

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

LUẬN VĂN THẠC SĨ KHOA HỌC

ỨNG DỤNG LOGIC MỜ TRONG
HỆ THỐNG THÔNG TIN ĐỊA LÝ (GIS)

NGÀNH: CÔNG NGHỆ THÔNG TIN

MÃ SỐ:.....

TRẦN VĂN ĐOÀI

Người hướng dẫn khoa học: TS. TRẦN ĐÌNH KHANG

HÀ NỘI 2006

LỜI CAM ĐOAN

Các kết quả nghiên cứu trong luận văn ngoài những vấn đề mang tính phổ biến mà tác giả đã đề cập tới dưới dạng các định nghĩa và khái niệm là hoàn toàn mới những vấn đề tham khảo cũng được trích dẫn cụ thể. Các hình vẽ, minh họa và kết quả thực nghiệm do chính tác giả thực hiện. Nội dung đề tài tác giả chưa công bố trên các công trình nghiên cứu khác. Tác giả xin chịu hoàn toàn trách nhiệm về nội dung của luận văn này.

Tác giả

Trần Văn Đoàn

LỜI CẢM ƠN

Luận văn của em sẽ rất khó hoàn thành nếu không có sự truyền đạt kiến thức quý báu và sự hướng dẫn tận tình của thầy giáo Trần Đình Khang.

Em xin chân thành cảm ơn những ý kiến phản biện quý báu của các thầy, cô giáo đã phản biện luận văn này. Và xin chân thành cảm ơn các ý kiến tham luận của các thầy, cô giáo và các bạn trong hội nghị khoa học lần thứ 20 trường Đại học Bách khoa Hà Nội. Qua đó em nắm bắt sự quan tâm của mọi người và các hướng cần làm rõ để người đọc có thể hiểu được ý đồ của tác giả.

Em xin chân thành cảm ơn quý thầy, cô trong Khoa Công nghệ Thông tin trường Đại học Bách khoa Hà Nội đã tận tình giảng dạy, truyền đạt những kiến thức quý báu trong suốt khóa học làm nền tảng cho em hoàn thành luận văn này. Em cũng xin cảm ơn các thầy, cô trong Trung tâm Đào tạo sau Đại học đã tạo mọi điều kiện để em hoàn thành khóa học và luận văn này.

Mặc dù đã cố gắng nỗ lực hết mình, song chắc chắn luận văn không khỏi còn thiếu sót. Em rất mong nhận được sự thông cảm và chỉ bảo tận tình của các thầy, cô giáo và các bạn cũng như những ai quan tâm tới lĩnh vực mà luận văn này thực hiện.

Hà Nội, ngày tháng 10 năm 2006

Tác giả

Trần Văn Đoàn

MỤC LỤC

	Trang
LỜI CAM ĐOAN.....	2
LỜI CẢM ƠN	3
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT	6
DANH MỤC CÁC BẢNG.....	7
DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ.....	8
MỞ ĐẦU.....	10
CHƯƠNG 1 - TỔNG QUAN.....	13
CHƯƠNG 2 - HỆ THỐNG THÔNG TIN ĐỊA LÝ (GIS).....	17
2.1 Khái niệm.....	17
2.1.1 Một số định nghĩa.....	18
2.1.2 Lịch sử phát triển của GIS.....	20
2.2 Thu thập dữ liệu.....	21
2.2.1 Thu thập dữ liệu không gian.....	22
2.2.2 Thu thập dữ liệu thuộc tính	22
2.3 Thao tác dữ liệu	22
2.4 Quản lý dữ liệu	22
2.5 Truy vấn và phân tích dữ liệu	23
2.6 Hiển thị dữ liệu.....	24
2.7 Mô hình dữ liệu	25
2.8 Các đối tượng trong GIS.....	26
2.9 Kết nối dữ liệu không gian và dữ liệu thuộc tính	34
2.10 Chồng xếp và phân tích trong GIS	35
CHƯƠNG 3 - ỨNG DỤNG LOGIC MỜ TRONG HỆ THỐNG THÔNG TIN ĐỊA LÝ	37
3.1 Giới thiệu chung.....	37
3.1.1 Nguyên lý mở rộng các hệ thống GIS.....	40
3.1.2 Tính không rõ ràng và hạn chế của Logic rõ trong GIS.....	40
3.1.3 Tính chất mờ trong các hệ thống GIS.....	43
3.2 Logic mờ trong GIS.....	44
3.2.1 Khái niệm về tập hợp rõ và tập hợp mờ.....	44
3.2.2 Hệ mờ trong GIS	51
3.2.3 So sánh giữa Logic mờ và logic rõ (logic kinh điển)	56
3.3 Mô hình dữ liệu không gian và các phép toán	57
3.3.1 Mô hình dữ liệu không gian	57
3.3.2 Phân lớp các phép toán GIS.....	58
3.4 Mở rộng mô hình dữ liệu với Logic mờ	61
3.5 Mở rộng các phép toán với Logic mờ	61

3.5.1 Phép toán phân lớp mờ (Fuzzy Reclassification).....	62
3.5.2 Phép toán vùng đệm mờ (Fuzzy Buffer).....	63
3.5.3 Khoảng cách mờ (Fuzzy Distance)	66
3.5.4 Chồng xếp mờ (Fuzzy Overlay)	68
3.5.5 Lựa chọn mờ (Fuzzy Select), tìm kiếm mờ	69
3.5.6 Suy luận mờ.....	70
3.6. Lựa chọn vị trí dựa trên một chuỗi các phép toán GIS.....	73
3.6.1 Lựa chọn vị trí sử dụng logic mờ	74
3.6.2 Bài toán ra quyết định không gian và logic mờ.....	75
CHƯƠNG 4 - GIẢI MỘT SỐ BÀI TOÁN BẰNG ỨNG DỤNG LOGIC MỜ	
TRONG GIS.....	79
4.1 Tìm vị trí mở rộng thành phố Thái Bình	79
4.1.1 Phát biểu bài toán	79
4.1.2 Phương pháp tiến hành	79
4.1.3 Kết quả đạt được.....	83
4.2 Bài toán xác định đường đi ngắn nhất sử dụng logic mờ	88
4.2.1 Phát biểu bài toán	88
4.2.2 Phương pháp tiến hành	88
4.2.3 Kết quả đạt được.....	90
4.3 Bài toán tìm vị trí xây dựng nhà máy xi măng	90
4.3.1 Phát biểu bài toán	90
4.3.2 Phương pháp tiến hành	91
4.3.3 Kết quả đạt được.....	94
KẾT LUẬN	96
TÀI LIỆU THAM KHẢO.....	97

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

GIS - Là từ viết tắt của: Geographic Information System

Logic mờ : Fuzzy Logic.

Tính không rõ ràng: Uncertainty.

WebGIS : Công nghệ đưa bản đồ lên mạng.

DBMS : Hệ quản trị cơ sở dữ liệu.

CSDL: Cơ sở dữ liệu.

Layer: Các đối tượng có cùng tính chất nào đó được nhóm với nhau.

Entities: Các thực thể không gian.

Table: Bảng dữ liệu (không gian và thuộc tính).

IFF: Nếu và chỉ nếu.

DANH MỤC CÁC BẢNG

Bảng 2.1. So sánh mô hình dữ liệu Vector và Raster.	26
Bảng 3.1. Bảng các phép toán logic rõ và mờ	49
Bảng 3.2. Bảng ví dụ mô tả các mục phân lớp	52
Bảng 3.3. Bảng minh họa ví dụ giải mờ	56
Bảng 3.4. Bảng so sánh Logic mờ và Logic rõ.....	57
Bảng 3.5. Bảng phân lớp các phép toán trong GIS.....	61
Bảng 3.6. Bảng minh họa độ thuộc về địa tầng.....	63
Bảng 3.7. Bảng minh họa độ thuộc về độ dốc	63
Bảng 4.1. Bảng mờ hóa lớp thông tin đất	80
Bảng 4.2. Bảng mờ hóa lớp thông tin địa tầng	81

DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ

Hình 2.1. Mô hình một hệ thống thông tin địa lý	19
Hình 2.2. Lịch sử phát triển của GIS.	21
Hình 2.3. Mô tả phân tích liên kề.....	24
Hình 2.4. Mô tả phân tích chồng xếp theo thời gian.....	24
Hình 2.5. Các đối tượng điểm trong GIS.....	27
Hình 2.6. Các đối tượng dạng đường trong GIS.....	27
Hình 2.7. Các đối tượng dạng vùng trong GIS.	28
Hình 2.8. Các đối tượng dạng lưới trong GIS.....	29
Hình 2.9. Phân tách bản đồ thành các lớp.....	31
Hình 2.10. ảnh vệ tinh cũng được xử lý trong GIS.....	31
Hình 2.11. Mô tả quan hệ lân cận hai Polygon P1 và P2	33
Hình 2.12. Mô tả quan hệ bao hàm(polygon đảo).....	34
Hình 2.13. Mô tả quan hệ giao nhau của hai polygon	34
Hình 2.14. Bảng mô tả các trường dữ liệu trong GIS.....	35
Hình 2.15. Chồng xếp chuỗi các lớp bản đồ trong GIS.....	36
Hình 3.1. Nguyên lý mở rộng các hệ GIS	40
Hình 3.2. Tính không rõ ràng trong GIS (Zhang & Goodchild 2002)	41
Hình 3.3. Phân loại tính chất không rõ ràng trong GIS.....	41
Hình 3.4. Tính chất không rõ ràng phát sinh khi xác định ranh giới	41
Hình 3.5. Một số hàm mờ và phạm vi tập rõ.....	45
Hình 3.6. Hàm mờ tuyến tính.....	46
Hình 3.7. Hàm mờ hình sin	47
Hình 3.8. Hàm mờ Gaussian	47
Hình 3.9. Tập mờ B bao hàm tập mờ A.....	48
Hình 3.10. Minh họa các phép toán tập hợp mờ	49
Hình 3.11. Hệ mờ áp dụng trong GIS.....	51
Hình 3.12. Phân tích với tập mờ (trái) và tập rõ (phải)	57
Hình 3.13. Mô hình mở rộng đối với các bảng dữ liệu	61
Hình 3.14. Các ví dụ về vùng đệm (điểm, đường, vùng)	63
Hình 3.15. Phép toán khoảng cách mờ giữa 2 vị trí(a);vị trí với vùng mờ(b) .	66
Hình 3.16. Mô tả chồng xếp các lớp	68
Hình 3.17. Mô tả chồng xếp mờ có trọng số.....	69
Hình 3.18. Phép toán lựa chọn mờ	70
Hình 4.1. Hàm mờ sử dụng lớp thông tin mở mang.....	80
Hình 4.2. Hàm mờ sử dụng cho lớp thông tin giao thông	80
Hình 4.3. Hàm mờ sử dụng cho lớp thông tin ô nhiễm	81
Hình 4.4. Phương trình chồng xếp mờ tính toán trên các trường	82
Hình 4.5. Thuộc tính sau khi chồng xếp	83
Hình 4.6. Vùng đệm mờ hóa lớp thông tin mở mang thành phố.....	83
Hình 4.7. Vùng đệm mờ hóa về lớp thông tin giao thông.....	84

Hình 4.8. Vùng đệm mờ hóa lớp thông tin ô nhiễm	84
Hình 4.9. Mờ hóa lớp thông tin địa tầng đất yếu	85
Hình 4.10. Mờ hóa lớp thông tin hiện trạng sử dụng đất	85
Hình 4.11. Kết quả sau khi chồng xếp	86
Hình 4.12. Giải mờ lát cắt $\alpha = 0.75$	86
Hình 4.13. Giải mờ lát cắt $\alpha = 0.7$	87
Hình 4.14. Giải mờ lát cắt $\alpha = 0.65$	87
Hình 4.15. Đồ thị G có hướng V- mờ.....	89
Hình 4.16. Đường đi ngắn nhất mờ của đồ thị mờ G	90
Hình 4.17. Hàm mờ sử dụng lớp thông tin gần mỏ than.....	91
Hình 4.18. Hàm mờ sử dụng lớp thông tin gần mỏ đất sét	92
Hình 4.19. Hàm mờ sử dụng cho lớp thông tin giao thông.....	92
Hình 4.20. Hàm mờ sử dụng lớp thông tin gần mỏ đá vôi.....	93
Hình 4.21. Hàm mờ sử dụng lớp thông tin gần cảng	93
Hình 4.22. Hàm mờ sử dụng cho lớp thông tin ô nhiễm	94
Hình 4.23. Giải mờ với lát cắt $\alpha = 0.33$	95
Hình 4.24. Giải mờ lấy lát cắt $\alpha = 0.36$	95

MỞ ĐẦU

Hệ thống thông tin địa lý (Geographic Information System - GIS) ra đời trên cơ sở phát triển của khoa học máy tính và được ứng dụng rộng rãi trong nhiều ngành khoa học có liên quan đến xử lý dữ liệu không gian. GIS được hình thành từ những năm 70 của thế kỷ trước và phát triển mạnh mẽ trong một hai chục năm trở lại đây. GIS đã trở thành công cụ hỗ trợ ra quyết định trong hầu hết các hoạt động kinh tế - xã hội, an ninh - quốc phòng, trong quản lý, quy hoạch, thăm dò, khai thác...

Đối với GIS, các dữ liệu thu thập thường không đầy đủ, không rõ ràng, không chắc chắn và mập mờ, điều đó dẫn đến dữ liệu và thông tin trong GIS là dữ liệu “*không rõ ràng*” hay dữ liệu “*mờ*”.

Phân tích dữ liệu không gian bằng cách kết hợp nhiều nguồn dữ liệu được khai thác từ các hệ thống thông tin địa lý là mục tiêu cao nhất của hầu hết các dự án GIS để diễn tả, phân tích các ảnh hưởng lẫn nhau, đưa ra các mô hình dự báo và hỗ trợ ra quyết định. Khái niệm “*không rõ ràng - mờ*” là một đặc trưng vốn có của dữ liệu địa lý và có thể sinh ra do: Thông tin tương ứng với chúng không đầy đủ; sự xuất hiện không ổn định khi thu thập, tập hợp các dữ liệu thuộc tính; việc sử dụng các diễn tả định tính đối với các giá trị thuộc tính và các mối quan hệ giữa chúng. Các hệ GIS thường không sẵn sàng cho việc xử lý với các dữ liệu mờ vì thế cần phải có sự mở rộng cả về mô hình dữ liệu, các phép toán và lập luận để giải quyết với dữ liệu mờ trong GIS làm cho hệ thống trở lên mềm dẻo hơn trong việc giải các bài toán không gian mà dữ liệu của chúng là các dữ liệu dạng mờ.

Theo phương pháp truyền thống khi xử lý, phân tích dữ liệu trong GIS các thao tác dữ liệu thực hiện một cách cứng nhắc đối với các thủ tục lập luận và phân tích. Quyết định tổng thể được thực hiện theo từng bước cụ thể và quy về kết quả ngay lập tức. Những ứng viên nào thoả điều kiện được giữ lại và các ứng viên không thoả điều kiện sẽ bị loại bỏ phụ thuộc vào giá trị ngưỡng.

Thêm vào đó các quyết định đưa ra là bắt buộc để biểu diễn các ràng buộc của chúng dưới dạng các điều kiện số học và các ký hiệu toán học trong các quan hệ rõ, chúng không cho phép sử dụng các điều kiện cú pháp dưới dạng ngôn ngữ tự nhiên. Mặt khác kết quả lựa chọn dựa trên các điều kiện được xác định là ngang nhau, không có giá trị trọng số của các đối tượng.

Một trong các phương pháp toán học nghiên cứu tính chất “*không rõ ràng*” của không gian là lý thuyết tập mờ Zadeh (1965-1988). Nó sử dụng độ thuộc để diễn tả một cá thể tham gia trong một tập hợp. Sự kết hợp lý thuyết tập mờ và GIS là các đối tượng không gian “*mờ*” đều có một đặc trưng chung là chúng có ranh giới “*không rõ ràng*” so với các đối tượng không gian “*rõ*”.

Lý thuyết tập mờ là giải pháp thích hợp nhất cho việc mô hình hóa dữ liệu “*không rõ ràng*” và đưa ra cơ sở lý thuyết để hỗ trợ các lập luận trên dữ liệu này. Trong luận văn này tác giả đề cập tới hướng ứng dụng của logic mờ trong GIS nhằm mục đích mở rộng và tăng cường các chức năng của hệ thống GIS. Làm cho hệ thống GIS trở lên mềm dẻo hơn và ứng dụng thuận lợi trong việc giải quyết các bài toán về không gian mà dữ liệu của nó là “*không rõ ràng*” hay còn gọi là dữ liệu “*không gian mờ*”.

Tác giả đã có thời gian làm việc về hệ thống thông tin địa lý - GIS trên 15 năm. Đã tham gia xây dựng, phân tích và xử lý nhiều hệ GIS. Nghiên cứu và phát triển các chức năng trên các hệ thống như: GeoConcept, Mapinfo, ArcInfo, GeoMedia...

❖ Mục tiêu của luận văn này chia làm các phần chính sau:

- Tổng quan, giới thiệu vai trò của GIS trong các hoạt động kinh tế xã hội, khuynh hướng phát triển và hướng nghiên cứu của đề tài.
- Giới thiệu vắn tắt về hệ thống GIS lịch sử phát triển, các chức năng và các khái niệm được sử dụng trong đề tài.
- Ứng dụng logic mờ trong GIS - phân tích tính mập mờ của dữ liệu trong GIS và các giới hạn của các hệ thống GIS thương mại hiện nay

và việc cần thiết phải mở rộng nó cả về mô hình dữ liệu và các phép toán trong GIS để phù hợp với tính mờ của dữ liệu.

- Ứng dụng trong giải quyết các bài toán không gian
- Kết luận đánh giá kết quả nghiên cứu trong đề tài.

❖ Đối tượng nghiên cứu của đề tài là các hệ thống GIS và ứng dụng trong các bài toán quy hoạch không gian.

❖ Phạm vi nghiên cứu của đề tài được giới hạn đối các đối tượng trong không gian mà ta coi như vị trí không gian của các đối tượng này là những đối tượng rõ. Tức là vị trí của nó ta coi như là chính xác. Còn sự không rõ ràng đối với vị trí của nó (Điểm mờ, đường mờ, vùng mờ) là hướng nghiên cứu trong tương lai.

❖ Ý nghĩa khoa học của đề tài: Lý thuyết tập mờ Zadeh (1965-1988) ra đời trên 40 năm và đã có cơ sở khoa học vững chắc. Nhưng việc ứng dụng logic mờ trong các hệ thống thông tin địa lý mới chỉ được một số nhà khoa học nghiên cứu mà chủ yếu là trên một số phép toán để giải quyết các vấn đề phát sinh từ một dự án nào đó. Theo chủ quan của tác giả việc nghiên cứu logic mờ trong GIS là hướng nghiên cứu hoàn toàn mới tại Việt Nam.

❖ Ý nghĩa thực tiễn của đề tài: Giữa lý thuyết và thực tiễn thông thường phải có thời gian nhất định để áp dụng những kết quả nghiên cứu lý thuyết vào thực tiễn. Đối với tiến trình phát triển xã hội hiện nay cần thiết phải rút ngắn tối đa thời gian giữa lý thuyết và thực tiễn. Sản phẩm của đề tài đã được áp dụng cụ thể trong bài toán mở rộng Thành phố Thái Bình. Hiện nay vị trí mở rộng Thành phố Thái Bình do đề tài chỉ ra đã trở thành khu đô thị mới của Thành phố theo hướng hiện đại. Sản phẩm của đề tài này sẽ được ứng dụng rất nhiều trong các bài toán quy hoạch không gian như mở rộng thành phố, quy hoạch dân cư, quy hoạch các khu vực kinh tế trọng điểm...

CHƯƠNG 1 - TỔNG QUAN

Trong tiến trình phát triển kinh tế xã hội, để đạt được một mục đích nào đó, con người cần phải có những quyết định chính xác và kịp thời. Những quyết định đó thường được thực hiện sau khi thu thập thông tin, dữ liệu của thế giới thực và phân tích xử lý nó theo một quan điểm nào đó.

Những quyết định này tác động trực tiếp hoặc gián tiếp trở lại thế giới thực theo khuynh hướng của người xử lý và ra quyết định. Nếu quyết định ấy tác động đến thế giới thực tạo ra nhiều kết quả có lợi cho con người thì quyết định ấy được đánh giá là tốt. Ngược lại, nếu quyết định tác động lên thế giới thực sinh ra nhiều hậu quả có hại cho con người hơn thì quyết định ấy được đánh giá là xấu.

Theo quan điểm thông tin, tiến trình nói trên thể hiện một sự tuần hoàn của dữ liệu: dữ liệu từ thế giới thực được thu thập, lưu trữ, phân tích, xử lý và ra quyết định. Trên luồng dữ liệu ấy, kết quả của bước sau phụ thuộc vào kết quả của bước trước: quyết định phụ thuộc vào kết quả phân tích và quan điểm của người ra quyết định, kết quả phân tích phụ thuộc vào chất lượng dữ liệu và khả năng của người phân tích. Chất lượng dữ liệu được đề cập ở đây bao gồm: độ chính xác, tính thời gian của dữ liệu. Chất lượng dữ liệu phụ thuộc vào thiết bị, công nghệ, khả năng và tinh thần trách nhiệm của người thu thập dữ liệu, phụ thuộc vào công nghệ, khả năng của thiết bị lưu trữ, bảo quản dữ liệu.

Cho đến nay, phương tiện truyền thống để hiển thị và lưu trữ dữ liệu địa lý là bản đồ. Trên bản đồ, các thực thể trong thế giới thực được biểu diễn bằng đường nét, hình vẽ, ký hiệu, v.v., vị trí địa lý của các đối tượng được xác định trong một hệ thống tọa độ Đề-Các hai chiều. Với bản đồ giấy truyền thống, các phép phân tích đơn giản như đo chiều dài, tính diện tích có thể được thực hiện bằng những dụng cụ đơn giản như thước đo cạnh, góc, đếm ô vuông... Các bài toán phân tích vùng cũng có thể thực hiện bằng cách chồng xếp, cắt dán các bản đồ chuyên đề được vẽ lên giấy trong suốt, giấy can hoặc giấy mờ.

Ngày nay, với sự phát triển của công nghệ thông tin, đặc biệt là từ khi xuất hiện ngành đồ họa vi tính cũng như sự gia tăng vượt bậc những khả năng phân cứng, hệ thống thông tin địa lý (GIS) đã phát triển nhanh chóng cả về mặt công nghệ cũng như ứng dụng. Hệ thống thông tin địa lý đã chứng tỏ khả năng ưu việt hơn hẳn các hệ thống tin bản đồ truyền thống nhờ vào khả năng tích hợp cao, cập nhật dễ dàng cũng như khả năng phân tích, tính toán của nó. Do đó, hệ thống thông tin địa lý đã nhanh chóng trở thành công cụ hỗ trợ ra quyết định cho tất cả các ngành từ qui hoạch đến quản lý, tất cả các lĩnh vực từ tài nguyên thiên nhiên, môi trường, đất đai, hạ tầng kỹ thuật đến xã hội nhân văn. Sự phát triển nhanh của công nghệ thông tin cùng với những kết quả của các thuật toán tối ưu, nhận dạng, xử lý ảnh, logic tính toán, trí tuệ nhân tạo và cơ sở dữ liệu quan hệ đã tạo điều kiện cho công nghệ thông tin địa lý ngày càng phát triển.

Hiện nay, trên thế giới đã hình thành nhiều cơ quan nghiên cứu GIS với qui mô lớn, nhiều hướng tiếp cận và mục tiêu khác nhau:

Khuynh hướng phát triển về lý thuyết

Về lý thuyết, hiện nay nhiều nhà khoa học đang theo đuổi nghiên cứu phương pháp biểu diễn dữ liệu không gian trong các hệ thống thông tin địa lý, sự liên quan các loại dữ liệu bao gồm dữ liệu không gian, dữ liệu thuộc tính, dữ liệu thời gian. Mối quan hệ giữa những bài toán phân tích không gian theo thời gian thực. Phân tích thống kê dữ liệu không gian. Thiết kế mô hình dữ liệu và cấu trúc dữ liệu thích hợp. Nghiên cứu phương pháp và kỹ thuật thiết kế cơ sở dữ liệu không gian. Nghiên cứu đầy đủ hơn về công nghệ bản đồ, truyền thông bản đồ.

Khuynh hướng phát triển phần cứng

Trong lĩnh vực GIS, những thành tựu sau đây của máy tính đã có tác động lớn đối với sự phát triển khoa học và công nghệ thông tin địa lý.

- (1) Tốc độ xử lý của máy tính và khả năng hiển thị độ phân giải cao.

- (2) Xử lý song song.
- (3) Xử lý phân tán trên mạng.
- (4) Khả năng lưu trữ dung lượng lớn, độ tin cậy cao .
- (5) Các thiết bị phần cứng đặc biệt như: Server, thiết bị mạng, gia tốc đồ họa, đồng xử lý và đặc biệt là những thiết bị hiển thị như datashow.
- (6) Thiết bị ngoại vi như: máy in màu độ phân giải cao, scanner màu hoặc đen trắng khổ lớn, các thiết bị multimedia, v.v. . .
- (7) Các thiết bị nhúng (GPS, thiết bị giám sát mục tiêu di động, PDA...Cho phép cài đặt các bản đồ trên các thiết bị này).

Khuynh hướng phát triển phần mềm

- (1) Hệ quản trị cơ sở dữ liệu.
- (2) GIS với kỹ thuật đa phương tiện.
- (3) GIS thông minh: Những công cụ thông minh sẽ được phát triển trong các hệ thống thông tin địa lý làm nhiệm vụ hỗ trợ ra quyết định. Các kỹ thuật thông minh nhân tạo bao gồm: logic mờ, hệ chuyên gia, mạng nơ-rôn nhân tạo, nhận dạng sẽ là thành phần quan trọng để phát triển kỹ thuật GIS. Xử lý ngôn ngữ tự nhiên cũng có thể được phát triển trong GIS để tăng khả năng giao tiếp với người sử dụng.
- (4) Mô phỏng và hỗ trợ ra quyết định: Xây dựng những hệ thống thông tin địa lý thời gian thực là một trong những mục tiêu quan trọng mà tất cả những nhà khoa học, kỹ thuật của nhiều ngành có liên quan như: kỹ thuật phần cứng, toán và kỹ thuật phần mềm, đang theo đuổi và hỗ trợ cho nhau. Mô hình hóa theo không gian và dự báo trong nhiều lĩnh vực khác nhau là những bài toán được quan tâm trong các hệ thống thông tin địa lý trong thời gian tới.
- (5) Khung công việc: Nhu cầu dịch vụ thông tin địa lý gia tăng, hiệu quả của hệ thống thông tin địa lý càng cao khi khả năng trao đổi và tích hợp

dữ liệu càng lớn. Trong thời gian tới, sẽ hình thành nhiều khung công việc theo từng địa phương, từng quốc gia, từng khu vực và toàn thế giới.

Khuynh hướng phát triển ứng dụng

Với nhu cầu sử dụng thông tin địa lý ngày càng tăng trong hầu hết các lĩnh vực xã hội, dịch vụ thông tin địa lý đã ra đời để đáp ứng nhu cầu của các cơ quan quản lý hành chính nhà nước và của tất cả mọi người, mọi tổ chức. Có thể chia làm hai nhóm dịch vụ thông tin địa lý là dịch vụ nhà nước và dịch vụ công cộng(WebGIS). Dịch vụ thông tin địa lý nhà nước nhằm cung cấp những thông tin tích hợp theo không gian phục vụ cho tiến trình ra quyết định trong quản lý hành chính nhà nước trên một địa bàn lãnh thổ nhất định. Dịch vụ thông tin công cộng nhằm cung cấp những thông tin phục vụ yêu cầu dân biết và phục vụ nhu cầu thiết kế, qui hoạch, kế hoạch của các thành phần kinh tế, của các nhà đầu tư, nhà kinh doanh, nhà thương mại trong nhiều lĩnh vực khác nhau.

