDL phi không gian như ORACLE, DB2, SQL Server... luôn cung cấp dịch vụ để bất kỳ hệ thống nào cũng có thể truy nhập và đọc dữ liệu trong phạm vi quyền sử dụng của mình.

Quan hệ trong cơ sở dữ liệu

Trong CSDL HTTTĐL, cần định nghĩa ngắn gọn và súc tích, mô hình dữ liệu mô tả mối quan hệ giữa những thực thể.

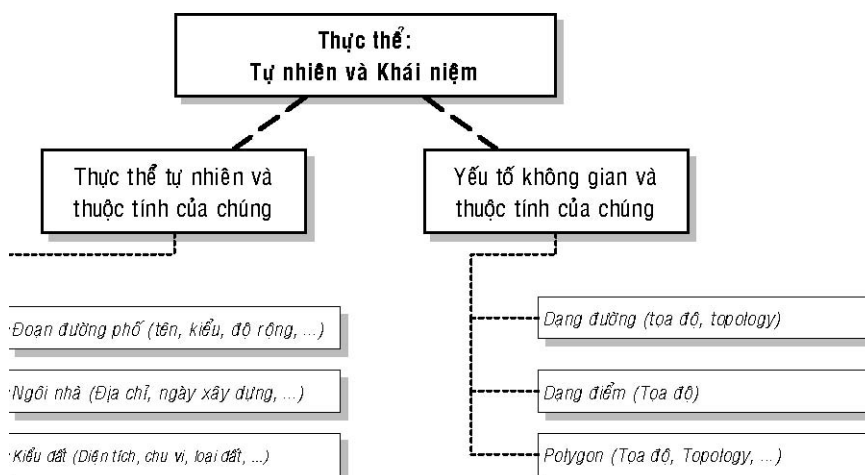
Ví dụ: Trong Lâm nghiệp

1. Mối quan hệ giữa loại rừng và Tỉnh là "thuộc vị trí"
2. Loại rừng-Thuộc vị trí-Tỉnh A Những mối quan hệ không trực tiếp:
3. Tỉnh - gồm rất nhiều -Loại rừng

Cần xác định mỗi thực thể chỉ có một mối quan hệ hay hơn một. Mối quan hệ có thể:

1. Một-tới-một
2. Một-với-nhiều
3. Nhiều-với-nhiều.
4. Có---> Tỉnh (một) <---gồm----->(nhiều) Loại rừng

Dữ liệu không gian không giống như những dữ liệu có qui luật trong CSDL máy tính, cần định nghĩa những thực thể không gian và những mối quan hệ của chúng. Xác định những thực thể dữ liệu địa lý gồm xác định tính chất tóm tắt thực thể và định nghĩa biểu diễn thực thể không gian (ví dụ: polygon đại diện cho một hồ nước). Sự định nghĩa những thực thể không gian là một trong những khác biệt HTTTĐL với hệ CSDL khác.

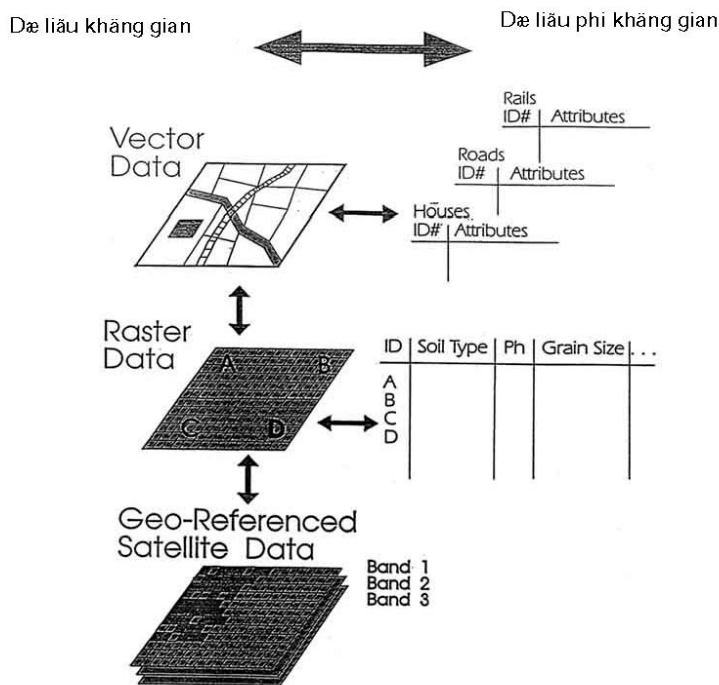


Hình 4.3 Ví dụ thực thể và thực thể không gian

3.7. Mô hình dữ liệu không gian

CSDLKG miêu tả các vật thể hay hiện tượng (gọi chung là đối tượng bản đồ) từ thực tế dưới dạng: -Vị trí của đối tượng theo một hệ tọa độ nào đó (vd: hệ lat/long theo độ, phút, giây hay theo hệ UTM bắc hoặc nam) -Các tính chất liên quan đến đối tượng tương ứng (vd: giá cả, màu sắc,...) -Mối liên hệ giữa các đối tượng xung quanh (mối liên hệ hình học- mô tả sự nối kết hay không nối kết,...)

-Thời gian xảy ra hiện tượng, hay thời điểm đo đạc CSDLKG bao gồm nhiều lớp DLKG giống như lớp bản đồ. Mỗi lớp DLKG chỉ thể hiện một dạng thông tin (lớp mưa, lớp sử dụng đất, lớp nguồn ô nhiễm không khí,...)



Ví dụ: Các lớp dữ liệu phục vụ nông nghiệp

LAYERS OF INFORMATION
Any information beneficial to land management decisions



Lập dữ liệu về đất



Lập dữ liệu hàm lượng N

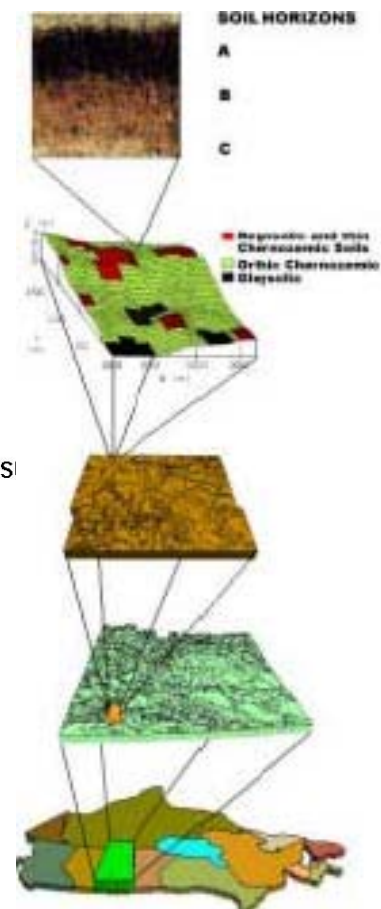


Lập dữ liệu ẩm độ



Lập dữ liệu về năng suất

Các cấp độ mô tả thông tin không gian:



Như chúng ta đã biết, bản đồ là một hình thức thể hiện dữ liệu không gian (spatial data) quen thuộc mà chúng ta thường gặp nhất. Bản đồ trình bày các nhóm điểm, đường và vùng, chúng được đặt ở vị trí địa lý (tọa độ) nào đó. Bản đồ thường được thể hiện ở dạng hai chiều. Các chú thích trên bản đồ cho biết những thông tin hay định nghĩa các điểm, đường và vùng mà nó thể hiện, những thông tin, định nghĩa đó mang tính phi không gian (non-spatial).

Bản đồ dùng để lưu trữ dữ liệu và cung cấp dữ liệu cho người sử dụng nó. Tuy nhiên ta không thể thể hiện nhiều thông tin trên một bản đồ cùng một lúc vì sẽ làm người dùng khó hiểu. Để giải quyết vấn đề này, trên cùng một khu vực cần thể hiện, người ta vẽ nhiều bản đồ, mỗi bản đồ thể hiện một số thông tin riêng, các bản đồ này còn gọi là bản đồ chuyên đề (thematic map) ví dụ như bản đồ đất đai, bản đồ khí hậu, v.v.... Tuy nhiên, tìm kiếm hay phân tích các dữ liệu không gian trên các bản đồ thuộc tính khác nhau thường tốn nhiều thời gian và bất tiện. Hơn nữa, việc vẽ các bản đồ bằng tay cũng tốn rất nhiều thời gian và công sức

Trong GIS, việc lưu trữ và thể hiện dữ liệu không gian (DLKG) được phân ra riêng biệt. Dữ liệu có thể được lưu trữ dưới mức độ chi tiết cao và sau đó được thể hiện ở mức độ kém chi tiết hơn và theo tỉ lệ thích hợp với mục tiêu sử dụng của người dùng. Ngoài ra, GIS còn cho phép người dùng thể hiện dữ liệu không gian dưới nhiều hình thức khác nhau như bản đồ chuyên đề cùng với biểu đồ, văn bản mô tả, v.v.... Mỗi cách được điều chỉnh tùy theo mục đích sử dụng.

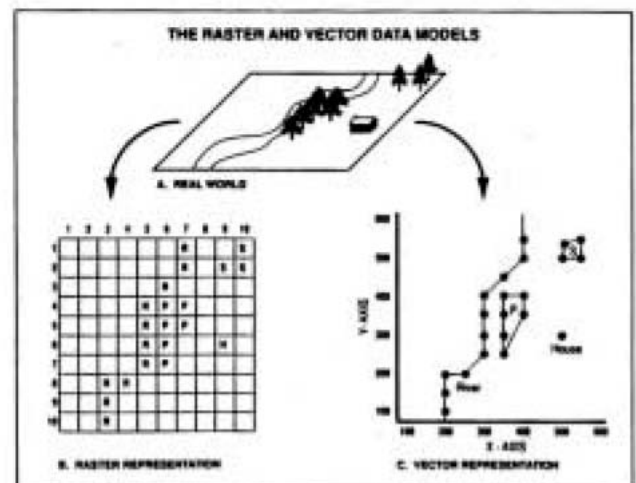
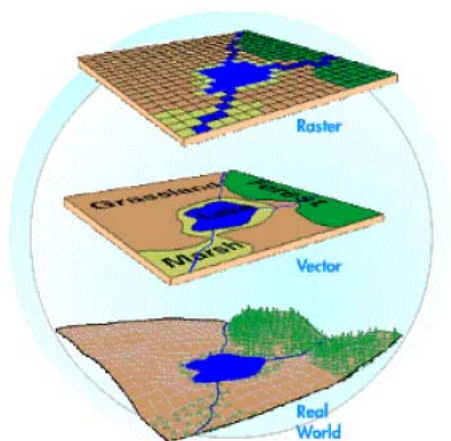


Figure 6.10 Comparison of the Raster and Vector Models. The landscape in A is shown in a raster representation (B) and in a vector representation (C). The pine forest stand (P) and spruce forest stand (S) are area features. The river (R) is a line feature, and the house (H) is a point feature.

Hình III.1 Mô hình dữ liệu không gian

Trong GIS, DLKG được thể hiện dưới dạng điểm, đường và vùng tương tự như bản đồ thông thường. Tuy nhiên, để dễ quản lý bằng máy tính, DLKG

được tổ chức lưu trữ khác với bản đồ. Thông tin về thực thể không gian trong GIS được mô tả bằng 4 thành phần:

- Vị trí địa lý của đối tượng được mô tả
- Mối liên hệ của đối tượng đó trong không gian
- Tính chất của đối tượng (phi không gian)
- Thời gian

Vị trí địa lý; “ở đâu?” và vị trí của đối tượng trong không gian được thể hiện một cách thống nhất theo một hệ thống tọa độ địa lý nào đó . Trong GIS, các DLKG của cùng 1 cơ sở dữ liệu (CSLD) phải cùng một hệ thống tọa độ. DLKG có thể được lưu trữ ở nhiều tỉ lệ (mức độ chính xác) khác nhau.

Thuộc tính: Tính chất thứ hai của dữ liệu không gian là thuộc tính “nó là cái gì?”. Trong GIS, các thuộc tính được lưu trữ và thể hiện dưới dạng bảng biểu. Mỗi trường thể hiện một thuộc tính của đối tượng. Ví dụ để thể hiện tính chất của các con kênh, ta mô tả bằng tên con kênh, năm đào, cấp kênh, năng lực tưới, năng lực tiêu, lưu lượng trung bình,...

Mối liên hệ không gian: Các đối tượng địa lý luôn có mối liên hệ không gian với nhau. Các liên hệ này có thể là: nằm trong, bên cạnh, cắt nhau, ở trên, ở dưới,..., ví dụ như con đường nằm cạnh bờ kênh, khu nông trường nằm trong huyện A, con đường B cắt ngang con đường C , ...

Thời gian: Một số sự vật, hiện tượng có sự thay đổi theo thời gian như sử dụng đất nông nghiệp, thời tiết,... do đó khi mô tả các sự vật hiện tượng này người ta luôn thể hiện thời điểm thu thập (đo đạc) dữ liệu.

3.7.1. Mô hình dữ liệu raster

Đây là hình thức đơn giản nhất để thể hiện dữ liệu không gian, mô hình raster bao gồm một hệ thống ô vuông hoặc ô chữ nhật được gọi là pixel (hay một phần tử của ảnh). Vị trí của mỗi pixel được xác định bởi số hàng và số cột. Giá trị được gán vào pixel tượng trưng cho một thuộc tính mà nó thể hiện. Ví dụ một căn nhà được thể hiện bằng 1 pixel có giá trị là H, con sông được thể hiện bằng nhiều pixel có cùng giá trị là R, tương tự khu rừng cũng được thể hiện bằng một nhóm pixel có cùng giá trị là D (cây dừa) hoặc S (cây soài).

Kích thước của pixel càng nhỏ thì hình ảnh nó thể hiện càng sắc nét, thông số thể hiện độ sắc nét gọi là độ tương phản (resolution). Ảnh có độ tương phản cao, thì độ sắc nét càng cao, kích thước pixel nhỏ. Tuy nhiên, hai ảnh raster có cùng kích thước, nếu ảnh nào có độ tương phản cao thì file dữ liệu chứa nó sẽ lớn hơn. Ví dụ nếu 1 pixel thể hiện một diện tích là 250m x 250m mặt đất trên thực tế, thì để thể hiện một khoảng cách 1km ta cần 4 pixel, để thể hiện một diện tích 1km x 1km ta cần 16 pixel. Khi ta giảm kích thước

pixel xuống còn 100m x 100m, để thể hiện một khoảng cách 1km ta cần 10 pixel, để thể hiện một diện tích 1km x 1km ta cần 100 pixel. Vì kích thước của file dữ liệu liên quan tới số lượng pixel nên ta thấy rằng kích thước của file tăng lên đáng kể khi ta tăng độ tương phản của ảnh raster.

