

MỤC LỤC

MỤC LỤC.....	1
CHƯƠNG 1. TỔNG QUAN VỀ HỆ THỐNG THÔNG TIN ĐỊA LÝ.....	3
1.1. Tổng quan về hệ thống tin địa lý.....	3
1.2. GIS là gì	4
1.3. Các lĩnh vực khoa học công nghệ liên quan tới GIS.....	6
1.4. Các lĩnh vực ứng dụng GIS.....	6
1.5. Các thành tố của hệ thống thông tin địa lý.....	7
1.6. Chức năng của hệ thống thông tin địa lý.....	9
Câu hỏi ôn tập.....	15
1.Định nghĩa GIS. Các thành phần của GIS.....	15
CHƯƠNG 2. HỆ THỐNG THAM CHIẾU KHÔNG GIAN.....	16
2.1. Hệ tọa độ địa lý.....	16
2.2. Hệ tọa độ quy chiếu.....	17
2.3. Các phép chiếu cơ bản	17
Câu hỏi ôn tập.....	20
1.Trình bày các phép chiếu cơ bản.....	20
CHƯƠNG 3. MÔ HÌNH VÀ CẤU TRÚC DỮ LIỆU KHÔNG GIAN.....	21
3.1. Các khái niệm cơ sở.....	21
3.2. Mô hình thông tin không gian.....	21
3.3. Mô hình dữ liệu raster.....	22
3.4. Mô hình dữ liệu vectơ.....	25
3.5. Mô hình dữ liệu lưới tam giác không đều.....	26
3.6. Biến đổi từ vectơ sang raster và ngược lại.....	27
CHƯƠNG 4. XÂY DỰNG CƠ SỞ DỮ LIỆU GIS.....	32
4.1. Thu thập thông tin địa lý và các phương pháp nhập dữ liệu địa lý.....	32
4.2. Cơ sở dữ liệu thông tin địa lý.....	37
Câu hỏi ôn tập.....	41
1.Nêu các phương pháp thu thập dữ liệu địa lý.....	41
MỘT SỐ ĐỀ THI MẪU.....	42

CHƯƠNG 1. TỔNG QUAN VỀ HỆ THỐNG THÔNG TIN ĐỊA LÝ

1.1. Tổng quan về hệ thống tin địa lý

Địa lý (geography) được hình thành từ hai khái niệm: trái đất (geo-earth) và tiến trình mô tả (graphy). Như vậy, địa lý được xem như tiến trình mô tả trái đất. Khi mô tả trái đất, các nhà địa lý luôn đề cập đến quan hệ không gian. Chìa khóa của nghiên cứu các quan hệ không gian là bản đồ. Theo Hiệp hội Bản đồ Quốc tế thì bản đồ là biểu diễn bằng đồ họa tập các đặc trưng trừu tượng và các quan hệ không gian trên bề mặt trái đất. Nói một cách khác bản đồ là quá trình chuyển đổi từ thông tin bề mặt trái đất sang bản đồ giấy.

Hệ thống thông tin là tập các tiến xử lý dữ liệu thô để sản sinh ra các thông tin có ích cho công tác lập quyết định. Chúng bao gồm các thao tác dẫn chúng ta đi từ lập kế hoạch quan sát và thu thập dữ liệu tới lưu trữ và phân tích dữ liệu, tới sử dụng các thông tin suy diễn trong việc lập quyết định. Theo quan niệm này thì bản đồ cũng là một loại hệ thống tin. Bản đồ là tập hợp các dữ liệu, các thông tin suy diễn từ nó được sử dụng vào công việc lập quyết định. Hệ thống tin địa lý là hệ thống tin được thiết kế để làm việc với dữ liệu quy chiếu không gian hay tọa độ địa lý. Khái niệm hệ thống tin địa lý được hình thành từ ba khái niệm: địa lý, thông tin và hệ thống được viết tắt là GIS. Ý nghĩa của chúng được diễn giải như sau:

Geographic Information System (Mỹ).

Geographical Information System (Anh, Úc, Canada).

Geographic Information Science (nghiên cứu lý thuyết và quan niệm của hệ thống tin địa lý và các công nghệ thông tin địa lý).

Geographic Information Studies (nghiên cứu về ngữ cảnh xã hội của thông tin).

Khái niệm “địa lý” được sử dụng tại đây vì GIS trước hết liên quan đến các đặc trưng địa lý hay không gian. Các đặc trưng này được ánh xạ hay liên quan đến các đối tượng không gian. Chúng có thể là các đối tượng vật lý, văn hóa hay kinh tế trong tự nhiên. Các đặc trưng trên bản đồ là biểu diễn ảnh của các đối tượng không gian trong thế giới thực. Biểu tượng, màu và kiểu đường được sử dụng để thể hiện các đặc trưng không gian khác nhau trên bản đồ 2D.

Khái niệm “thông tin” đề cập đến khối dữ liệu khổng lồ do GIS quản lý. Các đối tượng thế giới thực đều có tập riêng các dữ liệu chữ-số thuộc tính hay đặc tính (còn gọi là dữ liệu phi hình học, dữ liệu thống kê) và các thông tin vị trí cần cho lưu trữ, quản lý các đặc trưng không gian.

Khái niệm “hệ thống” đề cập đến cách tiếp cận hệ thống của GIS. Mỗi trường hệ thống GIS được chia nhỏ thành các mô đun để dễ hiểu, dễ quản lý nhưng chúng được tích hợp thành hệ thống thống nhất, toàn vẹn. Công nghệ thông tin đã trở thành quan trọng, cần thiết cho tiệm cận này và hầu hết các hệ thống thông tin đều được xây dựng trên cơ sở máy tính.

Khái niệm “công nghệ thông tin địa lý” (geographic information technology) là công nghệ thu thập và xử lý thông tin địa lý. Chúng bao gồm ba loại cơ bản sau:

Hệ thống định vị toàn cầu (global positioning system-GPS): đo đạc vị trí trên mặt đất dựa trên cơ sở hệ thống các vệ tinh.

Viễn thám (remote sensing): sử dụng vệ tinh để thu thập thông tin về trái đất.

Hệ thông tin địa lý (GIS).

1.2. GIS là gì

Thông tin địa lý bao gồm dữ liệu về bề mặt trái đất và các diễn giải dữ liệu để con người dễ hiểu. Nhìn chung thì thông tin địa lý được thu thập từ bản đồ hay được thu thập thông qua đo đạc, viễn thám, điều tra, phân tích hay mô phỏng. Thông tin địa lý bao hàm hai loại dữ liệu: không gian và phi không gian.

1.2.1. Định nghĩa GIS

Hiện nay có rất nhiều định nghĩa về GIS, nhưng đều có điểm giống nhau như: bao hàm khái niệm dữ liệu không gian, phân biệt giữa hệ thông tin quản lý và GIS. So với bản đồ thì GIS có lợi thế là lưu trữ dữ liệu và biểu diễn chúng là hai công việc tách biệt nhau. Do vậy GIS cho khả năng quan sát từ các góc độ khác nhau trên cùng tập dữ liệu. Sau đây là một số định nghĩa GIS hay sử dụng:

Định nghĩa của dự án The Geographer's Craft, khoa địa lý, trường đại học Texas

GIS là CSDL số chuyên dụng trong đó hệ trục tạo độ không gian là phương tiện tham chiếu chính. GIS bao gồm các công cụ để thực hiện các công việc sau đây:

Nhập dữ liệu từ bản đồ giấy, ảnh vệ tinh, ảnh máy bay, số liệu điều tra và các nguồn khác.

Lưu trữ dữ liệu, khai thác, truy vấn CSDL.

Biến đổi dữ liệu, phân tích, mô hình hóa, bao gồm các dữ liệu thống kê và dữ liệu không gian.

Lập báo cáo, bao gồm các bản đồ chuyên đề, các bảng biểu, biểu đồ và kế hoạch.

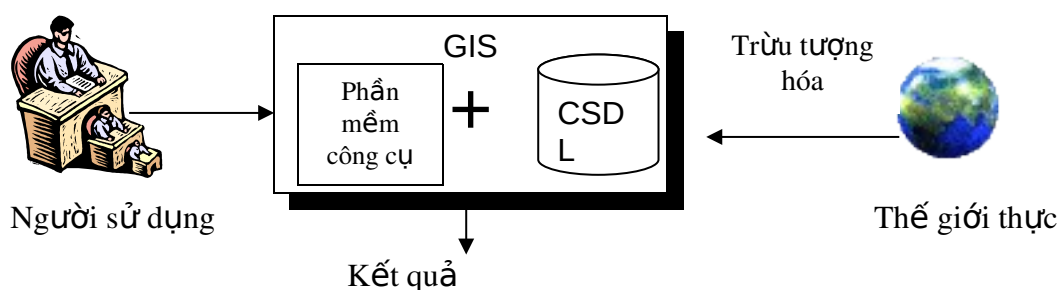
Từ định nghĩa trên ta thấy rõ ba vấn đề sau của GIS. Thứ nhất, GIS có quan hệ với ứng dụng CSDL, toàn bộ thông tin trong GIS đều liên kết với tham chiếu không gian, CSDL GIS sử dụng tham chiếu không gian như phương tiện chính để lưu trữ và xâm nhập thông tin. Thứ hai, GIS là công nghệ tích hợp, hệ GIS đầy đủ có các khả năng phân tích bao gồm phân tích ảnh máy bay, ảnh vệ tinh hay tạo lập mô hình thống kê, vẽ bản đồ... Cuối cùng, GIS được không chỉ xem như tiến trình phần cứng, phần mềm rời rạc mà còn được sử dụng vào trợ giúp quyết định.

Định nghĩa của David Cowen, Mỹ

GIS là hệ thống phần cứng, phần mềm và các thủ tục được thiết kế để thu thập, quản lý, xử lý, phân tích, mô hình hóa và hiển thị các dữ liệu qui chiếu không gian để giải quyết các vấn đề quản lý và lập kế hoạch phức tạp.

Độ phức tạp của thế giới thực là không giới hạn, song con người luôn mong lưu trữ, quản lý các dữ liệu về thế giới thực thế nên phải có CSDL lớn vô hạn để lưu trữ thông tin chính xác về chúng. Do vậy, để lưu trữ được dữ liệu không gian của thế giới thực và máy tính thì phải giảm số lượng dữ liệu đến mức có thể quản lý được bằng tiến trình đơn giản hóa hay trừu tượng hóa. Trừu tượng là đơn giản một cách thông minh, trừu tượng cho ta tổng quát hóa và ý

tưởng hóa vấn đề đang xem xét, loại bỏ các chi tiết dư thừa mà chỉ tập trung vào các điểm chính, cơ bản. Các đặc trưng đại lý phải được biểu diễn bởi các thanh phần rời rạc hay các đối tượng để lưu vào CSDL máy tính.



Hình 1.1: Hệ thống tin địa lý

Ý nghĩa chủ yếu của tin học hóa thông tin địa lý là khả năng tích hợp các kiểu và nguồn dữ liệu khác biệt. Mục tiêu của GIS là cung cấp cấu trúc một cách hệ thống để quản lý các thông tin địa lý khác nhau và phức tạp, đồng thời cung cấp các công cụ, các thao tác hiển thị, truy vấn, mô phỏng... Cái GIS cung cấp là cách thức suy nghĩ mới về không gian. Phân tích không gian không chỉ là truy cập mà còn cho phép khai thác các quan hệ và tiến trình biến đổi của chúng. GIS lưu trữ thông tin thế giới thực thành các tầng bản đồ chuyên đề mà chúng có khả năng liên kết địa lý với nhau.

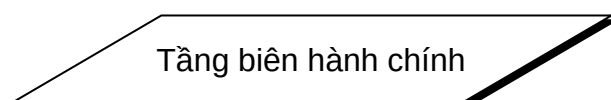
1.2.2. Cơ sở dữ liệu GIS

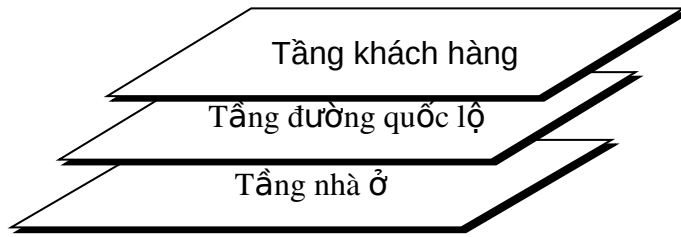
CSDL GIS là một nhóm xác định các dữ liệu trong một cấu trúc của một phần mềm quản lý CSDL, đó là tập hợp của các dữ liệu không gian và phi không gian.

- **Dữ liệu không gian:** là những mô tả số của hình ảnh bản đồ, chúng bao gồm tọa độ, quy luật và các ký hiệu dùng để xác định một hình ảnh cụ thể trên bản đồ.
- **Dữ liệu phi không gian (dữ liệu thuộc tính):** diễn tả đặc tính, số lượng, mối quan hệ của các hình ảnh bản đồ với vị trí địa lý của chúng.

Dữ liệu không gian	Dữ liệu phi không gian
Dạng thông tin	
Điểm, đường, vùng, ghi chú.	Thuộc tính, tham khảo địa lý, chỉ số địa lý, các quan hệ không gian.
Lưu trữ dạng	
Tọa độ, ký hiệu, chấm điểm, quy luật hiển thị.	Chữ số, ký tự...

Ví dụ tại thành phố Hải Phòng. Chúng ta có thể tách bản đồ ra thành các tầng như trong hình vẽ dưới đây:





Hình 1.2: Tầng bản đồ

1.3. Các lĩnh vực khoa học công nghệ liên quan tới GIS

GIS được xây dựng dựa trên các tri thức của nhiều ngành khoa học khác nhau để tạo ra một ngành khoa học mới. Trong đó:

Ngành địa lý: là ngành liên quan mật thiết tới vấn đề hiểu thế giới và vị trí của con người trong thế giới, cung cấp các kỹ thuật phân tích không gian.

Ngành bản đồ: bản đồ chính là dữ liệu đầu vào của GIS đồng thời cũng là khuôn mẫu quan trọng nhất của đầu ra GIS.

Công nghệ viễn thám, ảnh máy bay: ảnh viễn thám và ảnh máy bay là nguồn dữ liệu quan trọng của GIS. Viễn thám bao gồm cả kỹ thuật thu thập và xử lý dữ liệu mọi vị trí trên địa cầu với giá rẻ, ảnh máy bay và kỹ thuật đo chính xác của chúng là nguồn dữ liệu chính xác về độ cao bề mặt trái đất sử dụng làm đầu vào của GIS.

Bản đồ địa hình: cung cấp dữ liệu có chất lượng cao về vị trí của ranh giới đất đai, nhà cửa...

Ngành đo đạc, thống kê: cung cấp các vị trí cần quản lý và các phương pháp phân tích dữ liệu GIS. Ngành thống kê đặc biệt quan trọng trong việc hiểu các lỗi hoặc tính không chắc chắn trong dữ liệu GIS.

Khoa học tính toán: tự động thiết kế bằng máy tính cung cấp các kỹ thuật nhập, hiển thị, biểu diễn dữ liệu. Máy tính hoạt động như một chuyên gia trong việc thiết kế bản đồ, phát sinh các đặc trưng bản đồ.

Toán học: các ngành như hình học, đồ thị được sử dụng trong thiết kế và phân tích dữ liệu không gian.

1.4. Các lĩnh vực ứng dụng GIS

Công nghệ GIS ngày càng được sử dụng rộng rãi. GIS có khả năng sử dụng dữ liệu không gian và phi không gian từ các nguồn khác nhau khi thực hiện phân tích không gian để trả lời các câu hỏi của người dùng như:

Có cái gì ở...? Nhận diện: nhận biết tên hay các thông tin khác của đối tượng nào đó trên bản đồ.

... ở đâu? Vị trí: chỉ ra một hoặc nhiều vị trí thỏa mãn yêu cầu người dùng.

Cái gì thay đổi từ...? Xu thế: câu hỏi liên quan trực tiếp đến các dữ liệu không gian tạm thời, ví dụ như câu hỏi liên quan tới phát triển thành phố sẽ đưa ra các vùng qui hoạch chính trên bản đồ GIS.

Đường đi nào tốt nhất từ... đến...? Tìm đường đi tối ưu: dựa trên cơ sở mạng lưới của đường đi cho biết đường đi nào là rẻ nhất, ngắn nhất... mở rộng ra là đường đi qua một hệ thống điểm.

Giữa... và ... có quan hệ gì? Mẫu: câu hỏi này khác phức tạp tác động trên nhiều tập dữ liệu như quan hệ giữa vị trí nhà máy và địa phương, khí hậu và vùng sản xuất...