Hướng nghiên cứu logic mờ trong GIS là khuynh hướng phát triển phần mềm mà các chuyên gia đưa ra, cùng với các hệ chuyên gia mờ, trí tuệ nhân tạo, nhận dạng; nhằm thiết lập các công cụ hiệu quả trong các ứng dụng của GIS đối với các hoạt động kinh tế xã hội, an ninh quốc phòng.

CHƯƠNG 2 - HỆ THỐNG THÔNG TIN ĐỊA LÝ (GIS)

2.1 Khái niệm

Hệ thống thông tin địa lý (Geographic Information Systems - GIS) là các hệ thống dựa trên máy tính được thiết kế để hỗ trợ việc thu thập, quản lý, vận dụng, phân tích, mô hình hóa và hiển thị dữ liệu có tham chiếu không gian tại các thời điểm khác nhau. Ngày nay GIS được sử dụng rộng rãi trong các cơ quan chính phủ và các hoạt động riêng biệt. Ứng dụng của GIS được chia thành ba lĩnh vực chính:

Các ứng dụng về kinh tế xã hội: quy hoạch đô thị và vùng, đăng ký địa chính, khảo cổ học, tài nguyên thiên nhiên.

Các ứng dụng môi trường: lâm nghiệp, kiểm soát cháy và dịch bệnh.

Các ứng dụng trong quản lý: tổ chức các mạng đường ống và các dịch vụ khác như điện, điện thoại, định hướng thời gian thực cho tàu thuyền, máy bay, ô tô....

Trong các ứng dụng này GIS cung cấp các công cụ “ra quyết định” hiệu quả để giải quyết các vấn đề không gian phức tạp và thiếu thông tin hoặc bán cấu trúc.

GIS lưu trữ thông tin theo các tiêu chí khác nhau cho quy hoạch đô thị như: các bản đồ địa chính, độ cao, bản đồ quy hoạch, bản đồ sử dụng đất, thông tin phát triển kinh tế, thông tin dân cư và các phân tích thống kê khác nhau. Các bản đồ trong GIS là các lớp, mỗi lớp bao gồm các thông tin liên quan tới vùng chứa đựng trong bản đồ như các đối tượng không gian (đường giao thông, sông suối, hồ...), phân bố dân cư, phân chia đất... Đối với các đối tượng không gian, GIS chứa đựng đặc trưng đặc biệt cho phép chứng minh hóa trong dạng Vector hoặc với các dạng rời rạc (lưới và raster). GIS cũng có khả năng biểu diễn các loại đối tượng khác nhau (nhà, hồ, các bản vẽ đất) bằng cách sử dụng các dạng hình học đơn giản để diễn tả chúng như: điểm, đường, vùng. GIS đưa ra kỹ thuật để tích hợp, quản lý, phân tích dữ liệu và sản sinh

các báo cáo sức tích trên môi trường không gian. Công cụ này trợ giúp các nhà quy hoạch ra quyết định, phân tích và quy hoạch đô thị.

Một trong số các lợi ích chính của GIS là các tiến bộ cho phép quản lý, tổ chức và chia sẻ tài nguyên giữa các lĩnh vực khác nhau. Một cơ sở dữ liệu chia sẻ cho phép dữ liệu có thể thu thập một lần và sử dụng nhiều lần.

2.1.1 Một số định nghĩa

- ✓ Hệ thống thông tin địa lý là một công cụ máy tính để lập bản đồ và phân tích các sự vật, hiện tượng trên trái đất. Công nghệ GIS kết hợp các thao tác cơ sở dữ liệu thông thường và các phép phân tích thống kê, phân tích địa lý trong đó phép phân tích địa lý và hình ảnh được cung cấp duy nhất từ bản đồ. Những khả năng này phân biệt GIS với các hệ thống thông tin khác và làm cho GIS có phạm vi ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau (phân tích sự kiện, dự đoán tác động môi trường, hoạch định chiến lược...).
- ✓ Hệ thống thông tin địa lý là hệ thống tự động hóa quản lý các dữ liệu theo không gian và thời gian mà tích hợp của nó là thông tin địa lý.
- ✓ Hệ thống thông tin địa lý (Geographic Information System - GIS).

Geographic - Có nghĩa là vị trí của các khoản mục dữ liệu được xác định hoặc có thể được tính toán theo tọa độ địa lý (kinh độ, vĩ độ, cao độ).

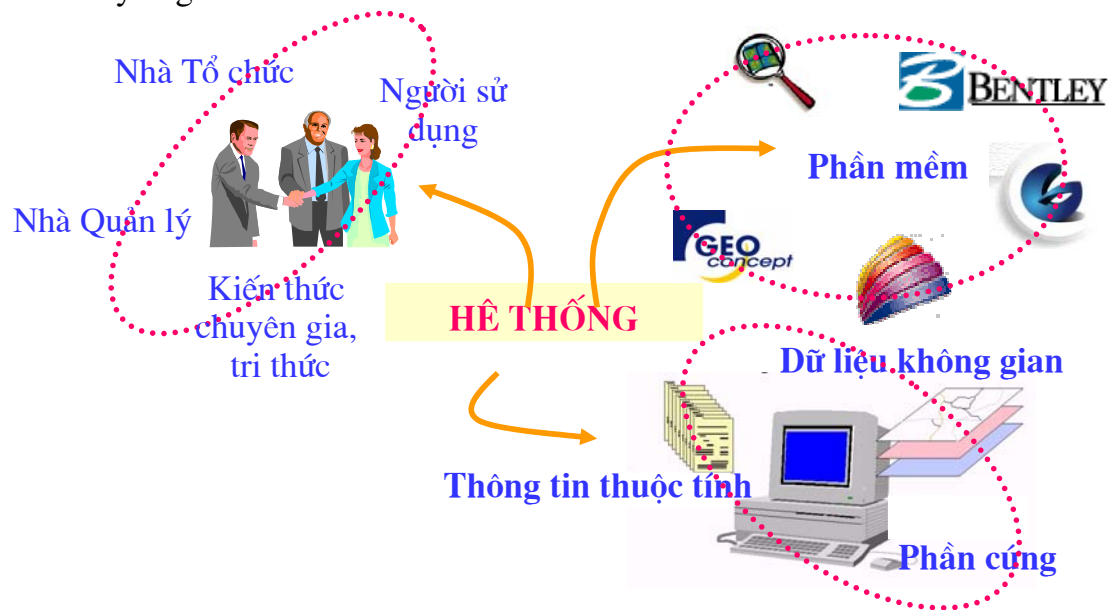
Information - Dữ liệu trong GIS được tổ chức để mang lại các tin tức có hiệu quả thông qua các phép xử lý và truy vấn khác nhau.

System - GIS được xây dựng bởi các chức năng khác nhau được liên kết lại "một cách hệ thống".

- ✓ GIS là một trường hợp đặc biệt của hệ thống thông tin ở đó cơ sở dữ liệu bao gồm các quan sát trên đặc trưng phân tán không gian, các hoạt động hoặc sự kiện mà có thể định rõ trong không gian như điểm, đường hoặc vùng. Một hệ thống thông tin địa lý vận dụng dữ liệu về các điểm,

đường, vùng này để nhận dữ liệu bằng cách hỏi đáp và phân tích đặc biệt.

Có nhiều định nghĩa về GIS, tùy theo cách tiếp cận. Xét từ góc độ hệ thống, GIS gồm các hợp phần: phần cứng, phần mềm, cơ sở dữ liệu và cơ sở tri thức chuyên gia.



Hình 2.1. Mô hình một hệ thống thông tin địa lý

Phần cứng là tất cả những gì mà một hệ thống có thể vận hành được. Phần cứng bao hàm các máy tính (Server/workstation), thiết bị lưu trữ, máy in, máy quét, máy vẽ, các thiết bị truyền thông...

Phần mềm bao gồm hai loại: phần mềm hệ thống và phần mềm ứng dụng, hiện nay trên thị trường tồn tại nhiều hệ mềm khác nhau trong lĩnh vực GIS như: ArcInfo, Mapinfo, GeoMedia, GeoConcept, CardCorp,... Mỗi phần mềm này đều được trang bị các công cụ hữu hiệu để thực hiện các chức năng của GIS.

Cơ sở tri thức chuyên gia là tập hợp các tri thức của lãnh đạo, nhà quản lý, các kiến thức chuyên ngành và kiến thức công nghệ thông tin. Tập hợp các tri thức chuyên gia này sẽ quyết định mô hình ứng dụng của GIS, xác định được các chức năng hỗ trợ quyết định của GIS, xác định được nội dung,

cấu trúc các hợp phần còn lại của hệ thống, các bước và phương thức cũng như mức đầu tư xây dựng và vận hành hệ thống.

Cơ sở dữ liệu là nơi tổ chức và lưu trữ dữ liệu (cả dữ liệu không gian và dữ liệu thuộc tính) nhằm cung cấp một cách hiệu quả các thông tin từ nó cho các truy vấn từ phía người sử dụng. Việc tổ chức và xây dựng cơ sở dữ liệu đòi hỏi sự thống nhất cao từ khâu thiết kế đến các ứng dụng thực tế và tuân thủ các chuẩn trong việc tổ chức và xây dựng cơ sở dữ liệu.

Mục đích chung của GIS thực hiện các nhiệm vụ sau:

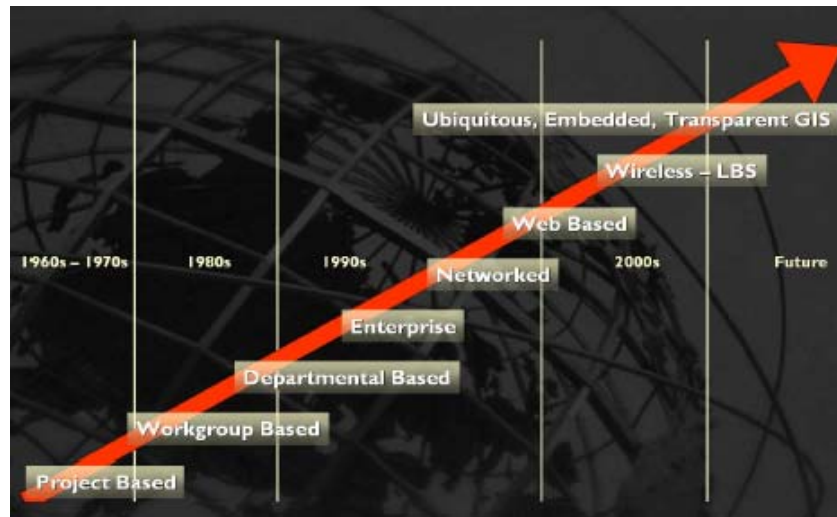
- Thu thập dữ liệu.
- Thao tác dữ liệu.
- Quản lý dữ liệu.
- Hỏi đáp, phân tích dữ liệu.
- Hiển thị, báo cáo.
- Công bố dữ liệu.

2.1.2 Lịch sử phát triển của GIS

GIS được hình thành từ những năm 70 của thế kỷ trước và phát triển rất nhanh cùng với sự phát triển của công nghệ thông tin. Đầu tiên GIS chỉ hình thành trên các dự án quy mô nhỏ có liên quan đến bản đồ và công nghệ ban đầu chủ yếu sử dụng để hiển thị bản đồ và các thông tin liên quan sau đó một số các ứng dụng khác nhau của GIS được khai phá như phân tích hỏi đáp tạo các dạng bản đồ chuyên đề...

Từ những hệ GIS đơn lẻ, GIS phát triển lên ở cấp cao hơn như: GIS làm việc theo nhóm, GIS làm việc trong các xí nghiệp, GIS trên mạng và đến những đầu của thế kỷ 21 công nghệ WebGIS phát triển mang lại sự ứng dụng có hiệu quả cao không chỉ cho các nhà chuyên môn về bản đồ mà đối với hầu hết người sử dụng cần tham khảo các thông tin liên quan tới lãnh thổ như: người bán hàng có thể suy xét tới mật độ dân cư để xây dựng cho mình kế hoạch bán hàng. Người kinh doanh bất động sản có thể thực hiện giao dịch

trên mạng mà không nhất thiết phải tới tận nơi. Các nhà chiến lược, quy hoạch quan tâm tới sự tập trung của dân số, giá trị đất đai... Trong tương lai không xa GIS phát triển để có thể cài đặt và tải trên các thiết bị cầm tay, thiết bị nhúng.



Hình 2.2. Lịch sử phát triển của GIS.

Các nhà nghiên cứu về GIS đưa ra GIS thế hệ thứ 5 bao gồm:

- *Tại mọi lúc, mọi nơi người sử dụng có thể truy cập các thông tin vị trí, các ứng dụng và hỗ trợ quyết định.*
- *Dữ liệu không gian cho phép thương mại và các tri thức phân tích.*
- *GIS như là trung tâm và các hệ thống thông tin thương mại như quản lý bất động sản, tài chính, con người, tài nguyên và quan hệ khách hàng.*
- *Các dịch vụ ứng dụng và dữ liệu không gian, khắp mọi nơi, trên các thiết bị nhúng và trong suốt.*

2.2 Thu thập dữ liệu

Trước khi dữ liệu địa lý có thể được dùng cho GIS, dữ liệu này phải được chuyển sang dạng số thích hợp. Quá trình chuyển dữ liệu từ bản đồ giấy, các bảng thống kê mô tả, các phiếu điều tra, các tài liệu về ảnh,..., sang các file dữ liệu dạng số được gọi là quá trình số hoá.

Công nghệ GIS hiện đại có thể thực hiện tự động hoặc bán tự động quá trình này với công nghệ quét ảnh cho các đối tượng lớn; những đối tượng nhỏ

hơn đòi hỏi một số quá trình số hoá thủ công (dùng bàn số hoá hoặc số hóa trên nền ảnh).

2.2.1 Thu thập dữ liệu không gian

Dữ liệu không gian có được từ nhiều nguồn khác nhau. Dữ liệu không gian có thể thu thập từ các dạng bản đồ giấy đã được sản xuất. Từ việc đo đạc ngoài thực địa, từ ảnh máy bay, ảnh vệ tinh, từ các sản phẩm được sản xuất trên các trạm đo vẽ ảnh số. Dữ liệu không gian cũng có thể được kết xuất, kế thừa từ các hệ thống phần mềm GIS khác, từ các khuôn dạng dữ liệu khác.

2.2.2 Thu thập dữ liệu thuộc tính

Dữ liệu thuộc tính được thu thập từ rất nhiều nguồn dữ liệu. Các dữ liệu điều tra cơ bản, các bảng biểu, số liệu thống kê, các sản phẩm được tính toán xử lý từ chính dữ liệu không gian, sản phẩm giải đoán từ ảnh vệ tinh, ảnh hàng không, điều tra ngoài thực địa và nhiều nguồn dữ liệu được kết xuất từ các hệ thống thông tin khác.

2.3 Thao tác dữ liệu

Có những trường hợp các dạng dữ liệu đòi hỏi được chuyển dạng và thao tác theo một số cách để có thể tương thích với một hệ thống nhất định. Ví dụ, các thông tin địa lý có giá trị biểu diễn khác nhau tại các tỷ lệ khác nhau (hệ thống đường phố được chi tiết hoá trong file về giao thông, kém chi tiết hơn trong file điều tra dân số và có mã bưu điện trong mức vùng). Trước khi các thông tin này được kết hợp với nhau, chúng phải được chuyển về cùng một tỷ lệ (mức chính xác hoặc mức chi tiết). Đây có thể chỉ là sự chuyển dạng tạm thời cho mục đích hiển thị hoặc cố định cho yêu cầu phân tích. Công nghệ GIS cung cấp nhiều công cụ cho các thao tác trên dữ liệu không gian và cho loại bỏ dữ liệu không cần thiết.

2.4 Quản lý dữ liệu

Đối với những dự án GIS nhỏ, có thể lưu trữ các thông tin địa lý dưới dạng các file đơn giản. Tuy nhiên, khi kích cỡ dữ liệu trở nên lớn và số lượng

người dùng cũng nhiều lên, thì cách tốt nhất là sử dụng hệ quản trị cơ sở dữ liệu (DBMS) để giúp cho việc lưu trữ, tổ chức và quản lý thông tin. Một DBMS chỉ đơn giản là một phần mềm quản lý cơ sở dữ liệu.

Có nhiều cấu trúc DBMS khác nhau, nhưng trong GIS cấu trúc quan hệ tỏ ra hữu hiệu nhất. Trong cấu trúc quan hệ, dữ liệu được lưu trữ ở dạng các bảng. Các trường thuộc tính chung trong các bảng khác nhau được dùng để liên kết các bảng này với nhau. Do linh hoạt nên cấu trúc đơn giản này được sử dụng và triển khai khá rộng rãi trong các ứng dụng cả trong và ngoài GIS.

2.5 Truy vấn và phân tích dữ liệu

Một khi đã có một hệ GIS lưu trữ các thông tin địa lý, có thể bắt đầu hỏi các câu hỏi đơn giản như:

- Ai là chủ mảnh đất ở góc phố?
- Hai vị trí cách nhau bao xa?
- Vùng đất dành cho hoạt động công nghiệp ở đâu?

Và các câu hỏi phân tích như:

- Tất cả các vị trí thích hợp cho xây dựng các toà nhà mới nằm ở đâu?
- Kiểu đất ưu thế cho rừng sồi là gì?
- Nếu xây dựng một đường quốc lộ mới ở đây, giao thông sẽ chịu ảnh hưởng như thế nào?

GIS cung cấp cả khả năng hỏi đáp đơn giản qua các giao diện và các công cụ phân tích tinh vi để cung cấp kịp thời thông tin cho những người quản lý và phân tích. Các hệ GIS hiện đại có nhiều công cụ phân tích hiệu quả, trong đó có hai công cụ quan trọng đặc biệt:

Phân tích liên kết

- Tổng số khách hàng trong bán kính 10 km khu hàng?
- Những lô đất trong khoảng 60 m từ mặt đường?

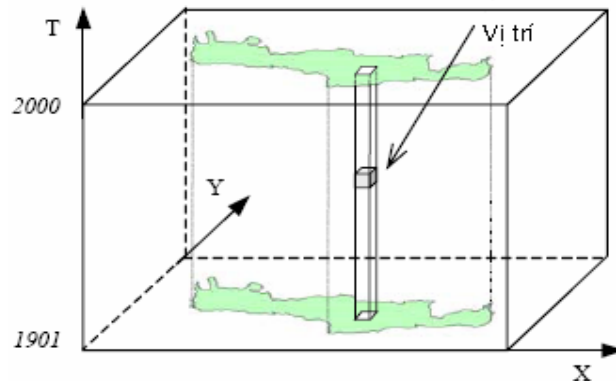
Để trả lời những câu hỏi này, GIS sử dụng phương pháp vùng đệm để xác định mối quan hệ liên kết giữa các đối tượng.



Hình 2.3. Mô tả phân tích liên kề.

Phân tích chồng xếp

Chồng xếp là quá trình tích hợp các lớp thông tin khác nhau. Các thao tác phân tích đòi hỏi một hoặc nhiều lớp dữ liệu phải được liên kết vật lý. Sự chồng xếp này, hay liên kết không gian, có thể là sự kết hợp dữ liệu về đất, độ dốc, thảm thực vật hoặc sở hữu đất với định giá thuế...và tạo ra sản phẩm là bản đồ mới có các đặc trưng từ những bản đồ đưa vào.



Hình 2.4. Mô tả phân tích chồng xếp theo thời gian

2.6 Hiện thị dữ liệu

Với nhiều thao tác trên dữ liệu địa lý, kết quả cuối cùng được hiển thị tốt nhất dưới dạng bản đồ hoặc biểu đồ. Bản đồ khá hiệu quả trong lưu trữ và trao đổi thông tin địa lý. GIS cung cấp nhiều công cụ mới và thú vị để mở rộng tính nghệ thuật và khoa học của ngành bản đồ. Bản đồ hiển thị có thể được kết hợp với các bản báo cáo, hình ảnh ba chiều, ảnh chụp và những dữ liệu khác (đa phương tiện).

GIS lưu giữ thông tin về thế giới thực dưới dạng tập hợp các lớp chuyên đề có thể liên kết với nhau nhờ các đặc trưng địa lý. Điều này đơn giản nhưng vô cùng quan trọng và là một công cụ đa năng đã được chứng minh là rất có giá trị trong việc giải quyết nhiều vấn đề thực tế, từ thiết lập tuyến đường phân phối của các chuyến xe, đến lập báo cáo chi tiết cho các ứng dụng quy hoạch, hay mô phỏng sự lưu thông khí quyển toàn cầu.

2.7 Mô hình dữ liệu

Hệ thống thông tin địa lý làm việc với hai dạng mô hình dữ liệu địa lý khác nhau về cơ bản - mô hình vector và mô hình raster. Trong mô hình Vector, thông tin về điểm, đường và vùng được mã hoá và lưu dưới dạng tập hợp các toạ độ (x,y).

Mô hình raster được phát triển cho mô phỏng các đối tượng liên tục. Một ảnh Raster là một tập hợp các ô lưới. Cả mô hình vector và raster đều được dùng để lưu dữ liệu địa lý với những ưu điểm, nhược điểm riêng. Các hệ GIS hiện đại có khả năng quản lý cả hai mô hình này. Bảng dưới đây so sánh giữa hai mô hình dữ liệu Vector và Raster:

Mô hình Vector	Mô hình Raster
<i>Ưu điểm</i> - Độ chính xác cao - Cấu trúc dữ liệu dạng nén mất ít dung lượng để lưu trữ - Cho phép các quan hệ hình học (topological) như tính liên kê, liên thông. - Gần gũi với thao tác vẽ bằng tay của con người.	<i>Ưu điểm</i> - Cấu trúc dữ liệu đơn giản - Hiệu quả trong tính toán - Các phép toán chồng xếp xử lý dễ dàng - Thích hợp cho việc thể hiện dữ liệu phức tạp, đa dạng - Thích hợp cho việc nâng cấp, xử lý ảnh
<i>Nhược điểm</i> - Cấu trúc dữ liệu phức tạp.	<i>Nhược điểm</i> - Quan hệ hình học khó nhận thức.

- Các phép toán chồng xếp xử lý khó khăn hơn. - Miêu tả mức cao biến đổi không gian khó khăn. - Không thích hợp cho việc thể hiện dữ liệu phức tạp, đa dạng. - Không thích hợp cho việc nâng cấp, xử lý ảnh.	- Khả năng nén thấp đòi hỏi dung lượng lưu trữ lớn. - Việc đưa ra tính thẩm mỹ không cao. - Miêu tả mức cao biến đổi không gian dễ dàng. - Thể hiện bản đồ không rõ nét nếu độ phân giải thấp. Nếu tăng độ phân giải sẽ dẫn đến kích thước file dữ liệu lớn.
---	---

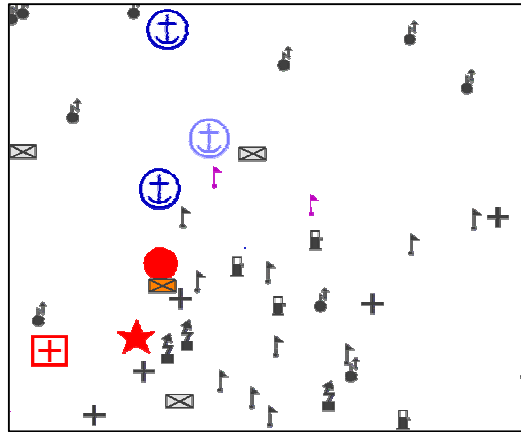
Bảng 2.1. So sánh mô hình dữ liệu Vector và Raster.

Trong lịch sử phát triển của GIS để xử lý đối với các dạng dữ liệu Vector cho đơn giản hóa mà khoảng thời gian từ 2000-2004 cấu trúc Topology đã bị lãng quên (bỏ qua). Tuy nhiên hiện nay các nhà nghiên cứu về GIS thấy được tầm quan trọng về cấu trúc Topology một trong các tính chất quan trọng nhất của GIS để giải quyết các vấn đề phức tạp trong không gian. Vì vậy cấu trúc Topology là một phần không thể thiếu trong GIS hiện nay.

2.8 Các đối tượng trong GIS

Khác với các hệ cơ sở dữ liệu khác, cơ sở dữ liệu GIS có một đặc thù riêng đó là có phần tham gia của dữ liệu không gian. Mỗi đối tượng trong cơ sở dữ liệu có mối quan hệ trực tiếp hoặc gián tiếp với các tọa độ (kinh độ, vĩ độ) để mô tả vị trí của đối tượng đó trong không gian.

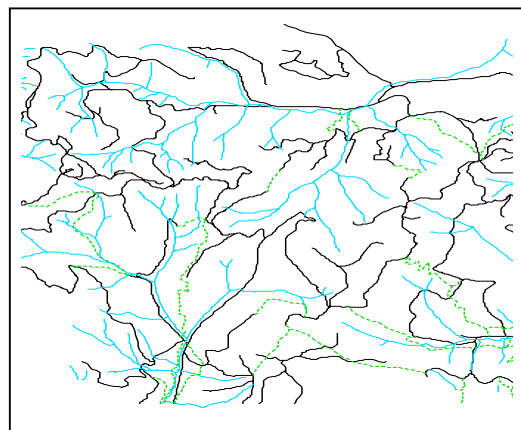
Điểm (Point - Multi Points) - Được xác định bởi tọa độ (x, y) trên bản đồ và các dữ liệu liên quan tới bản chất của đối tượng như màu sắc, ký hiệu, kích thước và các thuộc tính đi kèm khác như (tên đối tượng, mã đối tượng, hàm lượng... - như một bản ghi các thuộc tính kèm theo). Đối với các đối tượng là 3 chiều còn có thêm một tọa độ z (cao độ). Tuy nhiên đa số các dự án GIS đều quan tâm tới đối tượng là hai chiều và chiều thứ ba được coi như là một trường thuộc tính.



Hình 2.5. Các đối tượng điểm trong GIS.

Các đối tượng dạng điểm được mô tả như:

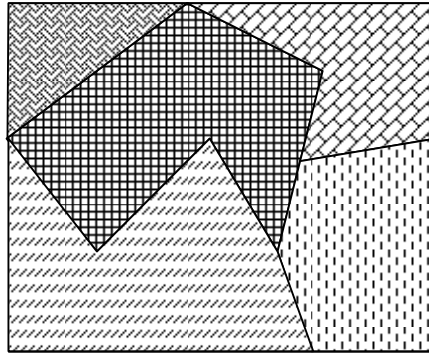
- Các nút giao thông.
- Các điểm độ cao.
- Cầu, cống, nhà, trường học, bệnh viện, nhà ga ...
- Các điểm bảo tồn động vật hoang dã, cây độc lập.
- Các trạm cấp nước điểm phát sóng, trạm ăng ten.
- Các khu vực là vùng nhưng đối với tỷ lệ bản đồ nhất định ta coi như nó suy biến thành các điểm.



Hình 2.6. Các đối tượng dạng đường trong GIS.

Đường (Line - Polyline) - Được đặc trưng bởi một dãy các toạ độ như sau: $(x_0, y_0)(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$ được nối với nhau trong trường hợp đặc biệt đường chỉ gồm hai điểm nối với nhau. Đối tượng đường có các thông tin thuộc tính đặc trưng như: Màu sắc, kích thước, kiểu đường và các thuộc tính liên quan khác (tên đối tượng, độ rộng, độ dài,...). Đối tượng đường dùng để mô tả cho các đối tượng dạng tuyến như:

- Đường giao thông.
- Các con sông, suối, đường điện, đường sắt, dây thông tin...
- Các đường bình độ.
- Đường tuần tra trên đất liền, trên biển.
- Các đường phân giới, bờ biển...