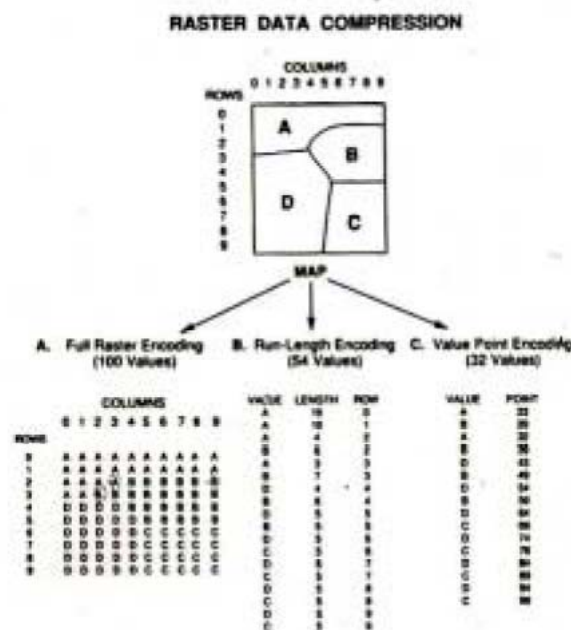


Một ảnh raster thông thường bao gồm hàng triệu pixel. Tuy nhiên, nhiều pixel gần nhau sẽ có cùng giá trị. Người ta dùng nhiều phương pháp nén (data compression) khác nhau để giảm kích thước file ảnh raster như là phương pháp Run-Length Encoding, phương pháp Value Point Encoding và phương pháp Quadrees. (hình III.2)

a. Phương pháp Run-Length Encoding: Trong phương pháp này, file dữ liệu sẽ lưu trữ thông tin về các run, mỗi run là một nhóm các pixel có cùng giá trị nằm liên tục nhau trên cùng một dòng, run được định nghĩa bởi giá trị (value), chiều dài (length) và dòng (row).

b. Phương pháp Value Point Encoding:

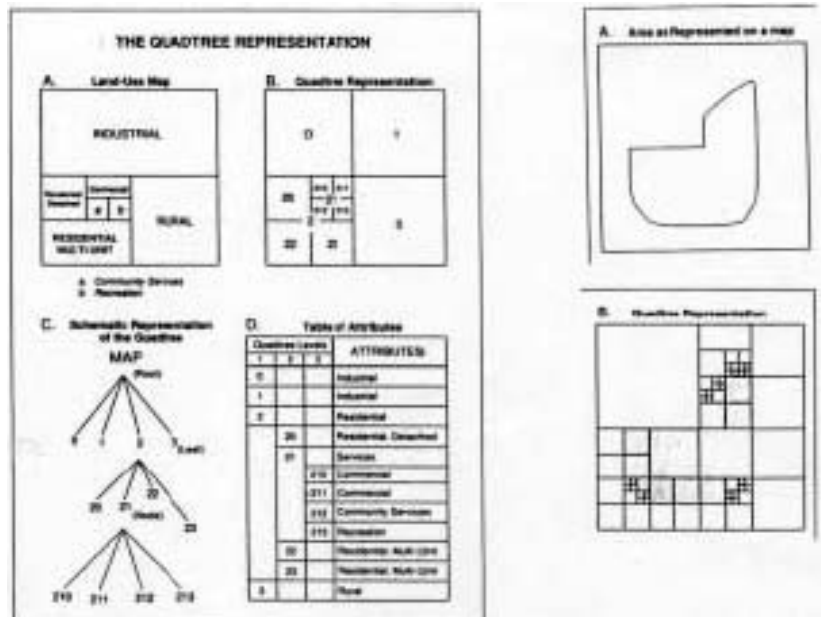
Trong phương pháp này, dữ liệu được lưu trữ là dòng và cột của điểm cuối cùng của một dãy các pixel có cùng giá trị.
dãy các pixel có cùng giá trị.



Hình III.2 Nén dữ liệu raster

c. Phương pháp Quadtrees: Ở phương pháp này, người ta chia ảnh thành các tiểu vùng. Mỗi một tiểu vùng phải có cùng một giá trị. Việc phân chia tiểu vùng được tiến hành như sau: - Chia ảnh ra làm bốn phần bằng nhau - Nếu phần ảnh nào nhiều giá trị khác nhau thì tiếp tục chia tiếp phần ảnh đó làm bốn.

- Tiếp tục xét các phần tư nhỏ và chia chúng ra làm tư nếu còn có sự khác biệt về giá trị trong nó.



Hình III.3. Phương pháp Quadtree

Đặc điểm của mô hình dữ liệu raster

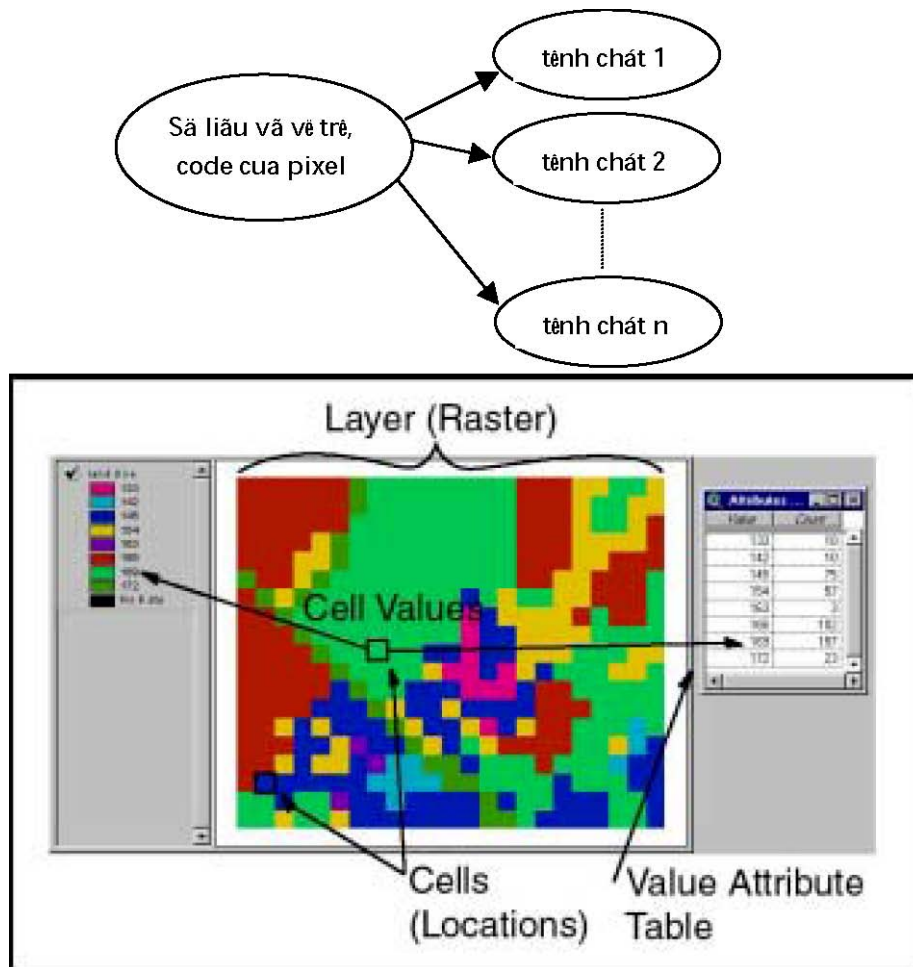
- Mỗi pixel là một đối tượng, có vị trí theo hàng, cột tương ứng trên ảnh, giá trị của pixel cho biết pixel đó thuộc loại đối tượng nào, tính chất của đối tượng đó được lưu trữ ở một cơ sở dữ liệu thuộc tính tương ứng.

- CSDLKG Raster có thể chứa hàng ngàn lớp DLKG (layer).

- Kiểu giá trị của pixel trong mỗi layer tùy theo việc mã hóa của người sử dụng, có thể là số nguyên, số thực hay ký tự alphabet. Thường giá trị số nguyên thường được dùng làm mã số để liên hệ với bảng DL khác hay làm chú giải để thể hiện bản đồ.

- Để thể hiện một bề mặt liên tục, người ta sử dụng mô hình raster, các bề mặt liên tục này thường thể hiện bề mặt địa hình, mưa, áp suất không khí, nhiệt độ, mật độ dân số,...

Như vật đối với CSDLKG raster các thông tin được tổ chức như hình dưới đây:

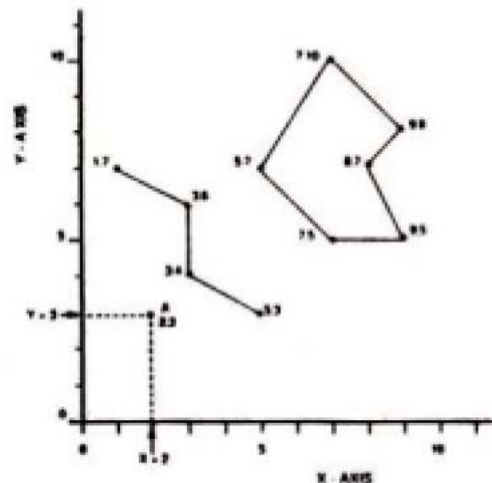
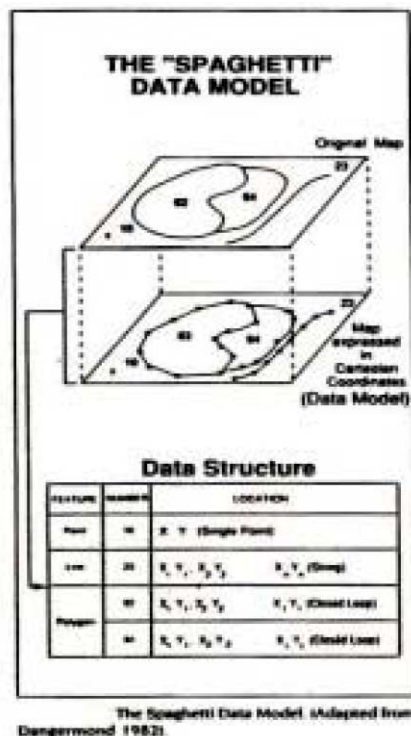


3.7.2. Mô hình dữ liệu vector

Mô hình dữ liệu vector thể hiện vị trí chính xác của vật thể hay hiện tượng trong không gian. Trong mô hình dữ liệu vector, người ta giả sử rằng hệ thống tọa độ là chính xác. Thực tế, mức độ chính xác bị giới hạn bởi số chữ số dùng để thể hiện một giá trị trong máy tính, tuy nhiên nó chính xác hơn rất nhiều so với mô hình dữ liệu raster.

Vật thể trên trái đất được thể hiện trên bản đồ dựa trên hệ tọa độ hai chiều x,y (Cartesian coordinate system), trên bản đồ vật thể có thể được thể hiện như là các điểm (point), đường (line) hay miền (area). Mô hình dữ liệu vector cũng tương tự như vậy, một vật thể dạng điểm (point feature) được chứa dưới dạng cặp tọa độ x,y; một vật thể dạng đường (line feature) được chứa dưới dạng một chuỗi các cặp tọa độ x,y; một vật thể dạng vùng (area feature) được chứa dưới dạng một chuỗi cặp tọa độ x,y với cặp đầu tọa độ bằng với cặp tọa độ cuối, hay còn gọi là đa giác (polygon). Trong hình III.4, các vật thể được số hóa (digitize) bằng các cặp tọa độ x,y. Vị trí của điểm A được thể hiện bởi tọa độ 2,3 và đường được thể hiện bởi chuỗi tọa độ 1,7; 3,6; 3,4; 5,3. Đa giác được thể hiện bởi một chuỗi tọa độ khác trong đó tọa

độ đầu và cuối bằng nhau: 7,10; 9,8; 8,7; 9,5; 7,5; 5,7; 7,10. Trong thí dụ này đơn vị của các tọa độ là tùy ý. Tuy nhiên trong GIS, vị trí thường được lưu trữ theo một hệ qui chiếu chuẩn như là hệ thống UTM, hệ thống quốc gia hay hệ kinh tuyến, vĩ tuyến.



Hình III.4 Thể hiện vật thể dạng điểm, đường, vùng theo chuỗi các cặp tọa độ

Trong mô hình dữ liệu vector, tùy theo cách lưu trữ dữ liệu, người ta chia ra thành các mô hình: Spaghetti Data Model, Topological Model, Triangulated Irregular Network (TIN), tạm dịch là mô hình dữ liệu kiểu mì ống, mô hình dữ liệu hình học và mô hình lưới tam giác bất qui tắc.

a. Spaghetti Data Model (SDM):

Trong SDM, tọa độ của các vật thể trên bản đồ được chuyển đổi và ghi nhận vào file dữ liệu theo từng dòng danh sách các cặp tọa độ. Như vậy các cặp tọa độ của cạnh chung của hai đa giác kề nhau phải được lập lại hai lần, mỗi lần cho một đa giác.

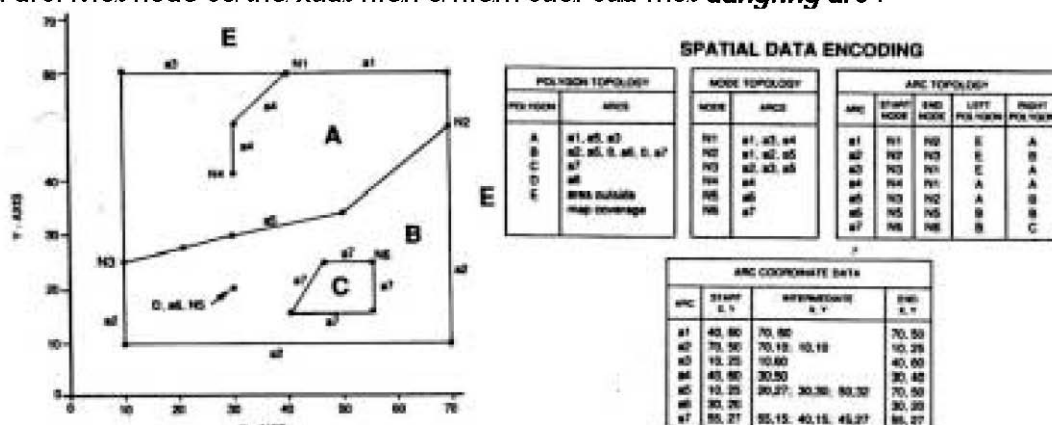
Cấu trúc của dạng mô hình này rất dễ hiểu, tuy nhiên mối liên hệ của các vật thể trong ảnh không được ghi nhận.

Hình III.5 Mô hình dữ liệu vector kiểu spaghetti

b. Topological Model: Đây là dạng mô hình dữ liệu được sử dụng rộng rãi nhất để lưu trữ dữ liệu không gian. Hình học là phương pháp toán học dùng cho việc xác định mối liên hệ giữa các vật thể trong không gian. Người ta

còn gọi mô hình này là mô hình Arc-node. Arc là một chuỗi các đoạn thẳng được bắt đầu và kết thúc bằng nút (node). Một node là một giao điểm của hai hay nhiều Arc. Một node có thể xuất hiện ở điểm cuối của một dangling Arc.

nhiều arc. Một node có thể xuất hiện ở điểm cuối của một *dangling arc*.



Hình III.6 Mô hình dữ liệu vector kiểu topology

Trong hình III.6 dữ liệu được lưu trữ trong 4 bảng. Bảng Polygon Topology liệt kê danh sách các arc tạo thành các đa giác (polygon), chú ý là nếu trong một polygon có các arc nằm trọn trong nó thì trong danh sách các arc tạo thành polygon đó trước và sau arc đó phải ghi số 0 như trường hợp polygon B. Vùng bên ngoài bản đồ cũng được xem như là một polygon. Bảng Node Topology liệt kê các arc đi qua cùng một node hay các node của dangling arc, Chú ý rằng điểm đóng của một đa giác được coi là một node và một điểm vừa được coi là một node và vừa được coi là một arc (trường hợp N5 và N6). Bảng Arc Topology liệt kê node đầu, node cuối, polygon bên trái và polygon bên phải của các Arc trong bản đồ. Nhìn vào các bảng này ta có thể dễ dàng biết được mối liên hệ của một vật thể đối với các vật thể xung quanh. Ví dụ, ta xét điểm N5, trong bảng Arc Topology ta thấy left polygon và right polygon đều là B, như vậy điểm N5 nằm trong polygon B. Hay xét arc a5, ta thấy left polygon là A, right polygon là B, như vậy a5 nằm giữa hai polygon A và B. Bảng Arc Coordinate Data liệt kê danh sách tọa độ các node tạo nên arc, các tọa độ này được phân ra tọa độ bắt đầu, tọa độ trung gian và tọa độ kết thúc.