Cái gì xảy ra nếu...? Mô hình: đây là câu hỏi liên quan đến các hoạt động lập kế hoạch và dự án như khi nâng cấp hệ thống giao thông tại Hải Phòng thì ảnh hưởng thế nào tới mạng lưới cung cấp điện, điện thoại, nước, dân cư...

Dưới đây là một vài ứng dụng chủ yếu của GIS trong thực tế:

Quản lý và lập kế hoạch mạng lưới đường phố: bao gồm các chức năng tìm kiếm địa chỉ, tìm vị trí khi biết địa chỉ đường phố; điều khiển đường đi, lập kế hoạch lưu thông xe cộ; phân tích vị trí, chọn địa điểm xây dựng các công trình công cộng; lập kế hoạch phát triển đường giao thông.

Giám sát tài nguyên thiên nhiên, môi trường: bao gồm chức năng quản lý tài nguyên, phân tích tác động môi trường...

Quản lý đất đai: lập kế hoạch dùng, miễn sử dụng đất; quản lý tưới tiêu...

Quản lý và lập các dịch vụ công cộng: bao gồm các chức năng tìm địa điểm cho các công trình ngầm; quản lý, bảo dưỡng công trình...

Phân tích tổng điều tra dân số, lập bản đồ các dịch vụ y tế, bưu điện và các dịch vụ công cộng khác...

1.5. Các thành tố của hệ thống thông tin địa lý

Hệ thống GIS bao gồm năm thành tố chính: con người, phương pháp, công cụ phần cứng, phần mềm và dữ liệu.

1.5.1. Con người

Con người ở đây là các chuyên viên tin học, chuyên gia GIS, thao tác viên GIS, phát triển ứng dụng GIS bao gồm:

Người sử dụng hệ thống: là những người sử dụng GIS để giải quyết các vấn đề không gian. Nhiệm vụ chủ yếu của họ là số hóa bản đồ, kiểm tra lỗi, soạn thảo, phân tích các dữ liệu thô và đưa ra các giải pháp cuối cùng để truy vấn dữ liệu địa lý. Những người này phải thường xuyên được đào tạo lại do GIS thay đổi liên tục và yêu cầu mới của kỹ thuật phân tích.

Thao tác viên hệ thống: có trách nhiệm vận hành hệ thống hàng ngày để người sử dụng hệ thống làm việc hiệu quả. Công việc của họ là sửa chữa khi chương trình bị tắc nghẽn hay là công việc trợ giúp nhân viên thực hiện các phân tích có độ phức tạp cao. Họ còn làm việc như quản trị hệ thống, quản trị CSDL, an toàn, toàn vẹn CSDL tránh hư hỏng, mất mát dữ liệu.

Nhà cung cấp GIS: cung cấp các phần mềm, cập nhật phần mềm, phương pháp nâng cấp cho hệ thống.

Nhà cung cấp dữ liệu: là các cơ quan nhà nước hay tư nhân cung cấp các dữ liệu sửa đổi từ nhà nước.

Người phát triển ứng dụng: là những lập trình viên, họ xây dựng giao diện người dùng, giảm khó khăn các thao tác cụ thể trên hệ thống GIS...

Chuyên viên phân tích hệ thống GIS: là nhóm người chuyên nghiên cứu thiết kế hệ thống, được đào tạo chuyên nghiệp có trách nhiệm xác định các mục tiêu của hệ GIS trong cơ quan, hiệu chỉnh hệ thống, đề xuất kỹ thuật phân tích đúng đắn...

1.5.2. Dữ liệu

Dữ liệu thống kê gắn theo các hiện tượng tự nhiên với những mức độ chính xác khác nhau. Hệ thống thước đo của chúng bao gồm các biến tên, số thứ tự, khoảng và tỉ lệ. Trong đó:

Biến tên: những biến chỉ có tên, không theo một trật tự nào cả, ví dụ như các loại đất (công viên, vùng dân cư, đất công nghiệp...) , loại cây trồng (ngô, khoai, sắn)...

Biến thứ tự là danh sách các lớp rời rạc nhưng có trật tự như trình độ học vấn (tiểu học, trung học, đại học, sau đại học), độ lớn (nhỏ, trung bình, lớn)... các giá trị ở đây chỉ là phản ánh một cách tương đối không chính xác số lượng vì vậy không thể thực hiện các phép tính toán được.

Biến khoảng cũng có trình tự tự nhiên nhưng khoảng cách của chúng có ý nghĩa như nhiệt độ, diện tích.

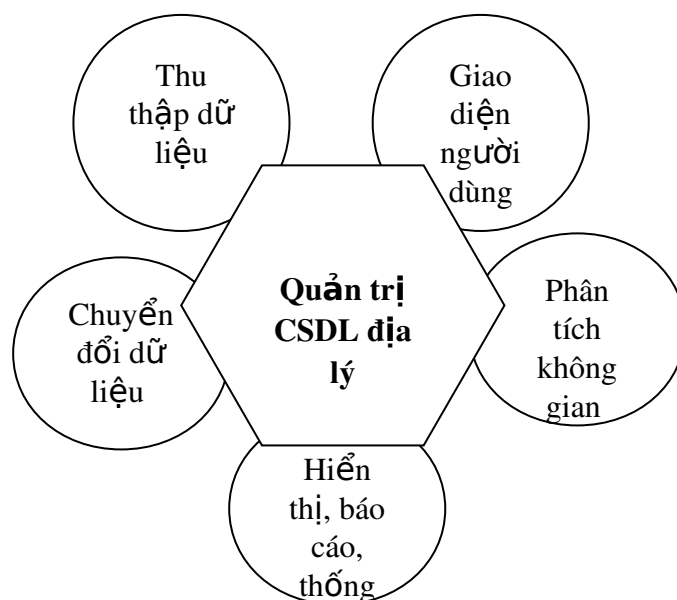
Biến tỷ lệ có cùng đặc tính như biến khoảng nhưng chúng có giá trị 0 tự nhiên hay điểm bắt đầu như lượng mưa, dân số.

Ngoài bốn loại dữ liệu trên GIS còn phân chia dữ liệu thành hai lớp khác nhau là không gian và phi không gian. Ví dụ như nhà hát lớn Hải Phòng, giá trị cấp kinh độ, vĩ độ là dữ liệu không gian dạng đơn giản nhất và các thông tin khác như khối lượng khí lưu thông, kết cấu thép... là dữ liệu thuộc tính hay phi không gian. Mỗi hệ GIS đều có kết nối giữa hai loại dữ liệu này.

Hệ GIS cần phải hiểu được dữ liệu trong các khuôn mẫu khác nhau không chỉ riêng khuôn dữ liệu riêng của hệ thống. Ví dụ như đường biên bản đồ có thể trong khuôn mẫu tệp DXF của AutoCad hay BNA của AtlasGis. Thông thường, GIS hiểu ngay khuôn mẫu DXF mà không cần sửa đổi đồng thời GIS phải hiểu ngay khuôn mẫu DBF của các thuộc tính được lưu trữ kèm theo. Phần mềm GIS lý tưởng đọc được các dữ liệu raster (DEN, GIFF, TIFF, JPEG, EPS) và khuôn mẫu vectơ (TIGER, HPGL, DXF, DLG, Postscript) một số phần mềm GIS chỉ có chức năng nhập dữ liệu vào các cấu trúc dữ liệu đơn giản như cấu trúc thực thể, cấu trúc tô pô. Với dữ liệu ba chiều, phần lớn phần mềm GIS trợ giúp lưới tam giác không đều (TIN), một số khác trợ giúp cấu trúc raster trên cơ sở lưới bao vây và cây tứ phân, số còn lại xây dựng một khuôn mẫu riêng cho mình tùy vào nhà sản xuất phần mềm nhưng thường là theo khuôn mẫu chuẩn quốc gia, quốc tế như SDTS (Spatial Data Transfer Santard) hay DIGEST.

1.5.3. Phần mềm

Một hệ thống GIS bao gồm nhiều môđun phần mềm. Khả năng lưu trữ, quản lý dữ liệu không gian bằng hệ quản trị CSDL địa lý là khía cạnh quan trọng nhất của GIS. Một phần mềm xử lý GIS tốt phải cung cấp cho người sử dụng các công cụ quản lý, phân tích không gian dễ dàng, chính xác.



Hình 1.3: Phần mềm GIS

1.5.4. Phần cứng

GIS đòi hỏi các thiết bị ngoại vi đặc biệt như bàn số hóa, máy vẽ, máy quét ảnh vào/ra. Các thiết bị có thể được nối với nhau thông qua thiết bị truyền tin hay mạng cục bộ.

1.5.5. Giao diện người dùng

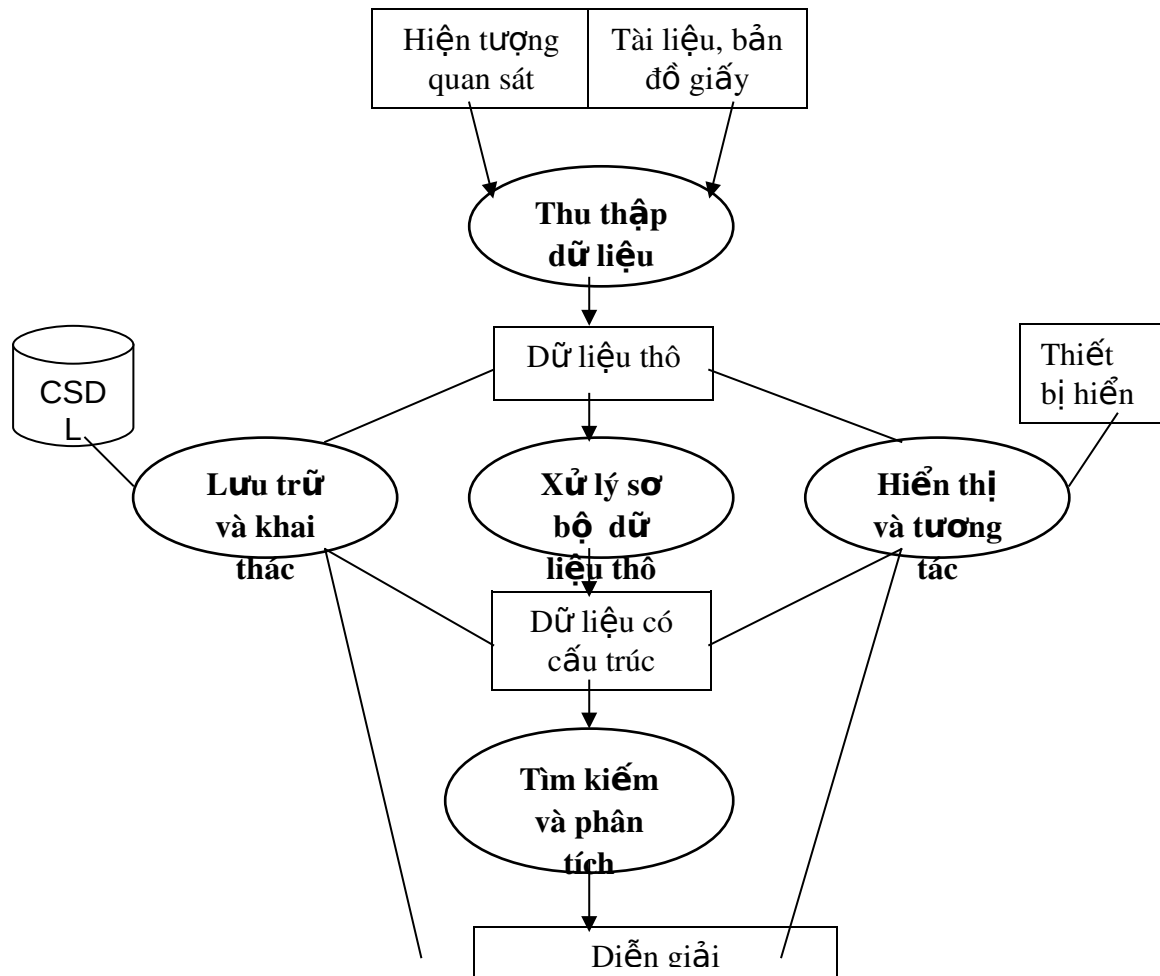
Giao diện đồ họa cho phép người dùng dễ dàng thực hiện các thao tác địa lý và các thao tác khác như truy nhập CSDL, làm báo cáo...

1.6. Chức năng của hệ thống thông tin địa lý

Các chức năng của GIS có thể chia làm năm loại như sau:

- Thu thập dữ liệu.
- Xử lý sơ bộ dữ liệu.
- Lưu trữ và truy nhập dữ liệu.
- Tìm kiếm và phân tích không gian
- Hiển thị đồ họa và tương tác.

Sức mạnh của hệ thống GIS khác nhau là khác nhau, kỹ thuật xây dựng các chức năng trên cũng rất khác nhau. Hình 1.4 mô tả quan hệ giữa các nhóm chức năng và cách biểu diễn thông tin khác nhau của GIS.



Hình 1.4: Các nhóm chức năng của GIS

1.6.1. Thu thập dữ liệu

Thu thập dữ liệu là quá trình thu nhận dữ liệu theo khuôn mẫu áp dụng được cho GIS. Mức độ đơn giản nhất của thu thập dữ liệu là chuyển đổi khuôn mẫu dữ liệu có sẵn từ bên ngoài. Trong trường hợp này GIS phải có mô đun chương trình hiểu được các khuôn mẫu dữ liệu chuẩn như DLG, DXF hay các dữ liệu đầu ra của GIS như Mapinfo, Arc Info, MapObject... GIS còn phải có khả năng nhập các ảnh bản đồ trong khuôn mẫu GIFF, JPEG... Trên thực tế nhiều kỹ thuật trắc đạc được áp dụng để thu thập dữ liệu như qua vệ tinh, máy bay, số hóa những bản đồ giấy...

Phần lớn dữ liệu không gian là các bản đồ giấy, GIS phải số hóa chúng mới sử dụng được, trình tự số hóa bao gồm:

Mã hóa dữ liệu: là tiến trình gắn thuộc tính của đặc trưng vào toàn bộ đối tượng hình học trên bản đồ, chúng có thể là điểm, đường, vùng... công việc này thường được thực hiện qua nhập bàn phím.

Kiểm chứng và sửa lỗi là so sánh hình vẽ từ dữ liệu số hóa với tài liệu nguồn. Phải đảm bảo mọi đặc trưng trên bản đồ được số hóa với độ chính xác cần thiết.

Nhìn chung công việc thu thập dữ liệu là nhiệm vụ khó khăn và nặng nề nhất trong quá trình xây dựng một ứng dụng GIS.

1.6.2. Xử lý dữ liệu thô

Hai khía cạnh chính của xử lý dữ liệu thô bao gồm:

Phát sinh dữ liệu có cấu trúc tô pô.

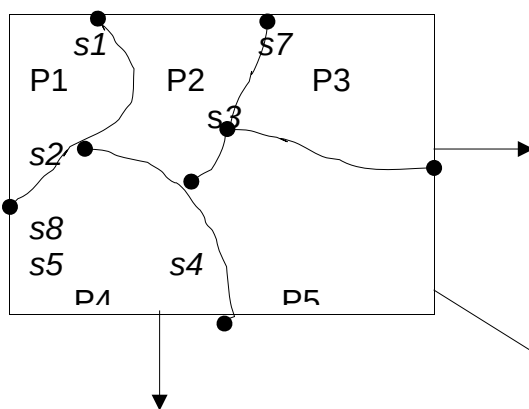
Với dữ liệu ảnh vệ tinh thì phải phân lớp các đặc trưng trong ảnh thành các hiện tượng quan tâm.

Mô hình quan niệm của thông tin không gian bao gồm mô hình hướng đối tượng, mạng và bề mặt. Quá trình phân tích trên cơ sở khác nhau đòi hỏi dữ liệu phải được biểu diễn và tổ chức cho phù hợp. Điều này đòi hỏi không chỉ chức năng tọa lập mô hình dữ liệu vectơ có cấu trúc tô pô và mô hình dữ liệu raster, mà còn có khả năng thay đổi cách biểu diễn, thay đổi phân lớp và sơ đồ mẫu, làm đơn giản hóa hay tổng quát hóa dữ liệu, biến đổi giữa hệ thống trục tọa độ khác nhau và biến đổi các phép chiếu bản đồ.