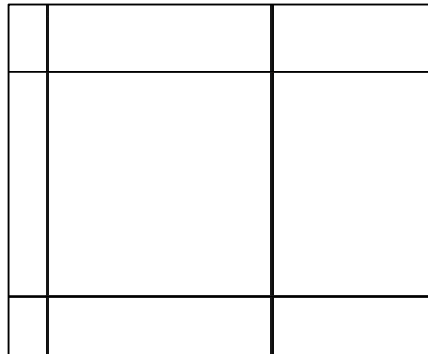
Hình 2.7. Các đối tượng dạng vùng trong GIS.

Vùng (Polygon - Region) - Được mô tả bởi một dãy các điểm toạ độ nối với nhau như sau: $(x_0, y_0)(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$. và được đóng kín, các thông tin liên quan tới vùng như màu sắc, kích thước, kiểu tô màu và các thuộc tính liên quan khác. Đối với vùng có thể bao hàm thêm khái niệm đảo, multi Polygon...

Các đối tượng vùng có thể là:

- Ranh giới thửa đất.
- Bãi đỗ xe.
- Sân vận động.

- Đường bao xã, huyện, tỉnh, quốc gia.
- Các con sông lớn, các vùng đệm của con sông.
- Các hồ chứa nước.
- Vườn quốc gia, khuôn viên...



Hình 2.8. Các đối tượng dạng lưới trong GIS.

Lưới (Grid) - Được mô tả một dãy các ô đều nhau mỗi mắt lưới cách nhau một khoảng cách nhất định. Các ô lưới có kích thước có thể chia theo mét (kích thước thường 1000 m x 1000 m). Lưới chia theo độ có thể có kích thước (1 độ x 1 độ, 0.5 độ x 0.5 độ). Lưới phẳng có thể chia theo km hoặc m có thể chia theo kích thước (1 km x 1 km, 100 m x 100 m)...

Ngoài các đối tượng nêu trên một số các hệ GIS còn có thêm một số các đối tượng đặc biệt khác như cung, hình tròn, hình chữ nhật, text,...để tạo ra các bản đồ có tính thẩm mỹ cao. Tuy nhiên các phép phân tích và chồng xếp bản đồ người ta thường quan tâm tới ba dạng đối tượng đặc trưng nhất: điểm, đường, vùng.

Lớp (Class - Layer) - Là một nhóm các đối tượng có cùng tính chất được tổ chức cùng với nhau chẳng hạn:

- Lớp các đường quốc lộ, đường tỉnh lộ
- Lớp thông tin thủy văn
- Lớp thông tin hành chính
- Lớp các thông tin về dân số

- Lớp thông tin về rừng
- Lớp thông tin về cầu phà
- Lớp thông tin về đường sắt.

Phân lớp dữ liệu

Để có được một cơ sở dữ liệu GIS trước tiên phải phân chia các đối tượng thực (Entities) thành các nhóm đối tượng có những thuộc tính tương tự nhau. Mục đích của việc phân nhóm này làm đơn giản hoá các bước quản trị của hệ thống cũng như các đánh giá phân loại...

Cơ sở dữ liệu GIS là tổng hợp của các đối tượng:

$$\text{CSDL GIS} = \cup \text{Entities}_i$$

Cơ sở dữ liệu GIS là tổng hợp của các bảng mỗi bảng là một nhóm các đối tượng có chung các thuộc tính nào đó:

$$\text{CSDL GIS} = \cup \text{Table}_i$$

Các đối tượng trong một nhóm dữ liệu nào đó được đặc trưng bởi:

$$\text{Table}_i = \Sigma \text{Entities}_j (\text{Att}_k = A) \quad (\text{Att}_k - \text{một thuộc tính phân loại nào đó})$$

Mỗi một Entities bao gồm m trường thông tin

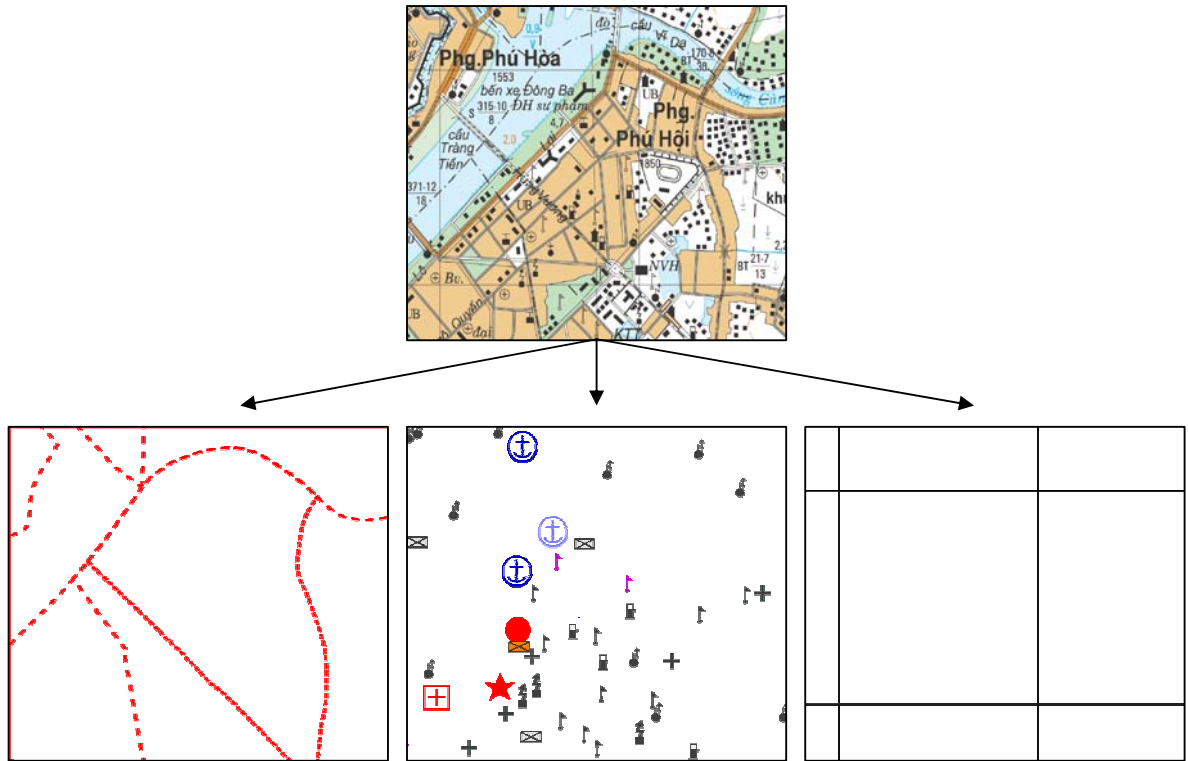
$$\text{Entities}_j = (F_1, F_2, \dots, F_m)$$

Các trường thông tin của đối tượng có thể là dữ liệu không gian định dạng cho đối tượng đó và các dữ liệu thuộc tính đặc trưng cho đối tượng đó.

Một lớp thông tin là một nhóm các đối tượng có cùng một thuộc tính phân loại nào đó. Tùy thuộc vào tính chất của thuộc tính đối tượng mà ta có thể phân loại được chúng theo một tiêu chuẩn xác định.

Đối với lớp thông tin về hành chính ta có thể phân loại các đối tượng thuộc về ranh giới hành chính (Quốc gia, Tỉnh, Huyện, Xã) tuy nhiên ta cũng có thể phân loại chúng sâu hơn thành lớp tỉnh, lớp huyện, lớp xã... Tuy khái niệm này là tương đối nhưng với một cơ sở dữ liệu việc phân loại chúng đóng vai trò quan trọng trong các xử lý sau này.

Việc phân lớp có thể theo một trường hoặc có thể theo một nhóm trường nào đó hoặc theo các đặc điểm tự nhiên mà người thiết kế ban đầu qui định.



Hình 2.9. Phân tách bản đồ thành các lớp



Hình 2.10. Ảnh vệ tinh cũng được xử lý trong GIS

Ảnh (Raster) - ảnh là dạng dữ liệu Raster được chia thành n hàng, m cột. Mỗi ô trên nền ảnh gọi là một pixel. Ảnh có thể thu được qua thiết bị thu ảnh vệ tinh, chụp bằng thiết bị bay chụp ảnh hàng không. Dựa vào ảnh sau khi được định vị về hệ toạ độ sử dụng, người sử dụng có thể giải đoán các đối tượng trên ảnh. Bằng việc kết hợp các đối tượng trên nền ảnh có thể mang lại những thông tin có ích cho người sử dụng.

Trong công nghệ xử lý ảnh bằng cách chụp ảnh lập thể công nghệ đo vẽ thành lập bản đồ dựa trên kỹ thuật đồng dạng điểm ảnh là công nghệ sử dụng hiệu quả trong đo vẽ và thành lập bản đồ hiện nay.

Dữ liệu raster hay còn gọi là các lưới của các cell có được từ nguồn sau:

- Phản xạ quang phổ.
- Dữ liệu thu bức xạ.
- Dữ liệu đã được phân lớp.
- Dữ liệu vector đã được raster hoá.
- Mô hình số độ cao.
- Dữ liệu ảnh quét.

Quan hệ giữa các thực thể trong cơ sở dữ liệu GIS

Các đối tượng trong cơ sở dữ liệu GIS đều có mối quan hệ tương quan với nhau. Các mối quan hệ đó có thể là quan hệ không gian hoặc quan hệ thuộc tính. Giữa hai đối tượng trong cơ sở dữ liệu có thể có mối quan hệ không gian và quan hệ thuộc tính. Tức là các thông tin của đối tượng này có quan hệ với thông tin của đối tượng khác và ngược lại. Ta có thể biểu diễn ngữ nghĩa theo mối quan hệ sau đây:

- Đối tượng X ($F_{i1}, F_{i2}, F_{i3} \dots F_{in}$) - Các trường có thể là thuộc tính hoặc đồ hoạ.
- Đối tượng Y ($F_{j1}, F_{j2}, F_{j3} \dots F_{jm}$) - Các trường có thể là thuộc tính hoặc đồ hoạ.

Khi đó $X \cap Y = K$ ($F_{k1}, F_{k2}, \dots, F_{kl}$) là mối quan hệ chung giữa hai đối tượng trong cơ sở dữ liệu GIS. Với các đối tượng trong cơ sở dữ liệu có cùng

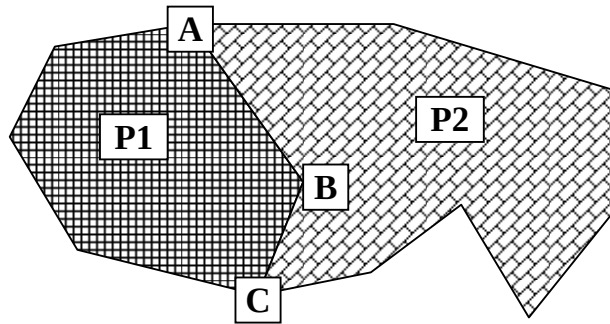
mối quan hệ $K (F_{k1}, F_{k2}, \dots, F_{kl})$ được phân chia thành cùng nhóm đối tượng hay một lớp thông tin.

Quan hệ không gian

Các đối tượng trong không gian có mối quan hệ với nhau gọi là quan hệ Topology. Giữa hai đối tượng trong một lớp hoặc hai đối tượng trong hai lớp khác nhau có thể có mối quan hệ không gian. Giữa lớp thông tin này và lớp thông tin khác đều có mối quan hệ không gian.

Các quan hệ không gian bao gồm:

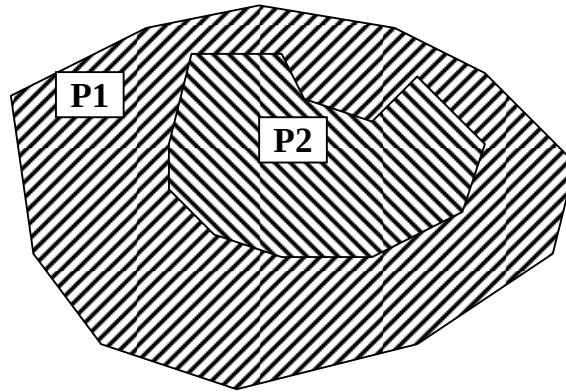
Quan hệ lân cận (hay quan hệ láng giềng) được minh họa theo hình vẽ dưới đây:



Hình 2.11. Mô tả quan hệ lân cận hai Polygon **P1** và **P2**

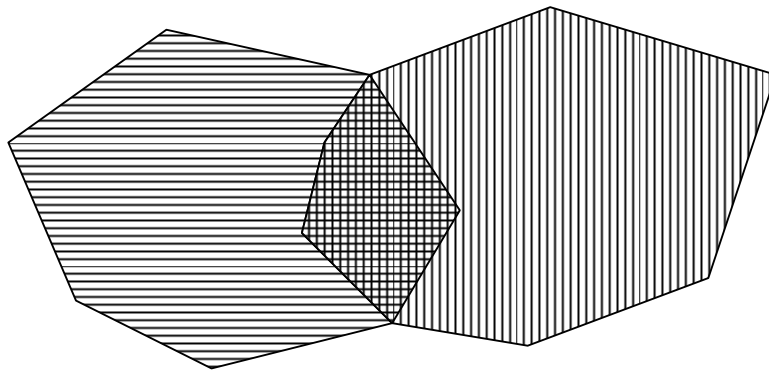
Hai Polygon P1 và P2 được miêu tả như trên được gọi là quan hệ lân cận (quan hệ láng giềng) vì nó cùng chung nhau một cạnh (A,B,C). Vector ABC có quan hệ 2 polygon P1 - Phải và P2 -Trái; P1 và P2 kề nhau bởi Vector ABC. Đây là đặc tính phổ biến nhất trong hệ thông tin địa lý như lớp ranh giới hành chính (tỉnh, huyện, xã).

Quan hệ bao hàm là quan hệ mà đối tượng này được bao kín bởi đối tượng khác chẳng hạn một huyện nằm trong một tỉnh, một hòn đảo nằm giữa hồ nước...



Hình 2.12. Mô tả quan hệ bao hàm (polygon đảo)

Quan hệ giao nhau là quan hệ mà hai đối tượng có một phần chung mà có thuộc tính giống như hai đối tượng đó.



Hình 2.13. Mô tả quan hệ giao nhau của hai polygon

Quan hệ nằm lên nhau là quan hệ của đối tượng thuộc lớp này nằm trên đối tượng của lớp khác. Đây là quan hệ rõ nhất đối với bài toán chồng xếp bản đồ.

2.9 Kết nối dữ liệu không gian và dữ liệu thuộc tính

GIS sử dụng mô hình dữ liệu Vector hoặc Raster để mô tả vị trí, còn dữ liệu phi không gian (thuộc tính) hầu hết được lưu trong các file dữ liệu riêng biệt có cấu trúc hoặc cơ sở dữ liệu quan hệ. Mỗi liên kết được biểu thị bằng cách gán cho các yếu tố địa lý một định nghĩa xác định, tên hay một trường chỉ số ID nào đó được xác định duy nhất. Dữ liệu thuộc tính được lưu trữ trên một hay nhiều file và liên kết với các đối tượng không gian theo chỉ số ID này.

Đối với các hệ GIS trước đây có một sự phân biệt rõ ràng dữ liệu thuộc tính và dữ liệu không gian theo mô hình dưới đây:

ID	Trường 1	Trường 2		Trường N
#1				
.....				
#100				
.....				

Hình 2.14. Bảng mô tả các trường dữ liệu trong GIS

Các dữ liệu được tổ chức riêng biệt và liên kết với nhau theo chỉ số ID được quản lý trực tiếp bằng phần mềm. Một số hệ thống đã có sự liên kết với các bảng dữ liệu thuộc tính được tổ chức trong các hệ cơ sở dữ liệu tuy nhiên dữ liệu trong cơ sở dữ liệu chỉ được lưu trữ và quản lý dưới dạng thuộc tính.

Ngày nay cùng với sự phát triển của các hệ thống thông tin các hệ GIS không có sự phân biệt rõ ràng giữa dữ liệu không gian và thuộc tính tất cả chúng đều được cấu trúc hóa và được quản lý trong cùng một hệ quản trị cơ sở dữ liệu. Nhu cầu sử dụng ngày càng cao, phạm vi ứng dụng của GIS không chỉ trên các máy tính cá nhân riêng lẻ mà hoạt động theo nhóm, xí nghiệp và trên mạng. Do đó các dữ liệu không gian và thuộc tính được tổ chức trong cùng một hệ quản trị cơ sở dữ liệu và người sử dụng có thể coi mỗi lớp thông tin như là một bảng dữ liệu thuộc tính đơn giản. Chính mô hình mới này làm cho vai trò và ứng dụng của GIS được mở rộng lên rất nhiều. Cho phép bảo mật thông tin, trao đổi thông tin và phân quyền sử dụng...

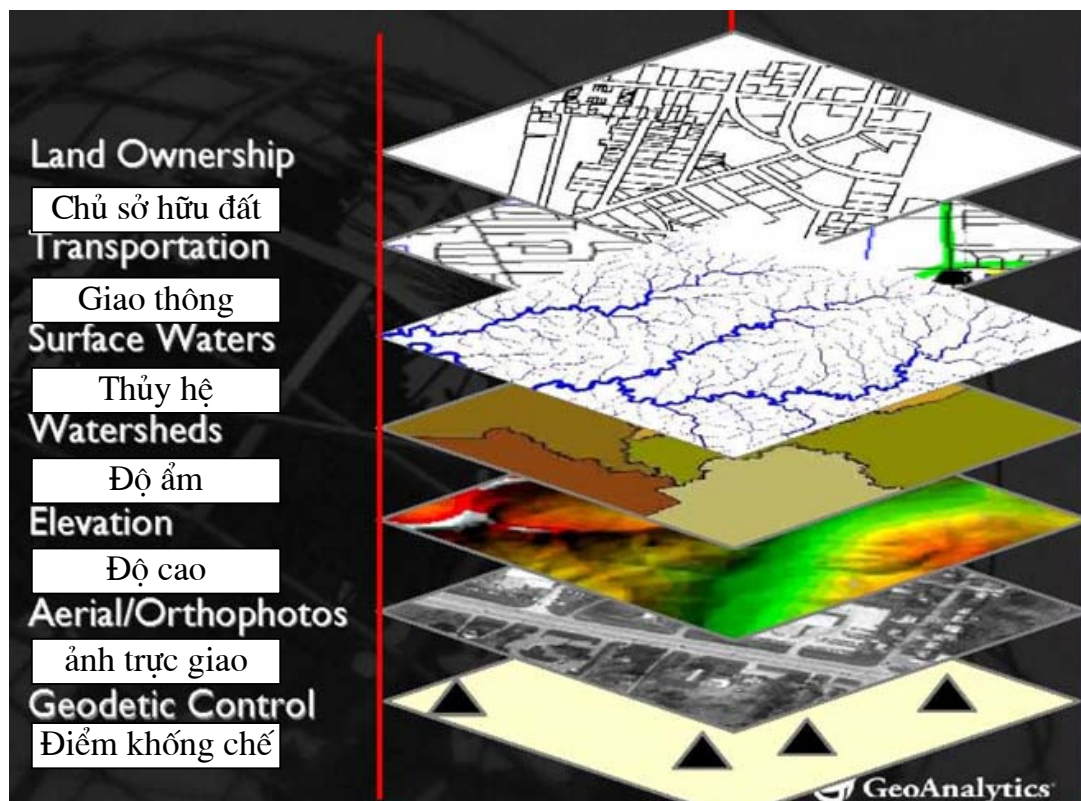
2.10 Chồng xếp và phân tích trong GIS

Chồng xếp bản đồ trong GIS là phép toán đại số thực hiện việc chồng xếp giữa lớp thông tin này và lớp thông tin khác. Kết quả cho ta lớp thông tin mới mà đã được chia cắt bởi các đối tượng trong hai lớp thông tin khi tiến hành chồng xếp. Đối tượng sinh ra ở lớp mới có thể sảy ra:

- Nằm trong hai đối tượng của hai lớp đã cho.
- Nằm trong đối tượng lớp 1 mà không nằm trong đối tượng lớp 2.
- Nằm trong đối tượng lớp 2 mà không nằm trong đối tượng lớp 1.
- Không nằm trong đối tượng của lớp 1 và lớp 2.

Trước khi tiến hành chồng xếp bản đồ các lớp đối tượng cần phải được đưa về cùng một hệ toạ độ, cùng múi chiếu...

Theo phương pháp truyền thống chồng xếp bản đồ được tiến hành trên từng cặp lớp riêng biệt các đối tượng bị cắt vụn để sinh ra các đối tượng mới sau đó thực hiện việc phân loại và tổng hợp lại thành lớp thông tin mới và thực hiện việc chồng xếp với lớp tiếp theo.



Hình 2.15. Chồng xếp chuỗi các lớp bản đồ trong GIS

CHƯƠNG 3 - ỨNG DỤNG LOGIC MỜ TRONG HỆ THỐNG THÔNG TIN ĐỊA LÝ

3.1 Giới thiệu chung

Tính chất “không rõ ràng” trong các hệ thống thông tin địa lý đã trở thành ngày càng phổ biến và được thừa nhận. Các lỗi sinh ra do các kết quả phân tích không gian có thể dẫn tới các kết luận nhầm lẫn, không có quyết định cuối cùng; dựa trên phân tích trong GIS có thể làm cho không có sự hiểu biết về độ chính xác của các giải pháp đưa ra. Khi đó độ tin cậy của ứng dụng qua các bước xử lý thu được không có đủ thông tin về sự không rõ ràng đã biết đối với các tập hợp dữ liệu nguồn.

Lý thuyết tập mờ đầu tiên được A. Zadeh đưa ra vào năm 1965. Lý thuyết này đáp lại sự không đầy đủ của logic kinh điển đối với nhiều vấn đề của thế giới thực. Các phép toán logic kinh điển chỉ thừa nhận hai trạng thái giá trị "0" và "1", trong khi đó phần lớn các thông tin trong thế giới thực là không chính xác, không đầy đủ, không rõ ràng và một trong các khả năng to lớn của con người là xử lý thông tin thực “không chính xác” và “mờ”.

“Đó là thích hợp để sử dụng các tập mờ bất kỳ khi nào chúng ta phải giải quyết với sự nhập nhằng, tính gần đúng và sự lưỡng lự trong các mô hình toán học hoặc quan niệm của các hiện tượng theo lối kinh nghiệm”. (BURROUGH 1989).

Hình dung khi đi bộ cắt ngang theo đường cây từ khu rừng tới đồng cỏ. Ta sẽ thừa nhận rằng có một ranh giới rõ nét giữa khu rừng và đồng cỏ; việc chuyển trạng thái là tương đối “từ từ” giữa hai loại thực vật này. Theo truyền thống sẽ có vấn đề khi tính toán cho việc thay đổi “từ từ” này và đường ngắt cứng (rừng = 0, đồng cỏ = 1) là không thích hợp. Thay vì, bỏ qua sự ngắt cứng đó một ý tưởng của trạng thái “lưỡng” giữa hai loại thực vật này và đưa ra nhiều trạng thái như: “trong rừng”, “phần lớn trong rừng”, “vẫn trong rừng

nhưng cũng trong đồng cỏ”, “*phần lớn trong đồng cỏ*” và “*trong đồng cỏ*”. Giữa hai loại thực vật “*rừng*” và “*đồng cỏ*” có một ranh giới “*mờ*” mà khi sử dụng đối với các tập hợp rõ sẽ gặp nhiều khó khăn trong việc xử lý các ranh giới như thế.

Khi chúng ta nói “*Nhà tôi ở cách xa đường giao thông chính*” và “*ở gần trường học*”, “*ở gần bệnh viện*”. Khái niệm “*xa*” và “*gần*” ở đây phụ thuộc hoàn toàn vào suy nghĩ của con người. Khái niệm “*gần*”, “*xa*” có thể liên quan tới bằng phương tiện nào chúng ta tới được, độ đo khoảng cách hình học và cả quan niệm về “*xa*” và “*gần*” của từng khu vực, lĩnh vực, tập quán... Những khái niệm đó hoàn toàn có thể quan niệm theo khái niệm “*mờ*” và rất nhiều các hiện tượng trong thế giới thực của chúng ta có sự mập mờ.

Một ứng dụng trong cuộc sống thực, chúng ta có thể tìm vị trí thích hợp để xây dựng một ngôi nhà. Tiêu chuẩn cho vị trí đó có thể được đưa ra bởi các điều kiện sau:

- Vùng đất để xây nhà có độ dốc vừa phải.
- Có hướng phù hợp.
- Có độ cao vừa phải.
- Gần hồ.
- Không gần đường giao thông chính.
- Không nằm trong khu vực cấm hoặc quy hoạch.

Tất cả các điều kiện nêu ra ở trên (ngoài trừ trường hợp không nằm trong khu vực cấm) là mập mờ hay không rõ ràng, nhưng phù hợp với cách mà chúng ta đưa ra các điều kiện trong ngôn ngữ và suy nghĩ của con người. Khi sử dụng phương pháp truyền thống các điều kiện đưa ra ở trên có thể chuyển thành các lớp rõ như:

- Độ dốc nhỏ hơn 10 độ.
- Hướng giữa 135 độ và 255 độ hoặc địa hình phẳng.

- Độ cao nằm trong khoảng 100 m và 200 m so với mặt nước biển.
- Nằm trong phạm vi 1000 m từ hồ.
- Không nằm trong phạm vi 300 m từ đường giao thông chính.

Nếu một vị trí rơi vào trong các tiêu chuẩn đưa vào, chúng ta có thể nhận được nó, ngược lại (thậm chí nếu nó có thể rất gần với tập ngưỡng) nó sẽ bị loại ra khỏi phân tích của chúng ta. Tuy nhiên nếu chúng ta cho phép độ thuộc theo các lớp, chúng ta cũng có thể điều chỉnh các vị trí đó mà chỉ bỏ qua một tiêu chuẩn bởi sự sai lệch một vài mét. Chúng sẽ chỉ nhận được độ thuộc thấp hơn và sẽ được kể đến trong phân tích.

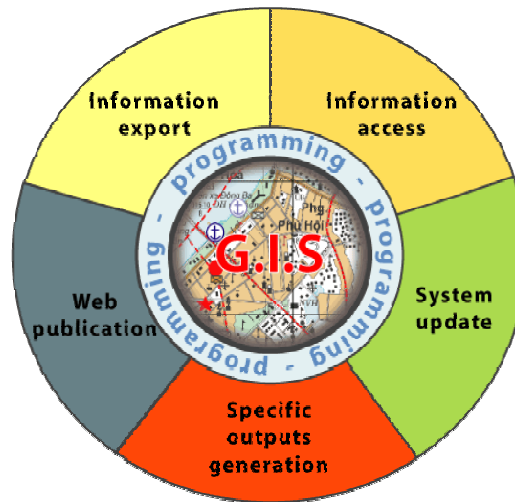
Các hệ thống GIS hiện tại có một số giới hạn làm ảnh hưởng đến hiệu quả trong việc ra quyết định không gian. Giới hạn lớn nhất là các hệ thống GIS thương mại đang lưu hành được thiết lập dựa trên logic kinh điển (logic rõ). Logic mờ (Fuzzy logic) là cơ sở logic thích hợp với một số khái niệm hiệu quả bổ xung cho việc xử lý dữ liệu không gian, quan sát tính mập mờ, mờ hồ trong thông tin, nhận thức, hiểu biết và suy nghĩ của con người. Điều này phù hợp hơn để đối xử với các vấn đề của thế giới thực.

Sự tiến bộ chính của lý thuyết tập mờ này là nó cho phép diễn giải tự nhiên, trong các mục dữ liệu dưới dạng ngôn ngữ, các vấn đề sẽ được giải đúng hơn so với các mục dữ liệu giá trị số chính xác của các quan hệ giữa chúng. Sự tiến bộ này thực hiện với các hệ thống phức tạp trong các phương pháp đơn giản, đó là lý do chính tại sao logic mờ được vận dụng rộng rãi trong kỹ thuật.

Logic mờ xuất hiện là phương tiện thiết kế các công cụ hiệu quả để ra quyết định không gian. Trong những năm gần đây, logic mờ đã được áp dụng thành công trong các xử lý GIS khác nhau. Các bổ sung quan trọng nhất là thực hiện trong các lĩnh vực phân lớp, phân tích, thu thập dữ liệu và trong xử lý ảnh.