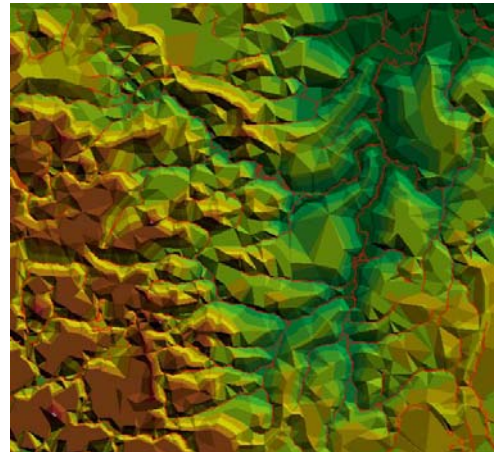
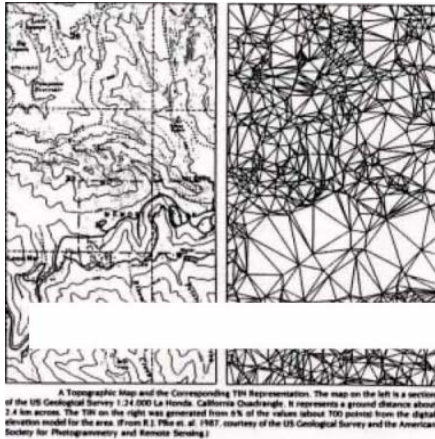
Topology model rất thích hợp cho các thao tác phân tích không gian (spatial analysis) vì đa số các bài toán phân tích không gian, người ta không dùng đến dữ liệu về tọa độ, mà chỉ cần dùng đến dữ liệu hình học.

c. Triangular Irregular Network (TIN):

TIN thường được dùng để thể hiện dữ liệu về địa hình (terrain data). TIN thể hiện bề mặt địa hình như là tập hợp các mặt tam giác liên kết với nhau (hình III.7). Mỗi đỉnh của tam giác được thể hiện bằng tọa độ địa lý x, y và z,

thường z thường được gán bằng với giá trị cao độ của địa hình.

Mỗi mặt của tam giác được gán cho một chữ cái và ba đỉnh của nó được gán bằng chữ số. Bảng Nodes (nút) thể hiện danh sách đỉnh của từng tam giác, bảng Edges (cạnh) thể hiện danh sách các tam giác nằm xung quang của từng tam giác, bảng X-Y coordinate thể hiện tọa độ của các đỉnh, bảng z coordinate thể hiện giá trị z của các đỉnh đó. TIN rất thích hợp trong việc tính toán các thông số của địa hình như độ dốc, hướng dốc.



Đặc điểm mô hình dữ liệu vector

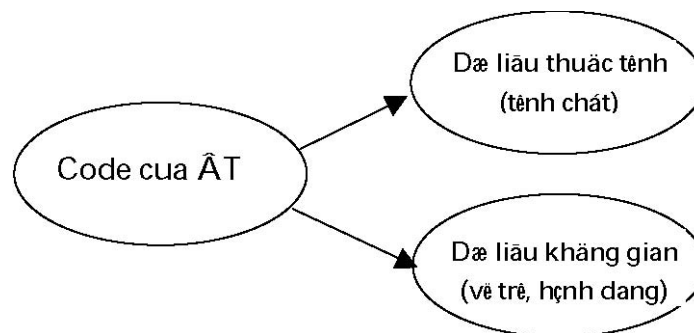
Để thể hiện các kiểu đối tượng (vd: như đường sá, sông ngòi, cao độ, thực vật,...) dưới dạng số (digital) trong CSDLKG, ta phải lựa chọn kiểu đối tượng không gian thích hợp: -Dạng 0 chiều (0D): Một đối tượng có vị trí trong không gian nhưng không có chiều dài. Dạng điểm -Dạng 1 chiều (1D): Một đối tượng có một chiều dài kết hợp với 1 hoặc nhiều đối tượng 0D. Dạng tuyến -Dạng 2 chiều (2D): Một đối tượng có một chiều dài và chiều rộng được bao quanh bởi ít nhất ba đối tượng 1D. Dạng diện tích-vùng -Dạng 3 chiều (3D): Một đối tượng có một chiều dài, chiều rộng và chiều cao/chiều sâu được bao quanh bằng ít nhất 2 đối tượng 2D. Dạng thể tích

-Dạng 4 chiều (4D): bao gồm 1 đối tượng 3D và yếu tố thời gian Đối với CSDLKG vector, người ta cũng lưu trữ theo từng lớp DLKG, mỗi lớp chỉ thể một loại thông tin và một dạng đối tượng (0D, 1D, 2D, 3D, 4D), Ví dụ trong một loại thông tin và một dạng đối tượng (0D, Ví dụ trong CSDLKG về qui hoạch nông nghiệp, người ta có các lớp DLKG về đất (dạng vùng, mỗi vùng thể hiện một loại đất), lớp DLKG về hệ thống kênh (dạng tuyến, thể hiện các con kênh), lớp công trình trên kênh (loại điểm, thể hiện cống, trạm bơm, đập tràn,...), lớp DLKG về trạm mưa (dạng điểm, thể hiện các trạm đo mưa),...

Như chúng ta đã biết trong mô hình vector, người ta có thể dùng nhiều phương pháp để lưu trữ vị trí của đối tượng, còn để liên hệ các đối tượng đó với tính chất của chúng trong

CSDL, người ta làm như sau: -Đối với dạng điểm: mỗi điểm được gán 1 mã số, mã số này sẽ tương ứng với mã số được ghi trong các bảng CSDL thuộc tính thể hiện các thuộc tính của các điểm tương ứng. -Đối với dạng tuyến: một đối tượng dạng tuyến được vẽ bằng nhiều đoạn nối (link) của các nút (node). Tương tự như đối tượng dạng điểm, mỗi đối tượng dạng tuyến cũng được gán một giá trị mã số (code), code này được dùng để liên hệ với các bảng DL thuộc tính mô tả đối tượng đó. Ngoài ra mỗi nút cũng có các thông tin thuộc tính như nút vượt, nút cắt, nút val (đường ống), nút đèn giao thông (đường). Các link cũng có các thuộc tính như hướng, kích thước (đường kính ống, chiều rộng đường), số làn xe, thời gian di chuyển từ đầu link đến cuối link, hoặc điện thế của đường tải điện, đường hầm, đường vượt,... -Đối với dạng vùng: thường là các vùng phân chia theo một tính chất nào đó hay biên giới hành chính, hồ, rừng núi, khu vực điều tra dân số,... biên của chúng được xác định bởi thuộc tính của chính đối tượng đó. Mỗi vùng cũng được gán code và liên hệ với bảng DL thuộc tính.

Như vật đối với CSDLKG vector các thông tin được tổ chức như hình dưới đây:



So sánh giữa mô hình dữ liệu Raster và Vector

Mô hình Raster	Mô hình Vector
Ưu điểm 1. Đơn giản 2. Thao tác chồng lấp (overlay) dễ dàng 3. Thích hợp cho việc thể hiện dữ liệu phức tạp (đa dạng) 4. Thích hợp cho việc nâng cấp, xử lý ảnh	Ưu điểm 1. Cấu trúc dữ liệu nén nhiều hơn so với mô hình raster 2. Thể hiện liên hệ hình học do đó thích hợp cho các phân tích về hình học hay phân tích về mạng lưới 3. Thích hợp cho việc số hóa các bản đồ được vẽ bằng tay
Nhược điểm 1 Khả năng nén kém 2. Không thể hiện rõ liên hệ hình học 3. Thể hiện bản đồ không rõ nét nếu ô tương phản thấp, nhưng nếu dùng độ tương phản cao sẽ làm tăng kích thước file ảnh.	Nhược điểm 1. Phức tạp 2. Thao tác chồng lấp phức tạp 3. Không thích hợp cho việc thể hiện dữ liệu phức tạp (đa dạng) 4. Không thích hợp cho việc nâng cấp, xử lý ảnh

Bài 4

BẢN ĐỒ VÀ CƠ SỞ XÂY DỰNG BẢN ĐỒ

4.1. Trái đất - quả cầu địa lý

4.1.1. Hình dạng - kích thước trái đất:

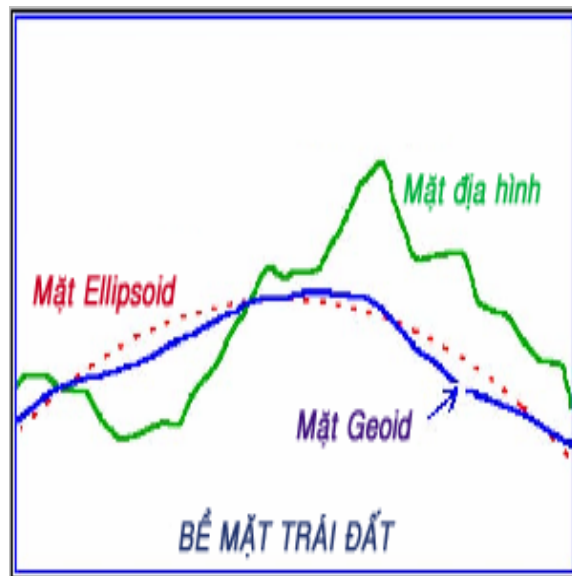
Bề mặt tự nhiên trái đất rất phức tạp về mặt hình học và không thể biểu thị nó bởi một qui luật xác định, hình dạng trái đất được hình thành và bị chi phối bởi hai lực là lực hấp dẫn và lực ly tâm tạo nên hình dạng ellipsoid của trái đất (hình 1)

Trong trắc địa người ta dùng mặt geoid, bề mặt này được tạo bởi mặt nước biển trung bình yên tĩnh kéo dài qua các lục địa và hải đảo tạo thành một mặt cong khép kín, có đặc điểm là ở bất kỳ điểm nào nằm trên pháp tuyến cũng trùng

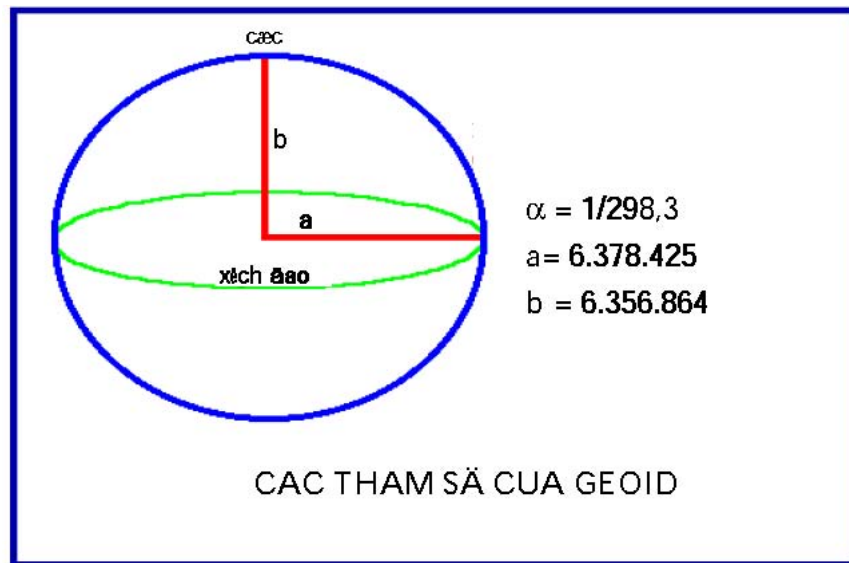
với phương dây dọi. Ngoài ra, do tác dụng của trọng lực, sự phân bố không đồng đều của vật chất có tỉ trọng khác nhau trong lớp vỏ của trái đất làm cho bề mặt geoid bị biến đổi phức tạp về mặt hình học.

Như vậy, bề mặt hoàn chỉnh của trái đất không phải là bề mặt đúng toán học, mà chỉ là mặt sẵn có của chính trái đất. Trong khoa học trắc địa bản đồ, để tiện lợi cho các bài toán đo đạc, người ta lấy mặt ellipsoid tròn xoay có hình dạng và kích thước gần giống mặt geoid làm bề mặt toán học thay cho mặt thật gọi là ellipsoid trái đất. Ellipsoid có khối lượng bằng khối lượng geoid, tâm của nó trùng với trọng tâm trái đất, mặt phẳng xích đạo trùng với mặt phẳng xích đạo trái đất. Kích thước và hình dạng của ellipsoid trái đất được xác định bởi giá trị các phần tử của nó (hình 2):

Độ dẹt (e) = $(BK \text{ trục lớn } a - BK \text{ trục nhỏ } b) / BK \text{ trục lớn } a$



Hình 1. Hình dạng ellipsoid của trái đất



Hình 2. Các tham số của geoid

Nhiều công trình nghiên cứu khoa học nhằm xác định α , a , b của ellipsoid trái đất nhưng kết quả không thống nhất, ở nước ta các trị số của F.N Kraxovski năm 1946 được dùng làm trị số chính thức đo đạc: $\alpha = 1/298,3$; $a = 6.378.425$; $b = 6.356.864$ Các số liệu kích thước trái đất được tính như sau:

Bán kính trung bình trái đất:	6.371,166 km
Độ dài vòng kinh tuyến:	40.008,5 km
Chu vi xích đạo:	40.075,5 km
Diện tích bề mặt trái đất:	510,2 triệu km ²
Thể tích trái đất:	1083 x 10 ²¹ km ³
Tỉ trọng trung bình:	5,52 g/cm ³
Trọng lượng của trái đất:	5,977 x 10 ²¹ tấn

Vì độ dẹt của ellipsoid trái đất nhỏ, nên trong trường hợp đo đạc khu vực nhỏ, người ta có thể coi trái đất như một khối cầu có bán kính gần trùng với trục quay của trái đất, R , theo F.N Kraxovski là 6371,116 km.

4.1.2. Các qui ước về điểm và đường cơ bản để xác định vị trí các đối tượng địa lý trên bề mặt trái đất

1. a. Cực trái đất: Giao điểm giữa bán kính trục nhỏ (trục trái đất) và mặt ellipsoid trái đất gọi là cực. Trái đất có hai cực là cực Bắc (P) và cực Nam (P').
2. b. Các kinh tuyến: Các mặt phẳng chứa trục trái đất và hai cực là mặt phẳng kinh tuyến. Giao tuyến giữa mặt phẳng kinh tuyến và mặt ellipsoid trái đất là kinh tuyến.
3. c. Các vĩ tuyến: Các mặt phẳng thẳng góc với trục trái đất được gọi là mặt phẳng vĩ tuyến. Mặt phẳng đi qua tâm trái đất chia trái đất thành hai bán cầu: bán cầu bắc và bán cầu nam, là mặt phẳng xích đạo. Mặt phẳng xích đạo cắt mặt

ellipsoid trái đất thành một vòng tròn lớn gọi là xích đạo. Các vòng tròn tạo nên bởi các mặt phẳng song song với mặt phẳng xích đạo gọi là vĩ tuyến.

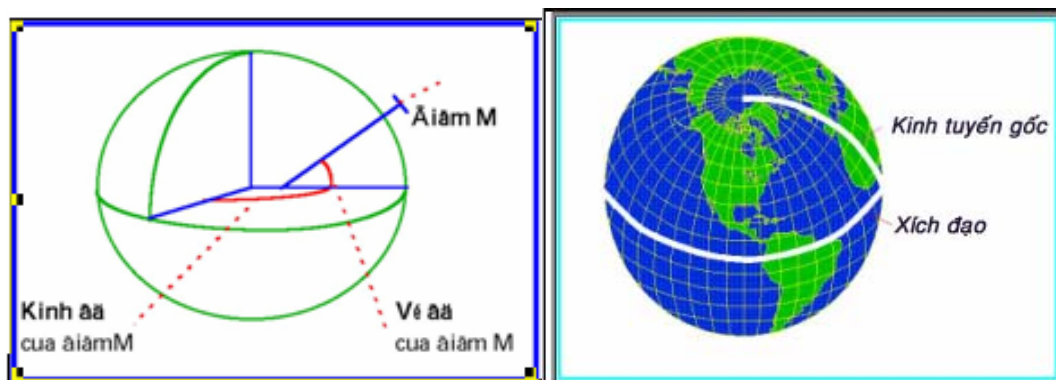
4.1.3. Tọa độ địa lý

Tất cả các điểm trên bề mặt ellipsoid trái đất đều được xác định vị trí bằng phương pháp tọa độ. Có nhiều hệ thống tọa độ, trong đó có hệ tọa độ địa lý.

Cơ sở để xác định tọa độ địa lý là kinh tuyến và vĩ tuyến. Tọa độ địa lý một điểm được xác định bằng vĩ độ và kinh độ của điểm đó. (hình.3) -Vĩ độ địa lý: của một điểm là góc hợp bởi đường dây dọi đi qua điểm đó và mặt phẳng xích đạo. Những vĩ độ được tính từ xích đạo (0°) về phía bắc đến 90° gọi là vĩ độ Bắc (N), và về phía nam đến 90° là vĩ độ Nam (S) .