Một số công cụ phân tích của GIS phụ thuộc chặt chẽ vào mô hình dữ liệu raster, do vậy nó đòi hỏi quá trình biến đổi mô hình dữ liệu vectơ sang raster, hay raster hóa. Biến đổi từ raster sang mô hình vectơ, hay vectơ hóa, đặc biệt cần thiết khi tự động quét ảnh. Raster hóa là tiến trình chia đường hay vùng thành các tế bào. Ngược lại, vectơ hóa là tập hợp các tế bào (pixel) lại thành đường hay vùng. Khi so sánh dữ liệu từ các nguồn khác nhau, vấn đề chung nảy sinh là sử dụng hai hay nhiều phân lớp hay sơ đồ mã hóa cho cùng hiện tượng. Để nhận ra các khía cạnh khác nhau của hiện tượng với dữ liệu có mức độ chi tiết khác nhau, cần phải có tiến trình xấp xỉ hóa hay đơn giản hóa để biến đổi về cùng một sơ đồ. Một vấn đề nữa nảy sinh khi tích hợp dữ liệu bản đồ là hệ thống tọa độ của chúng được đo vẽ trên cơ sở nhiều phép chiếu bản đồ khác nhau. Các dữ liệu không thể tích hợp trên cùng bản đồ nếu không biến đổi chúng về cùng hệ trục tọa độ nên phải đưa chúng cùng về một hệ tọa độ địa lý.

1.6.3. Lưu trữ và truy cập dữ liệu

Chức năng lưu trữ dữ liệu trong GIS liên quan đến tạo lập CSDL không gian. Nội dung của CSDL này có thể tổ hợp dữ liệu vectơ và/hoặc, dữ liệu raster, dữ liệu thuộc tính để nhận danh hiện tượng tham chiếu không gian. Thông thường dữ liệu thuộc tính của GIS trên cơ sở đối tượng được lưu trong bảng, chúng chứa chỉ danh duy nhất, tương ứng với đối tượng không gian, kèm theo rất nhiều mục dữ liệu thuộc tính khác nhau. Chỉ danh đối tượng không gian duy nhất được dùng để liên kết giữa dữ liệu thuộc tính và dữ liệu không gian tương ứng. Đôi khi mục dữ liệu trong bảng thuộc tính bao gồm cả giá trị không gian như độ dài đường, diện tích vùng mà chúng đã được dẫn xuất từ biểu diễn dữ liệu hình học. Ví dụ dưới đây là về dữ liệu đất:



ID	Số hiệu	Chủ đất	Thời gian
P1	123	Thành	1982
P2	523	Đạt	1988
P3	642	Nghĩa	1953
...

dữ liệu phi không gian

ID	Tọa độ đoạn thẳng
P1	(x1,y1) (x2, y2); (x2,y2) (x5, y5)
P2
P3
...	...

dữ liệu hình học (không gian)

Hình 1.5: Liên kết dữ liệu không gian và phi không gian

Trong mô hình dữ liệu raster thì các tệp thuộc tính thông thường chứa dữ liệu liên quan tới hiện tượng tự nhiên thay cho đối tượng rời rạc. Việc lựa chọn mô hình raster hay vectơ để tổ chức dữ liệu không gian được thực hiện khi thu thập dữ liệu vì mỗi mô hình tương ứng với cách tiếp cận, mô tả thông tin khác nhau. Tuy nhiên rất nhiều CSDL của GIS cho khả năng quản trị cả hai mô hình không gian nói trên, khi xây dựng CSDL không gian thì cần thiết phải liên kết bảng dữ liệu liên quan đến hiện tượng tự nhiên tương ứng.

Công nghệ CSDL truyền thống không thích hợp để quản lý dữ liệu địa lý. Một số hệ thống GIS xây dựng CSDL dựa trên tổ hợp các mô hình:

Mô hình quan hệ quản lý thuộc tính phi hình học.

Lược đồ chuyên dụng, phi quan hệ để lưu trữ, xử lý dữ liệu không gian.

Một số khác lợi dụng các phương tiện của lược đồ lưu trữ CSDL quan hệ để quản lý cả hai loại dữ liệu: hình học và phi hình học song tương đối khó khăn, phức tạp.

Phương tiện truy nhập trong CSDL GIS cần bao gồm cả phương tiện có sẵn của CSDL quan hệ chuẩn để truy vấn tới một thuộc tính của đối tượng nào đó, thông tin của đối tượng trong khoảng nào đó... Đặc biệt trong CSDL GIS là khả năng xác định dữ liệu theo vị trí và theo các quan niệm không gian, đây là nền tảng quá trình xâm nhập CSDI của hệ GIS.

1.6.4. Tìm kiếm và phân tích không gian

Đây là chức năng đóng vai trò rất quan trọng trong GIS. Nó tạo nên sức mạnh thực sự của GIS so với các phương pháp khác. Tìm kiếm và phân tích dữ liệu không gian giúp tìm ra những đối tượng đồ họa theo các điều kiện đặt ra hay hỗ trợ việc ra quyết định của người dùng GIS. Có rất nhiều các phương pháp tìm kiếm và phân tích dữ liệu không gian, các phương pháp khác nhau thường tạo ra các ứng dụng GIS khác nhau. Sau đây là một số phương pháp được dùng phổ biến nhất:

Tìm kiếm dữ liệu trong vùng không gian (Buffer)

Buffer được sử dụng trong việc xác định các đối tượng xung quanh một hay nhiều các điểm mốc. Quá trình thực hiện bao gồm việc tạo ra một vùng đệm quanh các điểm mốc đó và sau đó xác định các đối tượng căn cứ vào vị trí của chúng so với vùng đệm này. Một bài toán rất điển hình cho phương pháp Buffer này là bài toán về “Nhà máy hoá chất và các bệnh viện”. Mục đích của bài toán là xác định các vị trí thuận tiện nhất trên bản đồ cho việc di dời các bệnh viện trong trường hợp nhà máy hoá chất gặp sự cố. Các nhà máy hoá chất và bệnh viện được biểu

diễn trên bản đồ bằng các đối tượng điểm. Mỗi nhà máy bao gồm các thông tin chi tiết về loại hoá chất sản xuất và mức độ phát tán chất độc ra môi trường trong các điều kiện thời tiết khác nhau. Khi có sự cố, vùng nguy hiểm cần di dời sẽ được thể hiện trên bản đồ. Từ đó, chúng ta có thể biết được nên chuyển bệnh viện đến vùng nào là an toàn và thuận tiện nhất.

Tìm kiếm theo địa chỉ (Geocoding)

Một đối tượng trên bản đồ bao giờ cũng được biểu diễn bằng một kiểu dữ liệu đồ hoạ. Phần đồ hoạ này có thể thu được bằng cách số hoá hay quét ảnh bản đồ. Tuy nhiên, khi ta đã có bản đồ (bản đồ số), chúng ta cũng có thể xác định được phần đồ hoạ biểu diễn đối tượng hay là vị trí, hình dạng của đối tượng thông qua các dữ liệu mô tả vị trí của nó ví dụ: số nhà, tên đường, tên quận...

Geocoding (hay address matching) là một tiến trình nhằm xác định các đối tượng trên cơ sở mô tả vị trí của chúng. Đây là một kỹ thuật rất nổi tiếng, có mặt trong rất nhiều ứng dụng của GIS. Người ta gọi một geocoding service là quá trình chuyển đổi toàn bộ mô tả thuộc tính về vị trí sang mô tả không gian. Để tìm được vị trí thông qua địa chỉ, geocoding service phải tham chiếu đến ít nhất một nguồn dữ liệu bao gồm cả thông tin về địa chỉ (thuộc tính) và thông tin không gian (vị trí, hình dạng). Dữ liệu này được gọi là dữ liệu tham chiếu. Các geocoding service có thể thao tác trên nhiều kiểu dữ liệu tham chiếu khác nhau.

Sau khi đã geocoding dữ liệu tham chiếu (tức là ánh xạ mô tả thuộc tính vào mô tả không gian) ta có thể nhập địa chỉ của đối tượng cần tìm. Quy trình xử lý trải qua các bước sau:

Chuẩn hoá giá trị địa chỉ vừa nhập vào bằng cách tách nó thành các thành phần địa chỉ nhỏ.

Geocoding service sau đó sẽ tìm trong nguồn dữ liệu tham chiếu để xác định các đối tượng có các thành phần địa chỉ tương ứng với dữ liệu nhập vào.

Tập kết quả trả về sẽ được gán các trọng số (điểm) để tìm ra kết quả gần đúng nhất.

Geocoding service sẽ đánh dấu đối tượng vừa được tìm thấy trên bản đồ bằng một đối tượng đồ hoạ.

Phân tích mạng (Networks)

Networks là kỹ thuật được ứng dụng rất rộng rãi trong giao thông, phân phối hàng hoá và dịch vụ, vận chuyển nước hay xăng dầu trong các đường ống dài... Trong GIS, networks được mô hình dưới dạng các đồ thị một chiều hay mạng hình học. Mạng hình học này bao gồm các đối tượng đang được hiển thị trên bản đồ, mỗi đối tượng đóng vai trò là cạnh hoặc nút trong mạng.

Trong GIS để thiết lập nên mối quan hệ giữa nút - cạnh và cạnh - cạnh ta cần tạo các tô pô cho cơ sở dữ liệu. Tô pô được hiểu là mối quan hệ giữa các đối tượng trong bảng dữ liệu. Quan hệ tô pô giữa các đối tượng gần giống quan hệ relationship giữa các bảng. Chúng ta có hai kiểu luật liên kết là nút - cạnh và cạnh - cạnh. Nút - cạnh là luật liên kết được thiết lập giữa một nút của đối tượng A với một cạnh của đối tượng B. Cạnh - cạnh là luật liên kết giữa một cạnh của đối tượng A và một cạnh của đối tượng B qua một tập các nút. Khi đã tạo tô pô và xác

lập luật liên kết, một mạng logic đã được hình thành. Lúc này ta có thể áp dụng các thuật toán về mạng để giải quyết các bài toán đặt ra.

Phủ chòm hay chồng bản đồ (Overlay)

Overlay là quá trình chồng khít hai lớp dữ liệu bản đồ với nhau để tạo ra một lớp bản đồ mới. Đây là kỹ thuật khó nhất và cũng là mạnh nhất của GIS. Overlay cho phép ta tích hợp dữ liệu bản đồ từ hai nguồn dữ liệu khác nhau. Điều này tương tự như việc nhân hai ma trận để tạo ra một ma trận mới, truy vấn hai bảng cơ sở dữ liệu để tạo ra bảng mới, với overlay là gộp hai lớp trên bản đồ để tạo ra bản đồ mới. Overlay thực hiện điều này bằng cách kết hợp thông tin một lớp này với một lớp khác để lấy ra dữ liệu thuộc tính từ một trong hai lớp. Người ta chia overlay thành ba dạng phân tích khác nhau:

Point-in-polygon: chồng khít hai lớp point và polygon, đầu ra là lớp point

Line-in-polygon: chồng khít hai lớp line và polygon, đầu ra là lớp line.

Polygon-in-polygon: chồng khít hai lớp polygon và polygon, đầu ra là lớp polygon.

Hai lớp đưa vào overlay phải có sự thống nhất với nhau. Thống nhất về hệ quy chiếu, thống nhất về tỷ lệ, có được điều kiện này ta mới tiến hành overlay được. Quá trình overlay thường được tiến hành qua 2 bước: Xác định tọa độ các giao điểm và tiến hành chồng khít hai lớp bản đồ tại giao điểm này và kết hợp dữ liệu không gian, thuộc tính của hai lớp bản đồ. Các phép toán overlay bao gồm: phép hợp (Union), phép giao (Intersect) và phép đồng nhất (Identity).

Phân tích biên (Boundary)

Phân tích đường biên của các đối tượng như giao điểm, đôi khi việc xác định giao điểm giúp cho việc sửa lỗi song xác định giao điểm giữa các biên là khá khó khăn.

Tìm kiếm trong khoảng cận kề (Proximity): có 3 phương pháp.

Phương pháp thứ nhất được xem như mở rộng của tìm kiếm dữ liệu trong vùng, trong đó vùng tìm kiếm được xác định bởi xấp xỉ tới hiện tượng có sẵn. Việc tìm kiếm này được thực hiện trong vùng tạo bởi mở rộng đối tượng cho trước theo một khoảng cách cho trước. Trong GIS vùng này được gọi là vùng đệm, nó được xây dựng xung quanh đối tượng điểm, đối tượng đường hay đối tượng vùng.

Phương pháp thứ hai của tìm kiếm cận kề là tìm ra các vùng nối trực tiếp với đối tượng xác định trước, chẳng hạn như tìm các mảnh đất liền kề với mảnh đất sẽ xây dựng nhà máy.

Phương pháp thứ ba của tìm kiếm cận kề xảy ra khi cần phải tìm kiếm những vùng gần nhất tới tập các vị trí mẫu phân tán không đều. Các mẫu thường là các điểm. Tìm kiếm này thực hiện bằng cách tạo lập đa giác Thiessen, nó xác định các vùng xung quanh mỗi điểm mà gần điểm này hơn mọi điểm khác. Sơ đồ đa giác Thiessen còn được gọi là sơ đồ Voronoi, chúng được sử dụng để lập ra bản đồ sử dụng từ các mẫu đất cách biệt.

1.6.5. Hiện thị đồ họa và tương tác

Tầm quan trọng bản chất không gian của thông tin địa lý là đặc tả truy vấn và báo cáo kết quả là nhờ sử dụng bản đồ. Do vậy các chức năng lập bản đồ thường thấy ở trong GIS. Các

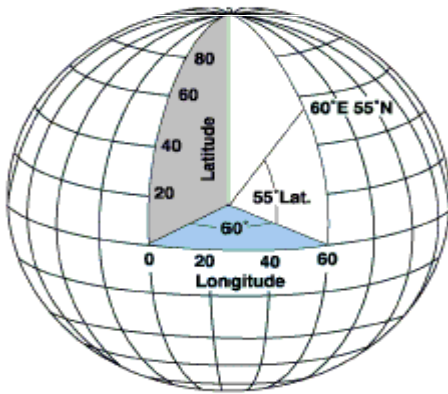
chức năng này được thể hiện bằng thực đơn như trong các trình vẽ bản đồ không phải là GIS thực thụ, để xác định màu, kiểu, mẫu của điểm, đường, vùng. Chúng có khả năng trang trí bản đồ bằng râu kí tự, chú giải với các thuộc tính khác nhau như loại, cỡ phông, hướng vẽ. Nhiều hệ GIS còn có khả năng biến đổi và vẽ theo các phép chiếu bản đồ khác nhau. Gần đây, biểu đồ diện tích được quan tâm nhiều đến để hiển thị dữ liệu. Trong đó, kích thước đặc trưng bản đồ được biến đổi tương ứng với các biến thuộc tính như mật độ dân số hay thu nhập. Kỹ thuật hiển thị khác giúp hiểu dễ dàng một số loại dữ liệu bằng thay đổi độ đậm nhạt hay sắc tố màu thay vì mô tả chính xác các đường biên. Biểu diễn thay đổi thời gian trên bản đồ cũng là nhu cầu thực tế, đôi khi chúng được đánh dấu trên cùng một bản đồ trong các thời điểm khác nhau và liên kết chúng bằng mũi tên. Các hệ GIS mới thường cố gắng xây dựng cơ chế phát sinh bản đồ hoạt ảnh, trong đó bản đồ tự thực hiện thay đổi sau khoảng thời gian ngắn. Một khả năng khác của GIS là khả năng hiển thị bản đồ 3D từ các điểm quan sát khác nhau, hiển thị 3D cho khả năng hiểu biết về bản chất của hiện tượng khi nghiên cứu thay đổi bề mặt địa hình.

Câu hỏi ôn tập.

1. Định nghĩa GIS. Các thành phần của GIS.
2. Nêu các chức năng của GIS.

CHƯƠNG 2. HỆ THỐNG THAM CHIẾU KHÔNG GIAN

Hình 2.1. Tọa độ cầu



Vị trí của vật thể trong không gian đều phải gắn liền với một hệ tọa độ. Trong GIS, để biểu diễn dữ liệu không gian người ta thường dùng 2 hệ tọa độ là: hệ tọa độ địa lý và hệ tọa độ quy chiếu. Hệ tọa độ địa lý là hệ tọa độ lấy mặt cầu ba chiều bao quanh trái đất làm cơ sở. Một điểm được xác định bằng kinh độ và vĩ độ của nó trên mặt cầu. Hệ tọa độ quy chiếu là hệ tọa độ hai chiều thu được bằng cách chiếu dữ liệu bản đồ nằm trên hệ tọa độ địa lý về một mặt phẳng.

2.1. Hệ tọa độ địa lý

Hệ tọa độ địa lý sử dụng bề mặt hình cầu để xác định vị trí của một điểm trên trái đất. Đơn vị đo của hệ là độ.