3.1.1 Nguyên lý mở rộng các hệ thống GIS

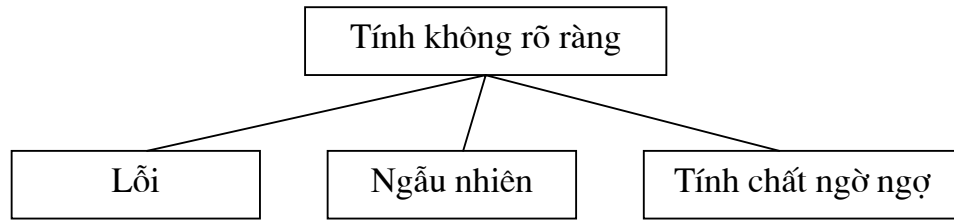
Các hệ thống GIS thương mại hiện nay đều có các bộ thư viện mở để thuận lợi cho việc phát triển và mở rộng các ứng dụng chuyên ngành. Một số hệ thống GIS có ngôn ngữ lập trình riêng dưới dạng Macro để người sử dụng có thể phát triển các ứng dụng đơn giản. Một số hệ thống GIS có các bộ Engine (nhân của hệ thống dưới dạng OCX, Dll hoặc Active) phục vụ cho việc phát triển các ứng dụng chuyên ngành bằng các ngôn ngữ lập trình thông dụng như C++, VB, Delphi, Java... Nguyên lý mở rộng của các hệ GIS được minh hoạ theo mô hình sau:



Hình 3.1. Nguyên lý mở rộng các hệ GIS

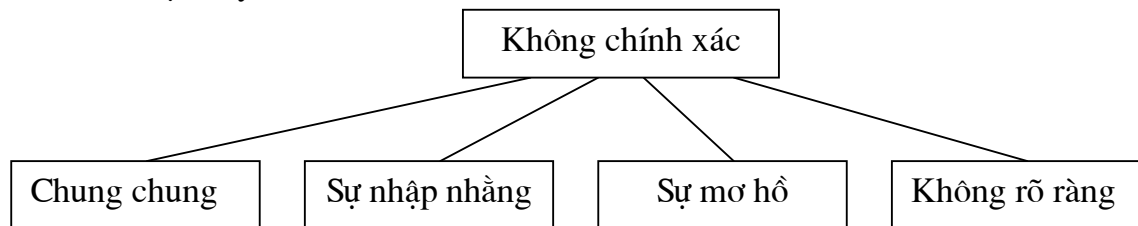
3.1.2 Tính không rõ ràng và hạn chế của Logic rõ trong GIS

Tính không rõ ràng, ngờ ngợ, mập mờ là bản chất của của các hệ thống thông tin địa lý. Các tính chất này sinh ra từ nhiều nguồn khác nhau: từ bản chất dữ liệu trong quá trình thu thập, điều tra; từ các sản phẩm qua các công đoạn tính toán; từ các hệ thống khác; từ bản chất quan niệm của con người; từ việc khoanh vùng, đo đạc nắn chỉnh hình học và từ bản chất hình học của các phép chiếu bản đồ... Tính không rõ ràng khác hẳn với lỗi cũng như tính đúng đắn (bởi vì ta chưa thể kết luận được nó là đúng hay sai). Nó thể hiện ở trạng thái tiềm năng (có thể đúng, có thể sai) mà con người chưa kiểm soát được nó.

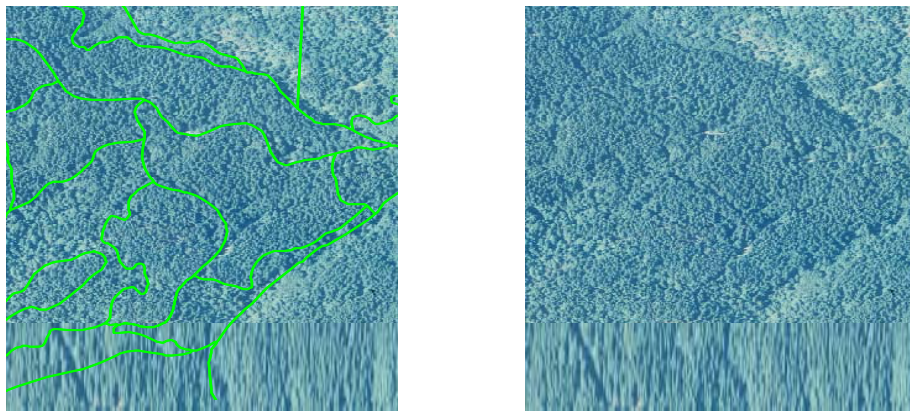


Hình 3.2. Tính không rõ ràng trong GIS (Zhang & Goodchild 2002)

Tính không rõ ràng có liên quan tới sự hiểu biết không hoàn chỉnh và không chính xác của chúng ta về thế giới thực. Chúng ta có thể phân biệt hai lớp không rõ ràng: dữ liệu và quy tắc (Eastman et al. 1993). Tính không rõ ràng về dữ liệu do sự quan sát của chúng ta về tự nhiên hoặc xã hội: chúng ta không chắc chắn được sự chính xác khi quan sát hoặc đo đạc. Tính không rõ ràng về quy tắc do việc lập luận của chúng ta về các quan sát này: chúng ta không chắc chắn được các kết luận do chúng ta có thể rút ra từ dữ liệu (thậm chí từ dữ liệu đầy đủ).



Hình 3.3. Phân loại tính chất không rõ ràng trong GIS



Hình 3.4. Tính chất không rõ ràng phát sinh khi xác định ranh giới

Tính không rõ ràng trong các hệ thống GIS có nghĩa là thông tin không hoàn hảo, không chính xác và mập mờ. Tính không rõ ràng là một đặc trưng vốn có của dữ liệu địa lý. Hiện nay các phương pháp sử dụng để diễn tả và phân tích thông tin địa lý là không đầy đủ, bởi vì chúng không có khả năng đối với tính không rõ ràng của dữ liệu. Điều này chủ yếu phù hợp với ứng dụng lý thuyết tập hợp kinh điển, ở đó một tập hợp có ranh giới được xác định chính xác và một yếu tố có tham gia đầy đủ hoặc không tham gia trong 1 tập hợp. Diễn tả dữ liệu dựa trên lý thuyết tập hợp kinh điển có ảnh hưởng trên các thủ tục lập luận và phân tích, thêm vào đó tất cả các vấn đề phân lớp dễ dàng và chính xác. Quyết định cuối cùng được làm sau các bước, mà mỗi bước này làm giảm trầm trọng các kết quả trung gian. Bất kỳ ràng buộc được chấp nhận với giá trị ngưỡng tuyệt đối và không chấp nhận được cho phép.

Dựa trên logic kinh điển một vị trí với độ dốc 10.001% sẽ bị loại bỏ khi ta lấy ngưỡng là độ dốc $< 10\%$, thậm chí nếu nó thoả mãn hoàn toàn tốt các ràng buộc khác được đưa ra bởi các tiêu chuẩn ra quyết định. Thêm vào đó tiêu chuẩn ra quyết định là bắt buộc để đưa ra các ràng buộc của chúng qua các số hạng số học và các ký hiệu toán học trong các quan hệ rõ (ví dụ, độ dốc $< 10\%$). Khi đó chúng không cho phép sử dụng các số hạng dưới dạng ngôn ngữ tự nhiên (chẳng hạn: *đất phẳng*, *đất dốc*, *đất thoải thoải*, *dốc đứng*). Cuối cùng, ảnh hưởng khác của lý thuyết tập hợp kinh điển là kết quả lựa chọn ở đó là như nhau, không có sự suy xét đến thứ bậc của các thực thể hợp lệ như quan tâm tới độ thuộc mà chúng tham gia đối với tập hợp các ràng buộc, tức là không xét đến trọng số của các điều kiện ràng buộc. Ví dụ khi ta làm nổi bật lớp dữ liệu đất “*khô-phẳng*” tất cả các vị trí mà thoả mãn các ràng buộc: đất khô (độ khô $< 20\%$) và đất phẳng (độ dốc $< 10\%$). Nhưng không có sự phân biệt rõ ràng giữa một vị trí có (độ ẩm = 10% và độ dốc = 3%) với vị trí khác có (độ ẩm = 15% và độ dốc = 7%).

3.1.3 Tính chất mờ trong các hệ thống GIS

Đối với các hệ thống GIS các dữ liệu thu thập thường không đầy đủ, không rõ ràng, không chắc chắn và mập mờ, điều đó dẫn đến dữ liệu và thông tin trong GIS là dữ liệu “không rõ ràng” hay còn gọi là dữ liệu “mờ”.

Các phương pháp sử dụng để diễn tả, chồng xếp và phân tích trong GIS là không đầy đủ bởi vì chúng không được rõ ràng trong việc làm tròn giá trị. Các phương pháp truyền thống tiến hành một cách cứng nhắc với các khái niệm về ngưỡng - giới hạn để phân định một trong hai trạng thái 0-1 (True-False, Yes/No).

Theo phương pháp truyền thống khi chồng xếp và phân tích dữ liệu trong GIS các xử lý được thực hiện một cách “áp đặt” đến các thủ tục lập luận và phân tích. Quyết định tổng thể được thực hiện theo từng bước cụ thể và quy về kết quả ngay lập tức. Những ứng viên nào thoả mãn điều kiện sẽ được giữ lại còn các ứng viên nào không thoả mãn điều kiện sẽ bị loại bỏ ngay tức khắc phụ thuộc vào giá trị ngưỡng (giá trị để phân biệt trạng thái 0-1, đúng-sai...)

Chính sự cứng nhắc của logic kinh điển kéo theo nhiều hạn chế nhất định khi đưa ra các quyết định không gian. Lý tưởng cho việc giải các bài toán không gian bằng logic kinh điển là có được kết quả cuối cùng qua một chuỗi các phép toán phân tích (nghĩa là có lời giải cuối cùng). Tuy nhiên vấn đề này không phải luôn luôn xảy ra. Do bản chất của dữ liệu trong GIS chúng ta có thể gặp tính hướng mà qua một chuỗi các xử lý tập các ràng buộc đưa vào và không nhận được kết quả đưa ra. Và quá trình thực hiện lại phải quay lại ban đầu từ việc phân ngưỡng. Trong khi đó nếu chúng ta giảm bớt một chút về một tiêu chuẩn nào đó ít quan trọng trong xử lý ra quyết định, chúng ta sẽ có được kết quả. Chính ý tưởng này mà việc vận dụng logic mờ vào các xử lý ra quyết định không gian luôn có được kết quả cuối cùng. Nó không tiến hành phân loại tức thì và cứng nhắc đối với tập dữ liệu đưa vào như với logic kinh điển mà nó tiến hành tính toán độ thuộc của chúng sẽ tham gia trong các xử lý sau

này. Và quyết định cuối cùng được xử lý với việc giải mờ sau cùng khi đã tiến hành các phép toán phân tích mờ.

Lý thuyết tập mờ là giải pháp thích hợp nhất cho các điều kiện mô hình hóa dữ liệu “không rõ ràng” và đưa ra cơ sở lý thuyết để hỗ trợ các lập luận dựa trên dữ liệu này. Ứng dụng của logic mờ trong các hệ thống GIS nhằm mục đích mở rộng và tăng cường các chức năng của hệ thống GIS. Làm cho hệ thống GIS trở lên mềm dẻo hơn và ứng dụng thuận lợi trong giải quyết các bài toán về không gian mà dữ liệu của nó là “không rõ ràng” hay còn gọi là dữ liệu không gian “mờ”.

3.2 Logic mờ trong GIS

Nhiều sự kiện chỉ ra độ ngờ ngợ hoặc không rõ ràng mà không thể biểu lộ một cách rõ ràng với các tập hợp rõ của lớp các ranh giới. Các đặc trưng không gian thường không có các ranh giới xác định rõ ràng, và các khái niệm như: “*dốc đứng*”, “*gần*”, hoặc “*phù hợp*” có thể biểu lộ với độ tham gia tới một tập mờ tốt hơn so với việc phân loại 0/1.

Trong suy nghĩ và ngôn ngữ của con người, chúng ta thường sử dụng các khái niệm không chắc chắn hoặc mập mờ. Suy nghĩ và ngôn ngữ của chúng ta không ở dạng nhị phân như ({đen, trắng} ; {0,1} ; {Yes, No} ; {True, False}). Trong cuộc sống thực chúng ta có nhiều thay đổi về sự suy xét và phân lớp dữ liệu của chúng. Các khái niệm mập mờ hoặc không rõ ràng được nói là mờ bắt gặp ở phần lớn mọi nơi trong cuộc sống hàng ngày của chúng ta.

3.2.1 Khái niệm về tập hợp rõ và tập hợp mờ

Định nghĩa 1: (hàm đặc trưng của tập rõ)

Cho A là 1 tập hợp con của tập hợp X hàm đặc trưng χ_A của A được định nghĩa. $\chi_A : X \rightarrow \{0,1\}$ với $\chi_A(x) = \begin{cases} 1 & \text{iff } x \in A \\ 0 & \text{iff } x \notin A \end{cases}$

Trong phương pháp này chúng ta luôn có thể chỉ ra một cách rõ ràng có hay không một phần tử thuộc một tập hợp hoặc không. Tuy nhiên nếu chúng

ta cho phép độ không rõ ràng như có hay không một phần tử thuộc một tập hợp, chúng ta có thể đưa ra độ tham gia của một phần tử tới một tập hợp.

Định nghĩa 2: (Tập mờ).

Một tập mờ A của không gian X được xác định bởi hàm mờ μ_A như sau:

$\mu_A: X \rightarrow [0,1]$ ở đó $\mu_A(x)$ là giá trị thành viên của x trong A . Không gian X luôn là tập rõ.

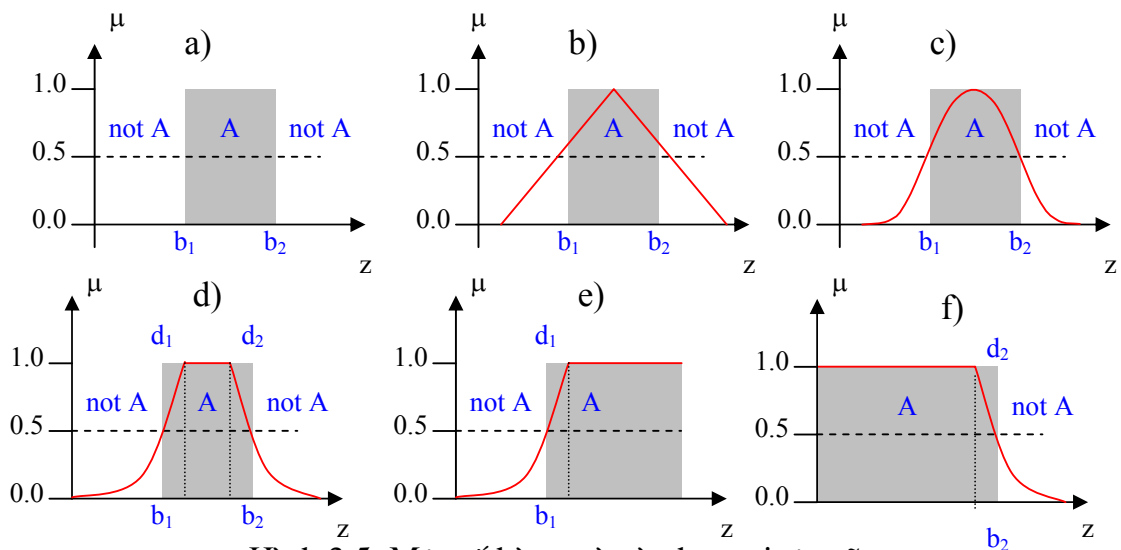
Nếu không gian được định nghĩa là một tập hợp xác định $X = \{x_1, x_2, \dots, x_n\}$ thì một tập mờ A trên X được biểu diễn như sau:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \dots + \mu_A(x_n)/x_n = \sum_{i=1}^n \mu_A(x_i)/x_i$$

$\mu_A(x_i)/x_i$ chỉ ra giá trị tham gia tới tập mờ A đối với x_i . Ký hiệu “/” được gọi là chia, hàm Σ và “+” như là tổng và nối của các khoản mục.

Nếu không gian là tập vô hạn $X = \{x_1, x_2, \dots\}$, thì tập mờ A trên X đưa ra: $A = \int_X \mu_A(x)/x$.

Lựa chọn hàm mờ hợp lệ cho một tập hợp là một trong các lĩnh vực quan trọng nhất của logic mờ. Nó thuộc về trách nhiệm của người sử dụng để lựa chọn một hàm mà diễn tả tốt nhất cho khái niệm mờ được mô hình hóa.

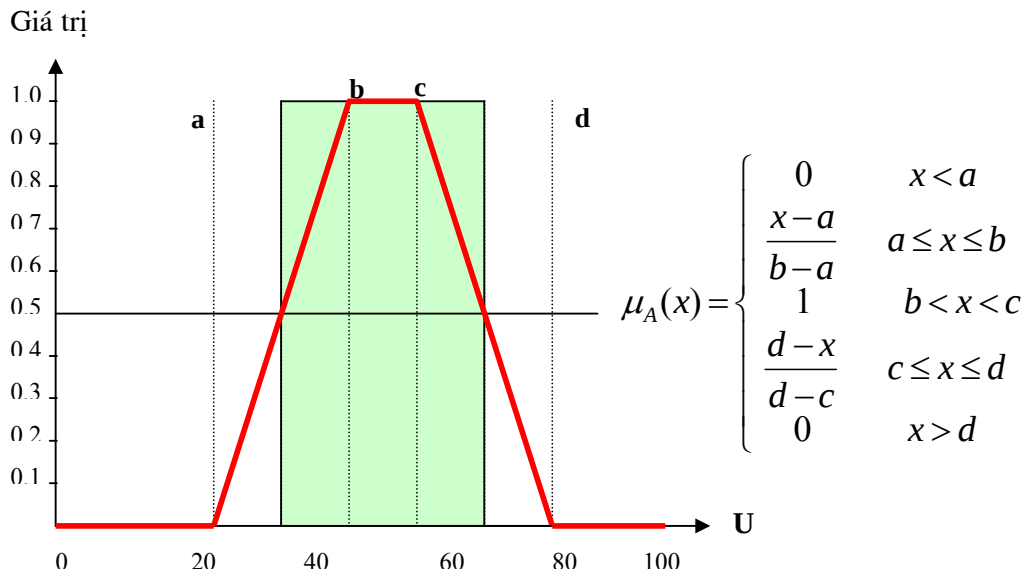


Hình 3.5. Một số hàm mờ và phạm vi tập rõ

Các tiêu chuẩn sau đây là hợp lệ đối với tất cả các hàm mờ:

- Hàm mờ phải là hàm có giá trị thực trong khoảng $[0,1]$.
- Các giá trị hàm mờ sẽ là 1 tại tâm của tập hợp.
- Hàm mờ sẽ suy biến khi 1 khoảng cách thích hợp từ tâm tới ranh giới.
- Các điểm với giá trị 0.5 (điểm cắt ngang) sẽ tại ranh giới của tập rõ, chẳng hạn nếu chúng ta vận dụng việc phân lớp rõ, ranh giới phân lớp sẽ miêu tả bởi các điểm cắt ngang.

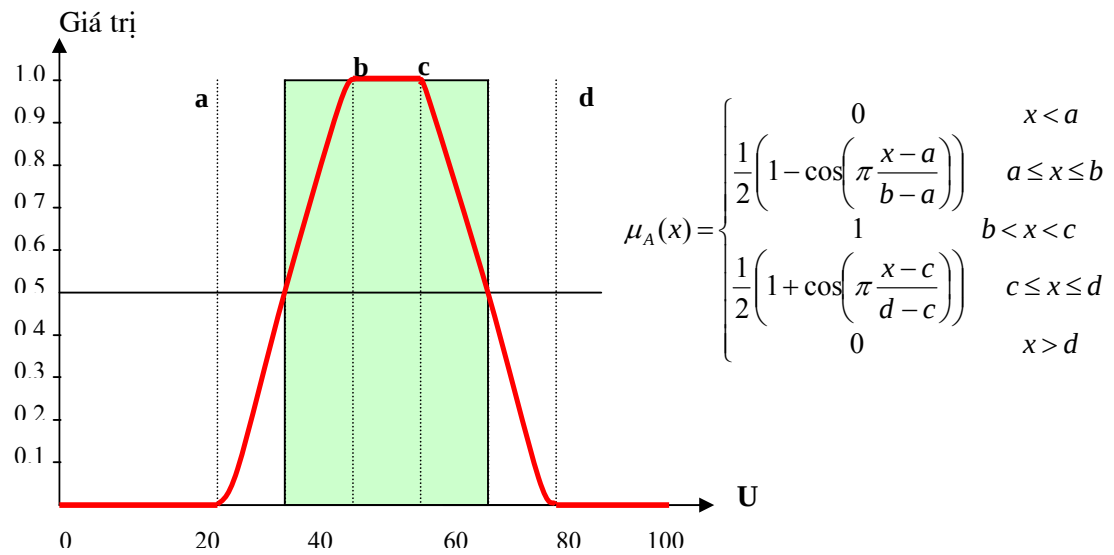
Chúng ta biết hai kiểu hàm mờ: Kiểu hàm mờ tuyến tính và kiểu hàm mờ hình sin. Hàm mờ tuyến tính có bốn tham số xác định hình dạng của hàm:



Hình 3.6. Hàm mờ tuyến tính

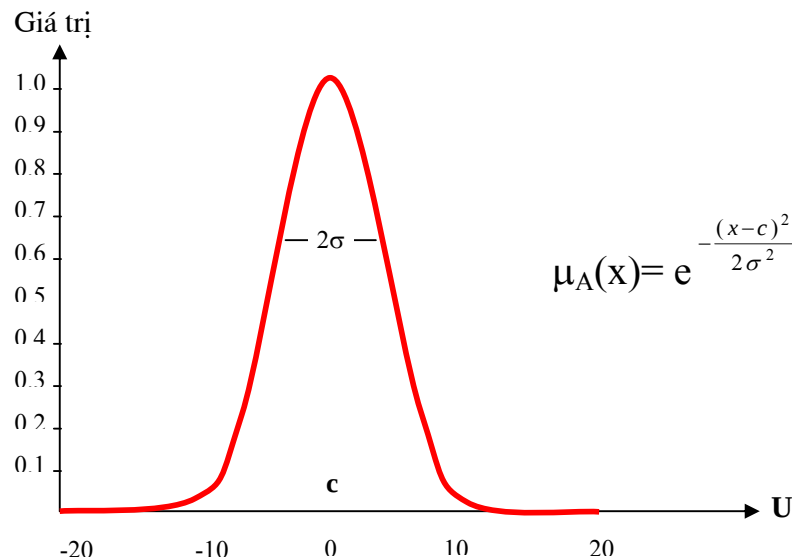
Bằng việc lựa chọn các giá trị thích hợp a, b, c, d chúng ta có thể tạo các hàm với các hình dạng khác nhau như: hình thang, hình tam giác, hình L, hình chữ S ...

Để chính xác hơn không bị gấp khúc tại các nút ta sử dụng hàm mờ hình sin. Giống như với hàm tuyến tính hình dạng của nó cũng có thể là hình chữ S, hình chữ L, hình chuông và cũng có 4 tham số thích hợp theo hình vẽ sau:



Hình 3.7. Hàm mờ hình sin

Trường hợp đặc biệt của hàm hình chuông là hàm Gaussian



Hình 3.8. Hàm mờ Gaussian

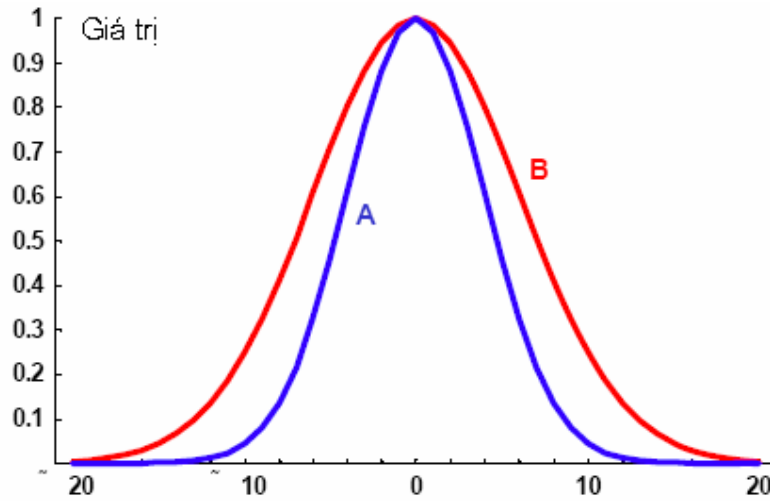
Các phép toán trên tập mờ được định nghĩa giống như đối với tập rõ. Tuy nhiên không phải tất cả các quy tắc cho tập rõ cũng là hợp lệ cho các tập mờ. Như đối với tập rõ chúng ta có tập con, hợp, giao và phân bù. Cộng thêm có các phép toán xen kẽ đối với hợp và giao của các tập mờ.

Định nghĩa 3. (Support). Tất cả các phần tử của không gian X có giá trị độ thuộc lớn hơn 0 đối với tập mờ A được gọi là *support* của A hay theo công thức: $\text{Supp}(A) = \{x \in X \mid \mu_A(x) > 0\}$.

Định nghĩa 4. (Height). *Height* của tập mờ A là giá trị lớn nhất trong A được viết là $\text{hgt}(A)$. Nếu $\text{hgt}(A) = 1$ khi đó tập được gọi là chuẩn.

Định nghĩa 5. (Equality). Tập mờ tương đương, hai tập mờ A và B là tương đương (được ghi là $A = B$) nếu đối với tất cả các thành viên của không gian X giá trị của chúng bằng nhau. $\forall x \in X, \mu_A(x) = \mu_B(x)$.

Định nghĩa 6. (bao hàm). Một tập mờ A bao hàm trong tập mờ B được ghi là $(A \subseteq B)$ Nếu mỗi phần tử của không gian các giá trị của A là nhỏ hơn hoặc bằng giá trị của B : $\forall x \in X, \mu_A(x) \leq \mu_B(x)$.