- Kinh độ địa lý: của một điểm là góc nhị diện hợp bởi mặt phẳng kinh tuyến gốc và mặt phẳng kinh tuyến đi qua điểm đó. Để tiện xác định vị trí các điểm trên địa cầu, người ta qui định trên địa cầu có 360 đường kinh tuyến các đều nhau. Khoảng cách giữa hai đường kinh tuyến là một cung tròn có góc ở tâm là 1° . Hội nghị thiên văn Quốc Tế họp ở Washington (1884) đã lấy đường kinh tuyến đi qua đài thiên văn Greenwich gần London, thủ đô Anh, làm kinh tuyến gốc (0°) thống nhất cho toàn thế giới.

Các kinh độ được tính từ kinh tuyến gốc về phía đông đến 180° là những kinh độ Đông (E), và về phía tây là những kinh độ tây (W). Thành phố Hà Nội có tọa độ là $105^\circ 52' E$ và $21^\circ 02' N$



Hình .3. Tọa độ của một điểm

4.2. Cơ sở toán học của bản đồ

Cơ sở toán học của bản đồ gồm có: cơ sở trắc địa (như hệ thống lưới tọa độ mặt bằng và độ cao chuẩn của nhà nước), lưới chiếu, tỉ lệ bản đồ, khung bản đồ, bố cục bản đồ, danh pháp và chia mảnh, ...

4.2.1. Tỷ lệ

Tỷ lệ (map scale) là tỉ số của khoảng cách trên bản đồ và khoảng cách thực tế mà nó thể hiện, thí dụ tỉ lệ bản đồ 1:25.000 thì là 1 cm trên bản đồ bằng 250 m ngoài thực địa. Các yếu tố được chú ý

khi chọn tỉ lệ bản đồ là:

- Mục tiêu sử dụng của bản đồ
- Yêu cầu của người sử dụng bản đồ
- Thành phần của bản đồ
- Kích thước của vùng được thể hiện
- Kích thước lớn nhất của bản đồ (xét yếu tố dễ sử dụng)
- Độ chính xác yêu cầu

Verbal Scale

One inch represents two thousand feet

Linear Scale



Fractional Scale

1 : 24,000

Một vài hạn chế trong việc lựa chọn tỉ lệ bản đồ cần chú ý là: -Tỉ lệ quá lớn: yêu cầu nhiều thông tin chi tiết cho thành phần chính của bản đồ dẫn đến tăng công việc vẽ bản đồ, tăng thời gian và giá thành sản phẩm -Tỉ lệ quá nhỏ: bản đồ khó đọc khi có nhiều thông tin cần trình bày, có thể làm người sử dụng bản đồ đọc sai thông tin.

4.2.2. Phép chiếu bản đồ

Hệ qui chiếu (map projection) có thể được định nghĩa như là sự sắp đặt một cách có hệ thống các kinh tuyến và vĩ tuyến, miêu tả bề mặt cong của hình cầu theo mặt phẳng.

Đối với một bề mặt có diện tích 30km x 30km, thì ta có thể xem bề mặt trái đất là phẳng(độ cong quả đất $< 1/1.000.000$). Do đó nếu một bản đồ phải thể hiện một diện tích nhỏ hơn diện tích này thì chúng ta có thể vẽ các thực thể trực tiếp lên mặt phẳng dựa trên số liệu đo đạc. Khi chúng ta phải thể hiện một vùng lớn hơn thì lúc đó chúng ta phải chọn hệ qui chiếu hợp lý. Sự lựa chọn hệ qui chiếu được dựa trên các yếu tố sau:

- Mục tiêu của bản đồ
- Yêu cầu của người sử dụng bản đồ
- Vị trí của vùng được thể hiện
- Hình dạng và kích thước của khu vực được thể hiện

Sự lựa chọn hệ qui chiếu cho một quốc gia phụ thuộc vào điều kiện mà nó đặt ra cho bản đồ, đó là hình dạng của quốc gia đó trên bản đồ phải giống thực tế, cùng diện tích bề mặt (theo tỉ lệ), các góc phải bằng nhau, khoảng cách cũng bằng nhau và khoảng cách ngắn nhất giữa hai điểm phải là đoạn thẳng. Tuy nhiên tất cả các điều kiện trên không thể cùng thỏa mãn một lúc, do đó chúng ta phải chấp nhận với một trong số các điều kiện đó

Hệ qui chiếu có thể được phân loại dựa trên tính chất của nó: -Hệ qui chiếu đồng góc(Conformal projections): góc đo được trên mặt đất bằng với góc

trên bản đồ. -Hệ qui chiếu đồng diện tích (Equivalent projections): Diện tích bề mặt trên mặt đất bằng diện tích trên bản đồ. -Hệ qui chiếu đồng khoảng cách (Equidistance projections): Khoảng cách từ tâm hệ qui chiếu đến các điểm khác trên bản đồ là thực. -Các hệ qui chiếu trung gian khác (không thuộc các hệ qui chiếu trên nhưng cho phép thể hiện một khu vực)

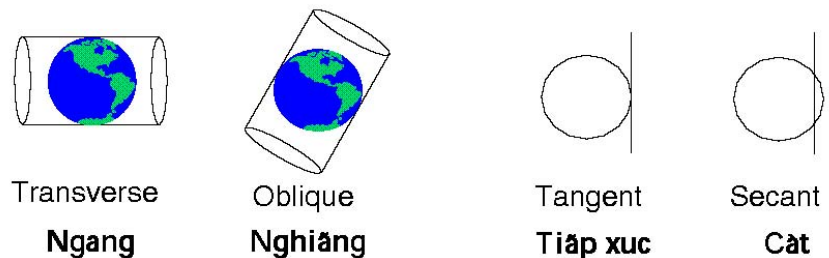
Dựa trên mặt chiếu hình hỗ trợ có các phép chiếu sau:
-Phép chiếu hình nón: là phép chiếu mà bề mặt hình học hỗ trợ là hình nón tiếp xúc (chiếu tiếp tuyến) hoặc cắt quả địa cầu (chiếu pháp tuyến)

CÁC DẠNG MẶT CHIẾU



CÁC DẠNG VẼ TRÊN MẶT CHIẾU

-Phép chiếu hình phương vị: là phép chiếu mà bề mặt hình học hỗ trợ là mặt phẳng tiếp xúc (chiếu tiếp tuyến) hoặc cắt quả địa cầu (chiếu pháp tuyến)



Hình 4. Các dạng mặt vẽ và mặt chiếu trái đất

- Phép chiếu hình trụ: là phép chiếu mà bề mặt hình học hỗ trợ là hình trụ tiếp xúc (chiếu tiếp tuyến) hoặc cắt quả địa cầu (chiếu pháp tuyến)

Căn cứ theo vị trí của mặt chiếu hình hỗ trợ với trục của quả địa cầu, có các phép chiếu: - Phép chiếu thẳng (hay phép chiếu đứng): Trục của mặt chiếu (mặt phẳng, nón hay trụ) trùng với trục quay của quả địa cầu.

- Phép chiếu ngang (hay phép chiếu xích đạo): Đối với phép chiếu phương vị, mặt chiếu hình hỗ trợ tiếp xúc ở một điểm hay một đường bất kỳ trên xích đạo. Ở phép chiếu hình nón và phép chiếu hình trụ, trục của mặt nón và mặt trụ nằm trong mặt phẳng xích đạo, vuông góc với trục quay của quả địa cầu.

- Phép chiếu nghiêng: Ở phép chiếu phương vị, mặt phẳng chiếu tiếp xúc với quả địa cầu tại một điểm nào đó giữa xích đạo và cực. Đối với phép chiếu hình nón và hình trụ, trục của mặt nón và mặt trụ có vị trí nghiêng so với mặt phẳng xích đạo

Người hoa tiêu, hay bộ đội pháo binh thường sử dụng bản đồ theo hệ qui chiếu đồng dạng, nhà kinh tế, hay nhà địa chất học muốn thể hiện các kết

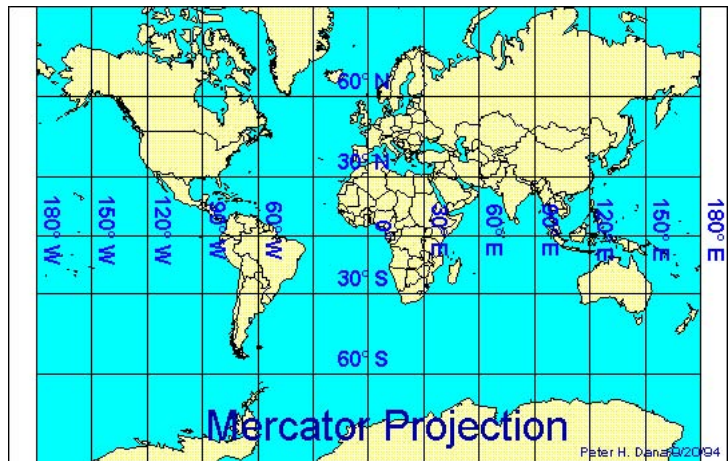
quả tính toán thống kê thì dùng bản đồ có hệ qui chiếu tương đương.

Phép chiếu cho hình cầu (hình II.5) là phép chiếu đồng dạng mặt cầu lên mặt trụ tiếp xúc theo xích đạo (mặt chiếu hình trụ, chiếu thẳng, tiếp tuyến, đồng góc), sau đó triển khai mặt trụ thành mặt phẳng.

Ở phép chiếu này tỉ lệ theo lưới chiếu các kinh tuyến và vĩ tuyến thay đổi như nhau, liên tục tăng dần từ xích đạo đến cực. Xích đạo có bề dài $2R$ và nó là vĩ tuyến duy nhất không có sai số về độ dài, từ xích đạo về cực các vĩ tuyến lần lượt bị kéo dài ra. Phép chiếu Mercator có tính đồng góc: góc trên bản đồ có độ lớn bằng góc tương ứng trên quả địa cầu. Vì thế các bản đồ theo lưới chiếu Mercator được dùng rộng rãi trong hàng hải và hàng không. (hình II. 6)

Phép chiếu Gauss cho hình cầu và hệ tọa độ vuông góc Gauss-Kruger:

Phép chiếu Gauss thực chất là một sự biến đổi của phép chiếu Mercator. Để giảm sự biến dạng, trước hết người ta chia mặt cầu thành các múi chiếu 60 theo các kinh tuyến (hình II. 7). Các múi được đánh số từ 1 đến 60 kể từ kinh tuyến gốc hết



Đông sang Tây bán cầu. Kinh tuyến gốc là giới hạn phía tây (trái) của múi thứ nhất. Mỗi múi được giới hạn bởi kinh tuyến phía Tây (trái) LT và kinh tuyến phía Đông (phải) LP. Kinh tuyến giữa của múi (kinh tuyến trục) là Lo . Các độ kinh này được tính như sau: $LT = 60(n-1)$; $LP = 60n$; $Lo = 60n-3$; với n là số thứ tự của múi.

Dựng mặt trụ nằm ngang ngoại tiếp với mặt cầu trái đất theo kinh tuyến trục của múi. Lấy tâm hình cầu làm tâm chiếu để chiếu múi này lên mặt trụ. Lần lượt chiếu các múi liên kề nhau bằng cách xoay cho kinh tuyến giữa của từng múi tiếp xúc với mặt trụ

4.2.3. Hệ thống phân mảnh và danh pháp bản đồ

Việc chia mảnh và đặt tên cho bản đồ chủ yếu do điều kiện ấn loát, in ấn, và giúp việc sử dụng bản đồ ngoài thực địa, treo tường, để bàn và bảo quản bản đồ được thuận tiện. Có hai cách chia mảnh và đánh số bản đồ:

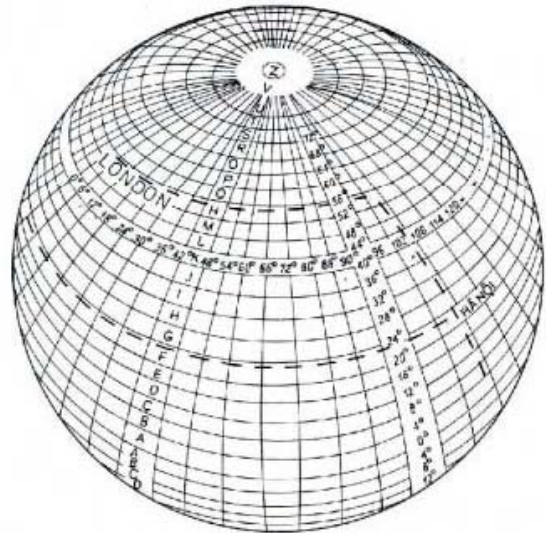
- Chia mảnh vuông góc: Khung của bản đồ hoặc trùng với đường của lưới tọa độ vuông góc hoặc theo đường phân chia khác. Bản đồ được chia thành các mảnh hình chữ nhật, đánh số thứ tự theo hàng ngang từ trái sang phải

và từ trên xuống dưới theo hàng dọc có sơ đồ kèm theo

- Hệ chia thành hình thang: Cách chia mảnh này dùng đường kinh tuyến và vĩ tuyến biên của mỗi mảnh bản đồ để làm khung. Hệ thống bản đồ cơ bản ở nước ta và các nước Đông Âu đều sử dụng hệ chia mảnh này.

Nguyên tắc chia mảnh hình thang được tiến hành như sau:

- Theo chiều kinh tuyến chia bề mặt trái đất thành 60 dải đánh số từ 1-60, mỗi dải cách nhau 6° . Thứ tự các dải được đánh số lần lượt bắt đầu từ kinh tuyến 180-174 T là dải số 1, 174-168 T là dải số 2... dải 60 từ 174 – 180 $^{\circ}$.
- Theo chiều vĩ tuyến từ xích đạo trở về hai cực, cứ 4° chia thành 1 đai có đánh số thứ tự bằng chữ in hoa A,B,C,D,...



Như vậy, bề mặt trái đất được chia thành các mảnh hình thang có độ chênh lệch kinh độ 6° và độ chênh lệch vĩ độ là 4° . Mỗi hình thang biểu thị trên một bản đồ 1:1.000.000. Danh pháp của nó được ghi rõ theo đai và dải.