Vì đây là hệ tọa độ gắn liền với trục trái đất nên để xác định vị trí của đối tượng người ta chia bề mặt trái đất thành các đường kinh tuyến và vĩ tuyến. Kinh tuyến là các đường cong cách đều nhau chạy qua hai điểm cực Bắc và Nam, vĩ tuyến là các đường tròn song song có tâm nằm trên trục của trái đất. Giao điểm giữa kinh tuyến và vĩ tuyến tạo thành các ô lưới. Trong số các kinh tuyến và vĩ tuyến có hai đường quan trọng nhất được lấy làm gốc tọa độ đó là: đường xích đạo và kinh tuyến chạy qua vùng Greenland nước Anh. Giao điểm giữa hai đường này là gốc tọa độ. Hai đường này cũng đồng thời chia trái đất làm 4 phần bằng nhau: nửa Bắc và Nam nằm phía trên và dưới của đường xích đạo; nửa Đông và Tây nằm ở phía bên phải và trái của kinh tuyến gốc.

Một điểm nằm trên mặt cầu sẽ có hai giá trị tọa độ là kinh độ và vĩ độ được xác định như trong hình vẽ trên. Giá trị này có thể được đo bằng độ theo cơ số 10 hoặc theo độ, phút, giây.

Miền giá trị của vĩ độ: -900 900

Miền giá trị của kinh độ: -1800 1800

Chỉ trên đường xích đạo thì khoảng cách một độ của vĩ tuyến mới bằng khoảng cách một độ trên kinh tuyến. Trên các vĩ tuyến khác khoảng cách này khác nhau rất nhiều. Người ta tính rằng một độ trên kinh tuyến dài khoảng 111,321 km trong khi 600 trên vĩ tuyến chỉ có độ dài 55,802 km. Chính vì sự khác nhau này nên ta không thể đo chính xác được chiều dài và diện tích của đối tượng khi dữ liệu bản đồ được chiếu lên mặt phẳng.

Trong hệ tọa độ địa lý có hai bề mặt hình cầu được sử dụng đó là: mặt cầu (tuyệt đối) và mặt Ellipsoid. Vì bề mặt của trái đất của ta không phải là hình cầu tuyệt đối mà nó gần với hình Ellipsoid nên mặt Ellipsoid thường được dùng để biểu diễn. Tuy nhiên đôi khi người ta cũng sử dụng mặt cầu để công việc tính toán dễ dàng hơn. Khi tỷ lệ bản đồ rất nhỏ (nhỏ hơn

1:5000.000), ở tỷ lệ này sự khác biệt giữa dữ liệu biểu diễn bằng mặt cầu và mặt Ellipsoid là không thể phân biệt được bằng mắt thường. Lúc này, mặt cầu được dùng. Nhưng khi tỷ lệ lớn hơn 1:1.000.000 thì người ta cần thiết phải dùng mặt Ellipsoid để đảm bảo độ chính xác. Do đó, việc lựa chọn mặt cầu hay mặt Ellipsoid phụ thuộc vào mục đích của bản đồ và độ chính xác dữ liệu.

Nếu mặt cầu dựa trên hình tròn thì mặt Ellipsoid lại có cơ sở là hình Ellip. Hình Ellip được xác định bởi hai bán trục mà ta hay gọi là: bán trục lớn và bán trục nhỏ. Ta cho Ellip xoay quanh bán trục nhỏ ta sẽ thu được hình Ellipsoid.

Kích thước và hình dạng của Ellipsoid được xác định bởi bán trục lớn a và bán trục nhỏ b , hay bởi a và hệ số dẹt $f = (a - b) / a$. Vì hệ số f rất nhỏ nên người ta thường dùng giá trị $1/f$ (1: bán kính xích đạo, f : bán kính cực).

2.2. Hệ tọa độ quy chiếu

Để thuận tiện cho sử dụng người ta phải nghiên cứu cách thể hiện bề mặt trái đất lên trên mặt phẳng của bản đồ. Do đó phải thực hiện phép chiếu bề mặt cong của trái đất lên mặt phẳng và hệ tọa độ quy chiếu ra đời. Hệ tọa độ này luôn lấy hệ tọa độ địa lý làm cơ sở.

Hệ tọa độ quy chiếu được đặc trưng bởi hai trục X, Y theo phương ngang và thẳng đứng. Gốc tọa độ là giao điểm của hai trục này. Hai trục giao nhau đồng thời chia mặt phẳng làm 4 phần tương ứng với 4 phần trong hệ tọa độ địa lý. Một điểm trên mặt được xác định được xác định bởi cặp giá trị (x, y) .

Có rất nhiều phép chiếu bề mặt cong của trái đất lên mặt phẳng song về cơ bản ta có thể hiểu như sau. Lấy một mảnh bìa cuộn xung quanh bề mặt cầu trong hệ tọa độ địa lý theo một hình trụ đứng. Từ tâm của bề mặt cong ta vẽ các tia cắt các điểm giao giữa kinh tuyến và vĩ tuyến, đồng thời kéo dài cắt mặt trụ. Thực chất của việc này là chiếu các ô lưới lên bề mặt phẳng. Mở tờ bìa ra ta có kết quả của phép chiếu. Nhìn vào tấm bìa ta nhận thấy, các ô lưới đã thay đổi khá nhiều (biến dạng), co lại hoặc dãn ra. Càng xa đường xích đạo thì sự biến dạng càng lớn. Điều này gây nên sự thay đổi về hình dạng, kích thước, khoảng cách của dữ liệu không gian. Sau đó người ta dùng các công thức toán học để tương ứng tọa độ của bề mặt cong lên tọa độ mặt phẳng chiếu. Các phép chiếu khác nhau gây ra các biến dạng bản đồ khác nhau nên việc sử dụng phép chiếu nào là dựa vào mục đích của bản đồ và độ chính xác của dữ liệu.

2.3. Các phép chiếu cơ bản

Trong phần này ta sẽ tìm hiểu ba phép chiếu cơ bản và thường được sử dụng nhất đó là phép chiếu với mặt chiếu: mặt hình nón, mặt hình trụ và mặt phẳng phương vị. Bước đầu tiên khi tiến hành phép chiếu này là tạo ra một hay một tập các điểm tiếp xúc. Các điểm tiếp xúc này được gọi là các tiếp điểm hay là tiếp tuyến. Các điểm này có vai trò rất quan trọng, vì độ biến dạng của phép chiếu trên những điểm này bằng không. Độ biến dạng sẽ tăng khi khoảng cách giữa điểm chiếu và điểm tiếp xúc tăng.

Mặt hình nón

Để thực hiện phép chiếu này người ta cho dùng một mặt hình nón “úp” lên bề mặt cầu. Đường thẳng tiếp xúc giữa mặt nón và mặt cầu là một vĩ tuyến và được gọi là vĩ tuyến chuẩn.

Các đường kinh tuyến sau khi chiếu mặt nón sẽ thành những đường thẳng đứng, các đường vĩ tuyến sẽ tạo thành những đường tròn. Sau khi thực hiện phép chiếu, người ta sẽ cắt hình nón dọc theo một kinh tuyến bất kỳ, lúc này ta sẽ được kết quả của phép chiếu trên bề mặt nón. Sự giao nhau giữa những đường thẳng và cung tròn sẽ tạo nên một mặt lưới. Đường thẳng đối diện với đường cắt được gọi là kinh tuyến trung tâm. Càng xa vĩ tuyến chuẩn độ biến dạng càng tăng. Do đó để tăng độ chính xác người ta cắt bỏ phần đỉnh của mặt nón hay ta không tiến hành chiếu lên vùng này. Phép chiếu này thường được dùng cho việc chiếu các vùng có các vĩ tuyến trung bình chạy qua và hướng theo chiều đông - tây.

Mặt hình trụ

Giống như phép chiếu mặt nón, phép chiếu này cũng có một đường thẳng tiếp tuyến. Khi sử dụng mặt trụ, người ta phân làm 3 loại tùy thuộc vào vị trí tương đối của mặt trụ so với mặt cầu:

Hình trụ được đặt theo phương thẳng đứng và tiếp xúc với mặt cầu theo một vĩ tuyến, thường là đường xích đạo. Gọi là phép chiếu Mercator.

Hình trụ được đặt theo phương nằm ngang, đường thẳng tiếp xúc là một kinh tuyến. Gọi là phép chiếu Transverse.

Hình trụ đặt xiên và tiếp xúc với mặt cầu theo một đường tròn có bán kính lớn nhất (bằng với bán kính đường xích đạo). Gọi là phép chiếu Oblique.

Phép chiếu thường được sử dụng nhất là Mercator. Trong phép chiếu này, các đường kinh tuyến sẽ được chiếu thành những đường thẳng đứng cách đều nhau, các đường vĩ tuyến sẽ trở thành những đường nằm ngang khoảng cách không đều nhau; tăng dần về phía hai cực. Do đó biến dạng sẽ tăng dần về phía hai cực. Sau khi thực hiện phép chiếu, người ta sẽ cắt mặt hình trụ dọc theo một kinh tuyến, trải ra trên mặt phẳng ta sẽ thu được kết quả.

Mặt phẳng phương vị

Đây là phép chiếu dữ liệu bản đồ lên một mặt phẳng tiếp xúc với mặt cầu. Điểm tiếp xúc này có thể là: nằm tại hai cực, tại đường xích đạo, hoặc tại một vị trí bất kỳ nằm giữa. Vị trí của điểm tiếp xúc cho ta biết vị trí tương đối của mặt phẳng chiếu với mặt cầu và tạo nên ba kiểu chiếu khác nhau: polar, equatorial và oblique.

Mặt phẳng chiếu tiếp xúc với cực của mặt cầu là kiểu chiếu đơn giản nhất và cũng hay dùng nhất. Trong phép chiếu này, các đường kinh tuyến sẽ được chiếu thành một chùm đường thẳng giao nhau ở điểm cực, vĩ tuyến là các đường tròn có cùng tâm là cực của mặt cầu. Góc giữa các đường kinh tuyến được bảo tồn.

2.3.1. Hệ thống phép chiếu UTM (Universal Transverse Mercator)

Có nguồn gốc từ phép chiếu Mercator ngang và thường được dùng để đo vẽ bản đồ có độ chính xác cao.

Đặc điểm:

+ Phép chiếu UTM chia bề mặt Trái đất thành 60 múi theo vĩ tuyến, mỗi múi 60 và đánh số từ 1 đến 60 kể từ kinh độ 180 Tây. Mỗi múi kéo dài từ vĩ độ 84 Nam đến vĩ độ 84 Bắc.

+ Chia bề mặt Trái đất thành các dải 80 theo kinh tuyến mở ra từ đường xích đạo và được gán các ký tự (từ C -> X, trừ I, O)

+ Các múi đặc biệt không theo qui luật trên:

- Kinh độ 0 đến kinh độ 36 và trên vĩ độ 72 Bắc.
- Múi 32 giữa vĩ độ 56 Bắc đến vĩ độ 64 Bắc.

Mục hõnh phộp chiếu UTM trong cộc trường hợp:

+ Tọa độ phẳng trong UTM qui định trục Y nằm trờn kinh tuyến trục ỏo và tăng theo hướng Bắc. Trục X vuờng góc với trục Y tại (ỏo , ỏo) và tăng theo hướng Đông. (ỏo , ỏo) là gốc tọa độ

+ Trái đất là hõnh cầu:

Từ kinh độ ỏ và vĩ độ ỏ ta tính được cặp tọa độ x, y theo công thức sau:

$$x = a * R * \ln[(1+B)(1-B)]$$

$$y = R * k_0 * (\arctg[\tg(\tilde{o})/\cos(\tilde{o}-\tilde{o}_0)] - \tilde{o}_0)$$

$$B = \cos(\tilde{o})\sin(\tilde{o}-\tilde{o}_0)$$

Trong đó: R = 6.371.116 m (bán kính địa cầu)

$k_0 = 0,9996$ là sai số hõnh chiếu trờn kinh tuyến trục ỏo

Từ (x, y) ta tính được kinh độ ỏ và vĩ độ ỏ theo cụng thức sau:

$$\tilde{o} = \arcsin(\sin(D) / \cos(x/R * k_0))$$

$$\tilde{o} = \tilde{o}_0 + \arctan(\sin(x/R * k_0) / \cos(D))$$

$$D = y/(R * k_0) + \tilde{o}_0$$

D được tính bằng độ

+ Trái đất là elip trũn xoay

Từ kinh độ ỏ và vĩ độ ỏ ta tính được cặp tọa độ x, y theo công thức sau:

$$x = k_0 N [A + (1-T+C)A^3 / 6 + (5 - 18T + T^2 + 72C - 58e'^2) / 120]$$

$$y = k_0 (M - M_0 + N * \tg \tilde{o} + [A^2 / 2 + (5-T+9C+4C^2)])$$

Trong đó: $k_0 = 0,9996$

$$e'^2 = e^2(1 - e^2)$$

$$N = a / (1 - e^2 \sin^2(\tilde{o}))$$

$$T = \tg^2(\tilde{o}) ; C = e'^2 \cos^2(\tilde{o})$$

$$A = (\tilde{o} - \tilde{o}_0) \cos(\tilde{o})$$

$$M = a [(1 - e^2/4 - 3e^2/256...) \delta - (3e^2/8 - 3e^4/32 - 45e^6/1024) \sin(2\delta)]$$

$$+ a [(15e^4/256 - 45e^6/1024 \dots) \sin(4\delta) - \dots]$$

+ Tại Bán cầu Bắc, gốc là giao của đường xích đạo và kinh tuyến trục, để tránh tọa độ âm, người ta dịch trục x về phía Tây khoảng 500 km. Khi đó hệ trục tọa độ phẳng sẽ là $x_0=500.000m$; $y_0=0$. Đối với phía Nam điểm gốc cũng sẽ thay đổi thành $x_0=500.000m$; $y_0=10.000.000$. Các giá trị (x_0, y_0) được sử dụng để hiệu chỉnh các giá trị x,y được tính toán.

+ (δ_0, ϕ_0) là gốc tọa độ địa lý, trong UTM thì $\delta_0=0$ cho mọi mýi cũn

$$\phi_0 = (zone + 0,5)\pi/30 - \pi$$

zone là mýi Utm cú giá trị từ 1->60.

2.3.2. Phép chiếu Robinson

Phép chiếu thuộc lớp giả trụ có nhiều lưới như sau:

+ Kinh tuyến gốc là đường thẳng vuông góc với đường xích đạo. Các kinh tuyến khác là những đường cong cách đều nhau.

+ Vĩ tuyến là các đường thẳng song song cách đều nhau từ 38o Nam đến 38o Bắc.

+ Tỷ lệ bản đồ đúng theo các vĩ độ 38o Nam và 38o Bắc. Sai số càng tăng khi tiến về hai cực.

+ Phép chiếu được sử dụng khá phổ biến trong việc in ấn bản đồ.

+ Ta có thể tính được tọa độ x, y từ (δ, ϕ)

$$x = 0,8487RX(\phi - \phi_0)$$

$$y = 1,3523RY$$

Trong đó X, Y được nội suy từ bảng hàng số Robinson

Câu hỏi ôn tập.

1. Trình bày các phép chiếu cơ bản.
2. Trình bày về hệ thống phép chiếu UTM (Universal Transverse Mercator).

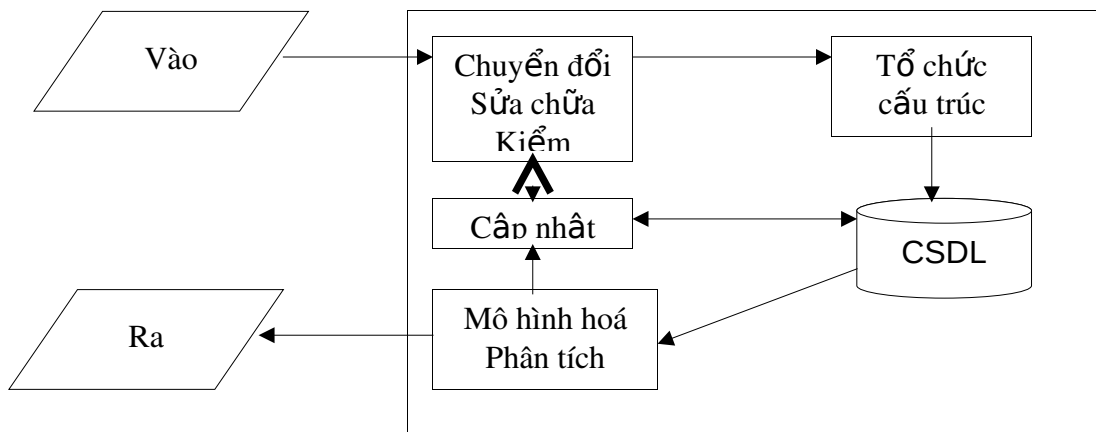
CHƯƠNG 3. MÔ HÌNH VÀ CẤU TRÚC DỮ LIỆU KHÔNG GIAN

3.1. Các khái niệm cơ sở

Dữ liệu: là các con số hay sự kiện được tập hợp có hệ thống cho một hay nhiều mục đích cụ thể. Chúng có thể tồn tại dưới nhiều hình thức khác nhau như ngôn ngữ tự nhiên, biểu tượng, biểu thức toán, tín hiệu sóng...