Hình 3.9. Tập mờ B bao hàm tập mờ A

Hợp: Có thể tính bằng một trong ba phép toán sau:

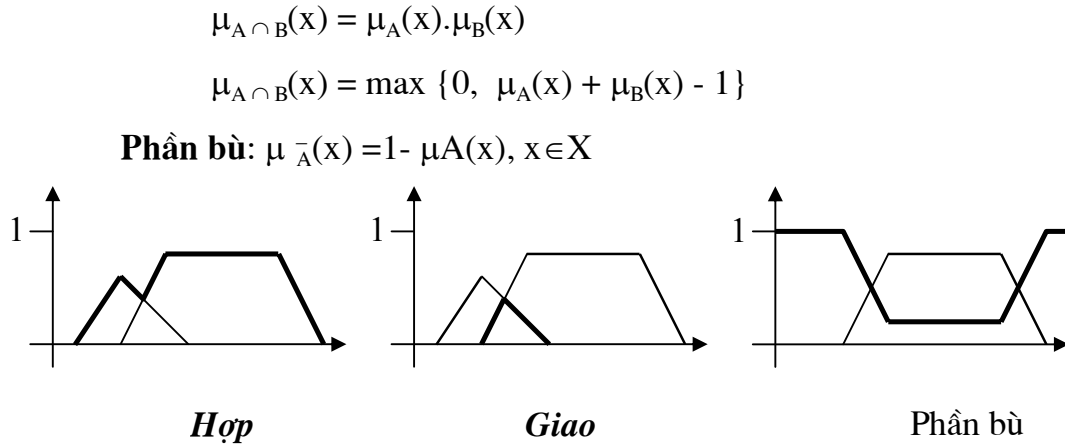
$$\mu_{A \cup B}(x) = \max \{ \mu_A(x), \mu_B(x) \}, x \in X$$

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)$$

$$\mu_{A \cup B}(x) = \min(1, \mu_A(x) + \mu_B(x))$$

Giao: Có thể được tính bằng một trong ba phép toán sau.

$$\mu_{A \cap B}(x) = \min \{ \mu_A(x), \mu_B(x) \}, x \in X$$



Hình 3.10. Minh họa các phép toán tập hợp mờ

Các quy tắc cho các phép toán tập hợp phù hợp cho cả logic rõ và logic mờ theo bảng sau đây:

$A \cup A = A$	<i>Đổi xứng hợp</i>
$A \cap A = A$	<i>Đổi xứng giao</i>
$(A \cup B) \cup C = A \cup (B \cup C)$	<i>Kết hợp</i>
$(A \cap B) \cap C = A \cap (B \cap C)$	<i>Kết hợp</i>
$A \cup B = B \cup A$	<i>Giao hoán hợp</i>
$A \cap B = B \cap A$	<i>Giao hoán giao</i>
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	<i>Phân phối giao</i>
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	<i>Phân phối hợp</i>
$\neg A \cup B = \neg(A \cap \neg B)$	<i>Phần bù hợp</i>
$\neg(A \cap B) = \neg A \cup \neg B$	<i>Phần bù giao</i>
$\neg \neg A = A$	<i>Phần bù của phần bù</i>

Bảng 3.1. Bảng các phép toán logic rõ và mờ

Quy tắc sau chỉ phù hợp cho tập hợp kinh điển:

$$A \cup \neg A = X \quad \text{và} \quad A \cap \neg A = \emptyset$$

Định nghĩa 7. (α -Cut). Lát cắt α (hoặc tập hợp mức α) A_α với $0 < \alpha \leq 1$ là tập tất cả các phần tử của không gian mà $A_\alpha = \{x \in X \mid \mu_A(x) \geq \alpha\}$ Một lát cắt α mạnh là $A_{\bar{\alpha}} = \{x \in X \mid \mu_A(x) > \alpha\}$

Trong các ứng dụng lựa chọn vị trí mỗi bản đồ tác nhân chúng ta có thể định nghĩa các lớp và các đơn vị không gian như tập con mà giá trị độ thuộc của nó trong vị trí thuận lợi được ánh xạ giữa 0 và 1 các phép toán Fuzzy AND, Fuzzy OR, Fuzzy Product, Fuzzy Sum và Fuzzy γ được sử dụng để tích hợp các bản đồ tác nhân với nhau.

Fuzzy AND

$$\mu_{\text{kết hợp}} = \text{MIN}(\mu_A, \mu_B, \mu_C, \dots).$$

Trong đó $\mu_{\text{kết hợp}}$ = Mỗi giá trị đơn vị không gian trong bản đồ đưa ra;

$\mu_{A,B,C}$ = các giá trị mờ thành phần.

Phép toán này sử dụng khi có hai hay nhiều tác nhân hoặc ràng buộc cùng nhau mà có thể trợ giúp giải quyết bài toán.

Fuzzy OR

$$\mu_{\text{kết hợp}} = \text{MAX}(\mu_A, \mu_B, \mu_C, \dots).$$

Trong đó $\mu_{\text{kết hợp}}$ = Mỗi giá trị đơn vị không gian trong bản đồ đưa ra;

$\mu_{A,B,C}$ = các giá trị mờ thành phần.

Phép toán này được sử dụng khi các tác nhân và điều kiện rõ ràng đầy đủ trong vùng nghiên cứu.

Fuzzy Product

$$\mu_{\text{kết hợp}} = \prod_{i=1}^n \mu_i.$$

Trong đó $\mu_{\text{kết hợp}}$ = Mỗi giá trị đơn vị không gian trong bản đồ đưa ra;

μ_i = Trọng số của bản đồ tác nhân thứ i .

Phép toán này được sử dụng khi làm giảm ảnh hưởng các bản đồ tác nhân cùng nhau.

Fuzzy Sum

$$\mu_{\text{kết hợp}} = 1 - \left(\prod_{i=1}^n (1 - \mu_i) \right).$$

Trong đó $\mu_{\text{kết hợp}}$ = Mỗi giá trị đơn vị không gian trong bản đồ đưa ra;

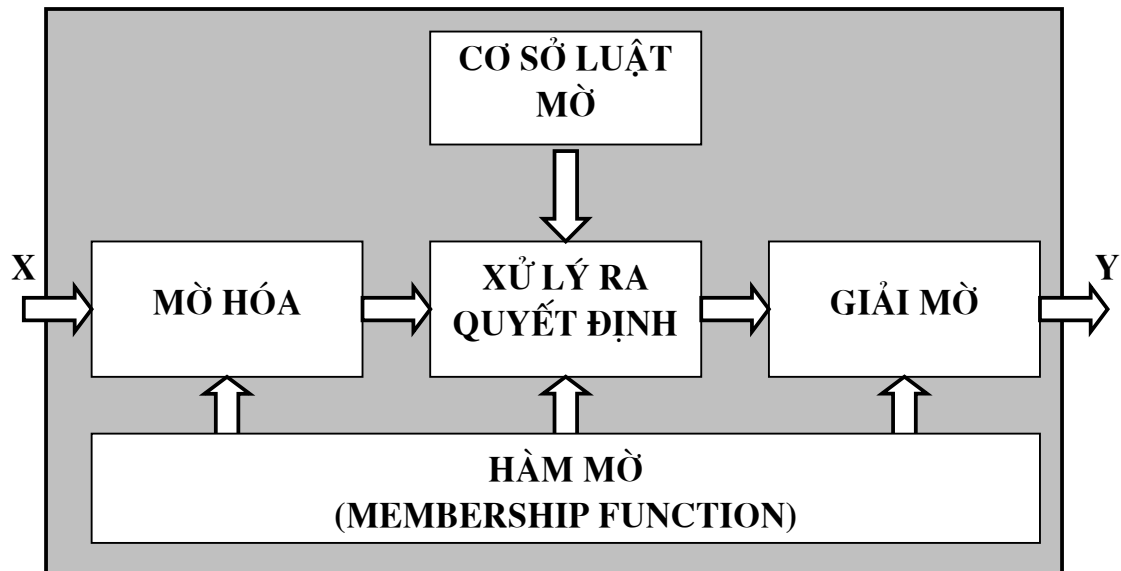
μ_i = Trọng số của bản đồ tác nhân thứ i.

Phép toán này được sử dụng khi làm tăng các ảnh hưởng của bản đồ tác nhân cùng nhau.

Fuzzy γ . Phép toán này là dạng tổng quát của các phép toán Fuzzy Sum và Fuzzy Product.

$$\mu_{\text{kết hợp}} = (Fuzzy Sum)^{\delta} * (Fuzzy Product)^{1-\delta}$$

3.2.2 Hệ mờ trong GIS



Hình 3.11. Hệ mờ áp dụng trong GIS

Bộ công cụ mờ cung cấp các công cụ cho việc xây dựng hệ thống suy luận mờ (Fuzzy Inference System - FIS). Kết luận mờ là xử lý công thức hoá, ánh xạ từ đầu vào tới đầu ra sử dụng logic mờ. Xử lý kết luận mờ bao gồm: các hàm mờ, các phép toán logic mờ và quy tắc IF-THEN. Có hai kiểu của hệ

thống suy luận mờ có thể là phương tiện trong các bộ công cụ logic mờ là kiểu Mamdani và kiểu Sugeno.

Phương pháp suy luận mờ Mamdani là phổ biến nhất cho ta thấy được lý thuyết tập mờ và nó mong chờ các hàm mờ đưa ra là tập mờ. Sau khi khối tập hợp xử lý, có 1 tập mờ cho mỗi biến đầu ra mà cần thiết giải mờ.

Có 5 phần xử lý suy luận mờ: *mờ hóa, chọn hàm mờ, xây dựng các quy tắc, ra quyết định và giải mờ.*

- **Mờ hóa**

Một vấn đề quan trọng khi ra quyết định là lập luận dựa trên các giá trị ngữ nghĩa được gán tới các thực thể vật lý (chẳng hạn: độ dốc giữa 4% và 10%). Một tập hợp các giá trị ngữ nghĩa sẽ được gán tạm thời tới các thực thể và độ đo được phân loại trong các mục dữ liệu. Mỗi giá trị ngữ nghĩa tương ứng với phạm vi của các giá trị vật lý. Mỗi tiêu chuẩn đưa vào sẽ được mờ hóa. Ví dụ độ dốc được phân chia thành 5 mục như sau:

Phân lớp độ dốc	Từ	Đến
Phẳng	0	2
Thoai thoải	2	4
Vừa phải	4	10
Dốc	10	20
Rất dốc	20	30

Bảng 3.2. Bảng ví dụ mô tả các mục phân lớp

- **Lựa chọn hàm mờ**

Một hàm mờ được xác định mỗi điểm trong khoảng đưa vào được ánh xạ tới giá trị mờ (độ thuộc) giữa 0 và 1. Khoảng đưa vào đôi khi quy cho toàn thể không gian mờ. Việc chọn hàm mờ, hình dạng và dạng của nó có ảnh hưởng lớn tới kết quả đưa ra bởi xử lý ra quyết định.

Chính vì sự mềm mại và công thức ngắn gọn của nó. Hàm mờ Gaussian là phương pháp phổ biến được áp dụng cho tập mờ.

Có một hàm chuyển đổi tương ứng cho mỗi giá trị ngôn ngữ, có nghĩa rằng số của các hàm bằng số của các giá trị ngữ nghĩa trung gian.

- **Mô tả không gian**

Mô hình dữ liệu không gian tổng quát trong không gian hai chiều theo lưới của các cells, hoặc đơn vị đất (các vùng trên bề mặt đất). Lưới này được tạo trong GIS, mỗi cell là một thực thể được liên kết với một bản ghi trong cơ sở dữ liệu. Đặc trưng quan trọng nhất của lưới là độ phân giải, bởi vì độ chính xác của kết quả phụ thuộc vào độ phân giải của nó. Đối với các vùng mỗi vùng là một thực thể được kết nối với một bản ghi dữ liệu. Độ chính xác của nó phụ thuộc vào tỷ lệ bản đồ và quá trình số hóa dữ liệu. Trong xử lý với dữ liệu vector bước quan trọng nhất là chuẩn hóa dữ liệu để tránh các lỗi bị cắt vụn trong quá trình xử lý.

- **Xây dựng các luật**

Các câu lệnh IF-THEN sử dụng thành công thức các câu lệnh điều kiện. Luật mờ IF-THEN đơn giản thừa nhận dạng: IF x IS A THEN y IS B.

Ở đó A và B là các giá trị ngôn ngữ được xác định lần lượt bởi tập mờ trên X và Y.

Phần IF của luật "x IS A" được gọi là giả thuyết, khi đó phần THEN của luật "y IS B" được gọi kết luận. Ví dụ của luật như thế có thể là:

IF *độ dốc* IS *nghiêng* THEN *vùng* IS *phù hợp*.

Đưa vào một luật IF-THEN là giá trị hiện thời để đưa vào biến (*độ dốc*) và đầu ra là 1 tập mờ nguyên thủy (*thích hợp*) Tập hợp này sẽ được giải mờ sau này, khi quy cho một giá trị đầu ra.

Khi thông dịch một luật IF-THEN chúng bao gồm các phần riêng biệt: thứ nhất khi đánh giá tiền định (bao gồm mờ hóa đầu vào và vận dụng bất kỳ các phép toán mờ) và thứ hai khi vận dụng kết quả đó tới kết quả cuối cùng.

Trong trường hợp “hai trị” hoặc logic nhị phân, các luật IF-THEN không gặp nhiều khó khăn. Nếu giả thuyết là đúng thì kết luận là đúng. Nếu giả thuyết đúng với nhiều độ thuộc, thì kết quả cũng đúng với cùng độ thuộc như thế.

Luật tiên định có thể có nhiều phần theo biểu thức logic như sau:

IF (độ dốc IS phẳng) AND (hướng IS nam) AND (khả năng đi lại IS gần) AND (độ cao IS thấp) AND (khả năng IS vùng nông nghiệp) THEN vùng IS phù hợp.

Trong trường hợp này tất cả các phần của luật tiên định được tính đồng thời và giải quyết thành số đơn giản khi sử dụng các phép toán logic.

Số trong ngoặc là trọng số của luật đó. Mỗi luật có một trọng số (một số giữa 0 và 1), mà vận dụng tới số đưa vào bởi giả thuyết (định nghĩa ban đầu). Nói chung trọng số này là 1 và vì thế nó không ảnh hưởng tại tất cả trên xử lý liên quan.

- **Ra quyết định**

Các thuật toán mờ được đánh giá khi sử dụng các quy ước tổng quát, đó là thủ tục điều khiển dữ liệu rút ra mà việc phân tích bao hàm hợp thành của các quan hệ mờ, thông thường hợp thành min-max được sử dụng. Hợp thành min-max dưới 1 phép toán liên quan đưa vào ảnh hưởng đến khía cạnh đúng của quy tắc trong một trạng thái chỉ ra (bằng cách phép toán phân đoạn với Mamdani hoặc theo tỷ lệ với Larsen). Tổng quát, quy ước tổng quát là một chuyển đổi của khía cạnh đúng của quy tắc bởi độ tương xứng với độ phủ quy tắc của nó và trong trạng thái đưa ra bởi phép toán liên quan được chọn. Ngoài ra các quy tắc với ELSE và sự kết hợp với các phép toán OR, AND có thể được sử dụng

Từ bảng các giá trị mờ, sử dụng các quy tắc đưa vào bây giờ có thể làm phân tích đa tiêu chuẩn hoặc ra quyết định đa tiêu chuẩn. Phương pháp dễ dàng nhất vận dụng dữ liệu trong các bảng là sử dụng với các câu lệnh SQL. Như thế trong công việc này nó đề xuất việc chuyển các quy tắc IF-THEN

thành câu lệnh SQL. Cuối cùng quy tắc IF- THEN có thể đưa ra trong cơ sở dữ liệu như:

```
SELECT
ID, đô thị tự trị
FROM
bảng tổng hợp
WHERE
độ dốc Is Not Null AND hướng nam Is Not Null AND độ gần Is Not Null
AND đất thấp Is Not Null AND [khu tự trị] Is Not Null;
```

Dựa trên truy vấn như thế trong GIS là lựa chọn tất cả các đơn vị đất cơ sở mà thoả mãn các điều kiện chỉ ra và tính toán tổng diện tích. Kết quả giống như việc thực hiện đối với logic kinh điển và không có bất kỳ khoảng dữ liệu.

Một vấn đề nảy sinh trong trường hợp này là chỉ một giá trị tham gia của các giá trị độ thuộc ảnh hưởng tới việc gán giá trị của nó tới tất cả các tiêu chuẩn quyết định. Trong phương pháp này sự đóng góp của các giá trị độ đo khác bị loại trừ.

Đối với tiêu chuẩn quyết định kết hợp của nhiều hơn một lớp và giá trị ngữ nghĩa sẽ được tính toán và gán tới các vị trí riêng biệt. Độ đo này nhận được từ sự suy xét độ đo trên 2 hoặc nhiều lớp. Ví dụ 1 tập mờ $A \in X$ với độ thuộc $\mu_A(x) \in [X]$ độ đo tổng thể có thể được đưa ra bởi hàm mũ sau: $\mu_E(x) =$

$$\sum_{i=1}^k [\mu_{A_i}(x)]^q.$$

Đối với bài toán về đất phẳng và khô độ đo tổng thể có thể sử

dụng công thức với $q = 2$ như sau: $\mu_{\text{phẳng-khô}}(l) = [\mu_{\text{phẳng}}(l)]^2 + [\mu_{\text{khô}}(l)]^2$. Trong bài toán của chúng ta lấy $q = 2$ ta có độ đo tổng thể theo câu lệnh SQL sau:

```
SELECT ID, đô thị tự trị, ([đất phẳng]^2 + [hướng nam]^2 + [độ gần]^2 +
[độ thấp]^2 + [khu tự trị]^2) AS Result
FROM TK
WHERE
đất phẳng Is Not Null AND hướng nam Is Not Null AND độ gần Is Not
Null AND độ thấp Is Not Null AND [khu tự trị] Is Not Null;
```

Đầu ra của hệ mờ là giá trị mờ. Có một lựa chọn khi sử dụng giá trị này không có bất kỳ sự sửa đổi (để lại công việc làm rõ cuối cùng cho thao tác của con người) hoặc để sử dụng lược đồ giải mờ và sản sinh ra đầu ra rõ.

Các lược đồ giải mờ chung nhất bao gồm các phương pháp của Tsukamoto's, Trọng tâm (Center of Area - COA) và Trung bình lớn nhất (Mean of Maximum - MOM).

Đầu ra được xác định trong bốn lớp như trong bảng sau. Các giá trị ngữ nghĩa này là từ thế giới thực và có các mục dữ liệu ra quyết định thường sử dụng trong công việc của họ. Vì vậy thậm chí không có sự chỉnh sửa nào kết quả vẫn đúng:

Các lớp hạng	Từ	Đến
Thích hợp lạ thường	75	100
Rất thích hợp	50	75
Thích hợp	25	50
Không thích hợp	0	25

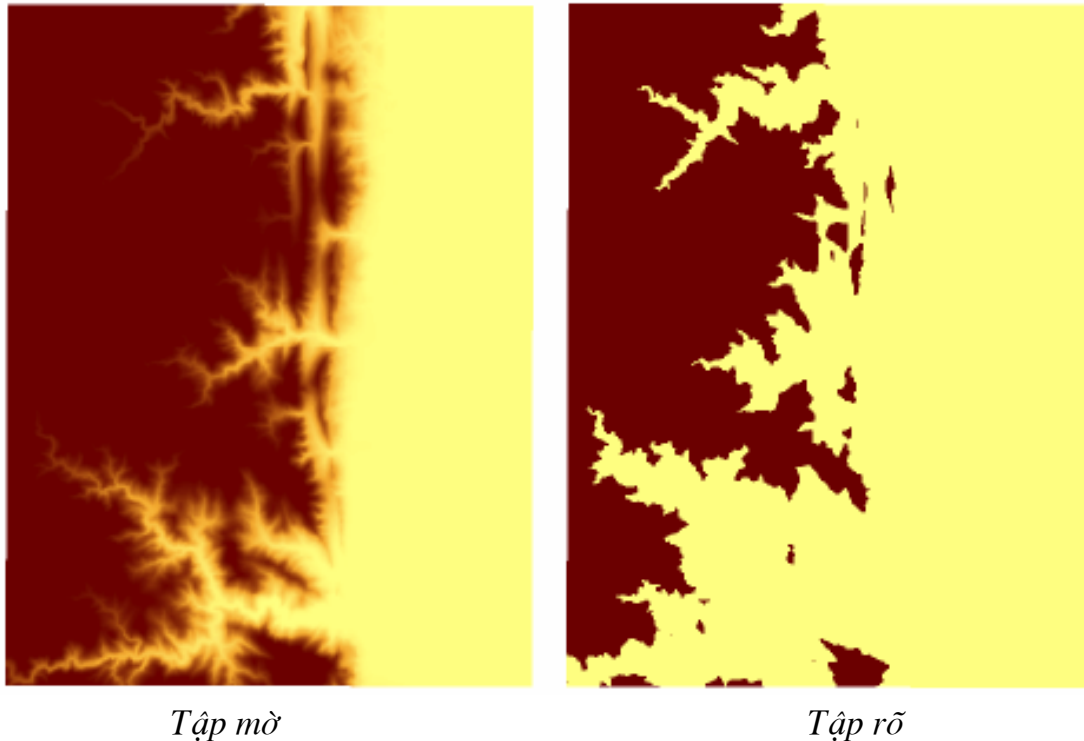
Bảng 3.3. Bảng minh họa ví dụ giải mờ

3.2.3 So sánh giữa Logic mờ và logic rõ (logic kinh điển)

Logic rõ	Logic mờ
<i>Nhận 1 trong 2 giá trị $\{0,1\}$; $\{Yes, No\}$; $\{True, False\}$</i>	<i>Các tập mờ $[0,1]$, $\{các\ giá\ trị\ ngữ\ nghĩa\}...$</i>
<i>Mọi thứ là phần của A hoặc không-A Nó không thể là A và không-A tại cùng thời điểm</i>	<i>Mọi thứ là phần của A và phần của không-A tại cùng thời điểm</i>
<i>ép buộc con người nghĩ rằng rất tốt</i>	<i>Cho phép con người nghĩ và quyết định rất tốt</i>
<i>- Phân lớp sắc nhọn - Vạch rõ sự khác biệt</i>	<i>- Các quyết định mờ - Thông tin mờ</i>

- Biểu thị sự chuyển trạng thái liên tục và các khác nhau tối	- Ngôn ngữ mờ - Biên giới mờ
---	---------------------------------

Bảng 3.4. Bảng so sánh Logic mờ và Logic rõ



Hình 3.12. Phân tích với tập mờ (trái) và tập rõ (phải)

3.3 Mô hình dữ liệu không gian và các phép toán

3.3.1 Mô hình dữ liệu không gian

Hệ thống thông tin địa lý - GIS là hệ thống cơ sở dữ liệu không gian bao gồm một thư viện các bản đồ (các lớp nói chung) mà tất cả đã được chuẩn hoá thống nhất (về toạ độ, đơn vị...). Mỗi lớp tương ứng với một chủ đề và chia thành các đối tượng: điểm, đường, vùng. Chẳng hạn lớp sử dụng đất được phân chia thành các vùng sử dụng đất như: *đầm lầy, sông ngòi, sa mạc, thành phố, công viên, nông nghiệp, dân cư*... Mỗi đối tượng trong một lớp ngoài các tính chất về không gian còn bao hàm các dữ liệu thuộc tính liên quan tới đối

tượng trong lớp đó. Các dữ liệu thuộc tính này có thể tương ứng một - một với từng đối tượng không gian, hoặc có thể liên quan tới các bản ghi dữ liệu thuộc các bảng dữ liệu khác được kết nối tới theo mô hình dữ liệu quan hệ. Đối với mô hình dữ liệu Raster, mỗi pixel trên bản đồ là chỉ số trỏ tới một bản ghi dữ liệu đặc trưng cho pixel đó trên bản đồ. Các hệ thống GIS hiện đại có khả năng kết hợp xử lý giữa dữ liệu raster và vector. Các đối tượng không gian ngoài các đặc trưng của chúng còn có mối quan hệ không gian của các đối tượng trong phạm vi của vị trí đối tượng đó (quan hệ topology)...

3.3.2 Phân lớp các phép toán GIS

Không có đại số chuẩn được định nghĩa trên dữ liệu địa lý. Điều này có nghĩa là không có tập hợp chuẩn của các phép toán cơ sở khi vận dụng đối với dữ liệu địa lý. Tập các phép toán trong GIS có thể khác nhau giữa hệ thống này với hệ thống khác dựa trên phạm vi ứng dụng. Tuy nhiên khả năng nguyên thủy của chúng không thay đổi bao gồm thực hiện bốn nhiệm vụ: lập chương trình, chuẩn bị dữ liệu, mô tả dữ liệu và các phép toán diễn tả dữ liệu.

Các phép toán lập trình: Chúng bao gồm một số các thủ tục ở mức hệ thống, như quản trị và ra lệnh các phép toán hệ thống và điều khiển sự liên lạc tới các thiết bị ngoại vi được nối với máy tính.

Các phép toán chuẩn bị dữ liệu: Chúng bao gồm các phương pháp khác nhau để thu thập dữ liệu từ các nguồn khác nhau (bản đồ số, bản đồ giấy, đo đạc thực địa...), chúng xử lý và gán một cách thích hợp trong cơ sở dữ liệu.

Các phép toán hiển thị dữ liệu: Chúng bao gồm các phương pháp khác nhau để diễn tả dữ liệu (như vẽ các bản đồ, biểu đồ, tạo báo cáo ...).

Các phép toán diễn tả: Các phép toán này chuyển dữ liệu thành thông tin và chúng được coi như là trung tâm của các hệ thống GIS.

Các phép toán diễn tả dữ liệu có thể được xem như là việc phân chia thành các cấp độ dữ liệu. Ở mức cao nhất là một thư viện các bản đồ (các lớp nói chung), tất cả chúng được chuẩn hóa (về cùng hệ tọa độ, cùng độ đo...).

Mỗi lớp được phân chia thành vùng, các vùng là tập hợp của các vị trí với giá trị thuộc tính chung. Ví dụ lớp sử dụng đất được chia thành các vùng sử dụng đất “đầm lầy”, “sông”, hoang mạc, thành phố, công viên và các vùng nông nghiệp; còn lớp mạng đường bao gồm các tuyến đường chạy qua không gian được bao phủ bởi lớp đó.

Các phép toán diễn tả dữ liệu trong các hệ GIS gồm:

- *Các phép toán với mỗi vị trí riêng biệt*
- *Các phép toán vị trí bên trong vùng lân cận*
- *Các phép toán vị trí bên trong một vùng*

Các phép toán được phân chia thành 3 lớp phép toán:

- *Lớp các phép toán cục bộ.*
- *Lớp các phép toán trung tâm.*
- *Lớp các phép toán vùng.*

Tất cả các xử lý dữ liệu được làm trên từng lớp dữ liệu cơ sở. Mỗi phép toán nhận một hoặc nhiều lớp như là đầu vào (các toán hạng) và sản sinh ra một lớp mới như là đầu ra (sản phẩm). Lớp sản phẩm này có thể đóng vai trò như là lớp đầu vào cho các xử lý tiếp theo.

Lớp các phép toán cục bộ: Bao gồm việc tính toán giá trị mới cho mỗi vị trí trên một lớp như là hàm của dữ liệu tồn tại liên quan cụ thể với vị trí đó. Dữ liệu được xử lý bởi các phép toán này có thể bao gồm các giá trị khu vực liên quan với mỗi vị trí trên một hoặc nhiều lớp.

Lớp các phép toán trung tâm: Bao gồm việc tính toán các giá trị mới cho mỗi vị trí như là một hàm lân cận của nó. Một lân cận được xác định như là tập bất kỳ của một hay nhiều vị trí mà hướng về một khoảng cách được chỉ ra hoặc một quan hệ hướng tới một vị trí riêng biệt, tiêu cự lân cận.

Lớp các phép toán vùng: Bao gồm việc tính toán giá trị mới cho mỗi vị trí như là hàm của các giá trị tồn tại tương ứng với một vùng chứa vị trí đó.