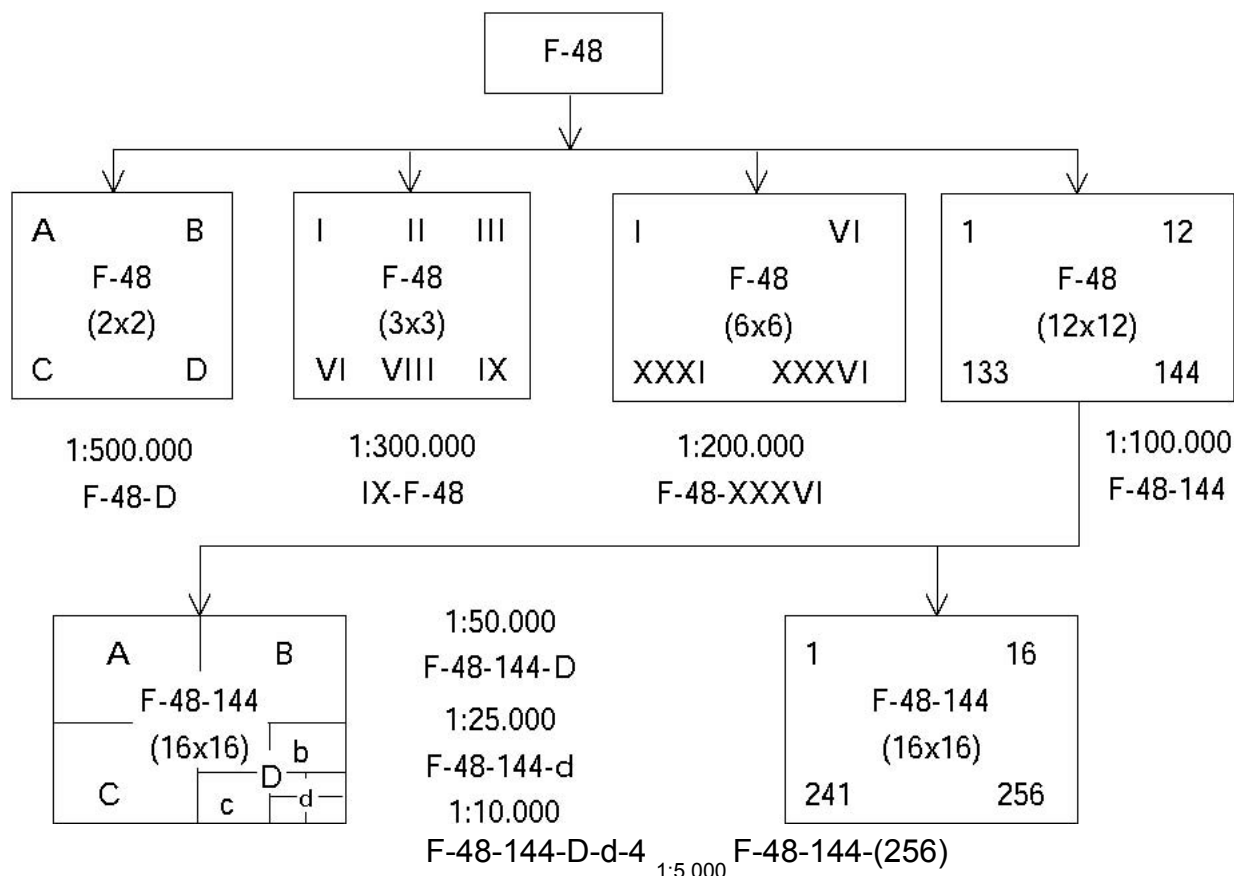
Ví dụ: Bản đồ nghi F-48 là tờ bản đồ có tỷ lệ 1:1.000.000, F biểu thị của đai từ 20-24 $^{\circ}$ vĩ độ, 48 là tên của dải thứ 48 từ kinh tuyến 102 $^{\circ}$ Đ đến 108 $^{\circ}$ Đ.

Nếu tờ bản đồ thể hiện phần bắc bán cầu thì ghi thêm chữ N (north) và ở nam bán cầu thì ghi thêm chữ S (south), ví dụ NF-48

Lãnh thổ Việt nam nằm ở trong các đai C,D,E,F và các dải 48,49. -Bản đồ tỷ lệ từ 1:500.000 đến 1:100.000 được chia mảnh và ghi số hiệu theo bản đồ 1:1.000.000. -Mảnh bản đồ tỷ lệ 1:50.000 đến 1:10.000 được chia mảnh và ghi số hiệu theo bản đồ 1:100.000. -Mảnh bản đồ tỷ lệ 1:5.000 đến 1:2.000 thể hiện vùng đất lớn hơn 20km² được chia mảnh và ghi số hiệu theo bản đồ 1:1.000.000.

- Đối với vùng đất nhỏ hơn 20km² ta có thể chia mảnh và ghi số hiệu theo tọa độ ô vuông với kích thước là 40x40 km cho bản đồ tỷ lệ 1:5.000 và 50x50km cho bản đồ tỷ lệ 1:2000 đến 1:500

Sơ đồ phân mảnh bản đồ:



4.3. Phương pháp thể hiện bản đồ

4.3.1. Phân loại bản đồ

Bản đồ đường nét (line maps) và bản đồ ảnh (photo and image map)
 Bản đồ “đường nét” là loại bản đồ thường sử dụng nhất, trong đó thể hiện các ký hiệu (symbols) và các ký tự chú giải. Nó được dùng để thể hiện các thông tin tóm lược về khu vực được vẽ (hình IV.1).



Hình IV.1 Bản đồ đường nét

Bản đồ ảnh là một dạng bản đồ mà người ta sử dụng ảnh chụp từ máy bay hay từ vệ tinh xuống khu vực cần thể hiện, sau đó người ta vẽ thêm vào các phần chính muốn thể hiện (như đường biên hành chính, đường đồng mức, lưới tọa độ,...) hoặc nhấn mạnh một số chi tiết (như đường, sông, sử dụng đất hay thực vật,...) cùng với tiêu đề, tỉ lệ và chú thích (Hình IV.2).

Ưu điểm của bản đồ ảnh là vẽ nhanh hơn bản đồ “đường nét” vì ta chỉ cần vẽ thêm vào ảnh chụp một số ký hiệu và ghi chú, hơn nữa bản đồ ảnh thể hiện chi tiết hơn do các vật thể được chụp trực tiếp chứ không phải là được vẽ lại. Bản đồ ảnh rất thích hợp cho những vùng không thể thể hiện tốt bởi bản đồ đường nét như vùng sa mạc, băng tuyết, đầm lầy,... Tuy nhiên nhược điểm của bản đồ ảnh là vấn đề giải đoán vật thể trên ảnh (như phân biệt các loại cây hay sử dụng đất) và đôi khi tán lá, mây hay các vùng núi dốc cao và vùng thành phố có nhà cao tầng có thể che khuất các vật thể ở dưới.



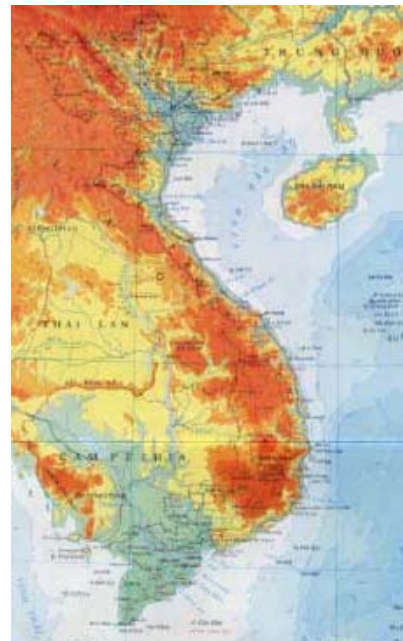
Do đó, bản đồ ảnh thích hợp cho một số trường hợp nhưng không thể hoàn toàn thay thế được bản đồ “đường nét”

Hình IV.2 Bản đồ ảnh

Bản đồ địa hình (topographic) và bản đồ chủ đề (thematic map)

Một cách phân loại bản đồ khác thường gặp là phân biệt giữa bản đồ địa hình và bản đồ chủ đề. Bản đồ chủ đề là bản đồ được thiết kế nhằm trình bày các thực thể hay các khái niệm cụ thể, bản đồ chủ đề thường được dùng khi muốn nhấn mạnh một hay nhiều chủ đề nào đó. Tùy theo nội dung bản đồ chủ đề thường được dùng trong việc:

- Tìm phương hướng, hoa tiêu
- Quy hoạch
- Dự đoán sự phát triển
- Khai thác tài nguyên, khoáng sản
- Quản lý
- Phân tích khoa học và so sánh
- Giáo dục, v.v...



Hình IV. 3. Bản đồ chuyên đề

Theo định nghĩa của hiệp hội bản đồ thế giới (ICA) thì bản đồ địa hình là: “bản đồ mà mục tiêu chính là để miêu tả và xác định các thực thể của bề mặt trái đất một cách trung thực nhất mà nó có thể trong Hình IV.3 Bản đồ địa hình sự giới hạn của tỉ lệ bản đồ”, các thực thể này có thể là thực thể tự nhiên hay thực thể nhân tạo. Các thực thể được trình bày trên bản đồ địa hình dưới dạng vị trí, hình dạng và cao độ. Bản đồ địa hình thông thường là

bản đồ sử dụng cho nhiều mục tiêu ví dụ như:

- Quản lý hành chính quốc gia
- Quân sự
- Du lịch và giải trí
- Quy hoạch
- Quản lý tài nguyên
- Địa chính hay định cư
- Giáo dục

4.3.2. Thành phần của bản đồ

Thành phần của bản đồ rõ ràng liên quan đến mục tiêu sử dụng của nó. Các thành phần trong bản đồ là: -Thành phần chính (chủ đề chính) -Thành phần thứ hai (bản đồ nền, thông tin cơ bản của bản đồ) -Thành phần phụ trợ (thông tin lẻ như chú thích, tỉ lệ, tiêu đề...)

Thành phần chính Là phần chủ đề của bản đồ, ví dụ như địa lý, địa chất, dân số. Đối với bản đồ địa hình, phần chính là tất cả các thông tin được vẽ, bao gồm cả tên của các vùng.

Thành phần thứ hai Đối với bản đồ chủ đề, thành phần này là phần địa hình, bao gồm lưới tọa độ.

Thành phần phụ trợ Bao gồm các thông tin lẻ như tiêu đề, chú thích, thanh tỉ lệ,...

4.3.3. Độ chính xác

Ba vấn đề của độ chính xác cần quan tâm là: -Chính xác về vị trí -Chính xác về chủ đề -chính xác về cách thể hiện

a. Chính xác về vị trí Độ chính xác của vị trí được vẽ trên bản đồ liên quan đến vị trí thực tế của nó trên thực tế. Độ chính xác này ảnh hưởng bởi: - Phép chiếu -Độ chính xác của việc thu thập dữ liệu và việc vẽ bản đồ -Tỉ lệ của bản đồ -Công cụ và độ ổn định của vật liệu được sử dụng trong việc vẽ bản đồ

Chính xác về chủ đề Độ chính xác về chủ đề liên quan đến thông tin chủ đề được thể hiện. Độ chính xác này ảnh hưởng bởi: -Việc thu thập thông tin thuộc tính: chất lượng của dữ liệu thống kê và phương pháp thống kê -Việc chuyển đổi dữ liệu: Dữ liệu của một phần của vùng đôi khi được dùng để thể hiện cho toàn vùng, ví dụ như trường hợp bản đồ mật độ dân số (một huyện có mật độ 50 người/km² không có nghĩa mọi km² của huyện đều có 50 người.

Chính xác về cách thể hiện Sự thể hiện của các biểu tượng trên bản đồ rất

quan trọng, nếu dùng sai biểu tượng thì có thể đánh lạc hướng của người sử dụng, hay làm mờ ranh giới của các vùng trên bản đồ.

4.3.4. Chú giải trên bản đồ - ngôn ngữ bản đồ

Ngôn ngữ bản đồ: Hệ thống ký hiệu bản đồ có thể xem như là một ngôn ngữ, nó thoả mãn ba chức năng cơ bản sau: -Dạng (hoặc cấu trúc): hình vẽ ký hiệu gợi ý cho ta liên tưởng đến đối tượng cần phản ánh -Bản thân ký hiệu phải chứa trong đó một nội dung nào đó về số lượng, chất lượng, cấu trúc, động lực phát triển của đối tượng cần phản ánh trên bản đồ

-Ký hiệu trên bản đồ phải phản ánh vai trò của đối tượng trong không gian và vị trí tương quan của nó đối với các yếu tố khác. Các ký hiệu được sắp xếp theo một quy định nhất định trong không gian.

Ngôn ngữ bản đồ là phương tiện tư duy và diễn đạt kết quả tư duy, là phương tiện liên lạc và trao đổi nhận thức, là công cụ để xây dựng mô hình thực tế khách quan, là phương tiện thể hiện những tri thức của loài người trên bản đồ.

Hệ thống ký hiệu qui ước bản đồ: các phương tiện chủ yếu được sử dụng trong ngôn ngữ bản đồ là các dạng đồ họa, màu sắc, chữ và con số. Ứng với những đặc tính của hiện tượng (chủ yếu là hiện tượng địa lý) với những tính chất và đặc điểm của đồ họa và màu sắc, các ký hiệu trên bản đồ thường ở các dạng ký hiệu điểm, ký hiệu tuyến và ký hiệu diện tích.

-Ký hiệu điểm (point): đối với đối tượng địa lý phân bố theo những điểm riêng biệt (cột mốc trắc địa, đối tượng diện tích nhỏ khi biểu hiện trên bản đồ không thể theo đường biên của nó như nhà, trụ điện,...) được thể hiện trên bản đồ bằng ký hiệu điểm. Các ký hiệu này dùng để xác định về vị trí của hiện tượng do đó lấy vị trí của hiện tượng làm tâm để vẽ, và kích thước ký hiệu đó không cần theo tỉ lệ.

-Ký hiệu tuyến (polyline): thường dùng thể hiện ranh giới (quốc gia, tỉnh, huyện, xã) hay đường giao thông, sông ngòi, đường dây điện... là loại đối tượng phân bố theo chiều dài. Ngoài ra còn có dạng tuyến tính đặc biệt như đường đẳng trị (đồng cao độ, đẳng mặn, đẳng nhiệt...), các ký hiệu này khi thể hiện trên bản đồ phải thể hiện đúng tỷ lệ theo chiều dài, chiều rộng có thể tăng rộng ra hai bên đường trung tâm.

-Ký hiệu diện tích (hay vùng: region): được dùng để thể hiện các hiện tượng phân bố theo vùng như các khu vực sử dụng đất đai (rừng, nông nghiệp, đầm lầy,...). Toàn khu vực phải vẽ theo tỷ lệ bản đồ và giới hạn bởi đường biên ngoài của nó, đường biên ngoài được thể hiện theo ký hiệu dạng đường (nét liền, chấm gạch, màu...). Trong khu vực thể hiện bằng một mẫu tô nào đó phản ánh tính chất của đối tượng (khu trồng lúa có mẫu tô hình cây lúa màu xanh, khu đô thị tô màu đỏ sậm,...)

Về mặt hình thức ký hiệu, người ta chia ra thành các dạng:

- Ký hiệu hình học và ký hiệu chữ.



Dùng các ký hiệu đặt vào vị trí của đối tượng mà nó thể hiện.

Hình thức của ký hiệu: Chữ, hình học, tượng hình...

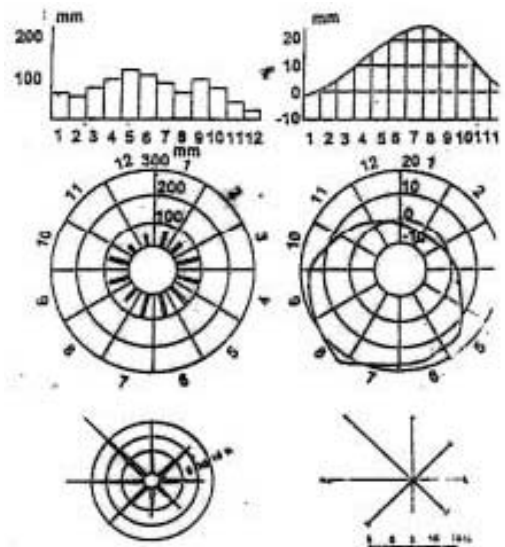
Thể hiện số lượng hiện tượng:

- + Thể hiện bằng độ lớn của ký hiệu
- + Thể hiện bằng màu sắc



- + Thể hiện bằng nét gạch trong ký hiệu
- + Thể hiện bằng độ dài của ký hiệu (ít sử dụng)
 - + Thể hiện qua diện tích ký hiệu
 - + Thể hiện động lực của hiện tượng (sự phát triển dân số):
 - + Ký hiệu tăng trưởng: các ký hiệu được đặt trùng lên nhau

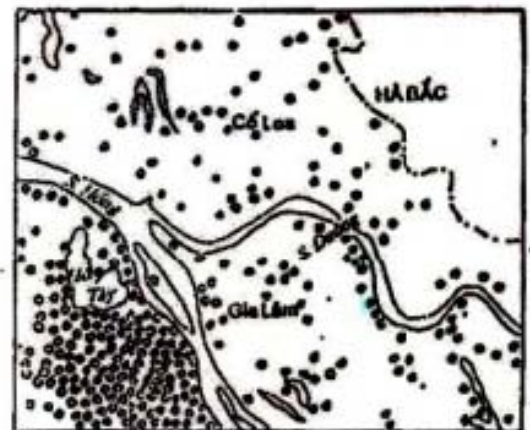
b. Phương pháp biểu đồ định vị: - Dùng thể hiện các hiện tượng phân bố liên tục hoặc trên toàn bộ mặt đất có sự biến đổi theo chu kỳ và việc nghiên cứu chúng được tiến hành ở những điểm nhất định như các hiện tượng trong khí quyển - Đặt biểu đồ thể hiện mối tương quan giữa các đặc trưng đo đạc hay tính chất của hiện tượng muốn thể hiện vào vị trí của hiện tượng đó. Thường dùng cho đối tượng dạng điểm.