Thông tin: được xem như dữ liệu đã được xử lý dưới khuôn mẫu hữu ích cho người dùng và là những giá trị nhận thức cho công tác lập quyết định. Dữ liệu là thành phần của thông tin, nhưng không phải mọi dữ liệu đều có ích, thông tin chỉ có ích khi nó tin cậy, chính xác, kịp thời, đầy đủ, dễ hiểu, dễ quản lý.

Hệ thống thông tin: có nhiệm vụ chuyển đổi dữ liệu thành thông tin theo các tiến trình khác nhau như biến đổi, tổ chức, cấu trúc hoá và mô hình hoá. Hai khái niệm tổ chức và cấu trúc đóng vai trò cốt yếu trong chức năng của một hệ thống thông tin, nếu không có chúng thì không có khả năng chuyển đổi dữ liệu thành thông tin.



Hình 3.1: Biến đổi dữ liệu thành thông tin

Mô hình dữ liệu địa lý: là các quy tắc được sử dụng để biến đổi đặc trưng địa của thế giới thực thành các đối tượng rời rạc. Mô hình dữ liệu được sử dụng để biểu diễn thực thể với mức độ phức tạp khác nhau.

3.2. Mô hình thông tin không gian

Dữ liệu là trung tâm của hệ thống GIS, hệ thống GIS chứa càng nhiều dữ liệu thì chúng càng có ý nghĩa. Dữ liệu của hệ GIS được lưu trữ trong CSDL và chúng được thu nhập thông qua mô hình thế giới thực. Dữ liệu trong hệ GIS còn được gọi là thông tin không gian. Đặc trưng thông tin không gian là có khả năng mô tả “vật thể ở đâu” nhờ vị trí tham chiếu, đơn vị đo và quan hệ không gian. Chúng còn khả năng mô tả “hình dạng hiện tượng” thông qua mô tả chất lượng, số lượng của hình dạng và cấu trúc. Cuối cùng, đặc trưng thông tin không gian mô tả “quan hệ và tương tác” giữa các hiện tượng tự nhiên. Mô hình không gian đặc biệt quan trọng vì cách thức biểu diễn thông tin sẽ ảnh hưởng đến khả năng hiển thị đồ họa của hệ thống

Hiện tượng tự nhiên được thể hiện bằng các thành phần rời rạc hay gọi là đối tượng trong CSDL. Đặc trưng địa lý được hình thành từ các thực thể. Thực thể là hiện tượng được quan tâm trong thế giới thực mà nó không bị chia nhỏ ra thành các hiện tượng cùng loại. Khi xây dựng CSDL ta phải quan tâm đến các mức của trừu tượng dữ liệu như sau đây:

Thế giới thực: là toàn bộ hiện tượng tự nhiên như chính nó tồn tại.

Mô hình dữ liệu quan niệm: là các thành phần và các quan hệ giữa chúng liên quan đến hiện tượng tự nhiên nào đó. Mô hình dữ liệu này độc lập với hệ thống cụ thể, cấu trúc dữ liệu, tổ chức và quản lý dữ liệu.

Mô hình dữ liệu vật lý hay cấu trúc tệp: là tập các qui luật để cài đặt cấu trúc dữ liệu trên các máy tính khác nhau.

Thông thường ta phải sử dụng một vài mô hình quan niệm trừu tượng không gian để giải thích và phân tích dữ liệu không gian. Mỗi mô hình sẽ nhấn mạnh một khía cạnh của hiện tượng không gian, nó phụ thuộc vào mục tiêu phân tích. Ở đây ta tập trung vào ba mô hình thường được sử dụng trong GIS để tổ chức và xử lý dữ liệu không gian, đó là mô hình trên cơ sở đối tượng, mạng và nền.

Mô hình không gian trên cơ sở đối tượng

Mô hình này tập trung vào các hiện tượng riêng rẽ sẽ được nghiên cứu độc lập hay cùng quan hệ của chúng với hiện tượng khác. Bất kỳ hiện tượng nào đều được xem như đối tượng, có thể tách rời khỏi các đối tượng láng giềng và phải có tính phân biệt, nhận dạng duy nhất. Đối tượng có thể bao gồm các đối tượng khác.

Mô hình không gian trên cơ sở mạng

Mô hình chia sẻ một vài khía cạnh của mô hình đối tượng vì chúng đề cập đến các đối tượng rời rạc nhưng đặc trưng chính của mô hình là xem xét sự tương tác giữa các đối tượng, thường là đường đi nối chúng. Hình dạng chính xác của đối tượng có thể không quan trọng, quan trọng ở đây là khoảng cách giữa các đối tượng cần xét. Ví dụ rõ nhất ở đây là mạng lưới giao thông đường bộ, đường biển và đường không.

Mô hình quan sát trên cơ sở nền

Mô hình phù hợp cho mô hình hóa hiện tượng, quan tâm đến các vật liên tục trải trên vài cùng không gian. Nền có thể được biểu diễn hai hay ba chiều là tùy vào loại ứng dụng. Một số dữ liệu có thể coi như nền hoặc đối tượng tùy vào cách thu thập dữ liệu. Ví dụ như nghiên cứu khu rừng, nếu thu thập từ ảnh thì là dữ liệu nền còn là dữ liệu đối tượng nếu xem xét tới đường biên khu rừng.

3.3. Mô hình dữ liệu raster

Đây là phương pháp biểu diễn các đặc trưng địa lý bằng các điểm ảnh. Được hình thành dựa trên cơ sở quan sát nền thế giới thực.

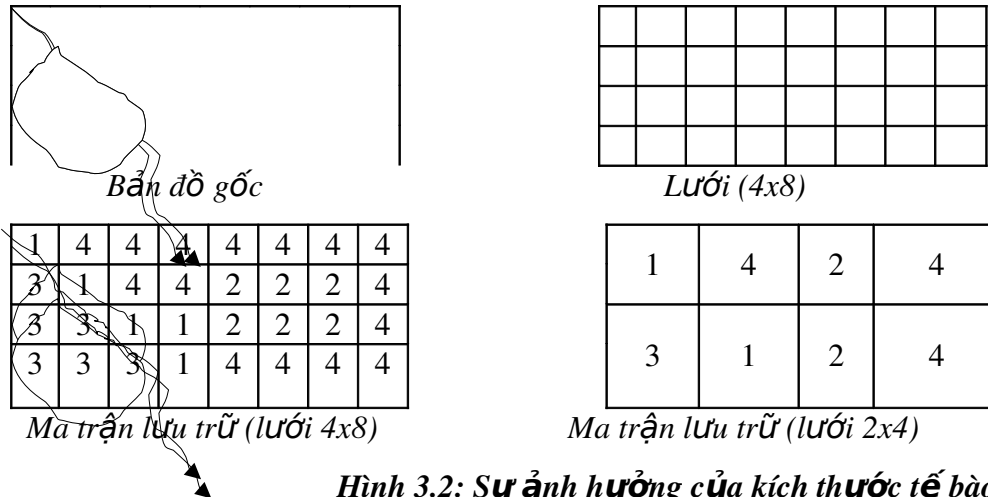
3.3.1. Trường raster đơn giản

Mô hình dữ liệu raster hay còn gọi là lưới tế bào hình thành nên cho một số hệ thống tin địa lý. Các hệ thống trên cơ sở raster hiển thị, định vị và lưu trữ dữ liệu đồ họa nhờ sử dụng các ma trận hay lưới tế bào. Độ phân giải dữ liệu raster phụ thuộc vào kích thước của tế bào hay

điểm ảnh, chúng có thể khác nhau từ vài đêximet đến vài kilômet. Tiến trình xây dựng lưới tế bào được mô tả như sau đây:

Giả sử phủ một lưới lên bản đồ, dữ liệu raster được lập bằng cách mã hóa mỗi tế bào bằng một giá trị dựa theo các đặc trưng trên bản đồ, độ chính xác của một đối tượng phụ thuộc vào kích thước hay độ phân giải của các tế bào lưới.

Ví dụ dưới sẽ chỉ rõ điều này:



Hình 3.2: Sự ảnh hưởng của kích thước tế bào

Bản đồ lại được chia thành nhiều tầng, mỗi tầng bao gồm hàng triệu tế bào vì vậy cần áp dụng các thuật toán nén dữ liệu mà vẫn cho khả năng khôi phục dữ liệu ban đầu.

Hình dạng bao phủ toàn bộ bản đồ gọi là khảm, khảm có thể là hình vuông, tam giác hay lục giác tuy nhiên khi sử dụng khảm hình tam giác hoặc lục giác sẽ khó khăn trong việc phân chia tế bào thành các tế bào nhỏ hơn, cách mã hóa, lưu trữ tế bào khó khăn hơn rất nhiều.

Lợi thế của hệ thống raster là mã dữ liệu đồ thành bản đồ trong máy tính nên các thao tác so sánh tế bào là dễ dàng song không thuận tiên cho việc biểu diễn đường, đoạn, điểm vì mỗi loại là một tập hợp các tế bào và có đoạn rộng, đoạn hẹp. Một vấn đề khó khăn nữa là xử lý tế bào trộn là tế bào có nhiều hơn một đặc trưng (như trong ví dụ trên một số ô có thể đánh một trong hai giá trị tùy theo hoạt động của chương trình hay quan niệm của người lập), vấn đề này dẫn tới khó khăn khi có nhu cầu phân tích là phủ bản đồ. Vì thế raster thích hợp hơn với hệ thống GIS hướng tài nguyên, môi trường vì hình thành trên sơ sở quan sát nền thế giới thực. Raster có nhiều tầng bản đồ (địa hình, đất đai..) hơn so với mô hình vectơ.

3.3.2. Cấu trúc dữ liệu raster

Số lượng dữ liệu của mô hình là rất lớn do đó cần có kĩ thuật quản lý khối dữ liệu sao cho dễ quản lý và dễ xâm nhập. Các phương pháp hay dùng là phân hoạch dữ liệu và xây dựng chỉ số xâm nhập nhanh.

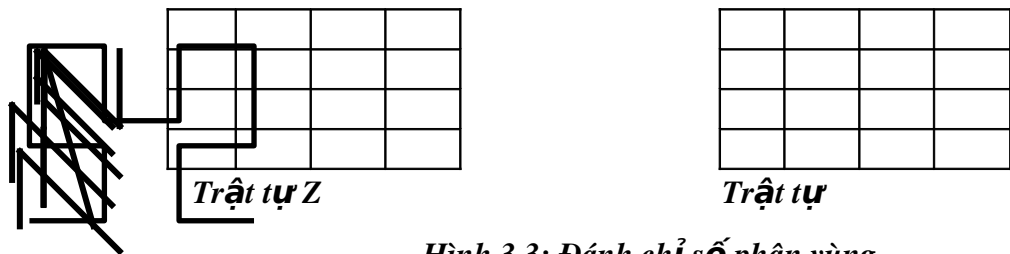
Phân hoạch dữ liệu

Phân hoạch CSDL không gian để quản lý khối lượng lớn thông tin bằng tập giới hạn các thao tác. Hai chiến lược được sử dụng đó là: mỗi phân hoạch chứa số lượng thông tin xấp xỉ

nhau và kích thước địa lý có thể khác nhau; phân hoạch địa lý cố định và lượng thông tin trong mỗi phân hoạch có thể khác nhau.

Xây dựng chỉ số xâm nhập nhanh (hay chỉ số không gian)

Ý tưởng chính của xây dựng chỉ số không gian là sử dụng khái niệm xấp xỉ qua phân hoạch số lượng thông tin xấp xỉ. Nếu mỗi đặc trưng có thêm trường chỉ số để chỉ ra nơi chứa chúng thì việc tìm kiếm địa chỉ sẽ giới hạn trong trường này. Tiến trình đánh số chỉ số thường sử dụng đó là phân hoạch đệ qui, kết quả là cho ta cấu trúc phân cấp gọi là cây tứ phân. Trong cây tứ phân mỗi vùng được chia thành bốn vùng và đánh chỉ số khác nhau, mỗi vùng lại chia tiếp thành bốn phần cho đến khi đạt được kích thước tế bào cần thiết tuy nhiên phân hoạch như vậy là không hợp lý và tiến trình phân hoạch chỉ cần tiến hành ở những vùng có nhiều đặc trưng hay được quan tâm tới nhiều tùy vào các chức năng của bản đồ. Cách đánh chỉ số ở đây thường áp dụng trật tự Peano hay Z (sử dụng cách đánh số như nhau ở bốn góc phần tư), trật tự Hilbert hay (sử dụng cách đánh số đối xứng ở bốn góc phần tư) như sau:



Hình 3.3: Đánh chỉ số phân vùng

Ví dụ:

11	13	31	33
10	12	30	32
01	03	21	23
00	02	20	22

Trật tự Z

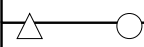

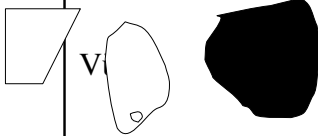
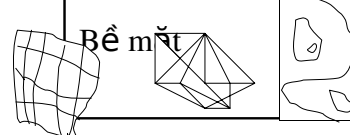
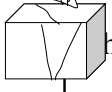
1	2	1	2
0	3	0	3
3	2	2	3
0	1	1	0

Trật tự

Ngoài ra có thể sử dụng sơ đồ R-tree hay R+-tree. R-tree là cây phân cấp, mỗi lá của cây là một hình chữ nhật và con trỏ, con trỏ trỏ tới đặc trưng của hình chữ nhật bao (vùng hay thùng chứa), còn hình chữ nhật xác định bởi tọa độ x,y cực đại và cực tiểu. Tập các hình chữ nhật được đánh chỉ số là lưu trữ trong tập các thùng. Tuy nhiên R-tree có hiện tượng hình chữ nhật trùng nhau nghĩa là hình chữ nhật cắt các thùng nhưng thùng được biểu diễn chỉ trong một hình chữ nhật, điều này được xử lý trong R+-tree bằng cách chia các hình chữ nhật dữ liệu để mỗi hình được biểu diễn trong hai vùng con.

3.4. Mô hình dữ liệu vectơ

Đây là phương pháp biểu diễn các đặc trưng địa lý bằng các phần tử đồ họa cơ bản (điểm, đường, đa giác, bề mặt ba chiều và khối trong 3D). Phương pháp vectơ hình thành dựa trên cơ sở quan sát đối tượng của thế giới thực.

Kiểu thành phần sơ cấp	Biểu diễn đồ họa	Biểu diễn vectơ
Điểm 	. + x	(x,y) trong 2D (x, y, z) trong 3D
Đường 		đanh sách tọa độ (tập các đoạn, đường thẳng trong trường hợp đường cong) hoặc các hàm toán học
Vùng 		đường có điểm đầu và cuối trùng nhau hay tập các đường nếu vùng có lỗ
Bề mặt 		Ma trận điểm, tập các tam giác, hàm toán học, đường bình độ
Khối 		Tập các bề mặt

Điểm là thành phần sơ cấp của dữ liệu địa lý trong mô hình vectơ, các điểm được nối với nhau bởi các đường để tạo thành các thực thể khác. Tùy theo tỷ lệ quan sát mà các thực thể được biểu diễn bằng điểm, đường hay vùng... Như vậy mô hình vectơ sử dụng các điểm hay đoạn thẳng để nhận biết các vị trí của thế giới thực nên cho phép thao tác nhiều hơn trên các đối tượng so với mô hình raster và chính xác hơn như tính chu vi, diện tích... đặc biệt là tìm đường đi nhanh và hiệu quả hơn rất nhiều.

Mô hình vectơ hợp với các hệ GIS định hướng các hệ thống quản trị CSDL, chúng có ưu việt trong lưu trữ số liệu bản đồ bởi chỉ lưu đường biên của các đặc trưng chứ không lưu cả vùng như trong mô hình raster nên truy nhập, tìm kiếm, hiển thị dễ dàng thông tin từ CSDL.


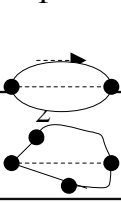
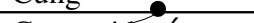
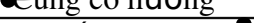

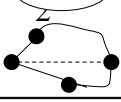
Mô hình vectơ sử dụng hai cấu trúc dữ liệu là toàn đa giác và cung-nút.