Lớp các phép toán	Minh họa các phép toán
Các phép toán cục bộ	
- Các phép toán tìm kiếm	Nhận thông tin liên quan tới các vị trí riêng biệt trên một lớp.
- Phân lớp và mã hóa lại	Tạo lại mã, tính toán lại, phân lớp lại
- Tổng quát hóa	Khái quát hóa, tóm lược
- Chồng xếp (liên kết không gian)	Chồng xếp, chồng lên nhau
Các phép toán trung tâm	
Các phép toán lân cận	Gán giá trị thuộc tính mới tới các vị trí riêng biệt trên một lớp, mô tả khoảng cách hoặc hướng của chúng trong một lân cận đối với tiêu cự lân cận
- Hỏi đáp theo cửa sổ và điểm	Zoom (in/out), điểm trong 1 polygon
- Topological	Rời nhau, gặp nhau, bằng nhau, chứa đựng, bên trong, bao phủ, chồng đè
- Hướng	Bắc, đông-bắc, yếu-giới hạn biên-bắc, cùng-mức
- Hình học (khoảng cách) và vùng đệm (buffer zone)	Gần, không xa, vùng đệm, hành lang
- Láng giềng gần nhất	Láng giềng gần nhất, k-láng giềng gần nhất
Nội suy	
- Các đặc trưng vị trí	Điểm-đường, (nghịch đảo) khoảng cách trọng số
- Các Polygon	Vùng, biểu đồ
Bề mặt	
- Hiển thị, hình dung	Đường bình độ, mô hình mạng tam giác
- Các đặc trưng vị trí	Độ cao, độ dốc, hướng dốc
Tính nối được	
- Đường đi và định vị	Tìm hành trình tối ưu, đường đi tối ưu, lan toả, tìm kiếm
- Tâm nhìn	Hiển thị, chiếu sáng, khung nhìn, trực giao, chiếu rọi
Các phép toán vùng khu vực	
- Các hỏi đáp dấu hiệu (lựa chọn không gian)	Hỏi đáp theo SQL, gọi lại

- Tìm kiếm	<i>Nhận thông tin đặc trưng các vị trí riêng biệt trên một lớp xảy ra với các vùng của lớp khác</i>
- Đo đạc	<i>Khoảng cách, diện tích, chu vi, thể tích</i>

Bảng 3.5. Bảng phân lớp các phép toán trong GIS

3.4 Mở rộng mô hình dữ liệu với Logic mờ

Trong lý thuyết tập mờ khái niệm độ thuộc (độ tham gia của các phần tử trong một tập hợp) được sử dụng để miêu tả các vị trí riêng biệt. Sự hợp nhất tính mờ thành mô hình dữ liệu không gian kéo theo việc định nghĩa lại các cấu thành của mô hình dữ liệu. Trong lý thuyết tập hợp rõ các vị trí riêng biệt trên một lớp được gán với các giá trị thuộc tính. Trong lý thuyết tập mờ chúng được gán các giá trị độ thuộc đối với mỗi giá trị thuộc tính. Các giá trị này được đưa vào bằng cách vận dụng cả các hàm mờ thích hợp và tri thức chuyên gia. Các dữ liệu được mờ hóa vào các trường mờ tương ứng với các đối tượng trong mô hình cơ sở dữ liệu.

Mô hình mở rộng mô hình dữ liệu được đưa ra bởi sơ đồ sau:

ID	← Các trường rõ →				← Các trường độ thuộc →			
	F_1	F_2		F_n	μ_1	μ_2	...	μ_n
#1				...	...	...	...	...
.....				...	...	...	...	...
#100				...	...	...	...	...
.....				...	...	...	...	...

Hình 3.13. Mô hình mở rộng đối với các bảng dữ liệu

3.5 Mở rộng các phép toán với Logic mờ

Sau khi mở rộng mô hình dữ liệu không gian với logic mờ, bước tiếp theo chúng ta tiến hành mở rộng với các phép toán. Mô hình dữ liệu sau khi mở rộng đã chứa các thông tin dữ liệu phù hợp với tính mờ trong GIS. Các

phép toán cũng phải có sự thay đổi để phù hợp với mô hình đã mở rộng ở trên. Điều này bao hàm sự hợp nhất của lý thuyết tập mờ vào trong các phép toán diễn tả dữ liệu cơ bản sẵn có trong các gói phần mềm GIS.

Ba lớp phép toán diễn tả dữ liệu được định nghĩa như sau để hợp nhất tính mờ:

Các phép toán cục bộ mờ: Chúng bao gồm việc tính toán giá trị mờ mới (giá trị độ thuộc) cho mỗi vị trí riêng biệt trên một lớp như một hàm mờ của dữ liệu mờ tồn tại kết hợp rõ ràng với vị trí đó. (phép toán chồng xếp mờ).

Các phép toán trung tâm mờ: Chúng bao gồm tính toán các giá trị mờ mới cho mỗi vị trí riêng biệt như là 1 hàm mờ lân cận của nó (phép toán khoảng cách mờ).

Các phép toán vùng mờ: Chúng bao gồm việc tính các giá trị mờ mới cho mỗi vị trí riêng biệt cho mỗi vị trí riêng biệt như 1 hàm mờ của các giá trị mờ tồn tại tương ứng với 1 vùng mờ chứa đựng vị trí đó (phép toán lựa chọn mờ).

3.5.1 Phép toán phân lớp mờ (Fuzzy Reclasification)

Phân lớp dữ liệu là phân chia các đối tượng theo các mức khác nhau phục vụ cho mục đích hiển thị hoặc các phân tích sau này. Các dữ liệu thu thập được cần được phân loại thành các chủ đề khác nhau đặc trưng cho một nhóm đối tượng nào đó (chẳng hạn đối với lớp rừng cho thể phân loại thành các loại rừng như: rừng già, rừng non, rừng nguyên sinh, rừng quốc gia cần được bảo vệ, rừng trồng, đất trống...). Phân lớp mờ cũng tương tự như phân lớp kinh điển. Chỉ khác nó có thể thực hiện được trên các dạng ngữ nghĩa khác nhau. Mỗi chủ đề trên một lớp được phân loại và sẽ được gán với độ thuộc mà chúng tham gia vào trong tập hợp. Trong ứng dụng mờ phân lớp theo khoảng được vận dụng nhiều trên các trường dữ liệu đối với các bài toán phân tích không gian. Bảng sau là một ví dụ về phân lớp mờ đối với chủ đề độ dày địa tầng, và độ dốc bề mặt:

Lớp	Fuzzy (độ thuộc)	Legend (Chủ đề lớp)
1	0.1	“1 mét”
2	0.3	“2 mét”
3	0.9	“3 mét”
4	0.9	“4 mét”
5	0.9	“5 mét”
6	0.9	“6 mét”

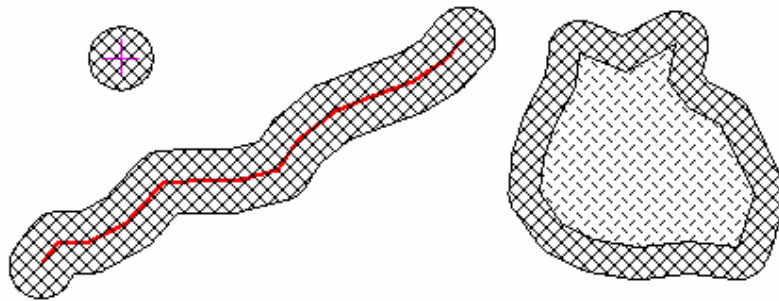
Bảng 3.6. Bảng minh họa độ thuộc về địa tầng

Lớp	Fuzzy (độ thuộc)	Legend (Chủ đề lớp)
1	0.9	“Thấp”
2	0.9	
3	0.7	
4	0.5	“Trung bình”
5	0.4	
6	0.1	
7	0.1	
8	0.1	“đốc đứng”

Bảng 3.7. Bảng minh họa độ thuộc về độ dốc

3.5.2 Phép toán vùng đệm mờ (Fuzzy Buffer)

Các phép toán vùng đệm (buffer) làm tăng kích thước của đối tượng bằng việc mở rộng ranh giới của nó.



Hình 3.14. Các ví dụ về vùng đệm (điểm, đường, vùng)

Nhận hoặc lựa chọn các đặc trưng bên trong hoặc bên ngoài ranh giới của vùng đệm.

Các phép toán vùng đệm có rất nhiều ứng dụng trong thực tế:

- Xác định các vị trí nằm ngoài các nhà máy hóa chất chẳng hạn nó không cách các nhà máy hóa chất dưới 10 km.

- Tìm tất cả các vùng bên trong 300 m của vùng đồn gỗ đưa ra

- Xác định các vùng ô nhiễm tiếng ồn xung quanh các con đường chính

- Các vùng đệm xung quanh vùng đất ô nhiễm để khoanh vùng bảo vệ nguồn nước ngầm.

- Các vùng dịch vụ (2000 m xung quanh tâm tái chế)

- Tạo các vùng bảo vệ tài nguyên (dự trữ tài nguyên thiên nhiên)

- Cụm bệnh dịch xung quanh các đặc trưng nào đó...

Các phép toán vùng đệm mờ bao gồm việc tính toán độ thuộc cho các vùng được mở rộng ranh giới bởi các đối tượng trên các lớp dữ liệu trong GIS. Đối với bản đồ vector xử lý với phép toán buffer đơn giản hơn. Nhưng đối với bản đồ raster phép toán buffer có sự khác biệt so với các phép toán khác.

Không như các phép toán tập hợp, các phép toán buffer raster không thể xác định bởi chính lưới cell trên bản đồ raster. Để xác định giá trị mới của một cell l trong bản đồ raster rõ, các giá trị của tất cả 4 cell lân cận của l được suy xét. Nếu ít nhất một giá trị là 1 thì giá trị của l thay đổi thành 1. Trong trường hợp khác giá trị mới của l là số lớn nhất của giá trị gốc của l và các giá trị của tất cả các cell lân cận của l . Bản đồ raster mờ có thể được làm tương tự: Giá trị của l được thay đổi bằng giá trị mờ lớn nhất trong lân cận của l , mà phải là giá trị trong khoảng $[0,1]$.

Hàm buffer là hàm tăng đơn điệu $\beta: [0, 1] \rightarrow [0, 1]$ mà ở đó giá trị không bao giờ vượt quá đầu vào của nó: $\forall m \in [0, 1]: \beta(m) \leq m$

Ví dụ đơn giản của hàm buffer mờ là $\beta(m) = \max\{0, m - 0,1\}$.

Nếu l_0 là lân cận của l_1 , khi đó độ thuộc của l_1 được xác định:

$$\mu(l_1) \leftarrow \max\{\mu(l_1), \beta(\mu(l_0))\}$$

Khi cập nhật độ thuộc của l_1 có một ảnh hưởng đến các cell lân cận của l_1 vì thế phải thực hiện lặp lại cho đến tình huống đạt được. Thuật toán buffer cho bản đồ raster được thực hiện như sau:

Brute-Force β -Buffering

Cho μ là hàm mờ của bản đồ

Cho β là hàm buffer

Cho L là tập tất cả các cell trong bản đồ để tạo buffer

Repeat Until μ là ổn định:

For each $l_0 \in L$ do:

For all neighbors l_i của l_0 do:

$$\mu(l_i) \leftarrow \max\{\mu(l_i), \beta(\mu(l_0))\}$$

β -Buffering by Local Propagation

Cho μ là hàm mờ của bản đồ

Cho β là hàm buffer

Cho L là tập tất cả các cell trong bản đồ để tạo buffer

While $L \neq \emptyset$ do:

Select $l_0 \in L$.

$L \leftarrow L - \{l_0\}$

For all neighbors l_i của l_0 do:

$$\mu(l_i) \leftarrow \max\{\mu(l_i), \beta(\mu(l_0))\}$$

If $\mu(l_i)$ bị thay đổi, then $L \leftarrow L \cup \{l_i\}$

β -Buffering With Ordered Cells

Cho μ là hàm mờ của bản đồ

Cho β là hàm buffer

Cho L là tập tất cả các cell trong bản đồ để tạo buffer

While $L \neq \emptyset$ do:

Select $l_0 \in L : \mu(l_0)$ là max trong L

$L \leftarrow L - \{l_0\}$

For all neighbors l_i của l_0 do:

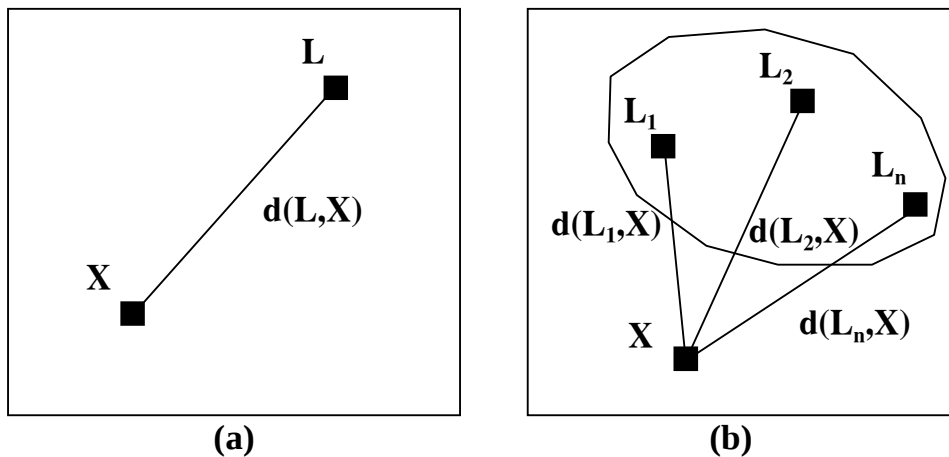
$\mu(l_i) \leftarrow \max\{\mu(l_i), \beta(\mu(l_0))\}$

3.5.3 Khoảng cách mờ (Fuzzy Distance)

Khoảng cách thường đòi hỏi để phân tích các quan hệ không gian giữa các đối tượng trong GIS. Có một số hệ đơn vị được sử dụng, việc lựa chọn hệ đơn vị phụ thuộc vào ứng dụng cụ thể và các đòi hỏi đưa ra bởi việc ra quyết định. Đối với hai điểm i và j khoảng cách Euclidean được đưa ra bởi công thức sau:

$$d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \text{ ở đó } (x_i, y_i) (x_j, y_j) \text{ là tọa độ của 2 điểm } i \text{ và } j.$$

Hai trường hợp khoảng cách mờ đưa ra: Trường hợp thứ nhất chúng chỉ ra các vị trí riêng biệt như thế nào trên lớp được phân loại dựa trên khoảng cách của chúng từ một vị trí đưa; Trường hợp thứ hai Chúng chỉ ra các vị trí riêng biệt như thế nào trên một lớp được phân lớp dựa trên khoảng cách của chúng từ một vùng mờ đưa vào. Để xác định một vị trí đặc trưng riêng biệt X dựa trên khoảng cách của nó từ vị trí L đưa vào.



Hình 3.15. Phép toán khoảng cách mờ giữa 2 vị trí(a);vị trí với vùng mờ(b)

Để mô tả một vị trí riêng biệt X dựa trên khoảng cách của nó từ một vị trí đưa vào L (hình .a) thủ tục sau được thực hiện. Thứ nhất khoảng cách Euclidean d từ L tới X được tính sử dụng phương trình $d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. Khi đó một hàm mờ được chọn để chuyển các khoảng cách thành các giá trị độ đo (mờ) trên các giá trị thuộc tính được xác định trước (dưới dạng giá trị ngôn ngữ) đặc trưng cho chủ đề “độ gần” (lân cận, gần, vừa phải, xa, quá xa). Cuối cùng, khoảng cách từ L tới X được chuyển thành các giá trị độ đo mờ. ở đây sản phẩm của phép toán khoảng cách mờ bao gồm tập của các lớp và mỗi lớp cung cấp các giá trị độ đo đối với một giá trị thuộc tính (lân cận, gần, vừa phải, xa, quá xa) đặc trưng cho chủ đề “gần với vị trí L”.

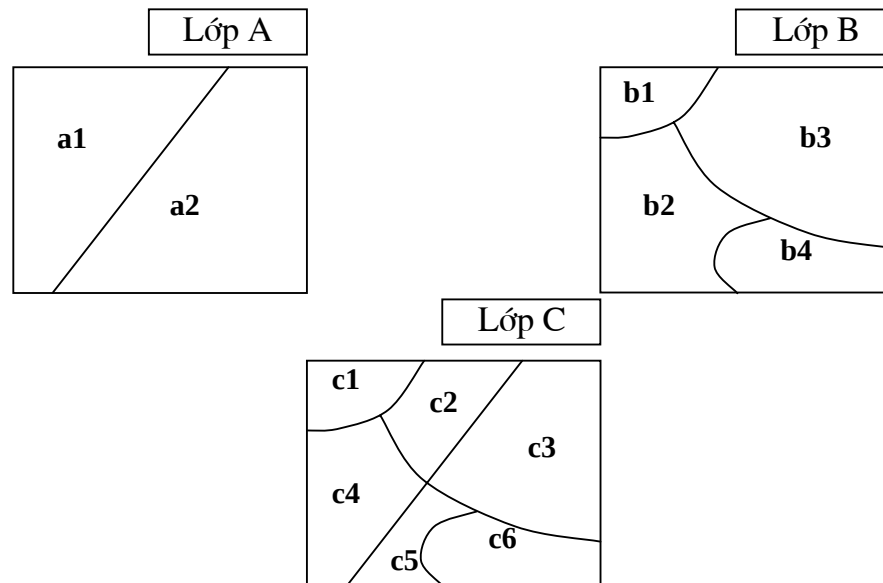
Để mô tả một vị trí riêng biệt X dựa trên khoảng cách từ một vùng mờ đưa vào mà bao gồm tập các vị trí riêng biệt $\{L_1, L_2, \dots, L_n\}$ với các giá trị độ đo khác nhau trong vùng mờ, thủ tục sau được thực hiện. Thứ nhất khoảng cách Euclidean d_i từ tất cả các vị trí $L_i (i=1,2,\dots,n)$ tới X được tính và chuyển thành các giá trị độ đo trên các giá trị thuộc tính được xác định trước đặc trưng cho chủ đề tính gần (ví dụ: *lân cận, gần, vừa phải, xa, quá xa*). Đối với mỗi giá trị thuộc tính A, vị trí riêng biệt X được gán với 1 tập các cặp $(MF_A^F(X), MF_Z^F(L_i))$, $(i=1,2,\dots,n)$, ở đó $MF(X)$ là giá trị độ đo đối với đặc trưng A chủ đề “tính gần”, và $MF_Z^F(L_i)$ là giá trị độ đo của vị trí L_i trong vùng mờ Z. Cuối cùng một hàm mờ được chọn bởi các chuyên gia được vận dụng để ánh xạ tập các cặp thành giá trị độ đo đơn giản (chẳng hạn độ đo tổng thể) đối với A đặc trưng cho chủ đề “gần với vùng mờ Z”.

Một vài hỏi đáp tương đối chung mà ở đó phép toán khoảng cách mờ được vận dụng trong kết hợp với phép toán lựa chọn mờ là: “*tìm tất cả các vùng gần với mạng đường giao thông đã tồn tại*”. “*tìm tất cả các vùng xa trường học*”...Tương tự như phép toán khoảng cách mờ các phép toán trung

tâm khác như hướng mờ (với giá trị ngữ nghĩa: *bắc, đông, nam, tây*); topological mờ (với các giá trị ngữ nghĩa: *liên thông, chồng đè*) có thể được xác định.

3.5.4 Chồng xếp mờ (Fuzzy Overlay)

Đối với bài toán chồng xếp không gian giống như phép toán **join** trong các hệ thống CSDL thông thường. Vấn đề khác biệt quan trọng lớn nhất là sử dụng các điều kiện quan hệ không gian. Chẳng hạn ta có mô hình chồng xếp hai lớp bản đồ như sau:



Kết quả chồng xếp:

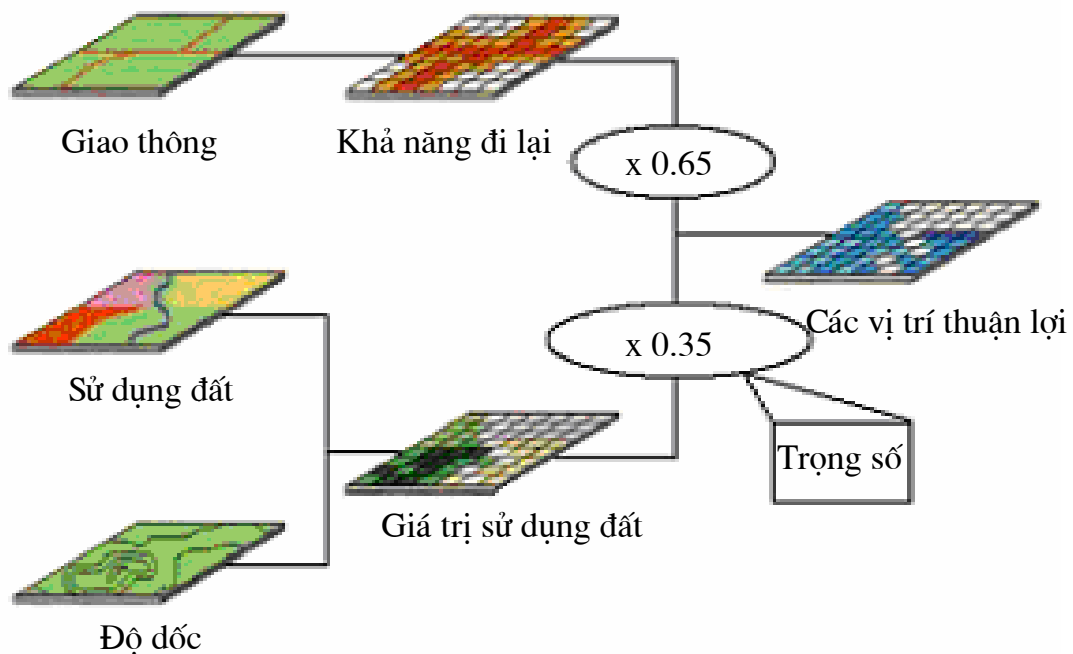
Lớp C	Lớp A	Lớp B
c1	a1	b1
c2	a1	b3
c3	a2	b3
c4	a1	b2
c5	a1	b4
c6	a2	b4

Hình 3.16. Mô tả chồng xếp các lớp

Phép toán chồng xếp mờ tương tự như bài toán chồng xếp bản đồ thông thường. Phép toán chồng xếp được định nghĩa như là việc gán các giá trị thuộc

tính mới tới các vị trí riêng biệt mà kết quả thu được từ việc kết hợp của hai hay nhiều lớp với nhau qua phép toán chồng xếp bản đồ.

Phép toán chồng xếp mờ lấy dạng tổng quát hơn và được định nghĩa như là việc tính toán và gán của một phép đo tổng thể (giá trị mờ) tới mỗi vị trí riêng biệt mà được đưa ra từ sự suy xét của các giá trị độ thuộc trên hai hoặc nhiều lớp đưa vào và thực hiện các phép toán mờ thích hợp. Độ đo tổng thể cũng được đưa ra trong phạm vi mờ $[0,1]$.



Hình 3.17. Mô tả chồng xếp mờ có trọng số.

3.5.5 Lựa chọn mờ (Fuzzy Select), tìm kiếm mờ

Phạm vi của phép toán lựa chọn mờ là làm nổi bật các vị trí riêng biệt trên một lớp dựa trên các giá trị mờ của chúng khi quan sát một đặc trưng thuộc tính đơn giản hoặc đa hợp hoặc một sự kết hợp của các lớp. Ở đây dựa trên các điều kiện đưa ra bởi các truy vấn, phép toán lựa chọn mờ có thể nổi bật:

- Các vị trí riêng biệt ở đó có giá trị mờ trong khoảng giá trị ngưỡng được xác định trước.

- Các vị trí n-riêng lẻ vượt trội với vị trí khác trên các giá trị mờ của chúng (khái niệm bậc).

Hình dưới đây minh họa ví dụ của phép toán lựa chọn mờ. Hình (a) miêu tả giá trị độ đo đối với thuộc tính gần với đường quốc lộ được gán tới các vị trí riêng biệt. Hình (b) làm nổi bật các vị trí với giá trị độ đo lớn hơn hoặc bằng 0.8. Hình (c) làm nổi bật các vị trí ở xa mạng đường đã tồn tại.

0.0	0.0	0.0	0.1	0.1	0.7	1.0	0.7
0.2	0.0	0.0	0.2	0.2	0.7	1.0	0.7
0.7	0.3	0.2	0.4	0.6	0.8	0.9	0.6
0.9	0.7	0.7	0.6	0.7	0.9	0.8	0.6
0.8	1.0	0.8	0.7	0.7	1.0	0.7	0.4
0.3	0.7	0.9	0.7	0.9	0.8	0.6	0.3
0.0	0.5	0.7	0.9	0.7	0.3	0.2	0.1
0.0	0.2	0.6	0.8	0.5	0.2	0.0	0.0

(a)

0.0	0.0	0.0	0.1	0.1	0.7	1.0	0.7
0.2	0.0	0.0	0.2	0.2	0.7	1.0	0.7
0.7	0.3	0.2	0.4	0.6	0.8	0.9	0.6
0.9	0.7	0.7	0.6	0.7	0.9	0.8	0.6
0.8	1.0	0.8	0.7	0.7	1.0	0.7	0.4
0.3	0.7	0.9	0.7	0.9	0.8	0.6	0.3
0.0	0.5	0.7	0.9	0.7	0.3	0.2	0.1
0.0	0.2	0.6	0.8	0.5	0.2	0.0	0.0

(b)

0.0	0.0	0.0	0.1	0.1	0.7	1.0	0.7
0.2	0.0	0.0	0.2	0.2	0.7	1.0	0.7
0.7	0.3	0.2	0.4	0.6	0.8	0.9	0.6
0.9	0.7	0.7	0.6	0.7	0.9	0.8	0.6
0.8	1.0	0.8	0.7	0.7	1.0	0.7	0.4
0.3	0.7	0.9	0.7	0.9	0.8	0.6	0.3
0.0	0.5	0.7	0.9	0.7	0.3	0.2	0.1
0.0	0.2	0.6	0.8	0.5	0.2	0.0	0.0

(c)

Hình 3.18. Phép toán lựa chọn mờ

Khác với phép tìm kiếm kinh điển trên các dữ liệu trong GIS. Các phép toán tìm kiếm mờ nhận thông tin dựa trên giá trị ngưỡng được xác định trước đối với độ đo tổng thể được gán tới các vị trí riêng biệt trên một lớp dữ liệu. Trong GIS sử dụng chức năng phân lớp chủ theo ngưỡng có thể lựa chọn các giá trị theo các ngưỡng đưa vào.

3.5.6 Suy luận mờ

Trong logic kinh điển chúng ta chỉ có 2 giá trị có thể cho biến logic, đúng hoặc sai, 1 hoặc 0. Các tập mờ cũng có thể vận dụng suy luận khi các khái niệm mập mờ được bao hàm.

Trong logic kinh điển khi suy luận dựa trên sự suy diễn của nó hoặc quy nạp. Trong lập luận mờ chúng ta sử dụng sự suy diễn mà được đọc như sau:

Giả thuyết₁ : **If x is A then y is B**

Giả thuyết₂ : **x is A'**

Kết luận : **y is B'**

Ở đây A, B, A', B' là các tập mờ A' và B' không chính xác giống như A và B .

Phương pháp điều khiển MAMDANI

Phương pháp Mamdani dựa trên suy diễn tổng quát sau:

$$p \Rightarrow q: \begin{cases} \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } z \text{ is } C_1 \\ \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } z \text{ is } C_2 \\ \dots \\ \text{If } x \text{ is } A_n \text{ and } y \text{ is } B_n \text{ then } z \text{ is } C_n \end{cases}$$

$$\frac{p_1: \quad \quad \quad x \text{ is } A', \quad y \text{ is } B'}{q_1: \quad \quad \quad z \text{ is } C'}$$

Giả thuyết₁ trở thành tập của các luật.

A, B, C là các tập mờ x, y là các biến giả thuyết z là các biến kết luận

If x is A and y is B then z is C

Giả thuyết

Kết luận

Xử lý suy diễn được tường minh theo thủ tục sau:

Cho x_0 và y_0 là đầu vào cho các biến giả thuyết.

- Vận dụng các giá trị đầu vào tới các biến giả thuyết cho mỗi luật và

tính min của $\mu_{A_i}(x_0)$ và $\mu_{B_i}(y_0)$:

Luật₁: $m_1 = \min(\mu_{A_1}(x_0) \text{ và } \mu_{B_1}(y_0))$

Luật₂: $m_2 = \min(\mu_{A_2}(x_0) \text{ và } \mu_{B_2}(y_0))$

...

Luật_n: $m_n = \min(\mu_{A_n}(x_0) \text{ và } \mu_{B_n}(y_0))$

- Cắt các hàm mờ của kết luận $\mu_{C_i}(z)$ tại m_i :

Kết luận của luật₁: $\mu_{C_1}(z) = \min(m_1, \mu_{C_1}(z))$

Kết luận của luật₂: $\mu_{C_2}(z) = \min(m_2, \mu_{C_2}(z))$

....