Các dạng biểu đồ và biểu đồ định vị khác nhau

c. Phương pháp chấm điểm: - Thể hiện sự phân tán của hiện tượng trên một vùng (phân bố dân cư). - Dùng cho các đối tượng dạng vùng - Mỗi điểm (hay một ký hiệu) sẽ tương đương với một số lượng hiện tượng qui ước. - Không đòi hỏi sự chính xác về mặt địa lý, chỉ thể hiện sự phân bố về mặt số lượng và sự phân bố của hiện tượng

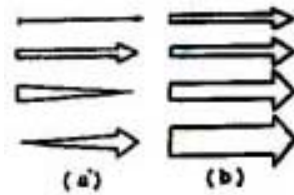
- Có thể dùng màu sắc hay các loại ký hiệu điểm khác nhau để thể hiện chất lượng và động lực của hiện tượng (chấm đỏ: nam, chấm xanh: nữ)



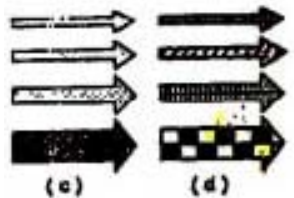
Phương pháp chấm điểm
một điểm (chấm) tương ứng 4. 500 người

- d. Phương pháp ký hiệu đường chuyển động: - Thể hiện sự di chuyển của các đối tượng, hiện tượng trên bản đồ (dòng sông, tuyến đường, gió)
- Dùng cho đối tượng dạng tuyến
 - Dạng vector (mũi tên)

- a. theo hình dạng b. theo độ lớn
c. theo màu sắc d. theo cấu trúc bên trong



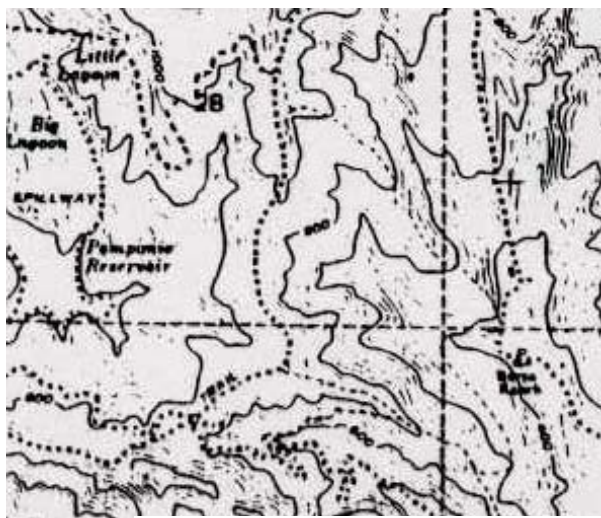
e. Phương pháp đường đẳng trị: -Đường nối các điểm có cùng chỉ số về số lượng của hiện tượng trên bản đồ (đồng cao độ, đẳng mưa, đẳng nhiệt...) -Kết hợp với màu sắc để tăng tính trực quan và thể hiện về mặt số lượng (cường độ màu phản ánh hướng và trình tự chuyển tiếp từ các trị số thấp nhất đến các trị số cao nhất)



-Điểm đẳng trị phải bảo đảm được tính chính xác về mặt địa lý.

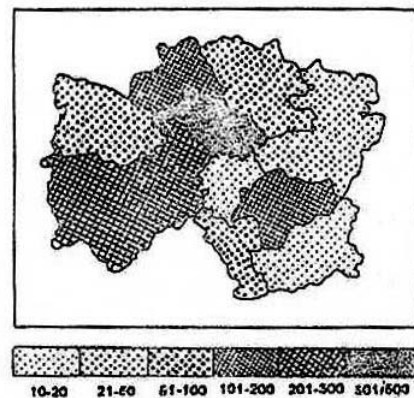
f. Phương pháp nền chất lượng:

-Thể hiện các hiện tượng phân bố liên tục trên bề mặt đất (lớp phủ thực vật, loại đất,...) hay các hiện tượng phân bố theo khối (dân cư, phân vùng lãnh thổ) -Thể hiện sự khác nhau về chất lượng của các hiện tượng



-Dùng màu sắc, mẫu tô hay đánh số khu vực trong ranh giới của hiện tượng
-Vị trí, hình dạng phải bảo đảm tính chính xác về mặt địa lý.

- Dùng cho đối tượng dạng vùng



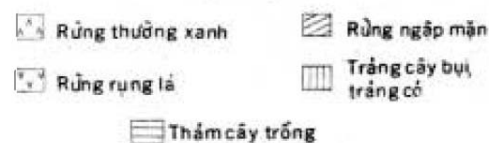
Phương pháp Cartogram

g. Phương pháp Cartogram: -Dùng cho đối tượng dạng vùng

- Đặt biểu đồ thể hiện mối liên quan của các đặc trưng của hiện tượng vào trong biên của hiện tượng đó

- Dùng cho nội tổng dạng vùng

THỰC VẬT



4.3. 6. Sự khái quát hóa (generalization) và sự phóng đại (axaggeration)

Vì bản đồ là sự thu nhỏ của thế giới thực nên ta không thể trình bày một cách chính xác tuyệt đối hình dạng và kích thước thực thể. Do đó thường người ta dùng hai kỹ thuật sau đây để thể hiện thực thể trên bản đồ:

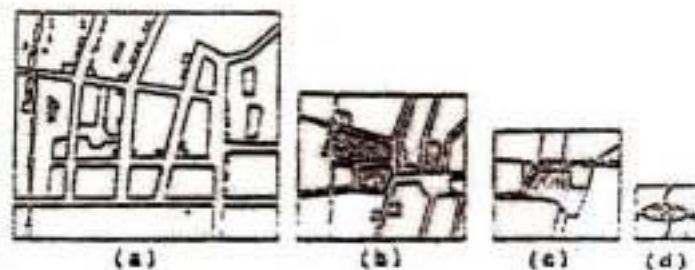
- Khái quát hóa là sự chọn lựa và đơn giản hóa sự thể hiện của thực thể trên bản đồ theo một tỉ lệ và mục đích thích hợp nhằm giúp cho bản đồ dễ đọc. - Sự phóng đại là kỹ thuật nhằm phóng to vật thể cần thể hiện hơn là tỉ lệ thực của nó nhằm giúp cho bản đồ dễ đọc hay nhằm nhấn mạnh vật thể đó.

Sự khái quát hóa yêu cầu chú ý đến các yếu tố sau: -Sự chọn lựa: Mục tiêu của bản đồ là yếu tố chính để chọn lựa thực thể nào nên vẽ trên bản đồ. Sự chọn lựa thường liên quan đến tỉ lệ bản đồ.

-Sự đơn giản hóa: Các thực thể phải được trình bày trên bản đồ nhưng qua nhỏ hay quá phức tạp mà không thể trình bày được chi tiết nếu không bỏ bớt hay đơn giản hoá. Tỉ lệ là yếu tố tham gia chính.

-Lược bỏ: Để duy trì tính dễ đọc và sạch sẽ của bản đồ, một vài thực thể sẽ không được thể hiện, ngay cả nó rõ ràng. Tỉ lệ vẫn là yếu tố ảnh hưởng chính nhưng yếu tố tự nhiên và địa hình cũng quan trọng.

Mối liên hệ giữa sự khái quát hóa và sự phóng đại rất gần, thực ra chính sự phóng đại hóa là sự khái quát hóa. Ví dụ trong trường hợp bản đồ đường xá tỉ lệ 1/50000, nếu ta vẽ theo đúng tỉ lệ con đường rộng 10m thì nét vẽ thể hiện con đường này chỉ rộng 0.2mm cho tất cả các đoạn rẽ hay đoạn xoắn, nhưng trong bản đồ chúng ta phải dùng nét vẽ 1mm để thể hiện. Tuy nhiên với nét vẽ này ta không thể thể hiện chính xác được các đoạn rẽ và đoạn xoắn



Khái quát hóa với các tỉ lệ: a. Tỉ lệ 1:25.000; b. 1:50.000; c. 1:100.000; d. 1:1.000.000

Như vậy sự lựa chọn biểu tượng để trình bày thực thể rất quan trọng. Nên tránh sự phóng đại quá đáng cũng như sự đơn giản hình dạng không cần thiết

NHẬP, QUẢN TRỊ VÀ XUẤT DỮ LIỆU CỦA HỆ THỐNG THÔNG TIN ĐỊA LÝ

Xây dựng CSDL HTTTĐL phức tạp hơn việc tái tạo bản đồ. Việc quan trọng xây dựng CSDL HTTTĐL là nhập dữ liệu từ các nguồn tài liệu. Nó gồm nhiều nguồn khác nhau được sử dụng, như ảnh hàng không, tài liệu bảng biểu, những dữ liệu số khác vv...

Trong bài này sẽ giới thiệu các phương pháp nhập dữ liệu, sửa chữa các dữ liệu và tổ chức thành CSDL trong GIS. Những phương pháp này ngày càng được cải tiến để nâng cao chất lượng dữ liệu nhập vào, cũng như giảm nhẹ các thao tác cho người nhập dữ liệu.

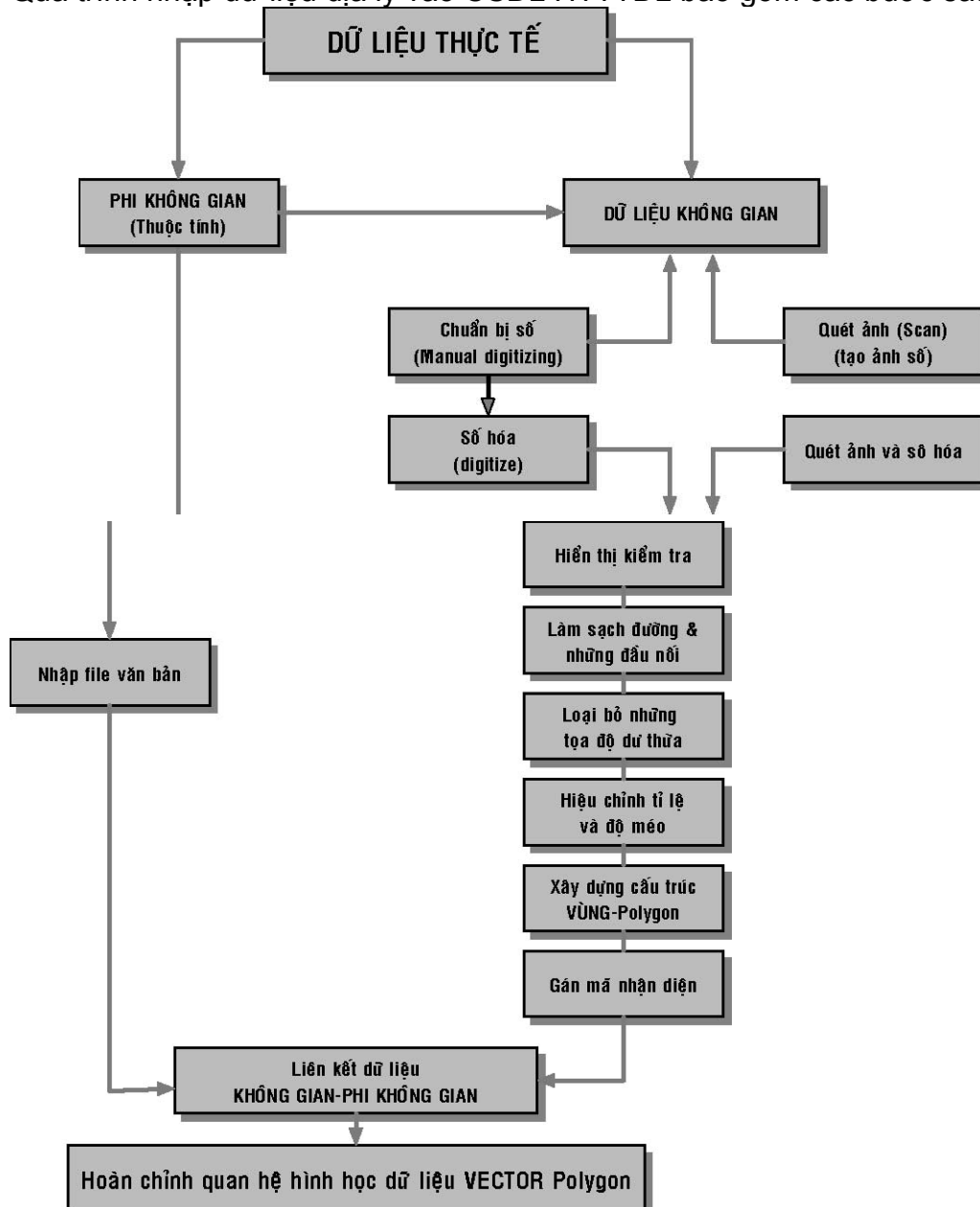
5.1. Những thủ tục thu thập dữ liệu

Dữ liệu từ nhiều nguồn khác nhau và tổ chức nó thành một CSDL có ích đòi hỏi yêu cầu phát triển những thủ tục bao gồm:

- Thu nhận dữ liệu: bao gồm thu thập dữ liệu đã có từ các nguồn, định lượng và kiểm tra tính hoàn thiện và chất lượng hay tạo ra dữ liệu mới bằng đo đạc và giải đoán ảnh máy bay hay đi thực địa khảo sát. Những dự án HTTTĐL cố gắng dựa trên những dữ liệu có sẵn hơn là tạo mới dữ liệu vì chi phí cao hơn dữ liệu nguyên gốc. Nhưng những dữ liệu có sẵn thường xây dựng trên mục đích khác sẽ có sự gượng ép khi sử dụng vào HTTTĐL. Do đó việc định lượng và kiểm tra sự thích hợp nguồn dữ liệu cho sử dụng HTTTĐL là rất quan trọng.
- Sửa chữa dữ liệu nguồn: Tập trung vào những tài liệu bản đồ, hoạt động này gọi là làm sạch bản đồ. Tùy thuộc vào kỹ thuật sử dụng chuyển đổi ảnh bản đồ thành dạng dữ liệu số của nó, những tài liệu nguồn thường có một chuẩn nào đó. Một số quá trình chuyển đổi yêu cầu dữ liệu tương đối hoàn hảo và tự động sửa chữa tất cả những cái cần thiết trên bản đồ. Những yêu cầu dữ liệu kết quả do phân tích HTTTĐL xác định. Những qui trình sửa chữa dữ liệu cần kiểm tra trong những dự án mẫu phạm vi nhỏ trước khi sử dụng trong dự án.
- Kiểm soát sự thay đổi, phần lớn loạt bản đồ thường không cố định và sẽ được sửa chữa, nâng cấp theo chu kỳ. Một phần của bản đồ sẽ gửi tới số hóa sửa chữa những thay đổi và xây dựng file dữ liệu số mới.

□ Xây dựng CSDL HTTTĐL, khi số hóa xong bản đồ, chúng ta sẽ có một tập hợp các file dữ liệu số, chưa được tổ chức trong CSDL. Quá trình xử lý tích hợp hệ thống là liên kết các file dữ liệu số và thiết đặt CSDL HTTTĐL trong dạng sẽ có ích cho người sử dụng. Một số điều cần cân nhắc bao gồm là cấu trúc CSDL, tích hợp HTTTĐL, bảo trì và sử dụng. (hình 4.11).