3.4.1. Cấu trúc dữ liệu toàn đa giác

Mỗi tầng trong CSDL của cấu trúc toàn đa giác được chia thành tập các đa giác. Mỗi đa giác được mã hóa thành trật tự các vị trí hình thành đường biên của vùng khép kín theo một hệ tọa độ nào đó và được lưu trữ độc lập. Do cách tổ chức nên ta không biết được các đa giác kề nhau hay nói cách khác là không có tổ chức tô pô trong cấu trúc này. Với cấu trúc này, một điểm có thể lưu lại nhiều lần nên việc cập nhật là rất khó khăn.

3.4.2. Cấu trúc dữ liệu cung-nút

Một khía cạnh quan trọng của mô hình vectơ là khả năng tách biệt các thành phần để thực hiện đo đạc và xác định quan hệ không gian giữa chúng. Quan hệ này là một quan hệ tô pô, mỗi đường hay vùng cắt nhau sẽ tạo ra các giao điểm (nút) và các vùng (đường) mới, các vùng (đường) mới này được xác định bởi các cung. Các cung, nút là các thành phần cơ bản của các tô pô trong bản đồ đã số hóa. Số chiều, hình dạng đồ họa hay dùng trên bản đồ của các tô pô cơ bản được mô tả trong bảng dưới đây.

Số chiều	Loại tô pô	Biểu diễn	Mô tả
0 	Nút		Biểu diễn điểm rời rạc, giao điểm
1 	Cung		Nối hai nút
	Cung có hướng		Cung có hướng xác định
	Xuyên		Mạch cung khép kín
	Bề mặt		Vùng khép kín bởi một hay nhiều cung
	Polyhedron		Cung khép kín bởi các bề mặt

Thông thường tô pô ba chiều có bề mặt bao bọc, đa giác được tạo từ một hay nhiều cung. Tô pô hai chiều chỉ gồm các đa giác, cung, nút. Đa giác được tạo ra nhờ sách các cung thành phần (các cung phải có hướng xác định), trong đó dấu trừ (-) được sử dụng cho hướng cung phù hợp. Với các đa giác ló lõm thì thường qui định chiều kim đồng hồ cho biên ngoài và ngược chiều kim đồng hồ cho biên trong.

Các đối tượng không gian chia sẻ trong một tầng sẽ chỉ phải ghi một lần. Trong cấu trúc dữ liệu tô pô đa tầng, các đối tượng của một lớp bản đồ sẽ được ghi tại một tầng cụ thể, như vậy sẽ có các cung lặp trong các tầng và việc phân tích phủ bản đồ cần phải xác định giao điểm của các cung để thông tin chính xác.

3.5. Mô hình dữ liệu lưới tam giác không đều

3.5.1. Phương pháp Voronoi

Ta xét bài toán phân bố các cột điện thoại của thành phố sao cho mỗi máy điện thoại được nối với cột gần nhất vì thế phải chia thành phố thành các vùng, mỗi vùng có duy nhất một cột và khoảng cách từ mỗi vị trí trong vùng đến cột trong vùng đó là ngắn nhất. Kết quả phân hoạch này là sơ đồ Voronoi.

Sau đây là định nghĩa sơ đồ Voronoi: gọi $P=\{p_1, p_2...p_n\}$ là tập các điểm trong mặt phẳng tọa độ. Các điểm này gọi là vị trí. Hãy phân hoạch mặt phẳng bằng cách gán từng điểm của nó cho vị trí gần nó nhất. Toàn bộ các điểm được gán cho p_i hình thành vùng Voronoi $V(p_i)$, $V(p_i)$ bao gồm mọi điểm gần vị trí p_i hơn bất kì vị trí nào khác: $V(p_i)=\{x: |p_i-x| \leq |p_j-x|, i \neq j\}$. Trong đó, một số điểm không chỉ có một vị trí (hay láng giềng) gần nhất.

Tập các điểm có nhiều láng giềng gần hơn hình thành từ sơ đồ Voronoi $V(P)$. Đường bao giữa hai vùng Voronoi gọi là sườn Voronoi. Một điểm là giao của ba hay nhiều hơn cạnh

Voronoi gọi là đỉnh Voronoi. Cặp vị trí đủ gần nhau để chia sẻ cạnh Voronoi và tồn tại vòng tròn đi qua hai vị trí nhưng không chứa bên trong chúng một vị trí nào khác thì gọi là cặp vị trí liên quan.

Sơ đồ Voronoi của hai vị trí: xét hai điểm p_1 và p_2 , gọi B_{12} là đường trung trực của đoạn thẳng $p_1 p_2$ thì B_{12} là sơ đồ Voronoi của p_1 và p_2 .

Sơ đồ Voronoi của ba vị trí không thẳng hàng: xét ba điểm p_1, p_2, p_3 . Ta xét ba sơ đồ Voronoi B_{12}, B_{23}, B_{31} là ba đường trung trực của ba đoạn $p_1 p_2, p_2 p_3, p_1 p_3$ ba đường này gặp nhau tại một điểm là tâm đường tròn ngoại tiếp tam giác $p_1 p_2 p_3$ điểm này chính là sơ đồ Voronoi của p_1, p_2, p_3 .

Sơ đồ Voronoi của nhiều vị trí (lớn hơn 3): ta tìm sơ đồ Voronoi 2 điểm bất kì, mạng lưới thu được là sơ đồ Voronoi.

Trong hệ thống tin địa lý, sơ đồ Voronoi được áp dụng để hình thành các chức năng như biến đổi đối tượng từ raster sang vectơ nhờ kỹ thuật tìm xương, xây dựng tam giác Delaunay trong mô hình TIN...

3.5.2. Cấu trúc địa lý TIN

TIN được nghiên cứu từ những năm 70 của thế kỷ XX, các hệ thống GIS thương mại có mô hình TIN xuất hiện vào những năm 80 của thế kỷ XX. TIN có khả năng biểu diễn bề mặt liên tục từ tập điểm địa lý rời rạc. Khái niệm hình học TIN là tập các đỉnh được nối với nhau thành các tam giác, các tam giác này hình thành bề mặt ba chiều. Bề mặt TIN được sử dụng để biểu diễn các vấn đề khác nhau như độ cao, lượng mưa... Trong GIS vectơ thì TIN được coi như các đa giác có thuộc tính là độ dốc, hướng và diện tích. Các đỉnh của chúng có thuộc tính độ cao, các cạnh có thuộc tính độ dốc và hướng. Mô hình này khá đơn giản và kinh tế.

Có nhiều phương pháp hình thành TIN, một phương pháp phổ biến là dựa trên cơ sở xây dựng tam giác Delauney. Năm 1934 Delauney đã chứng minh nếu nối từng cặp vị trí trong sơ đồ Voronoi có hai đa giác Voronoi của chúng cùng chia sẻ một cạnh Voronoi bởi đoạn thẳng thì ta có các tam giác, các tam giác này thỏa mãn tiêu thức Delauney khi vòng tròn ngoại tiếp của chúng không bao bất cứ điểm địa lý nào khác.

3.6. Biến đổi từ vectơ sang raster và ngược lại

So sánh mô hình raster và vectơ

Mô hình raster	Mô hình vectơ
Ưu điểm	
Mô hình hiệu quả	Thuận tiện biểu diễn hiện tượng tự nhiên
Dễ tổ hợp, nạp chồng	Mô hình cô đọng
Hướng ảnh vệ tinh	Có khả năng tạo lập tô pô lưới, tổng quát hóa, dễ sửa đổi
Có khả năng mô phỏng	Thao tác hình học dễ dàng, chính xác
Dễ phân tích dữ liệu	
Nhược điểm	
Chất lượng đồ họa hạn chế	Cấu trúc dữ liệu phức tạp

Khó mô hình hóa mạng	
Biến đổi phi tuyến phức tạp	

Phần lớn các hệ thống GIS trên cơ sở vectơ đều sử dụng thiết bị đồ họa trên công nghệ raster. Mỗi khi hiển thị dữ liệu vectơ trên màn hình, máy in... thì phải raster hóa nhờ các giải thuật biến đổi vectơ – raster.

Với hệ thống GIS trên cơ sở raster thì biến đổi raster - vectơ là thao tác cơ bản để sử dụng dữ liệu bản đồ.

3.6.1. Biến đổi vectơ sang raster

K/n: tìm tập pixel trong không gian raster trùng khớp với vị trí của điểm, đường hay đa giác trong biểu diễn vectơ. Tổng quát là tiến trình xấp xỉ vì với vùng không gian cho trước thì mô hình raster chỉ có khả năng địa chỉ hóa các vị trí nhờ tọa độ nguyên.

Raster hóa đường thẳng

Thuật toán raster hóa đoạn thẳng được thực hiện theo cách tăng dần, bắt đầu từ điểm cuối của đường.

Thuật toán cơ bản

+ Tổng số pixel tối thiểu tạo nên đoạn thẳng được xác định bởi vị trí pixel giữa hai đầu đoạn thẳng theo chiều x, y. Nếu Ab là đoạn thẳng cần raster hóa, A(x1, y1); B(x2, y2) trong hệ tọa độ raster thì khoảng cách x, y sẽ là:

$$Dx = \text{abs}(x1 - x2)$$

$$Dy = \text{abs}(y1 - y2)$$

Tổng số pixel cần vẽ là

$$D_{\text{max}} = \max(Dx, Dy)$$

$$n = D_{\text{max}} + 1$$

Giá trị dịch chuyển của x, y được xác định theo biểu thức sau:

$$\text{Incx} = Dx / D_{\text{max}} \quad \text{Incy} = Dy / D_{\text{max}}$$

+ Để tìm vị trí pixel tiếp theo ta phải làm tròn tọa độ thành số nguyên gần nhất sau khi tăng chiều x, y như sau:

For i = 1 to n-2 {

$$x = x + \text{incx}$$

$$y = y + \text{incy}$$

$$ix = \text{round}(x)$$

```

iy = round(y)

setpixel(ix, iy);

}

```

Thuật toán Bresenham dành cho đường thẳng có hệ số góc lớn hơn 0.

Dựa trên tương quan của Dx và Dy mà ta quyết định điểm nào sẽ được vẽ tiếp theo hay biến thiên theo chiều nào.

+ Dx > Dy ta cho y biến thiên theo x

+ Dy >= Dx ta cho x biến thiên theo y

Để tăng tốc độ thực hiện ta thay số thực bằng số nguyên và tránh phép toán nhân, chia.

Giả sử Dx:Dy= 8:12 thì y sẽ biến thiên theo x với tỉ lệ này, tức x tăng thêm 12 điểm thì y mới tăng 8 điểm hay khi x tăng 1 thì y tăng 8/12, ta không thực hiện chia ngay mà lưu lại trong biến d, đến khi nào d > Dx thì cho y tăng 1 và giảm d đi Dx.

Tương tự với hệ số góc nhỏ hơn 0 thì ta cho một chiều tăng, chiều kia giảm.

Raster hóa đa giác

Tiến trình raster hóa đa giác đòi hỏi phải tìm các pixel nằm trong nó.

Quá trình Raster sử dụng thuật toán biến đổi đường quét đa giác, thuật toán này xác định khá tốt các điểm nằm trong đa giác nhưng raster hóa đường biên sẽ phát sinh lỗi do kích thước điểm ảnh. Ở đây ta sử dụng hai phương pháp để xác định pixel có nằm trên biên hay không.

+ Phương pháp “Tâm điểm”: nếu tâm của pixel nằm trong đa giác thì nó thuộc đa giác.

+ Phương pháp “Đơn vị trội”: Nếu diện tích phần pixel thuộc đa giác lớn hơn phần còn lại thì nó thuộc đa giác.

Hai phương pháp này thường dẫn đến các kết quả khác nhau khi raster hóa.

Cố kết đường quét

Một cách raster hóa đa giác khác đó là tìm ra các hàng pixel nằm gần kề của đa giác thông qua thuật toán tìm các biên của tập hợp liên tục các pixel của một hàng hay một đường quét ảnh để tạo thành khối đa giác.

Thuật toán biến đổi đường quét của đa giác thực hiện như sau:

+ Khảo sát từng đường quét đi qua đa giác để tìm tọa độ giao điểm giữa chúng và cạnh đa giác. Các cạnh song song với đường quét sẽ được bỏ qua.

+ Sắp xếp tọa độ giao điểm theo thứ tự tăng dần và làm đầy đường quét giữa các cặp điểm.

Để tăng tốc độ thực hiện ta cần tăng tốc độ tìm giao điểm, điều này được giải quyết dễ dàng khi đã xác định được độ dốc của đa giác.

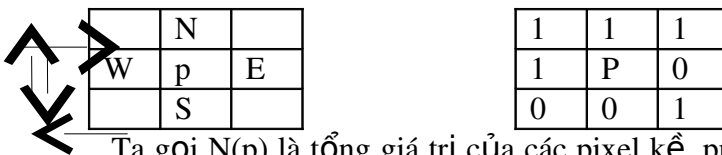
3.6.2. Biến đổi raster sang vectơ

Các bước thực hiện:

- + Chọn ngưỡng: chuyển đổi cường độ ảnh về ảnh hai màu.
- + Làm trơn: loại bỏ các biến dạng do nhiễu, các điểm đốm.
- + Làm mảnh: làm mảnh đường thẳng sao cho độ rộng của chúng bằng 1 pixel.
- + Mã xâu: chuyển ảnh vectơ thành tập các xâu pixel, mỗi xâu biểu diễn một đường.
- + Giảm thiểu vectơ: mỗi xâu pixel được chuyển vào dãy vectơ.

Làm mảnh

Ở đây ta cần chú ý đến thuật toán làm mảnh. Giá sử ta cần làm mảnh màu đen trong một ảnh nhị phân. Xét với 1 pixel P có tám pixel kề.



Ta gọi $N(p)$ là tổng giá trị của các pixel kề, p_n, p_e, p_s, p_w là giá các pixel trên, dưới, trái, phải. T_p là số lần biến đổi của pixel từ 0 \rightarrow 1 theo chiều kim đồng hồ. Trong ví dụ $N(p) = 5$; $p_n = 1$; $p_e = 0$; $p_s = 0$; $p_w = 1$; $T_p = 2$

Thuật toán được thực hiện như sau:

Bước 1: Với mỗi điểm ảnh P ta thực hiện

Nếu $2 \leq N(p) \leq 6$ và $T(p) = 1$ và $p_n.p_e.p_s = 0$ và $p_e.p_s.p_w = 0$ thì đánh dấu P

Bước 2: Gán giá trị 0 cho các điểm đánh dấu, nếu không có điểm đánh dấu thì dừng lại.

Bước 3: Quay lại bước 1.

Mã xâu hay tạo lập xâu

Xâu được hình thành từ các pixel mảnh, cần xác định xem mỗi pixel là nằm ở giữa, đầu hay cuối đoạn thẳng. Thuật toán tạo xâu sẽ tìm pixel tạo ra điểm cuối "sợi", sau đó duyệt theo các pixel trên đường và dừng lại ở điểm cuối hay giao điểm. Như vậy trật tự các pixel đã được tạo ra hay xâu pixel được tạo.

Câu hỏi ôn tập.

1. Thực hiện thuật toán raster hóa đường thẳng với dữ liệu như sau:
 - Nhập tọa độ điểm đầu và cuối $A(x_1, y_1)$, $B(x_2, y_2)$ của đoạn AB trong hệ tọa độ đề các. Nhập tỉ lệ kích thước lưới raster tương ứng với hai chiều của hệ tọa độ. Đưa ra phần lưới raster chứa đoạn thẳng AB.
2. Thực hiện thuật toán làm mảnh.

Cho file ảnh đen trắng, đưa ra ảnh sau khi làm mảnh.

CHƯƠNG 4. XÂY DỰNG CƠ SỞ DỮ LIỆU GIS

Đầu tiên ta phải thu thập dữ liệu, đây là khâu quan trọng nhất và khó khăn nhất trong quá trình lập một hệ thống GIS, dữ liệu thu thập được phải được xử lý sơ bộ trước khi thực hiện các thao tác trên chúng.

4.1. Thu thập thông tin địa lý và các phương pháp nhập dữ liệu địa lý

GIS có rất nhiều loại dữ liệu cần phải thu thập tùy vào mục tiêu phản ánh khác nhau của hệ thống. Trong thời gian ban đầu mới xuất hiện GIS, người sử dụng thường tự phát triển khuôn mẫu riêng để lưu trữ GIS, do vậy rất khó khăn khi chia sẻ và chuyển đổi giữa các hệ GIS với nhau. Cộng thêm việc nhập dữ liệu vào GIS rất tốn kém thời gian, thường tiêu tốn tới 80% ngân sách của dự án. Vì vậy, việc chia sẻ dữ liệu trở nên thông thường trong GIS và đẩy nhanh quá trình chuẩn hóa dữ liệu và phát sinh khái niệm dữ liệu về dữ liệu để mô tả nguồn gốc và độ chính xác của các lớp GIS. Nếu dữ liệu có sẵn không phù hợp thì phải tự số hóa bản đồ.