Kết luận của luật_n: $\mu_{c'1}(z) = \min(m_1, \mu_{c1}(z))$

Tính kết luận cuối cùng bằng cách xác định hợp các tất cả các kết luận riêng biệt từ bước trên:

$$\mu_c(z) = \max(\mu_{c'1}(z), \mu_{c'2}(z), \dots, \mu_{c'n}(z)).$$

Kết quả của kết luận cuối cùng là một tập mờ. Chúng ta cần thiết phải giải mờ. Có một vài luật để giải mờ một trong các luật đó là trọng tâm.

$$Z_0 = \frac{\sum \mu_c(z).z}{\sum \mu_c(z)}$$

Phương pháp đơn giản hóa

$$p \Rightarrow q: \begin{cases} \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } z \text{ is } c_1 \\ \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } z \text{ is } c_2 \\ \dots \\ \text{If } x \text{ is } A_n \text{ and } y \text{ is } B_n \text{ then } z \text{ is } c_n \end{cases}$$

$$\frac{p_1: \quad \quad \quad x \text{ is } A', \quad y \text{ is } B'}{q_1: \quad \quad \quad z \text{ is } c'}$$



- Vận dụng các giá trị đầu vào tới các biến giả thuyết cho mỗi luật và tính min của $\mu_{A_i}(x_0)$ và $\mu_{B_i}(y_0)$:

$$\text{Luật}_1: m_1 = \min(\mu_{A1}(x_0) \text{ và } \mu_{B1}(y_0))$$

$$\text{Luật}_2: m_2 = \min(\mu_{A2}(x_0) \text{ và } \mu_{B2}(y_0))$$

...

$$\text{Luật}_n: m_n = \min(\mu_{An}(x_0) \text{ và } \mu_{Bn}(y_0))$$

- Tính toán giá trị kết luận trên luật

$$\text{Kết luận của luật}_1: c'_1 = m_1 . c_1$$

$$\text{Kết luận của luật}_2: c'_2 = m_2 . c_2$$

....

Kết luận của luật_n: $c'_n = m_n \cdot c_n$

- *Tính toán kết luận cuối cùng như sau:* $c' = \frac{\sum_{i=1}^n c'_i}{\sum_{i=1}^n m_i}$

3.6. Lựa chọn vị trí dựa trên một chuỗi các phép toán GIS

Mục đích của việc lựa chọn vị trí dựa trên một chuỗi các phép toán là đưa ra tuần tự các phép toán diễn tả dữ liệu mà có thể sắp xếp các thủ tục để hoàn thành nhiệm vụ lựa chọn vị trí. Chẳng hạn lựa chọn vị trí cho việc phát triển khu tái định cư. Phương pháp cơ bản để làm điều này là tạo một tập hợp các ràng buộc được giới hạn bởi phạm vi quy hoạch và tập các điều kiện cho phép. Trong tình huống đơn giản xét tập các ràng buộc và điều kiện gồm:

- *Vùng đất trống.*
- *Đất khô.*
- *Vị trí bằng phẳng*
- *Gần mạng giao thông đã tồn tại*
- *Hướng dốc là hướng nam.*
- *Vùng đất quy hoạch có diện tích giữa 1 và 1.5 km².*

Trong 6 điều kiện trên điều kiện cuối cùng cho các vùng có kích thước phù hợp cho công tác quy hoạch được thực hiện sau cùng khi đã tiến hành xử lý với 5 điều kiện ban đầu và tính toán diện tích cho tất cả các vùng thỏa mãn 5 điều kiện đầu. Sau đó các nhà quy hoạch sẽ xem xét các vùng đất thỏa điều kiện ràng buộc thứ 6 đáp ứng cho mục đích quy hoạch.

Các đòi hỏi trên sử dụng ba lớp dữ liệu đầu vào của vùng nghiên cứu:

- Lớp thông tin địa hình địa chất (mô hình số độ cao của vùng).
- Lớp thông tin đô thị: Bao gồm cơ sở hạ tầng đã tồn tại của vùng (đường xá, các toà nhà,...)
- Lớp độ ẩm: Bao gồm độ ẩm đất của vùng (hồ, đầm lầy, đất khô,...).

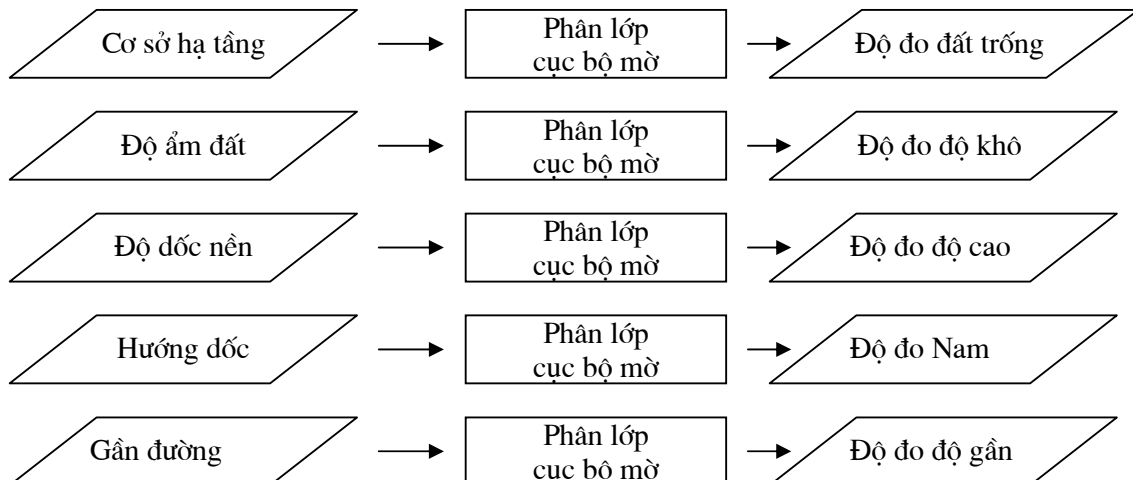
3.6.1 Lựa chọn vị trí sử dụng logic mờ

Đối với bài toán lựa chọn vị trí cho việc phát triển khu dân cư đã nêu ra ở trên, nhiều tiến bộ đã được đưa ra bởi các phép toán diễn giải dữ liệu mờ có thể được coi là điểm sáng. Để các điều kiện đưa ra quyết định các giá trị ngữ nghĩa có thể được suy xét như sau:

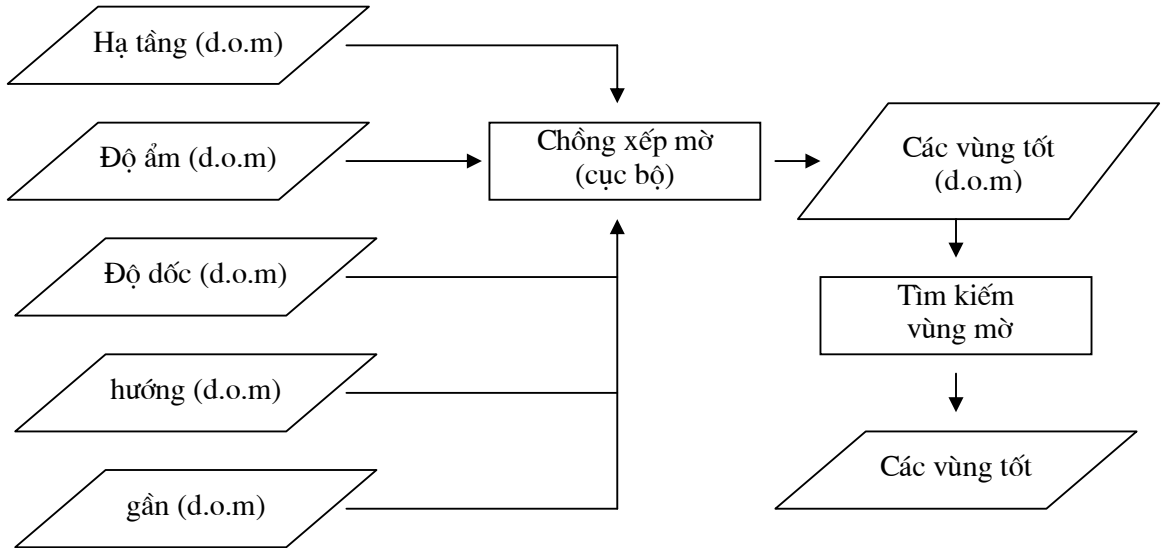
- Độ dốc nền { *phẳng, thoải, vừa phải, dốc đứng* }
- Tính phát triển { *hoang, nửa phát triển, đã phát triển* }
- Độ ẩm đất { *khô, vừa phải, ẩm, nước* }
- Tính thuận lợi về giao thông { *lân cận, gần, vừa phải, xa, quá xa* }
- Hướng dốc { *bắc, đông, nam, tây* }

Các hàm chuyển đổi sẽ chấp nhận ánh xạ các độ đo nền tới các giá trị d.o.m đặc trưng cho các vị trí riêng biệt của vùng nghiên cứu. Bằng cách thực hiện một phân lớp mờ, một lớp d.o.m sẽ được sản sinh đối với mỗi giá trị ngữ nghĩa đặc trưng cho 1 chủ đề. Ví dụ các layer tương ứng với các giá trị ngữ nghĩa quan tâm (*hoang vắng, khô, phẳng, gần, nam*) một phép chồng xếp mờ sẽ đưa ra một lớp mới mà phân lớp tất cả các vị trí riêng biệt của vùng nghiên cứu dựa trên độ tham gia của chúng đối với các điều kiện đưa ra bởi quyết định. Phép toán lựa chọn mờ sẽ làm nổi bật tất cả các vị trí tốt nhất cho hoạt động quy hoạch.

Sản sinh các lớp với các giá trị mờ



Sản sinh lớp các vùng tốt và độc lập thỏa mãn điều kiện quy hoạch



3.6.2 Bài toán ra quyết định không gian và logic mờ

Lý thuyết tập mờ có những ưu thế để miêu tả và vận dụng sự mập mờ mà quan hệ tới việc phân lớp của các vị trí riêng biệt theo các giá trị thuộc tính của chúng. Thay cho các giá trị số của các thực thể thế giới thực và các giá trị đo được gán bằng các giá trị ngôn ngữ. “*Chẳng hạn vị trí là xa với đường quốc lộ*”. Câu lệnh này có các đặc trưng không rõ ràng. Sự không rõ ràng quan hệ tới nhận thức về khoảng cách giữa vị trí và mạng đường. Nhận thức khoảng cách có thể được tạo thành bởi độ đo khoảng cách từ mục tiêu tới đường quốc lộ gần nhất chẳng hạn 20 km, cảm giác và nhận thức của sự quan sát. Khái niệm không rõ ràng miêu tả mức độ thuộc của một đối tượng trong một tập hợp. Độ đo này được đưa ra như là độ thuộc. Độ thuộc thường là giá trị trong khoảng $[0,1]$ và được gọi như là lĩnh vực mờ.

Các giá trị ngôn ngữ được gán tới các thực thể tương ứng với khoảng giá trị vật lý ($xa \Rightarrow \text{khoảng cách} \in [15 \text{ km}, \infty]$). Việc chuyển các giá trị vật lý thành giá trị mờ được thiết lập qua công việc các hàm chuyển đổi theo dạng:

$$f : R \rightarrow [0,1].$$

Thủ tục chuyển đổi các giá trị vật lý thành giá trị mờ được gọi là mờ hóa và các giá trị mờ là đơn vị mờ tương ứng giá trị vật lý thuộc tập hợp biểu thị bởi giá trị ngữ nghĩa.

Một vấn đề quan trọng với việc ra quyết định là lập luận dựa trên các giá trị ngữ nghĩa được gán tới các thực thể vật lý. Theo lược đồ đưa ra một tập hợp các giá trị ngữ nghĩa sẽ không có thật để phân lớp các thực thể và các độ đo trong các khoản mục. Mỗi giá trị ngữ nghĩa tương ứng tới một giới hạn của các giá trị vật lý khi các hàm chuyển đổi được đưa ra để ánh xạ các giá trị vật lý đối với các giá trị mờ. Có một hàm chuyển đổi được gán tới mỗi giá trị ngữ nghĩa. Ở đây số các hàm chuyển đổi bằng số các giá trị ngữ nghĩa. Có các dạng hàm chuyển đổi sau:

- *Tuyến tính tăng* : Nó được sử dụng trong các trường hợp ở đó ánh xạ thẳng các giá trị vật lý tới phạm vi mờ là cần thiết. Hàm tuyến tính tăng được mô tả bởi phương trình:

$$LI(x) = (x-c_0)/(c_I-c_0), \forall x \in [c_0, c_I]$$

- *Tuyến tính giảm*: Nó biểu diễn bởi phương trình:

$$LD(x) = (x-x_0)/(c_0-c_I) + 1, \forall x \in [c_0, c_I]$$

- *Tam giác*: Tập các giá trị vật lý được phân chia thành k phần: $[c_0, c_1]$, $[c_1, c_2]$, ..., $[c_{k-1}, c_k]$. Hàm chuyển đổi các giá trị vật lý thành giá trị mờ :

$$TR_1(x) = (x- c_0)/(c_0- c_1) + 1, \forall x \in [c_0, c_1]$$

$$TR_2(x) = 2(x- c_i)/(c_{i+1}- c_i), \forall x \in [c_i, (c_i+c_{i+1})/2]$$

$$TR_3(x) = (x- c_0)/(c_1- c_0), \forall x \in [c_{k-1}, c_k]$$

Suy xét phân lớp của các vị trí riêng biệt trên một lớp dựa trên các giá trị độ dốc của đất (các giá trị vật lý). Bốn giá trị ngữ nghĩa được sử dụng: *[phẳng, thoải, vừa phải, dốc]*. Hàm chuyển đổi tuyến tính giảm và tăng cho trường hợp đầu và cuối. Chú ý rằng phương pháp quy ước để phân lớp độ dốc bao gồm các lớp riêng rẽ với giới hạn chỉ ra khi thu thập phân lớp mờ.

Việc chuyển dần dần giữa các lớp, khi đưa ra một phương pháp tốt hơn tới việc phân loại các khái niệm mờ hồ như thoải thoải và dốc. Dựa trên phân lớp mờ 1 vị trí với độ dốc 6% được gán bằng 0.6 đối với mức bằng phẳng, 0.1 đối với thoải thoải, 0 đối với vừa phải và 0 đối với dốc đứng.

Các vị trí riêng biệt của vùng nghiên cứu có thể chỉ ra trong cách tương tự dựa trên sự ngừng lại của tiêu chuẩn đưa ra bởi ra quyết định. Đối với các ràng buộc lựa chọn vị trí tái định cư nêu ra ở trên các giá trị ngữ nghĩa có thể được suy xét:

- *Đô thị*: [đất trống, đang quy hoạch, đã quy hoạch]
- *Mức độ ẩm đất*: [khô, vừa phải, đầm lầy, nước]
- *Độ dốc nền*: [phẳng, thoải thoải, vừa phải, dốc]
- *Gần đường giao thông*: [liền kề, gần, vừa phải, xa, quá xa]
- *Hướng dốc*: [bắc, đông, nam, tây]

Tiêu chuẩn quyết định là kết hợp của nhiều hơn một lớp và giá trị ngữ nghĩa (*nền phẳng và đất khô*) độ đo tổng thể sẽ được tính và gán tới các vị trí riêng biệt. Độ đo này được đưa ra bằng cách suy xét độ thuộc trên hai hay nhiều lớp. Đối với tập mờ $A \in X$ với hàm mờ $\mu_A(x) \in [0,1]$, độ đo tổng thể có thể đưa ra bởi hàm tiềm năng theo công thức sau:

$$e(A) = \sum E[\mu_A(x)] \text{ với mọi } x \in X, \text{ ở đây } E: \mu_A[0,1] \rightarrow [0,1]$$

Một hàm như thế được sử dụng chung nhất là:

$e(A) = \sum \mu_A^q(x)$ ở đây q là số nguyên dương. Hàm như thế với giá trị trọng số lớn nó chiếm ưu thế còn với các giá trị nhỏ gần như không được đánh giá.

Với ví dụ trên, nếu có một đòi hỏi làm nổi bật các vị trí *phẳng và khô* độ đo tổng thể được đưa ra bởi: $e(\text{phẳng-khô}) = \mu_{\text{phẳng}}^2(x) + \mu_{\text{khô}}^2(x)$ cho mỗi vị trí riêng biệt x .

Lập luận dựa trên các giá trị ngữ nghĩa bao hàm các phép toán phân lớp, chống xếp và tìm kiếm cục bộ và lý thuyết logic mờ sẽ được hợp nhất trong chúng như sau:

- *Các phép toán phân lớp mờ*, gán độ thuộc cho mỗi giá trị ngữ nghĩa tới các vị trí riêng biệt trên một layer. Độ thuộc đưa ra bởi việc vận dụng hàm chuyển đổi thích hợp.

- *Các phép toán chống xếp mờ*: tính toán và gán độ đo tổng thể tới mỗi vị trí riêng biệt được đưa ra từ việc suy xét độ thuộc trên 2 hay nhiều layer. Độ đo mờ cũng đưa ra phạm vi mờ $[0,1]$.

- *Các phép toán tìm kiếm mờ*: nhận thông tin dựa trên giá trị ngưỡng xác định trước đối với các độ đo tổng thể được gán tới các vị trí riêng biệt trên một lớp.

Thủ tục lựa chọn vị trí tái định cư dựa trên tập các ràng buộc được đưa ra trong dạng ngữ nghĩa chẳng hạn (*vùng đất trống, khô, phẳng gần đường giao thông, hướng dốc nam*) có thể bao gồm các phép toán sau:

<i>Trống</i>	= Local (<i>phân lớp mờ</i>) của <i>layer</i> <i>đô thị</i>
<i>Khô</i>	= Local (<i>phân lớp mờ</i>) của <i>layer</i> <i>độ ẩm</i>
<i>Phẳng</i>	= Local (<i>phân lớp mờ</i>) của <i>layer</i> <i>độ dốc</i>
<i>Gần</i>	= Local (<i>phân lớp mờ</i>) của <i>layer</i> <i>độ dốc</i>
<i>Nam</i>	= Local (<i>phân lớp mờ</i>) của <i>layer</i> <i>lân cận giao thông</i>
<i>Vị trí tốt</i>	= Local (<i>chống xếp mờ</i>) của <i>trống, khô, phẳng, gần, nam</i>
<i>Các vị trí tốt nhất</i>	= Local (<i>tìm kiếm mờ</i>) của <i>vị trí tốt</i>

CHƯƠNG 4 - GIẢI MỘT SỐ BÀI TOÁN BẰNG ỨNG DỤNG LOGIC MỜ TRONG GIS

4.1 Tìm vị trí mở rộng thành phố Thái Bình

4.1.1 Phát biểu bài toán

Theo quyết định của thủ tướng chính phủ cho phép chuyển Thị xã Thái Bình thành Thành phố loại 2 (thành phố trực thuộc tỉnh). Các điều kiện mở rộng thành phố được các chuyên gia đô thị đưa ra như sau:

- (1). Phải liên kê thành phố (thị xã cũ) hiện tại.
- (2). Khu đất phải có độ dày địa tầng tối thiểu và không đứt gãy.
- (3). Nơi mà không bị úng lụt (trong 100 năm trở lại) (độ cao tương đối 5 m so với mặt nước biển).
- (4). Là đất nông nghiệp, không phải đất thành phố và khu công nghiệp.
- (5). Không phải là đất đã được xếp loại ưu tiên qui hoạch kinh tế trọng điểm.
- (6). Cách đường giao thông chính một khoảng cách nhất định.
- (7). Không phải là khu vực nhạy cảm về môi trường.
- (8). Yêu tiên phương án sông nằm giữa thành phố tương lai.

4.1.2 Phương pháp tiến hành

Nhận xét bài toán.

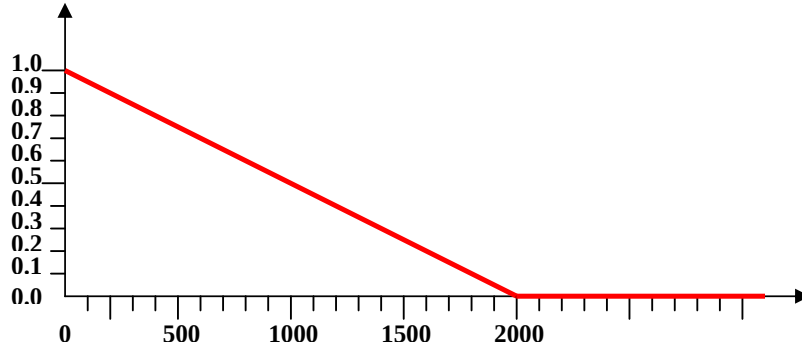
Do đặc thù địa hình thị xã Thái bình bằng phẳng và không bị úng lụt trong vòng 100 năm trở lại đây cho nên tiêu chí thứ 3 không cần phải xét tới. Điều kiện thứ 8 không cần xét đến vì kết quả sau khi phân tích việc ưu tiên được đánh giá sau cùng, điều kiện 4 và 5 có thể ghép lại thành một. Các bước tiến hành như sau:

* Lớp thông tin về đất bao gồm các loại đất sau: (*Đất thổ cư, Đất đô thị, Đất chuyên lúa, Sông hồ*). Với bốn loại đất theo dạng ngôn ngữ tự nhiên nêu trên sử dụng luật IF THEN ta sẽ gán độ thuộc như sau:

Loại đất	Giá trị mờ
Đất thổ cư	0.7
Đất đô thị	0
Đất chuyên lúa	1
Sông hồ	0

Bảng 4.1. Bảng mờ hóa lớp thông tin đất

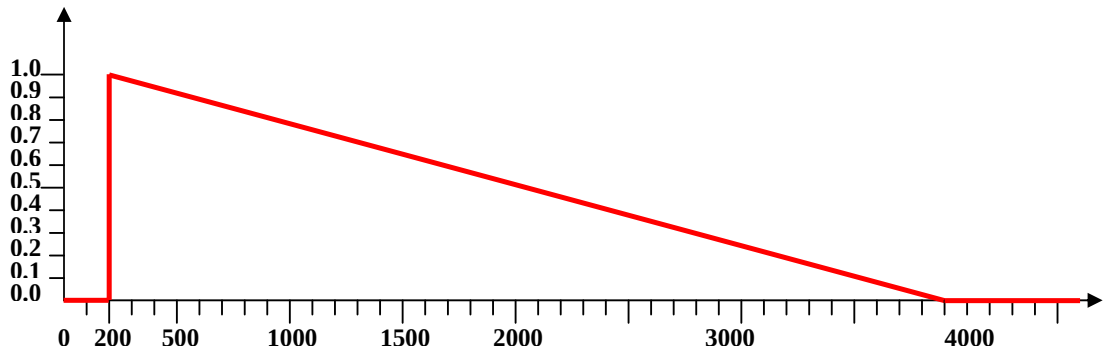
*Lớp thông tin về sự mở mang được tạo bằng cách tạo vùng đệm cho khu vực đô thị cũ. Ta sẽ tạo các vùng đệm cách nhau 100 m và sử dụng hàm mờ dạng tuyến tính giảm để tính giá trị mờ cho mỗi vùng đệm được tính. Hàm mờ được sử dụng để mờ hoá như sau:



Hình 4.1. Hàm mờ sử dụng lớp thông tin mở mang

$$\text{Hàm mờ sử dụng } \mu_{\text{mở mang}}(x) = \begin{cases} (2000 - x) / 2000 & 0 \leq x \leq 2000 \\ 0 & x > 2000 \end{cases}$$

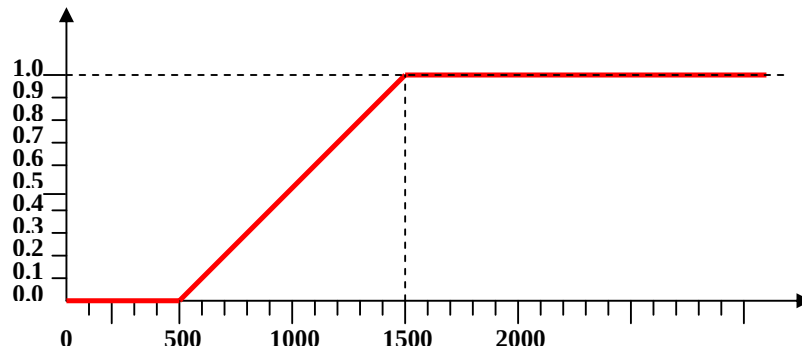
*Lớp thông tin về giao thông theo quy định phải có hành lang giao thông của các tuyến đường. Do đó các giá trị trong phạm vi 200 m về mỗi bên của mỗi tuyến đường là ranh giới phân định không cho phép. Hàm mờ tuyến tính giảm được xác định như sau:



Hình 4.2. Hàm mờ sử dụng cho lớp thông tin giao thông

$$\text{Hàm mờ sử dụng } \mu_{\text{giao thông}}(l) = \begin{cases} (4000 - x)/3800 & 200 \leq x \leq 4000 \\ 0 & x > 4000; 0 \leq x < 200 \end{cases}$$

* Lớp thông tin về ô nhiễm do ảnh hưởng của nhà máy gạch sự ô nhiễm phụ thuộc vào khoảng cách từ nhà máy gạch tới khu mở rộng Thành phố. Các vị trí càng gần nhà máy độ ảnh hưởng càng cao, càng ở xa sự ô nhiễm càng giảm. Do đó sử dụng hàm mờ tuyến tính tăng để xác định sự ảnh hưởng của các vị trí mở rộng thành phố. Các vị trí ở xa sự ảnh hưởng càng thấp khi đó độ thuộc càng cao. Các vị trí ở gần độ thuộc càng nhỏ hàm tuyến tính sau được sử dụng. Theo kinh nghiệm chuyên gia vùng bị ảnh hưởng nhiều nhất trong vòng bán kính 500 m và vùng ngoài vùng ảnh hưởng là 1500 m. Hàm mờ sử dụng có dạng sau.



Hình 4.3. Hàm mờ sử dụng cho lớp thông tin ô nhiễm

$$\text{Hàm mờ sử dụng } \mu_{\text{ô nhiễm}}(l) = \begin{cases} (x - 500)/1000 & 500 \leq x \leq 1500 \\ 1 & x > 1500 \\ 0 & x < 500 \end{cases}$$

* Đối với lớp thông tin về địa chất có hai loại theo ký hiệu địa chất vùng có độ dày thích hợp và vùng đất yếu không phù hợp cho việc phát triển các khu cao tầng ở đây ta có thể xác định hai loại giá trị (1 cho vùng đất có độ dày bền vững và 0 cho vùng đất yếu.

Địa tầng	Giá trị mờ
amQ...-† TM š	1
aQ...-† TM š	0

Bảng 4.2. Bảng mờ hóa lớp thông tin địa tầng

Thực hiện chồng xếp 5 lớp thông tin trên sử dụng công thức:

$$\mu E(l) = \sum_{i=1}^k [\mu A_i(l)]^q \text{ . Lấy } q = 2 \text{ Ta có:}$$

$$\mu E(l) = \{ [\mu_{\text{thổ cư}}(l)]^2 + [\mu_{\text{mở mang}}(l)]^2 + [\mu_{\text{giao thông}}(l)]^2 + [\mu_{\text{địa tầng}}(l)]^2 + [\mu_{\text{ô nhiễm}}(l)]^2 \} / 5;$$

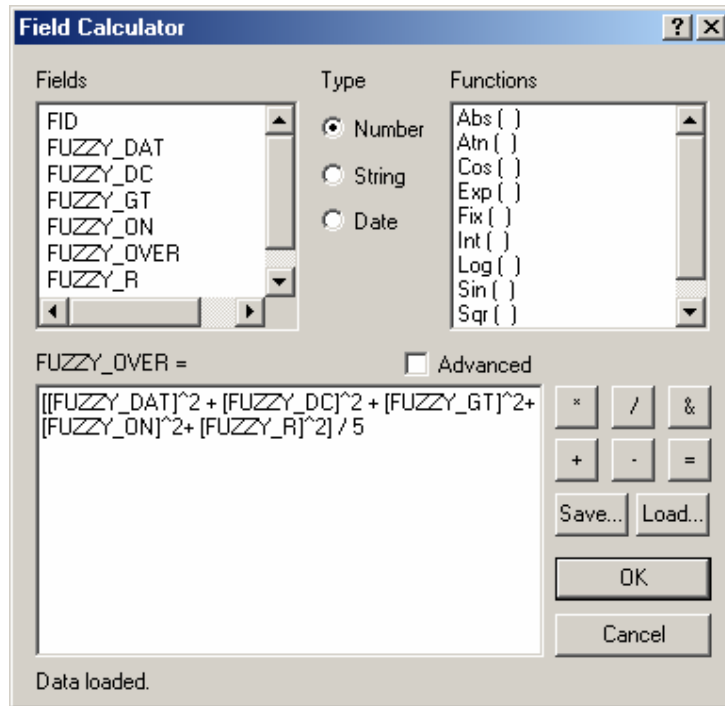
ở đây ta có 5 lớp tham gia chồng xếp để bảo đảm giá trị sau khi tính toán vẫn nằm trong khoảng $[0,1]$ ta nhân với $1/5 = 0.25$ như là trọng số ngang bằng cho năm lớp nêu trên. μE được gọi là độ đo tổng thể từ 5 lớp tham gia trong quá trình chồng xếp.