Quá trình nhập dữ liệu địa lý vào CSDL HTTTĐL bao gồm các bước sau:



Hình 4.12 Các bước tạo, sửa chữa dữ liệu Vector polygon

5.2. Nhập dữ liệu

5.2.1. Nhập dữ liệu địa lý

Trong các hệ tự động hóa vẽ bản đồ, thuộc tính của các đối tượng hình học được thể hiện thông qua màu sắc, kiểu đường, kiểu tô màu vv... Nhưng trong HTTTĐL, các dữ liệu thuộc tính phi không gian được lưu trữ một cách định lượng và được tổ chức quản lý theo các tiêu chuẩn chặt chẽ của một CSDL. Có thể được phân tách thành ba công việc:

- ☐ Nhập dữ liệu không gian
- ☐ Nhập dữ liệu thuộc tính
- ☐ Kết nối hai loại dữ liệu trên

Dữ liệu được đưa vào CSDL thông qua các chức năng đầu vào của HTTTĐL. Nhập dữ liệu là một quá trình mã hóa, lưu trữ và tổ chức dữ liệu thành CSDL. Đây là một quá trình rất quan trọng của hệ thống. Nó ảnh hưởng đến độ chính xác và tính logic của dữ liệu trong CSDL, do đó các thiết bị nhập số liệu cũng như các phần mềm nhập số liệu phải đảm bảo độ chính xác.

Có 5 phương pháp nhập dữ liệu thường dùng trong GIS:

- ☐ Nhập từ bàn phím
- ☐ Nhập từ bàn số hóa (digitizer)
- ☐ Nhập bằng máy quét (scanner)
- ☐ Nhập trực tiếp từ tập tin hiện hữu
- ☐ Từ các dữ liệu viễn thám

5.2.2. Nhập dữ liệu không gian

Có nhiều phương pháp để nhập dữ liệu không gian vào HTTTĐL. Dữ liệu đầu vào thường là các bản đồ giấy, các bản vẽ tay, các ảnh hàng không, ảnh vũ trụ, điểm lấy mẫu và các dữ liệu khác đo được gián tiếp bằng các thiết bị cảm ứng. Phương pháp nhập số liệu phụ thuộc vào cấu trúc dữ liệu và phần mềm CSDL của HTTTĐL sử dụng.

Các phương pháp sau:

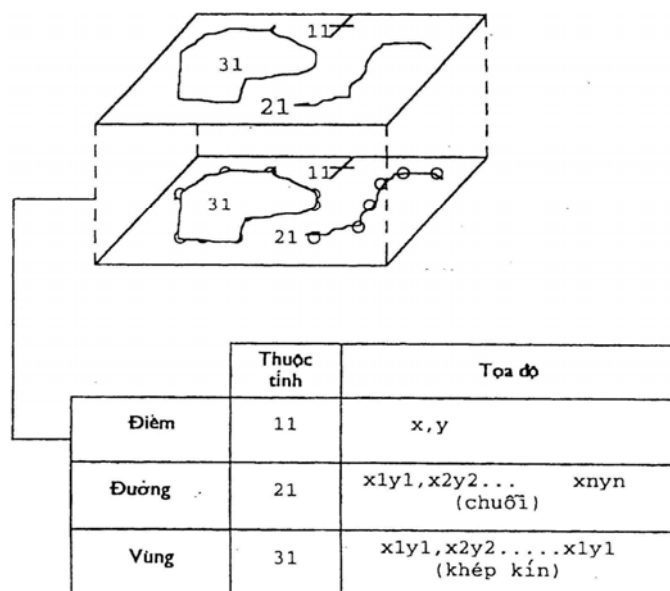
Nhập dữ liệu thủ công từ bàn phím:

Vector: Dữ liệu này thể hiện là điểm, đường hay vùng được nhập từ bàn phím, theo các tọa độ chính xác hoặc nhập vào tệp dữ liệu ASCII.

Raster: Đối với hệ thống này, mọi điểm, đường, vùng đều được biến thành các cell. Phương pháp thông dụng nhất được diễn ra như sau: Đầu tiên

chọn kích cỡ lưới ô, sau đó chồng lên bản đồ. Giá trị tại từng ô nhận được từ bản đồ sẽ được ghi lại vào máy tính.

Số hóa thủ công (Số hóa bằng bàn số hóa)



Hình 5.2 Tập tin chứa tọa độ x, y chuyển từ bản đồ vào
Số hóa thủ công tiến hành sử dụng bàn số hóa và công cụ con trỏ mã hóa những điểm trên bản đồ.

Để số hóa được dữ liệu bản đồ giấy, đặt bản đồ lên bàn số hóa, khai báo các mốc tọa độ khống chế và kích chuột vào các đối tượng trên bản đồ để nhận được tọa độ.

Bàn số hóa sẽ cho phép nhập 3 kiểu dữ liệu chính (Điểm, Đường, Text). Các đối tượng vùng được xây dựng trên cơ sở cấu trúc topo mạng đa giác. Công việc này đòi hỏi nhiều thời gian và tập trung cao độ. Kết hợp giữa kỹ năng và sự nhanh nhạy, mắt giữ chữ thập ở vị trí chính xác và đầy đủ yếu tố bản đồ.

Theo thông thường sự chuyển đổi dữ liệu vector thực hiện số hóa thủ công những bản đồ giấy. Xây dựng dữ liệu Vector thường tiến hành với chi phí cao.

Nhập bằng máy quét

Nhập bằng máy quét (scanner)

Dùng scanner để nhập dữ liệu vào máy tính sẽ nhanh hơn dùng bàn số hóa, đầu ra của scanner là ảnh số (digital image) dạng raster, kích thước của điểm ảnh (pixel) thường dùng là 0.02mm. Ảnh raster được xử lý để nâng cao chất lượng hình ảnh. Để sản xuất dữ liệu dưới dạng vector phải thực hiện quá trình vector hóa.

Các thiết bị này chuyển đổi tín hiệu ánh sáng phản xạ từ bản đồ thành giá trị pixel và ghi chúng lại theo một khuôn dạng nhất định.

Chất lượng dữ liệu raster quét phụ thuộc vào thiết bị quét, về độ phân giải theo mỗi chiều của mặt phẳng quét và độ phân giải theo tín hiệu phản xạ từ bản đồ. Thông thường trên thị trường Việt nam hiện có các máy quét với độ phân giải từ 600-2400 dpi (dot per inch), nghĩa là có thể phân biệt được từ 0.05-0.01 mm trong khi đó mắt thường chỉ phân biệt được 0.1 mm. Có 2 loại scanner chính: Loại scanner đen trắng và loại scanner màu. Loại đen trắng chỉ nhận được mức năng lượng phản xạ lại từ bản đồ, loại kia thì có bộ phận phân tích thành 3 dải màu cơ bản và ghi chúng thành các cấp độ khác nhau. Thường thì mỗi dải màu được phân biệt 256-1024 cấp độ.

Thiết bị quét ảnh là thiết bị rất chính xác, do đó các scanner khổ lớn có công nghệ quét và hiệu chỉnh sai số hoàn toàn khác so với các scanner khổ nhỏ. Trong quá trình phân tích, scanner thường được cài đặt chương trình phần mềm làm tăng cường chất lượng của ảnh quét.

Chuyển đổi Vector sang Raster

Các dữ liệu bản đồ lưu trữ trong CSDL HTTTĐL thường được lưu dạng cấu trúc vector. Nguyên nhân chính của cách là này là độ chính xác, cập nhật dễ dàng và chiếm ít không gian lưu trữ.

Số hóa trực tiếp từ bản đồ thành raster khó sửa chữa và phụ thuộc rất nhiều vào độ phân giải của chúng. Sử dụng cấu trúc dữ liệu raster thì dữ liệu phụ thuộc kích thước cell trong lưới ô kết quả. Nếu chọn kích cỡ cell lớn thì tiết kiệm không gian lưu trữ dữ liệu nhưng dữ liệu sẽ kém chính xác, cần độ chính xác dữ liệu cao thì lưới ô kết quả sẽ có dung lượng lớn.

Có rất nhiều thuật toán để chuyển đổi dữ liệu raster sang vector.

Nhận dạng vùng

Nhận dạng vùng thường được ứng dụng để chuyển raster sang vector. Với mục đích lưu trữ, tra cứu và in ấn, không cần phải xây dựng topology cho dữ liệu. Để xây dựng thành những lớp dữ liệu vector dạng vùng cần xây dựng cấu trúc topology cho dữ liệu (theo hình 4.12).

Nhận dạng Đường

Nhận dạng Đường thường được ứng dụng để nhập số liệu từ bản đồ quét thành dữ liệu vector. Quá trình nhận dạng này bao giờ cũng bắt đầu từ ảnh

quét. Có 2 kiểu nhận dạng đường:

- ☐ Nhận dạng tự động: Nhận dạng tự động là quá trình nhận dạng ảnh quét trong đó không có sự tham gia chỉnh sửa của con người. Kiểu nhận dạng này thường dùng để nhận dạng các bản đồ quét có chất lượng tương đối tốt, không cần sự để tâm đặc biệt của con người. Quy trình nhận dạng như sau:
 - ☐ Làm tăng cường chất lượng ảnh quét (xoá các pixel thừa, làm trơn ảnh,...),
 - ☐ Lọc xương ảnh để nhận dạng Đường,
 - ☐ Chuyển đổi xương ảnh thành vector.
 - ☐ Nhận dạng các ký tự.
- ☐ Nhận dạng ký hiệu bao hàm cả nhận dạng ký tự, chủ yếu dùng để nhận dạng các bảng số, văn bản. Nhận dạng ký hiệu ít khi được dùng để nhận dạng các ký hiệu hay ký tự trên bản đồ.

Nhận dạng tự động có 2 nhược điểm chính:

- ☐ Yêu cầu chất lượng ảnh quét cao, đòi hỏi quá trình sơ xử lý công phu,
- ☐ Không cho phép hiệu chỉnh thông số trong quá trình xử lý nhận dạng, do đó vẫn phải kiểm tra, sửa chữa trên kết quả nhận dạng.
- ☐ Để khắc phục các nhược điểm trên, người ta thường dùng phương pháp nhận dạng bán tự động, cụ thể người ta số hóa trực tiếp trên ảnh quét.

Nhận dạng bán tự động: Quá trình số hóa trên ảnh quét được trợ giúp bởi một số công cụ phần mềm nhận dạng. Người dùng hiển thị ảnh bản đồ quét lên trên màn hình và kích chuột vào Đường trên bản đồ ảnh, phần mềm sẽ tự động số hóa dọc theo Đường đó đến khi nào cắt Đường khác thì dừng lại. Lúc đó người dùng chỉ hướng cho chương trình tiếp tục nhận dạng.

5.2.3. Dữ liệu viễn thám

Tư liệu ảnh chụp từ vệ tinh và ảnh máy bay được sử dụng để sản xuất bản đồ Ảnh máy bay:

- ☐ Dùng làm bản đồ địa hình.
- ☐ Sử dụng các thiết bị trắc địa ảnh để vẽ đường đồng mức địa hình.
- ☐ Lập bản đồ tài nguyên thiên nhiên như: bản đồ đất, rừng, địa chất, hiện trạng sử dụng đất và lớp phủ đất.

Ảnh vệ tinh:

- ☐ Các vệ tinh viễn thám như: Landsat MSS, TM; SPOT, ERS..
- ☐ Dữ liệu ảnh vệ tinh thường ở dạng số.
- ☐ Dùng lập bản đồ sử dụng đất, ngập lụt, chất lượng nước, rừng...
- ☐ Ảnh vệ tinh SPOT có thể dùng để lập mô hình số độ cao, bản đồ địa hình tỉ lệ 1:50.000.

5.2.4. Nhập dữ liệu thuộc tính phi không gian

Dữ liệu thuộc tính phi không gian (còn gọi là thuộc tính đối tượng) là những tính chất, đặc điểm riêng của thực thể không gian, nó cần lưu trữ và thể hiện trong HTTTĐL.

Ví dụ: một con Đường cần được số hóa như một tập các pixel nối với nhau trong cấu trúc dữ liệu raster hoặc là một thực thể dạng đường trong cấu trúc vector. Đường trong HTTTĐL lại được thể hiện với một màu nào đó hoặc ký hiệu hoặc những thông tin kèm theo (kiểu của Đường, dạng bề mặt Đường, phương pháp xây dựng, ngày xây dựng). Tất cả các số liệu này đều được gán cho thực thể. Các dữ liệu này được gán mã khóa mà khi cần, lần theo mã khóa đó, người ta sẽ có toàn bộ số liệu về thực thể.

Dữ liệu thuộc tính phi không gian thông thường được các Hệ quản trị CSDL (HQTCSĐL) quản lý như: phần mềm MapInfo, ORACLE, DB2, INFORMIX, SYSDATABASE, SQL SERVER, ACCESS để quản lý số liệu thuộc tính phi không gian của mình.

Đối với các HQTCSĐL quan hệ, người dùng sẽ nhập số liệu tuân thủ các quy tắc của một HQTCSĐL quan hệ. Quá trình nhập số liệu diễn ra như sau:

- ☐ Thiết lập CSDL mới nếu chưa có CSDL. Nếu đã có CSDL, khởi động CSDL,
- ☐ Mở các bảng tương ứng để nhập số liệu,
- ☐ Kiểm tra và cập nhật các mã khóa,
- ☐ Cập nhật kết nối (nếu phát sinh).

5.3. Dữ liệu dạng số

Ở các nước tiên tiến như Canada, Mỹ các dữ liệu dạng số rất đa dạng:

- ☐ Thông tin dạng số đã được thiết lập cho mục đích thương mại
- ☐ Thông tin địa lý Liên bang được sản xuất bởi cơ quan bản đồ quốc gia và cơ quan dân số và thống kê.
- ☐ Ở cấp bang và tỉnh thành, chương trình số hóa thông tin địa chính là ưu tiên hàng đầu.
- ☐ Thông tin tài nguyên thiên nhiên được chuyển sang dạng số ở cấp Liên bang, bang và tỉnh thành.
- ☐ Ở cấp quận huyện, thông tin sơ thừa đất dạng số là thành phần trung tâm của CSDL GIS.

Dữ liệu địa lý được phân chia thành 4 nhóm chính:

- ☐ Dữ liệu bản đồ nền.
- ☐ Dữ liệu tài nguyên thiên nhiên.

☐ Dữ liệu độ cao số.

☐ Dữ liệu thống kê. Dạng dữ liệu cần được chuẩn hóa cho mục đích sử dụng rộng rãi.

5.4. Quản trị dữ liệu.

5.4.1. Nối kết dữ liệu không gian và phi không gian

Nguyên tắc chung kết nối hai loại dữ liệu này là:

- ☐ Với dữ liệu không gian: trong quá trình số hóa hay biên tập dữ liệu mỗi đối tượng được gán mã thuộc tính, người dùng sẽ phải nhập các mã khóa một cách trực tiếp cho từng đối tượng.
- ☐ Với dữ liệu thuộc tính: được cấu trúc thông tin và nhập mã cho từng mã này. Đối với dữ liệu thuộc tính phi không gian thì người dùng nhập mã khóa
- ☐ Sử dụng phần mềm HTTTĐL nối kết thông tin thuộc tính.

5.4.2. Kiểm tra sai số và làm chính xác dữ liệu.

Dữ liệu không gian và phi không gian đều có sai số. Việc thao tác nối kết chúng là cách tốt nhất để kiểm tra chúng.

- ☐ Đặc điểm không gian có thể sai về vị trí hoặc hình dạng.
- ☐ Sự thiếu cung và vùng hở dẫn đến polygon có hai nhẵn.
- ☐ Dữ liệu không gian liên kết sai với dữ liệu phi không gian.
- ☐ Dữ liệu phi không gian chưa hoàn thành.

5.4.3. Tiềm xử lý dữ liệu không gian

- ☐ Tiềm xử lý topology và định vị tâm hoặc nhẵn của polygon
- ☐ Vector hóa và làm sạch các dữ liệu quét
- ☐ Chuyển cấu trúc dữ liệu từ raster sang vector và ngược lại
- ☐ Chuyển các dạng format dữ liệu

Dữ liệu trong GIS bao gồm nhiều dạng khác nhau. Những tập tin dữ liệu này phải được chuyển đổi thành cấu trúc dữ liệu và dạng tập tin phù hợp với các hệ GIS.

Tập tin raster được nhập vào hệ GIS raster có thể không đòi hỏi định dạng lại, chỉ cần thay đổi dữ liệu gốc một số thông tin (header) để xác định lại tên, gốc, kích thước và những thông số khác được dùng với hệ thống.