Dữ liệu trong GIS bao gồm dữ liệu hình học và phi hình học (các thuộc tính). Dữ liệu được thu thập trong hệ thống thông tin địa lý thường:

Dưới dạng số như bản đồ số hóa, CSDL, bảng tính, ảnh vệ tinh...

Phi số hóa như bản đồ giấy, ảnh chụp, các bản vẽ...

Dữ liệu trắc địa

Tổng hợp từ các nguồn khác

Dữ liệu được thu thập qua hai phương pháp chính. Thứ nhất là thu thập từ chính các đối tượng, phương pháp này cho độ chính xác cao nhưng tốn kém, chúng bao gồm:

Trắc địa mặt đất: dùng để thu thập dữ liệu tô pô có tỷ lệ lớn như đo vẽ địa hình, đất đai. Kết quả của trắc địa là các vectơ hai hoặc ba chiều, độ chính xác trong khoảng vài xentimet đến đêximet.

Phương pháp định vị bằng vệ tinh: thực hiện nhờ hệ thống định vị toàn cầu (GPS). Máy thu GPS hiển thị vị trí bằng chữ và số của lưới tọa độ chọn trước. Lưới tọa độ là mẫu vuông hay chữ nhật đặt trên bản đồ. Lưới sử dụng để mô tả vị trí nhờ tổ hợp chữ và số. Khi sử dụng GPS thì ta có địa lý vectơ hai và ba chiều. Độ chính xác của GPS được nâng cao khi có hai máy cùng xác định một điểm thường nằm trong khoảng 1-2 xentimet đến 10-20 met.

Chụp ảnh bằng máy bay hay vệ tinh: sử dụng các thiết bị hay cảm biến để thu thập từ xa các quan hệ phổ và không gian của đối tượng quan sát. Các ảnh do vệ tinh chụp được gửi về qua sóng radio và phải được xử lý theo để thu về thông tin cần thiết. Độ chính xác phương pháp phụ thuộc vào chất lượng ảnh truyền và cách xử lý dữ liệu ảnh.

Thứ hai là thu thập dữ liệu từ nguồn số hóa. Phương pháp này cho dữ liệu ít chính xác hơn, chi phí cũng thấp hơn. Chúng bao gồm:

Số hóa bằng tay các bản đồ giấy.

Sử dụng các CSDL số hóa có sẵn.

Số hóa bản đồ giấy tự động bằng máy quét.

4.1.1. Tìm kiếm bản đồ phù hợp

Tìm kiếm bản đồ phù hợp phụ thuộc vào lựa chọn vùng, chủ đề và chu kỳ phù hợp, thời gian phù hợp. Phần lớn thư viện bản đồ chứa mô tả vùng, chủ đề, thời gian phát hành, tuy nhiên ít khi thời gian thu thập dữ liệu thích hợp với thời gian phát hành bản đồ. Như hệ thống GIS phục vụ trong giao thông cần quan tâm đến các lớp bản đồ chứa thông tin về đường biên hành chính, bờ biển, đường đi, đường sắt, các thành phố, quận huyện... Các lớp này có thể chia thành các lớp con khác như lớp đường sắt chia thành đường đơn và đôi, đường đi có lớp đường quốc lộ, đường nội thành (một hoặc hai chiều)...

4.1.2. Số hóa bản đồ

Khi chuẩn bị bản đồ để số hóa hay quét thì cần coi trọng phương pháp qui chiếu không gian. Tiến trình tham chiếu địa lý là tiến trình cung cấp địa chỉ không gian có thể thực hiện nhờ sử dụng hệ tọa độ địa lý, tọa độ lưới hay không dùng tọa độ. Khi sử dụng hệ tọa độ địa lý, trái đất được coi là hình cầu. Khi sử dụng hệ tọa độ lưới, trái đất coi như mặt phẳng. Hệ tọa độ địa lý Utm tổ hợp cả hệ tọa độ địa lý và tọa độ lưới vì nó là tập hợp các mặt phẳng bao quanh trái đất.

Nhiều hệ GIS không lưu trữ tọa độ theo khuôn mẫu độ, phút, giây mà lưu trữ tọa độ hệ thập phân. Do vậy, cần phải công thức chuyển đổi giữa chúng như sau:

$$\text{DecimalDegrees} = \text{Degrees} + (\text{Minutes} + \text{seconds}/60)/60$$

VD: chuyển đổi 30o35'45" sang độ thập phân

$$\text{DecimalDegrees} = 35 + (35 + 45/60)/60 = 35.59583$$

Số hóa bản đồ giấy bằng bàn số hóa hay nhập bản đồ giấy bằng máy quét là phương pháp thường được sử dụng. Số hóa bằng bàn số hóa là tiến trình ghi lại vị trí của trình tự các điểm đặc trưng dọc theo đường quét trên bản đồ. Phương pháp này cho kết quả dưới dạng bản đồ vectơ hai chiều. Kết quả nhập bản đồ số bằng máy quét là số liệu raster, nếu sử dụng công cụ phần mềm chuyển đổi raster sang vectơ thì cũng có số liệu bản đồ hai chiều. Độ chính xác của dữ liệu phụ thuộc trước hết vào độ chính xác bản đồ giấy.

Bàn số hóa

Bao gồm bàn nhỏ chứa lưới dây kim loại mịn đặt theo các trục Đềcác. Con trỏ chứa cuộn dây kim loại và được nối với bàn để người thao tác xác định điểm cần ghi. Vị trí chính xác được xác định nhờ dấu thập phân mỏng tại đầu con chạy. Trên con trỏ còn có các phím để nhập dữ liệu hay nhập mã lệnh. Công nghệ chung nhất của bàn số hóa dựa trên nguyên tắc điện từ. Lưới dây trong bảng nhỏ và cuộn dây của con trỏ hoạt động như bộ thu, bộ phát hay ngược lại. Nếu con trỏ là bộ phát thì bộ điều khiển của bàn sẽ quét lưới dây kim loại theo chiều X và Y để tìm ra vị trí chữ thập mỏng đầu con trỏ.

Mục tiêu chính của tiến trình chuyển đổi thông tin không gian từ dạng tuyến tính sang dạng số hóa là đảm bảo quan hệ bản đồ, bao gồm:

- Đảm bảo các liên kết tồn tại giữa các điểm.

- Các đường song song được bảo toàn.
- Bảo toàn các vị trí tương đối, tuyệt đối, tính liên kết.
- Các đường gần sát không được cắt nhau.

Tiến trình số hóa thường bắt đầu bằng cố định bản đồ gốc lên mặt bàn số hóa. Trước hết phải số hóa các điểm điều khiển hay lưới tạo độ để đăng ký hệ thống tọa độ, sau đó mới đến các đặc trưng. Một phần quan trọng của số hóa bản đồ là đảm bảo thông tin tham chiếu địa lý, biểu thị bằng lưới và các đường kinh vĩ tuyến cần được lưu trữ theo bản đồ số hóa. Phần lớn bản đồ tỷ lệ trung bình và tỷ lệ lớn (trên 1:50000) đều chứa lưới vùng cơ sở cho phép chuyển đổi tuyến tính từ tọa độ số hóa sang tọa độ lưới. Nếu giả sử bản đồ không bị biến dạng hay không đòi hỏi độ chính xác cao thì chỉ cần ba điểm điều khiển tạo độ không thẳng hàng, nếu cần chính xác cao thì cần nhiều điểm hơn (thường là 20 điểm).

Bàn số hóa thường có hai chế độ nhập tọa độ, đó là điểm và đường. Trong chế độ điểm, bàn số hóa chỉ phát sinh tọa độ mỗi khi nhấn phím trên con chạy. Chế độ này thường áp dụng để số hóa đặc trưng điểm như độ cao điểm, thành phố trong bản đồ tỷ lệ nhỏ và các đặc trưng đường hay đa giác như hệ thống sông ngòi, đường biên hành chính. Chế độ đường nhằm tăng tốc độ số hóa các đặc trưng đường hay đa giác vì chúng tự động ghi lại tọa độ các điểm khi thao tác viên dịch chuyển con trỏ theo đường số hóa. Việc tự động ghi tọa độ dựa trên cơ sở thời gian hay khoảng cách. Chế độ đường thường làm phát sinh một lượng lớn dữ liệu dư thừa vì thế chế độ điểm được sử dụng nhiều hơn.

Mã hóa đặc trưng bản đồ

Sau khi số hóa các đặc trưng điểm, đường và đa giác, ta phải cung cấp thông tin kèm theo để chỉ ra ý nghĩa của chúng. Chúng có thể là thông tin phân loại lớp như khu nhà, đường đi hay chỉ danh như tên đất, tên đường... Các thông tin này thường được nhập từ bàn phím.

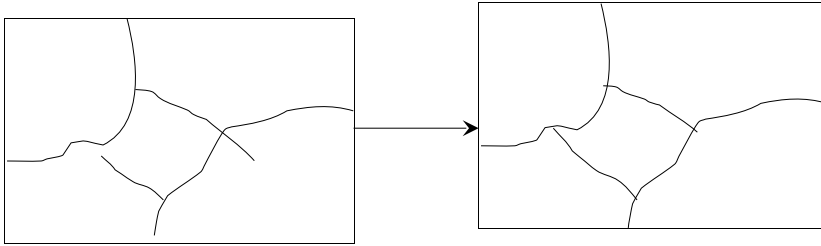
Mã hóa điểm, đường thường không gặp khó khăn nhưng với đa giác khá phức tạp vì mỗi đoạn, điểm liên kết có thể gắn với các đa giác khác. Nếu mỗi đa giác được mã hóa độc lập thì phải số hóa hai lần các đoạn thẳng chung nhau, một giải pháp tránh số hóa hai lần là liên kết với cặp mã đặc trưng cho chúng. Cặp mã này đại diện cho các đa giác trái, phải của đoạn thẳng (khái niệm trái, phải phụ thuộc vào hướng của đoạn thẳng). Sau khi mã hóa đoạn thẳng biên đa giác theo cách này ta có thể xây dựng lại đa giác qua việc tìm kiếm các liên kết đoạn thẳng trái, phải của chúng, các đoạn thẳng tìm được phải được sắp xếp để có được một khâu điểm đầu và điểm cuối. Mỗi liên kết chỉ có duy nhất một chỉ danh, dữ liệu địa lý được tạo ra nhờ tham chiếu đến danh sách các chỉ danh của liên kết.

Tìm kiếm và hiệu chỉnh lỗi

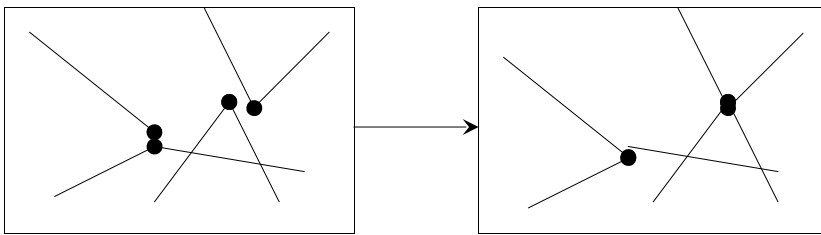
Mọi bản đồ số đều tồn tại các lỗi do xây dựng từ nguồn dữ liệu chính xác không cao hay do quá trình số hóa nên công việc sửa lỗi là cần thiết. Các lỗi do thao tác viên tạo nên thường dễ sửa chữa hơn là xuất phát từ dữ liệu gốc. Các lỗi thường gặp khi số hóa bản đồ để hình thành đa giác, bao gồm:

- Xâu đoạn thẳng lơ lửng do vẽ quá hay vẽ hụt.
- Không gán nhãn cho đa giác.

- Đụng độ nhãn trong đa giác.
- Xâu đoạn thẳng có cùng nhãn, đa giác trái, phải.
-



Hình 4.1: Lỗi đoạn thẳng lơ lửng



Hình 4.2: Chập nút

Trong mạng các xâu đoạn thẳng hình thành đa giác thì các xâu phải kết thúc tại các điểm nút. Khi số hóa chúng hay bị vẽ quá hoặc hụt, các lỗi này có thể phát hiện ngay khi số hóa và hiệu chỉnh qua việc kéo các đầu nút, trong hệ thống GIS chức năng này có sẵn được gọi là “chập”, có khả năng chập hai hay nhiều điểm cách nhau một khoảng cho trước lại.

Khi số hóa mỗi đa giác cần một điểm để đặt nhãn và lỗi trong quá trình này là thiếu nhãn hay thừa nhãn. Lỗi về đụng độ đa giác là do thao tác viên thực hiện không chính xác. Các lỗi này thường phải làm lại ngay trong quá trình số hóa.

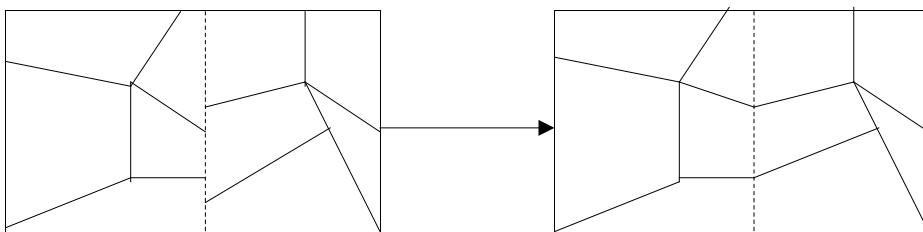
Lỗi khi số hóa còn xảy ra khi số hóa nhiều lần cùng một xâu, điều này có thể phát sinh các đa giác lạ hoặc không có điểm tham chiếu dẫn đến không xác định được chính xác đa giác, lỗi này thường là do đặt sai vị trí điểm số hóa hay sai trật tự số hóa.

Sử dụng tệp dữ liệu bản đồ có sẵn

Các dữ liệu có sẵn thường là các tệp đường biên, chúng mô tả hình học dưới dạng số của các đơn vị hành chính. Cần phải tìm các dữ liệu có sẵn phù hợp với hệ GIS cần xây dựng và lưu trữ dưới cấu trúc dữ liệu chuẩn.

Tích hợp dữ liệu bản đồ số hóa từ nhiều nguồn khác nhau

Khi số hóa một bản đồ vào nhiều thời điểm khác nhau hay sử dụng các hệ tọa độ khác nhau thường xảy ra hiện tượng biến dạng cạnh sườn, các đối tượng như đường đi, sông ngòi... hay các đặc trưng topology trải trên nhiều bản đồ hay mất tính liên tục.



Hình 4.3: Biến dạng cạnh sườn

Sau khi số hóa bản đồ dữ liệu thu được phần lớn dưới dạng mô hình vectơ song GIS sử dụng linh hoạt cả hai mô hình vectơ và raster do đó cần phải chuyển đổi giữa hai mô hình này. Mỗi dữ liệu vectơ trên màn hình, máy in... thì phải raster hóa nhờ các thuật toán biến đổi vectơ-raster trong hệ tọa độ. Với hệ thống GIS trên cơ sở raster thì thao tác biến đổi vectơ-raster là thao tác căn bản để hệ thống sử dụng được dữ liệu vectơ. Với hệ thống GIS trên cơ sở vectơ thì việc thu thập dữ liệu thông qua máy quét là dạng raster cần chuyển đổi để sử dụng.

4.2. Cơ sở dữ liệu thông tin địa lý

4.2.1. Tập cơ sở dữ liệu

Như ta đã biết địa lý trong GIS lưu dưới dạng CSDL không gian, nội dung các tập trong CSDL không gian được phân biệt bởi kiểu đối tượng hình học hay lớp các đặc trưng bản đồ. Do vậy, chúng có thể là tập thông tin tham chiếu điểm (ví trí thành phố, làng, nhà ở), tập thông tin tham chiếu đường (quốc lộ, sông ngòi...), ngoài ra có thể là tập lưu trữ các danh mục mã mô tả đặc trưng và tập chỉ số đến nội dung các tập khác trong hệ thống.