Trong trường hợp GIS vector có cấu trúc topology, quá trình chuyển đổi định dạng dữ liệu có thể tốn kém và mất nhiều thời gian và tùy thuộc vào lượng dữ liệu, khả năng phần cứng và phần mềm nếu dữ liệu thu tập không phù hợp với hệ thống GIS. Ví dụ: thông tin bản đồ số hóa bằng các phần mềm thiết kế (CAD) trong cấu trúc phi topology thường rất khó khăn khi chuyển đổi sang cấu trúc topology.

- ☐ Thao tác hình học: tịnh tiến, quay, thay đổi tỉ lệ, thay đổi hệ thống lưới

chiều.

Các lớp dữ liệu trong GIS được hiển thị cùng hệ tọa độ. Các phần mềm GIS thường cho phép chuyển dữ liệu từ lưới chiếu này sang lưới chiếu khác. Ví dụ; có thể chuyển bản đồ Việt nam từ hệ tọa độ địa lý sang lưới chiếu Gauss.

□ Nắn chỉnh (chủ yếu tài liệu ảnh).

Chuyển đổi hình học được dùng để gán tọa độ mặt đất vào lớp dữ liệu hoặc bản đồ trong GIS hoặc hiệu chỉnh hình học của một lớp dữ liệu trong một lớp khác cùng khu vực. Có hai cách dùng để chuyển đổi.

-Chuyển đổi dùng vị trí tương đối: một lớp dữ liệu (slave) được chuyển đổi theo lớp dữ liệu khác (master) dựa trên những địa vật cố định như ngã tư đường, điểm giao nhau của hai con suối.

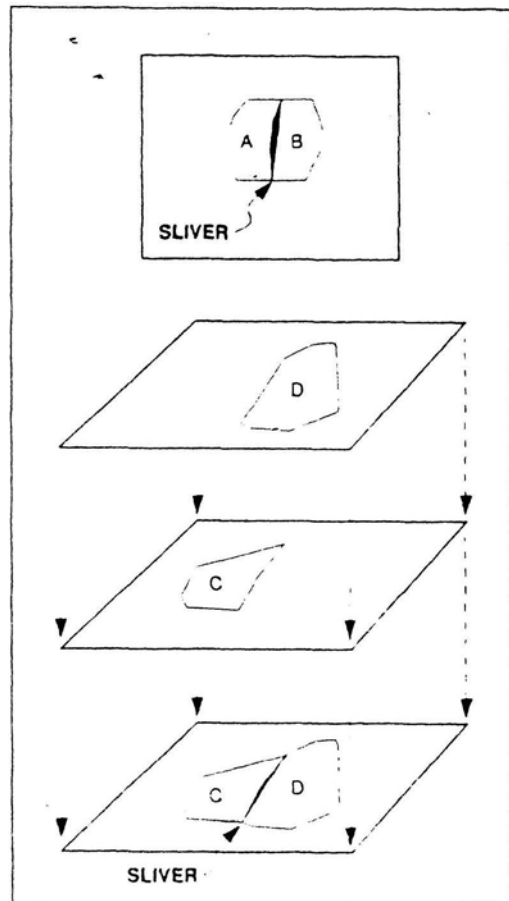
-Chuyển đổi dùng vị trí tuyệt đối: một phương pháp khác để chuyển đổi hình học của các lớp dữ liệu để hiệu chỉnh đến vị trí tuyệt đối của từng lớp. Ví dụ; như hệ tọa độ UTM.

□ Chức năng làm trùng khít (Conflation)

Việc làm trùng khít là quá trình làm trùng tọa độ của các đối tượng tương ứng trong các lớp dữ liệu khác nhau. Ví dụ; sai số nhỏ gây ra trong quá trình nhập dữ liệu vị trí địa lý của đối tượng trên bản đồ. Khi chồng hai lớp dữ liệu, cùng một đối tượng sẽ không trùng nhau một cách chính xác (sliver).

Chức năng này nhằm sửa chữa (chỉnh, thêm, bớt, thay đổi...) vị trí địa lý của đối tượng. Một số phần mềm (ARC/INFO) có khả năng xóa các sliver. (Xem hình 6.2).

Hình 6.2 Slivers có thể tạo ra trong quá trình số hóa



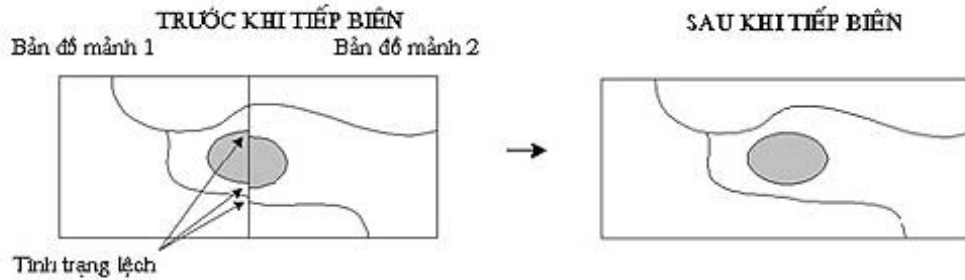
□ Ghép biên (Edge matching)

Điều chỉnh các sai số nhỏ tại biên các mảnh bản đồ. Sai số có thể do bản đồ gốc bị co giãn hoặc do số hóa. Xem hình 6.1

Khi khu vực nghiên cứu nằm trên nhiều mảnh bản đồ khác nhau thì thông tin nên được lưu giữ như những khoảng vị địa lý liên tục. Về mặt lý thuyết, dữ

liệu từ những bản đồ tiếp giáp nên gặp nhau một cách chính xác tại biên của bản đồ. Tuy nhiên những đặc tính cắt qua ranh giới có thể có sự gián đoạn nhỏ (xem hình 7). Phần không tương ứng của các đặc tính nên được gắn một cách hoàn hảo với nhau tạo ra cơ sở dữ liệu không gian phù hợp. Quá trình này được gọi là "tiếp biên" hoặc "kéo" và sẽ được thực hiện một cách tự động.

Khi tổ hợp các đối tượng từ các lớp dữ liệu khác nhau được lưu trữ trên một hệ quy chiếu chung và mô hình dữ liệu, vấn đề về sự gián đoạn sẽ nảy sinh.



Hình 7: Tiếp biên

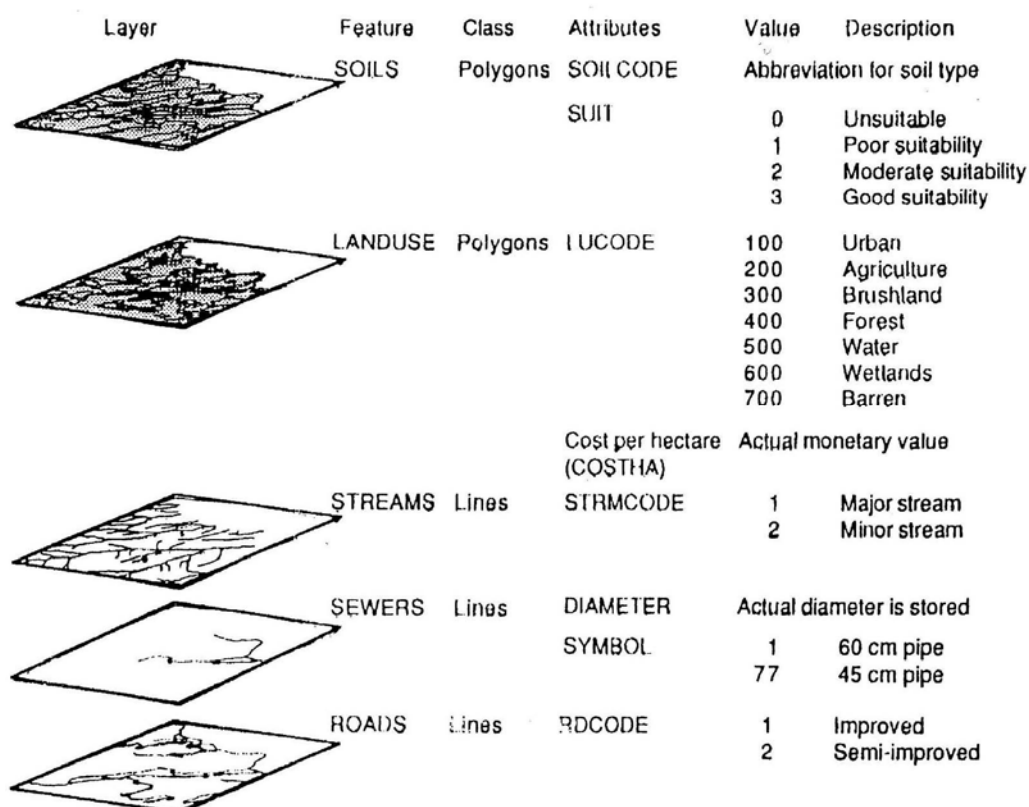
Chỉnh khớp là thủ tục khớp các đối tượng trên các lớp dữ liệu khác nhau. Ví dụ cùng một nhà thờ được lưu trữ trên 2 lớp dữ liệu khác nhau sẽ có thể được xác định trên 2 vị trí khác nhau vì lỗi do số hoá. Vẽ lại một vài đặc tính trên các lớp là phương pháp thủ công để giảm thiểu những lỗi này. Điều này được thực hiện bằng cách sử dụng một lớp tham chiếu làm mẫu vẽ trên các lớp khác. Những đối tượng không tương thích có thể được vẽ lại từ nền của bản đồ cơ sở, những đặc tính cần giữ lại có thể được sao chép từ bản đồ cơ sở sang các lớp khác. Một vài hệ thống tin địa lý cung cấp chức năng bán tự động để giải quyết vấn đề này. Sự tổ hợp dữ liệu trong các lớp khác nhau có thể đưa tới Slivers khi hai ranh giới tương ứng có một sự sai lệch nhỏ. Những sai lệch này sẽ được nhận biết và được sửa một cách tự động (Aronof 1989).

- ☐ Biên tập (Editing of graphic elements)
- ☐ Làm thưa tọa độ (Line coordinate thinning)

Thường tọa độ khi số hóa được nhập vào nhiều hơn cần thiết để xác định các đối tượng đường hoặc vùng. Một số phần mềm có chức năng làm giảm lượng tọa độ được lưu trữ trong hệ GIS. Xem hình 6.3.

5.4.4. Lưu trữ dữ liệu

Dữ liệu bản đồ số (digital map) gồm hai kiểu thông tin: về không gian và dữ liệu mô tả. Chúng được lưu trữ bằng một loạt File chứa đựng cả hai (dữ liệu không gian) và (mô tả về các đối tượng bản đồ) trong máy tính. Sức mạnh của GIS là sự liên kết hai kiểu dữ liệu và duy trì mối liên kết giữa các yếu tố.



Hình 5.3 Thông tin địa lý tổ chức theo lớp (nối kết dữ liệu không gian và thuộc tính)

5.4.5. Tổ chức dữ liệu

Các yếu tố bản đồ được tổ chức logic bằng một tập hợp các lớp. Một bản đồ được tổ chức thành các lớp thể hiện một thuộc tính giống nhau, như sông, địa chất, thổ nhưỡng, ô nhiễm nước vv...thêm vào những khu vực có diện tích nhỏ được phối hợp vào khu vực có diện tích lớn hơn như.

Tổ chức các lớp một cách khoa học sẽ giúp cho việc lưu trữ đơn giản, các File dữ liệu sẽ gọn hơn, máy tính sẽ chạy nhanh hơn, tổ chức phân tích và chất vấn dữ liệu đơn giản hơn. Việc tổ chức các lớp thông tin nói riêng, hay nói rộng ra là tổ chức một hệ thống cơ sở dữ liệu bản đồ theo từng chuyên đề là hết sức công phu, đòi hỏi nhiều kiến thức chuyên gia. Nhưng nó sẽ đem lại một hiệu quả to lớn cho người sử dụng. Khi nhấn vào một đối tượng quan tâm, ta sẽ có những thông tin cần thiết (cảnh ảnh, âm thanh, đoạn phim mô tả) vv...

Chất vấn dữ liệu (SQL) tạo ra các bản đồ dẫn xuất:

Trên cơ sở dữ liệu đã được xây dựng, cụ thể là bản đồ sẽ được thể hiện trên màn hình máy tính, gồm các lớp, các yếu tố, và các bảng thông tin mô tả. Người sử dụng có thể tìm hiểu các thông tin đầy đủ chỉ bằng cách nhấn mũi tên vào đối tượng hay đo khoảng cách, xem diện tích một hồ vv... Đây là

những thông tin đơn giản.

Đơn giản : Tìm tổng Số người làm việc với GIS ở Việt Nam.

Phức tạp hơn: (có điều kiện) tìm tất cả các điểm mỏ vật liệu xây dựng ở Tp. Hồ Chí Minh, có vĩ độ lớn hơn $10^{\circ}15' (?)$. Tìm những đô thị có diện tích lớn hơn $100\text{km}^2 (?)$ v v.... Để trả lời những câu hỏi này, máy tính sẽ phải loại bỏ các đối tượng trên cơ sở điều kiện đã nêu ra và tạo ra những bản đồ dẫn xuất.

Câu hỏi mô hình: Vẽ khu vực xung quanh con sông và cách bờ sông 1km, và tính diện tích khu vực đó.

Trên đây là một số câu hỏi và còn rất nhiều chất vấn khác ta có thể xây dựng các sản phẩm riêng, tùy vào phần mềm, và cách giải quyết vấn đề sẽ đem lại kết quả to lớn nhờ sự trợ giúp của GIS.

5.4.6. Cập nhật dữ liệu

Nhiều thông tin địa lý thay đổi theo thời gian như: ranh giới hành chính, biên sử dụng đất, đường giao thông, cơ sở hạ tầng...

Tùy thuộc vào mức độ thay đổi dữ liệu trong không gian và tầm quan trọng của sự sửa đổi mà việc cập nhật sẽ được thực hiện. Ảnh máy bay và ảnh vệ tinh sẽ giúp ích nhiều trong quá trình cập nhật.

5.4.7. Thao tác dữ liệu

Bao gồm việc đo đạc (khoảng cách giữa hai điểm, chu vi, diện tích, khối lượng). Phân tích thống kê (đường đi tối ưu, các phép nội suy). Chồng ghép bản đồ. Xem chương 8 Phân tích địa lý.

5.5. Xuất dữ liệu

5.5.1. Thành phần module đầu ra của GIS

- ☐ Cửa sổ (lựa chọn các diện tích để in bản đồ)
- ☐ Điều chỉnh tỷ lệ
- ☐ Tạo và sửa chữa ký hiệu
- ☐ Lựa chọn màu
- ☐ Làm chỉ dẫn
- ☐ Trình bày các bản vẽ

5.5.2. Thiết bị và sản phẩm đầu ra của GIS

Thiết bị đầu ra gồm màn hình, máy in và máy vẽ. Sản phẩm đặc trưng của GIS là tài liệu mà người dùng sử dụng trực tiếp bao gồm.

- ☐ Các bảng tài liệu
- ☐ Đồ Thị
- ☐ Bản đồ
- ☐ Ảnh
- ☐ Các mô hình 3D
- ☐ Dữ liệu xuất ra từ CSDL GIS trong dạng thích hợp đối với người sử dụng, có ba dạng sau:
 - ☐ Hardcopy: thông tin được in ra giấy, film,...
 - ☐ Softcopy: thông tin được xem trên màn hình máy tính (văn bản, đồ họa đơn sắc hoặc màu), softcopy có thể thay đổi nhưng việc xem thông tin bị hạn chế do kích thước màn hình.
 - ☐ Thông tin xuất trong dạng điện tử: gồm những tập tin máy tính.

5.5.3. Thiết bị xuất hardcopy gồm

- ☐ Pen plotter-máy vẽ dùng bút.
- ☐ Ink jet plotter – máy vẽ phun mực.
- ☐ Thermal plotter – máy vẽ nhiệt.
- ☐ Electrostatic plotter – máy vẽ tĩnh điện.
- ☐ Dot matrix printer – máy in kim.
- ☐ Lazer printer – máy in Lazer.
- ☐ Optical film writer – thiết bị ghi film.
- ☐ Screen copy device – thiết bị copy màn hình