Tập dữ
liệu địa lý

100	144297	607410	20000	Hà Nội
102				Hải Phòng
220				Đà Nẵng
...				
123				
145				
Chỉ danh	Tọa độ X	Tọa độ Y	Diện tích	Tên

Các bảng sau đây mô tả một cách thức lưu trữ thông tin địa lý trên tập:

Tên trường	Mô tả	Loại	Số byte
FEATMO	Số hiệu duy nhất của đặc trưng	Int	4
EASTING	Tọa độ phía đông	Int	4
NORTHING	Tọa độ phía bắc	Int	4
...			
AREA	Diện tích		
POPULATION	Dân số	Int	4
NAME	Tên thành phố	Char	30
COUNTRY	Tên quốc gia	Char	30

Tên trường	Mô tả	Loại	Số byte
Bản ghi đầu tệp			
RFEATMO	Số hiệu duy nhất của đặc trưng	Int	4
RCLASS	Lớp đặc trưng (sông ngòi, quốc lộ)	Char	2
RNAME	Tên đối tượng	Char	30
NCOORDS	Tổng số tọa độ	Int	2
Bản ghi tọa độ			
LAT (1)	Kinh độ	Float	4
LONG (1)	Vĩ độ	Float	4
...			
LAT (n)	Kinh độ	Float	4
LONG (n)	Vĩ độ	Float	4

Các bản ghi trên được hình thành từ trộn các trường số và ký tự. Dữ liệu số không phải lúc nào cũng lưu dưới dạng số nguyên hay thực, có thể sử dụng các chuỗi thay thế nhưng kém hiệu quả hơn do yêu cầu kích thước dữ liệu nhiều hơn. Do vậy, dữ liệu bản đồ thường được lưu dưới dạng số trong các tệp nhị phân. Còn các tệp độc lập với dữ liệu bản đồ lưu dưới dạng tệp ký tự.

Cũng giống như CSDL thông thường, các tệp của CSDL địa lý được tổ chức theo các dạng tệp tuần tự, tệp chỉ số, tệp chỉ số sắp xếp, danh sách kết nối, cây.

Tệp dữ liệu tuần tự:

...	Record 5	Record 4	Record 3	Record 2	Record 1	
-----	----------	----------	----------	----------	----------	--

Tệp dữ liệu chỉ số:

Khóa	Chỉ số
Hà Nội	1
Hải Phòng	2
Viên Chăn	3
...	

STT	Tên thành phố	Nước	Dân số	...
1	Hà Nội	Việt Nam		
2	Hải Phòng	Việt Nam		
3	Viên Chăn	Lào		
...				

Tập dữ liệu chỉ số sắp xếp:

Khóa	Chỉ số
Hải Phòng	2
Hà Nội	1
Viên Chăn	3
...	

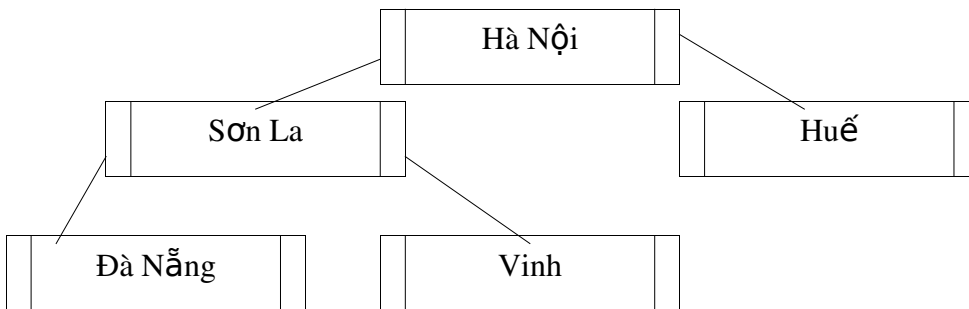
STT	Tên thành phố	Nước	Dân số	...
1	Hà Nội	Việt Nam		
2	Hải Phòng	Việt Nam		
3	Viên Chăn	Lào		
...				

Danh sách kết nối:

Khóa	Chỉ số
Việt Nam	1
Lào	2
Thái Lan	3
...	

STT	Tên thành phố	Nước	...	Con trỏ
1	Hà Nội	Việt Nam		2
2	Hải Phòng	Việt Nam		4
3	Viên Chăn	Lào		5
...				

Cây: (ví dụ sơ đồ đường không)

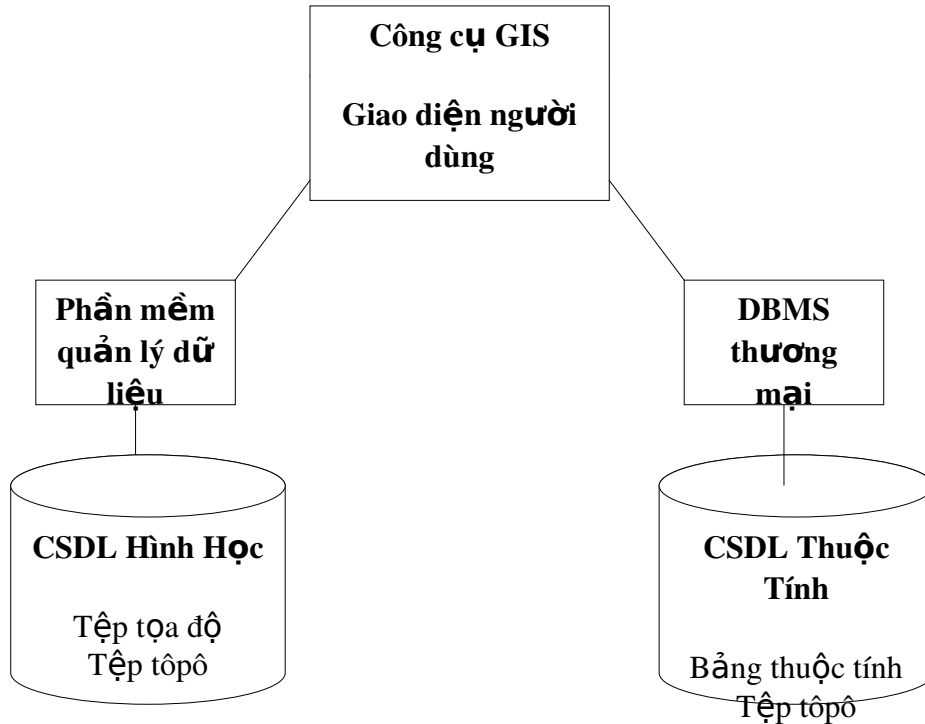


4.2.2. Ứng dụng CSDL trong GIS

Ngày nay, hầu hết các hệ thống thông tin được xây dựng xung quanh CSDL quan hệ. Tuy nhiên, GIS đòi hỏi CSDL của mình không những lưu trữ được các đối tượng mà còn có khả năng tìm kiếm trực tiếp và tính toán dữ liệu không gian. Do đó, hệ thống GIS thương mại đều được xây dựng dựa trên một trong ba kiến trúc sau:

Kiến trúc đối ngẫu

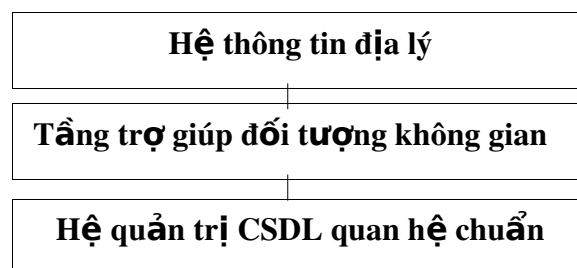
GIS có hai hệ thống CSDL tách biệt, một cho dữ liệu hình học và một cho dữ liệu thuộc tính. Ta gọi chúng là có kiến trúc đối ngẫu, kiến trúc này tách thành các hệ con để lưu trữ và truy nhập dữ liệu không gian, và thông tin thuộc tính được lưu trữ trong RDBMS. Các thành tố hình học và thuộc tính của các đối tượng được liên kết với nhau bởi chỉ danh duy nhất. Để xâm nhập đối tượng phải truy nhập cả hai hệ con sau đó tổ hợp kết quả. Lợi thế của kiến trúc này là từng phần trên cơ sở RDBMS chuẩn cho nên việc lưu trữ và truy nhập có hiệu quả cao nhưng bất lợi ở chỗ không bảo đảm tính toàn vẹn dữ liệu như trong trường hợp thực thể vẫn tồn tại trong hệ con lưu trữ không gian trong khi đã bị xóa trong RDBMS.



Hình 4.4: Kiến trúc đối ngẫu của GIS

Kiến trúc tầng

Kiến trúc tầng lưu cả dữ liệu không gian trong mô hình dữ liệu quan hệ, khi đó thực thể hình học phải được bẻ ra làm nhiều phần để lưu vào các bảng khác nhau. Truy nhập thực thể gốc phải được thực hiện kết nối các quan hệ, dẫn tới hệ thống chạy chậm và khó sử dụng, để tránh khó khăn ta hình thành giao dịch không gian trong tầng đỉnh của CSDL quan hệ chuẩn. Tầng này có trách nhiệm thông dịch câu lệnh truy vấn hình học sang câu lệnh SQL chuẩn, kết quả sẽ cho truy nhập không gian nhanh hơn nhưng truy vấn phức tạp hơn.

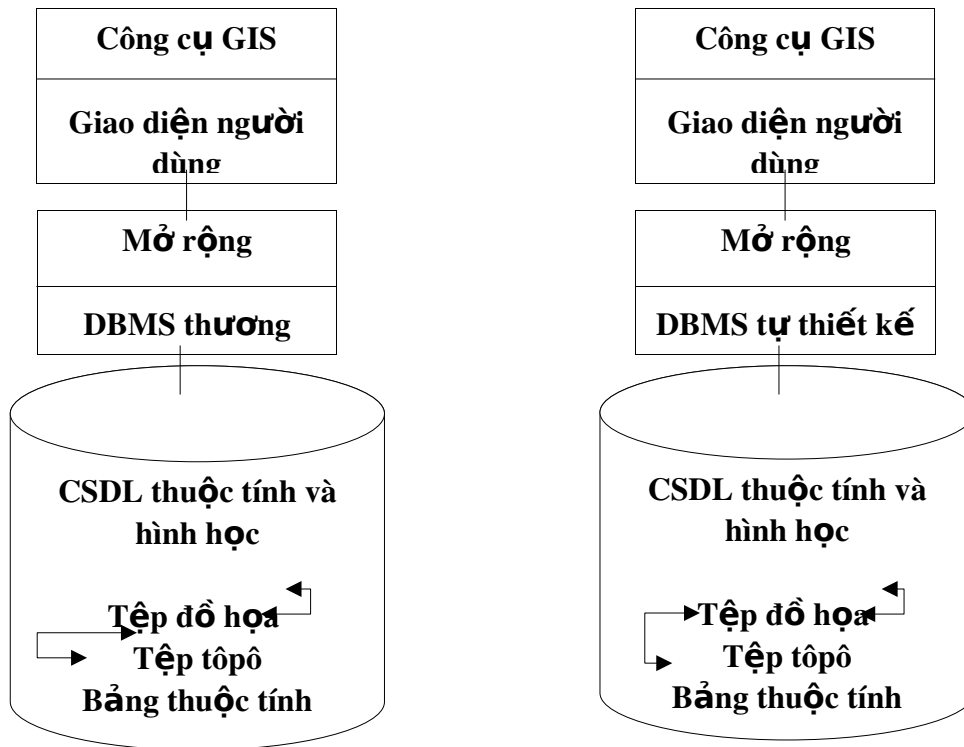


Hình 4.5: Kiến trúc phân tầng của GIS

Kiến trúc tích hợp

Phần mở rộng không gian được tích hợp vào DBMS, khi đó ngôn ngữ truy vấn được mở rộng bởi khả năng truy vấn các kiểu hình học và các toán tử không gian như tính toán khoảng cách, giao điểm, chu vi... Mô hình GIS kiểu này có hai giải pháp tích hợp là:

- Mở rộng quan hệ quản trị dữ liệu thương mại chuẩn cho dữ liệu không gian.
- Hệ quản trị CSDL mới theo hướng đối tượng.



Hình 4.6. Kiến trúc tích hợp của hệ GIS

Câu hỏi ôn tập.

1. Nêu các phương pháp thu thập dữ liệu địa lý.
2. Nêu quá trình số hóa bản đồ.
3. Cho biết các mô hình CSDL ứng dụng trong GIS .

MỘT SỐ ĐỀ THI MẪU

Trường Đại Học Hàng Hải Việt Nam
Khoa Công nghệ Thông tin
BỘ MÔN HỆ THỐNG THÔNG TIN
 -----***-----

THI KẾT THÚC HỌC PHẦN

Tên học phần: CÁC HỆ THỐNG THÔNG TIN SỐ	<u>Đề thi số:</u>	<u>Ký duyệt đề:</u>
Năm học: x	x	x
Thời gian: 60 phút		

Câu 1: (2 điểm)

GIS là gì. Mục đích và Ứng dụng của GIS?

Câu 2: (3 điểm)

Trình bày hiểu biết của anh chị về các kỹ thuật thu thập dữ liệu không gian?

Câu 3: (3 điểm)

Trình bày hiểu biết của anh chị về các hàm kết hợp dữ liệu không gian với các toán tử Vector?

Câu 4: (2 điểm)

Cho biết quá trình số hóa bản đồ?

-----***HẾT***-----

Lưu ý: - Không sửa, xóa đề thi, nộp lại đề sau khi thi

Trường Đại Học Hàng Hải Việt Nam
Khoa Công nghệ Thông tin
BỘ MÔN HỆ THỐNG THÔNG TIN
 -----***-----

THI KẾT THÚC HỌC PHẦN

Tên học phần: CÁC HỆ THỐNG THÔNG TIN SỐ	<u>Đề thi số:</u>	<u>Ký duyệt đề:</u>
Năm học: x	x	x
Thời gian: 60 phút		

Câu 1: (2 điểm)

Cho biết kiến trúc và chức năng của GIS?

Câu 2: (3 điểm)

Trình bày hiểu biết của anh chị về các hệ tham chiếu không gian?

Câu 3: (3 điểm)

Trình bày hiểu biết của anh chị về các hàm kết hợp dữ liệu không gian với các toán tử Raster?

Câu 4: (2 điểm)

Cho biết các hình thức của bản đồ?

-----***HẾT***-----

Lưu ý: - Không sửa, xóa đề thi, nộp lại đề sau khi thi

Trường Đại Học Hàng Hải Việt Nam
Khoa Công nghệ Thông tin
BỘ MÔN HỆ THỐNG THÔNG TIN
 -----***-----

THI KẾT THÚC HỌC PHẦN

Tên học phần: CÁC HỆ THỐNG THÔNG TIN SỐ	<u>Đề thi số:</u>	<u>Ký duyệt đề:</u>
Năm học: x	x	x
Thời gian: 60 phút		

Câu 1: (2 điểm)

GIS là gì. Trình bày quá trình tìm kiếm và phân tích không gian?

Câu 2: (3 điểm)

Trình bày quá trình thu thập dữ liệu, thuận lợi và khó khăn trong quá trình này?

Câu 3: (3 điểm)

Trình bày mô hình vectơ?

Câu 4: (2 điểm)

Trình bày quá trình vectơ hóa raster?

-----***HẾT***-----

Lưu ý: - Không sửa, xóa đề thi, nộp lại đề sau khi thi

Trường Đại Học Hàng Hải Việt Nam
Khoa Công nghệ Thông tin
BỘ MÔN HỆ THỐNG THÔNG TIN
 -----***-----

THI KẾT THÚC HỌC PHẦN

Tên học phần: CÁC HỆ THỐNG THÔNG TIN SỐ	<u>Đề thi số:</u>	<u>Ký duyệt đề:</u>
Năm học: x	x	x
Thời gian: 60 phút		

Câu 1: (2 điểm)

GIS là gì. Các thành phần của GIS.

Câu 2: (3 điểm)

Trình bày hiểu biết của anh chị hệ thống phép chiếu UTM?

Câu 3: (3 điểm)

Trình bày quá trình raster hóa vectơ?

Câu 4: (2 điểm)

Cho biết các mô hình CSDL trong GIS ?

-----***HẾT***-----

Lưu ý: - Không sửa, xóa đề thi, nộp lại đề sau khi thi

Trường Đại Học Hàng Hải Việt Nam
Khoa Công nghệ Thông tin
BỘ MÔN HỆ THỐNG THÔNG TIN
 -----***-----

THI KẾT THÚC HỌC PHẦN

Tên học phần: CÁC HỆ THỐNG THÔNG TIN SỐ	<u>Đề thi số:</u>	<u>Ký duyệt đề:</u>
Năm học: x	x	x
Thời gian: 60 phút		

Câu 1: (2 điểm)

Cho biết kiến trúc và chức năng của GIS?

Câu 2: (3 điểm)

Trình bày hiểu biết của anh chị về các hệ tham chiếu không gian?

Câu 3: (3 điểm)

Trình bày cấu trúc Raster?

Câu 4: (2 điểm)

Trình bày quá trình raster hóa vectơ?

-----***HẾT***-----

Lưu ý: - Không sửa, xóa đề thi, nộp lại đề sau khi thi