Sau khi chồng xếp mỗi đối tượng của bản đồ kết quả bao gồm các trường:

FUZZY_DAT, FUZZY_DC, FUZZY_GT, FUZZY_ON, FUZZY_R

Giá trị kết quả được tính và cập nhật trên trường FUZZY_OVER theo công thức sau đây:

$$\text{FUZZY_OVER} = (\text{FUZZY_DAT}^2 + \text{FUZZY_DC}^2 + \text{FUZZY_GT}^2 + \text{FUZZY_ON}^2 + \text{FUZZY_R}^2) / 5.$$



Hình 4.4. Phương trình chồng xếp mờ tính toán trên các trường

Attributes of Dat_DatYeu_ONhiem_Gthong_Rg

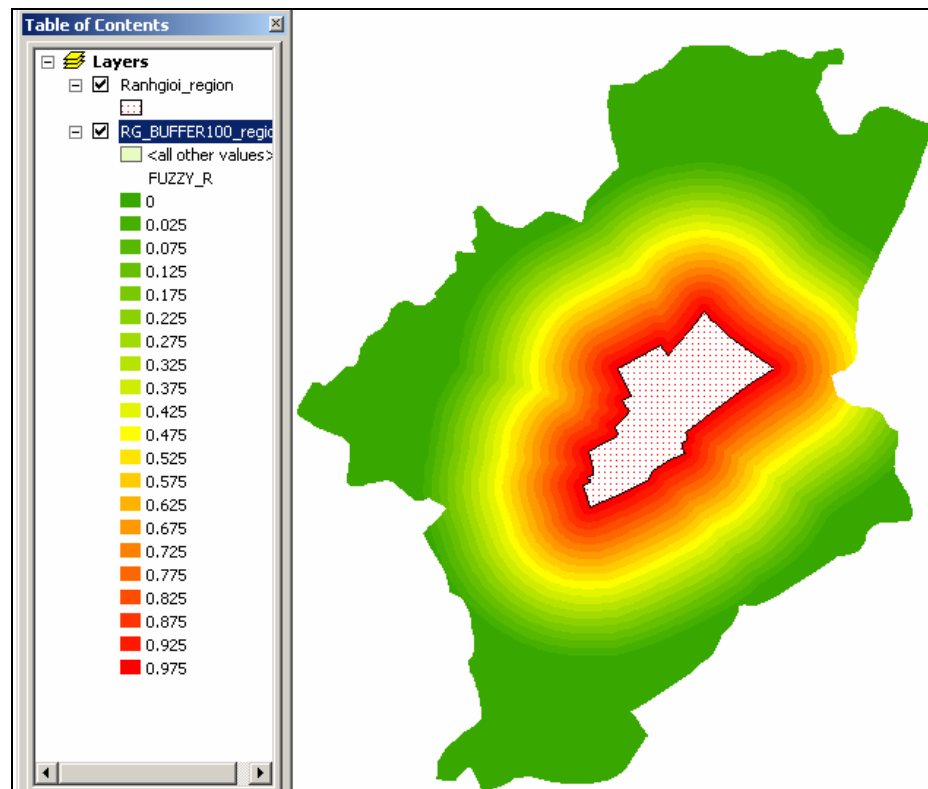
FID	FUZZY_DAT	FUZZY_DC	FUZZY_ON	FUZZY_GT	FUZZY_R	FUZZY_OVE
2064	1	1	1	0.8553	0	0.7463
2065	1	1	1	0.8553	0	0.7463
2066	1	1	1	0.8553	0	0.7463
2067	1	1	1	0.8816	0	0.7554
2068	1	1	1	0.8816	0	0.7554
2069	1	1	1	0.8816	0	0.7554
2070	1	1	1	0.8816	0	0.7554
2071	1	1	1	0.8816	0	0.7554
2072	1	1	1	0.8816	0	0.7554
2073	1	1	1	0.8816	0	0.7554
2074	1	1	1	0.8816	0	0.7554
2075	1	1	1	0.8816	0	0.7554
2076	1	1	1	0.9079	0	0.7649
2077	1	1	1	0.9079	0	0.7649
2078	1	1	1	0.9079	0	0.7649
2079	1	1	1	0.9079	0	0.7649
2080	1	1	1	0.9079	0	0.7649

Record: 0 Show: All Selected Records (0 out of 6044 Selected.)

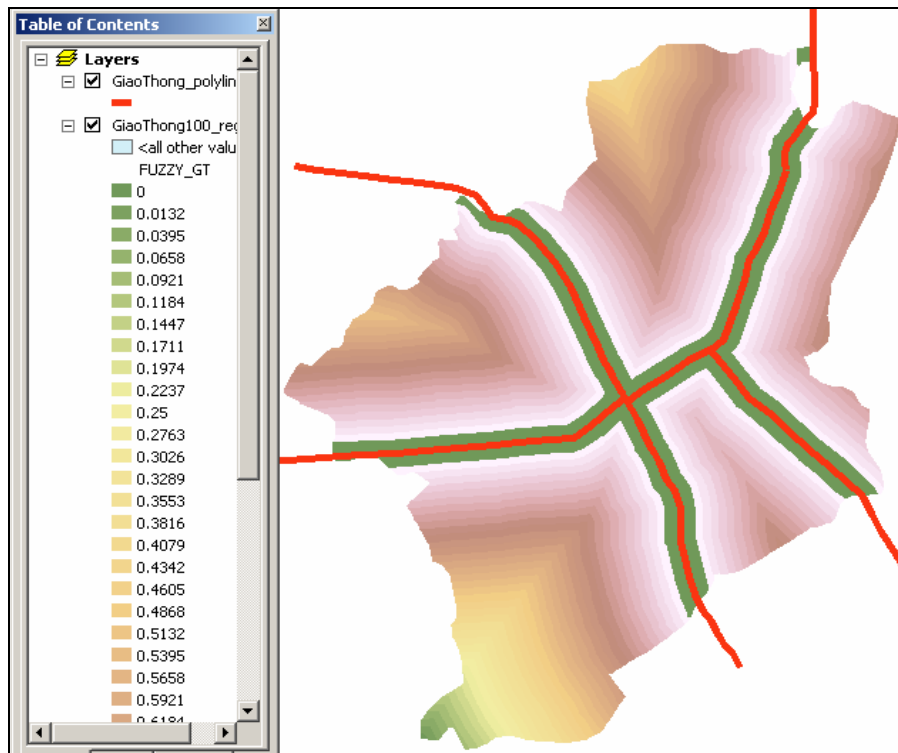
Hình 4.5. Thuộc tính sau khi chồng xếp

4.1.3 Kết quả đạt được

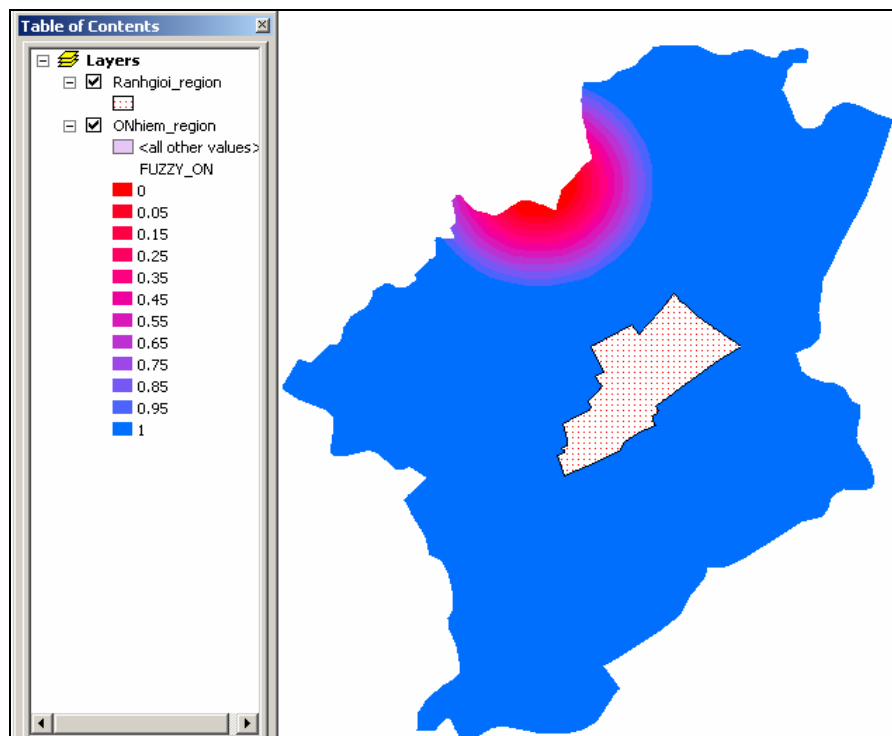
Mờ hóa



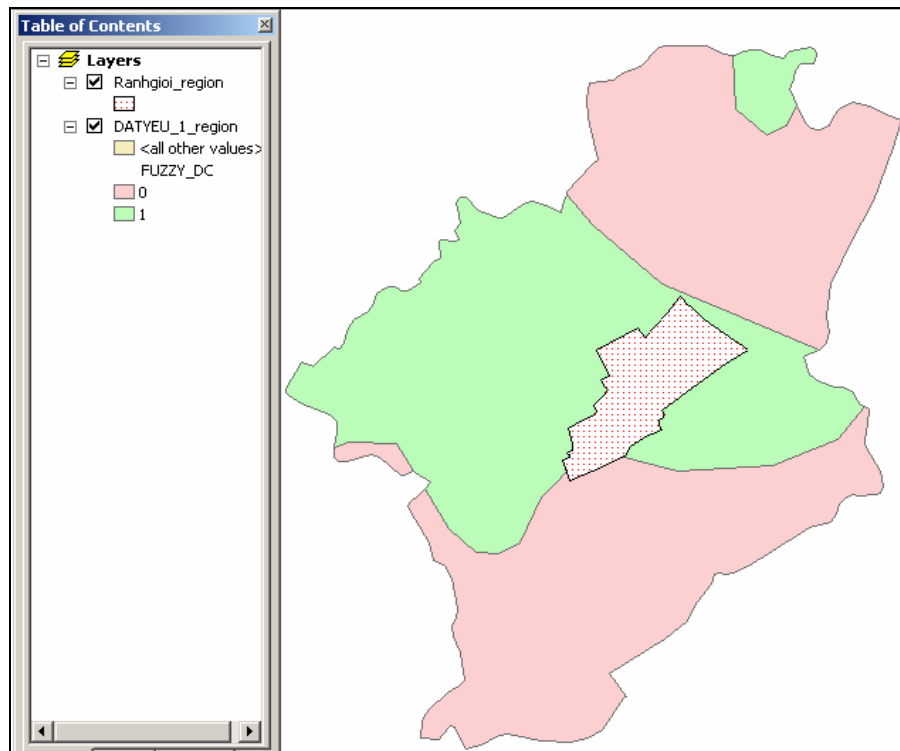
Hình 4.6. Vùng đệm mờ hóa lớp thông tin mở mang thành phố



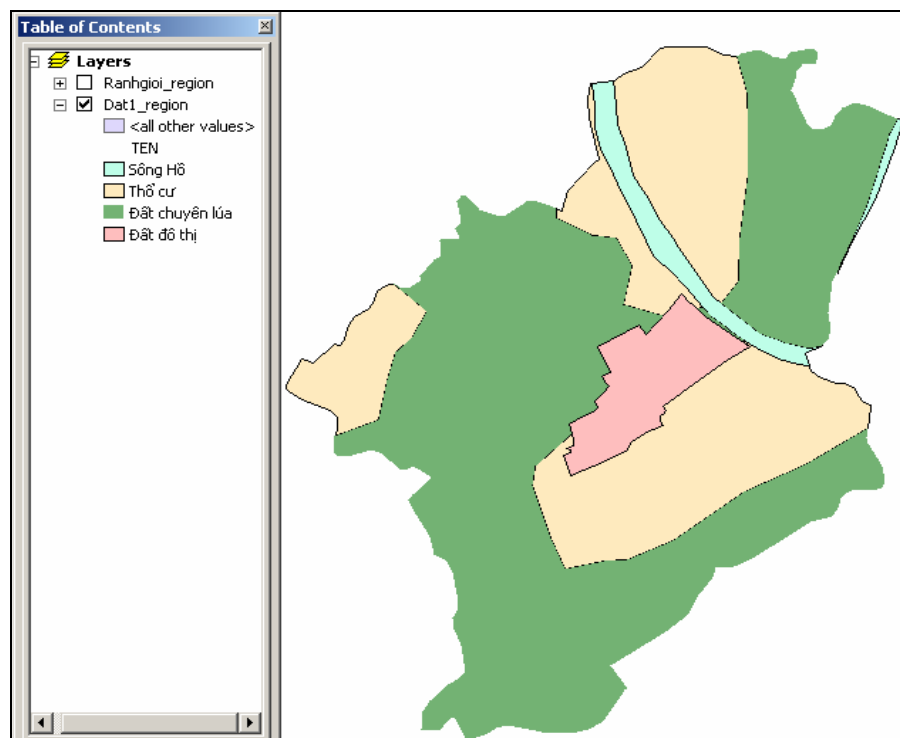
Hình 4.7. Vùng đệm mờ hóa về lớp thông tin giao thông



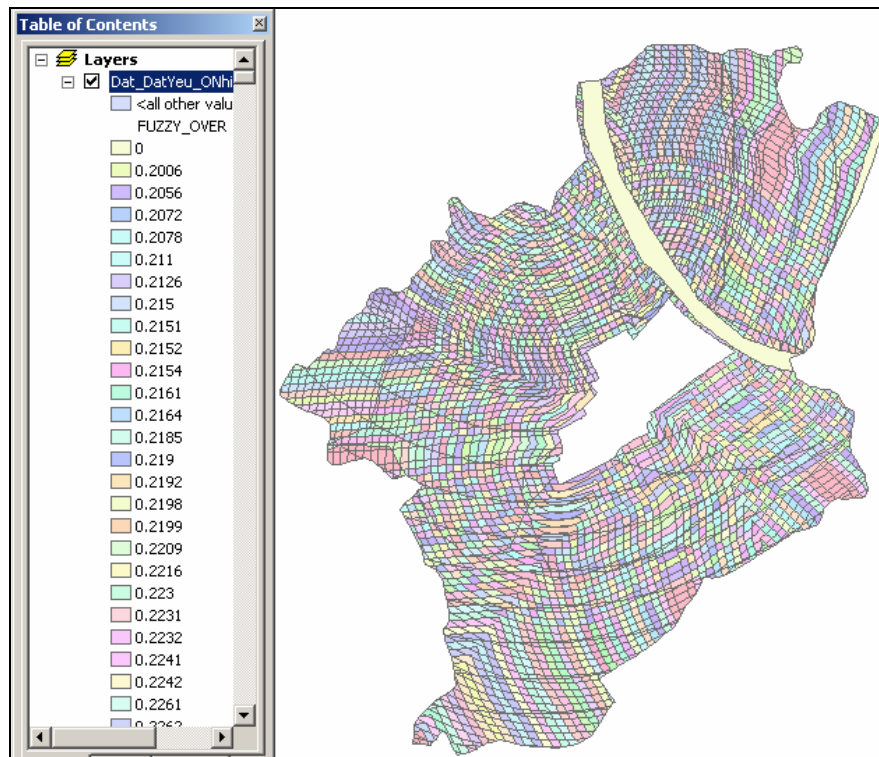
Hình 4.8. Vùng đệm mờ hóa lớp thông tin ô nhiễm



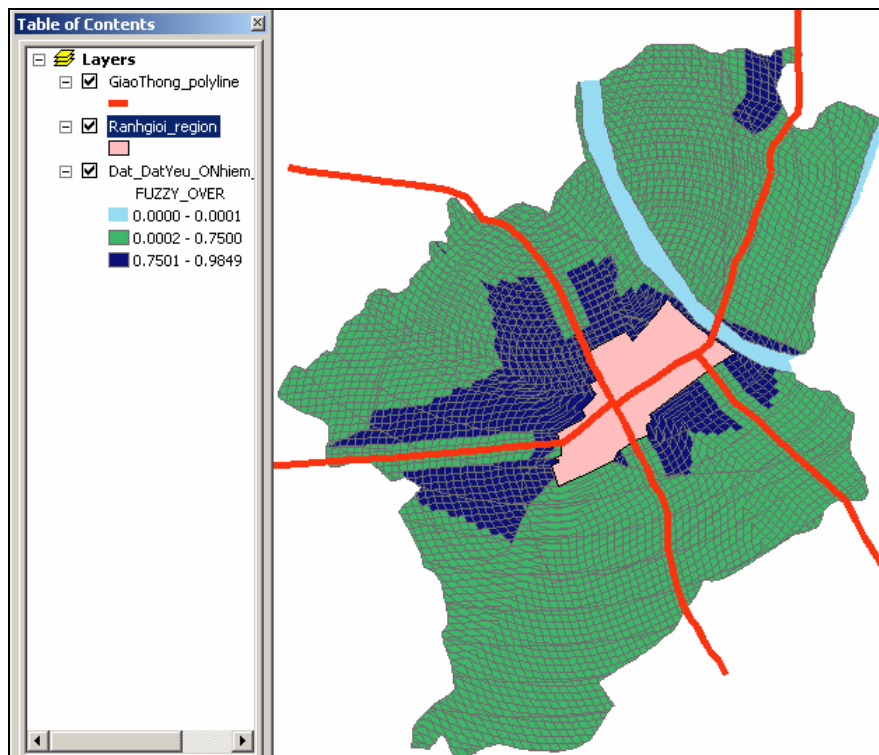
Hình 4.9. Mờ hóa lớp thông tin địa tầng đất yếu



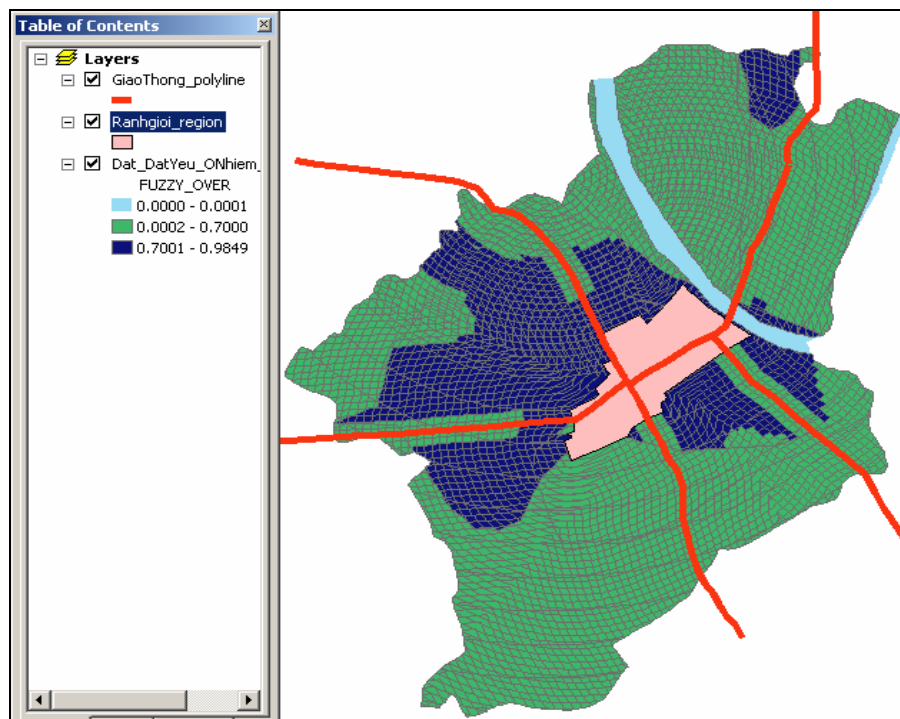
Hình 4.10. Mờ hóa lớp thông tin hiện trạng sử dụng đất



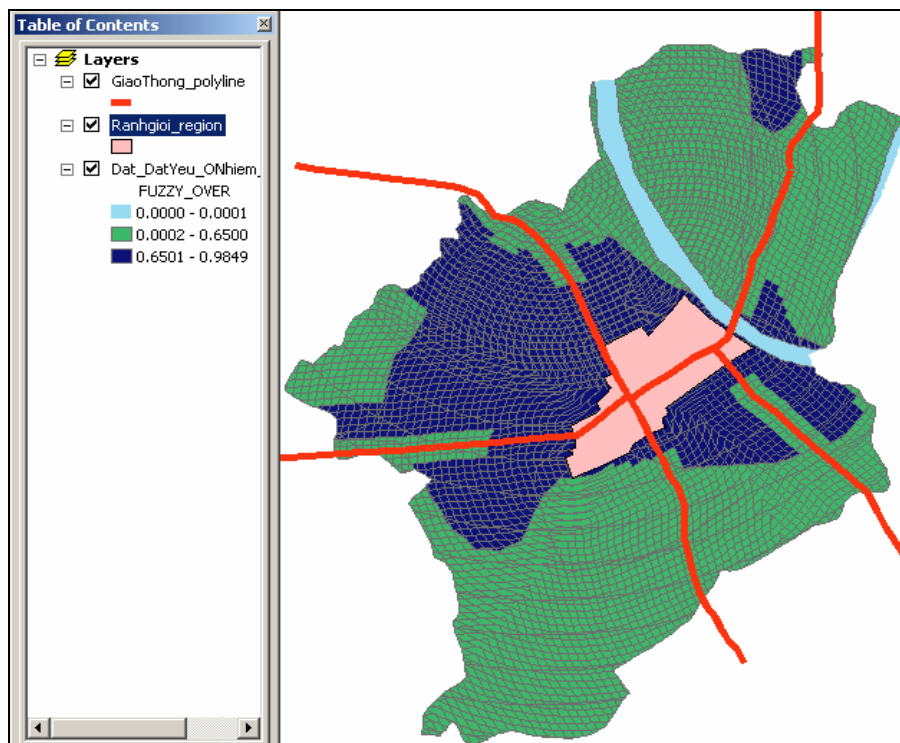
Hình 4.11. Kết quả sau khi chồng xếp



Hình 4.12. Giải mờ lát cắt $\alpha = 0.75$



Hình 4.13. Giải mờ lát cắt $\alpha = 0.7$



Hình 4.14. Giải mờ lát cắt $\alpha = 0.65$

Qua ba lựa chọn giải mờ trên vùng xanh đậm là các vị trí để có thể quy hoạch cho việc mở rộng thành phố. Dựa trên bản đồ trên mà các chuyên gia có thể lựa chọn các phương án cần thiết cho việc ra quyết định khu đất mở rộng thành phố.

4.2 Bài toán xác định đường đi ngắn nhất sử dụng logic mờ

4.2.1 Phát biểu bài toán

Một trong các công cụ sử dụng thường xuyên trong việc thu thập dữ liệu địa lý được triu tượng hoá trong GIS là các loại đồ thị khác nhau và các thay đổi được tạo ra với dự định sử dụng của chúng. Lý thuyết đồ thị có nhiều ứng dụng khác nhau trong phân tích hệ thống, kinh tế và giao thông vận tải. Trong nhiều trường hợp chúng ta phải sử dụng dữ liệu không rõ ràng mà chúng ta không thể suy xét chúng trong các tính toán khi sử dụng đồ thị bình thường. Logic mờ và lý thuyết đồ thị mờ cho chúng ta một công cụ thích hợp để sử dụng trong các trường hợp đó.

4.2.2 Phương pháp tiến hành

Xét một đồ thị mờ G với kiểu thuần chủng V mờ. Cho Π là tập tất cả các đường đi từ đỉnh v_a tới đỉnh v_b và cho chiều dài mờ của đường đi là :

$$l_p = \text{length}(P) = \sum_{e_k \in P} w_p, \text{ trong đó } P \in \Pi \text{ ở đây } e_k \text{ là các cạnh của } G.$$

Tập mờ của các đường đi ngắn nhất là tập mờ S trên Π với các thành viên π_S được đưa vào bởi :

$$\pi_S(P) = \min \{ \mu_{l_p \leq Q} \}, \text{ Trong đó } P \in \Pi, Q \in \Pi$$

Tính hỗ trợ bao gồm tất cả các đường đi mà có khả năng có chiều dài nhỏ nhất:

$$\text{supp}(S) = \{ P \in \Pi \mid \mu_{l_p \leq Q} > 0, \forall Q \in \Pi \}$$

Tập mờ của các đường đi ngắn nhất định nghĩa trên có thể thu lại thành tập mờ đường đi ngắn nhất, ở đó mỗi cạnh e_i có thành viên trong tập mờ S' :

$$\mu_{S,i} = \max_{e_i \in P, P \in \Pi} \{ \pi_S(P) \}, \text{ for } i = 1, \dots, n_E$$

Thuật toán FSA:

Bước 1:

Xây dựng đồ thị $\overline{G}$ và $\underline{G}$ đồng nhất với G và trọng số trên các cạnh của $\overline{G}$ và $\underline{G}$ có thể tính như sau:

$$\text{Đối với } \overline{G} : \overline{\tau}_a = \sup\{\text{supp}(\tau_a)\}$$

$$\text{Đối với } \underline{G} : \underline{\tau}_a = \inf\{\text{supp}(\tau_a)\}$$

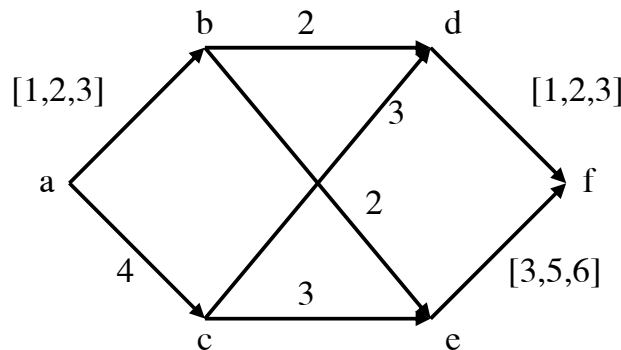
Bước 2:

Tìm đường đi ngắn nhất p từ v_a tới v_b trong $\overline{G}$. Đây là vấn đề đường đi ngắn nhất kinh điển và nhiều thuật toán tốt có thể sử dụng để giải nó. Biểu thị k là chiều dài của đường p .

$$\mu_{S,i} = \min \{ l_p \} \mid P \in \Pi \quad (25)$$

Bước 3:

Cho $\underline{S}$ là tập tất cả các đường đi từ V_a tới V_b trong $\underline{G}$, mà chiều dài nhỏ hơn k . Cho S là tập tất cả các đường đi trong G . Hình dạng của các đường đi trong cả S và $\underline{S}$ là đúng. Như thế S , là tập của tất cả các đường đi ngắn nhất mờ. Cuối cùng tính độ mờ cho mỗi đường đi từ S trong sự suy xét của k .

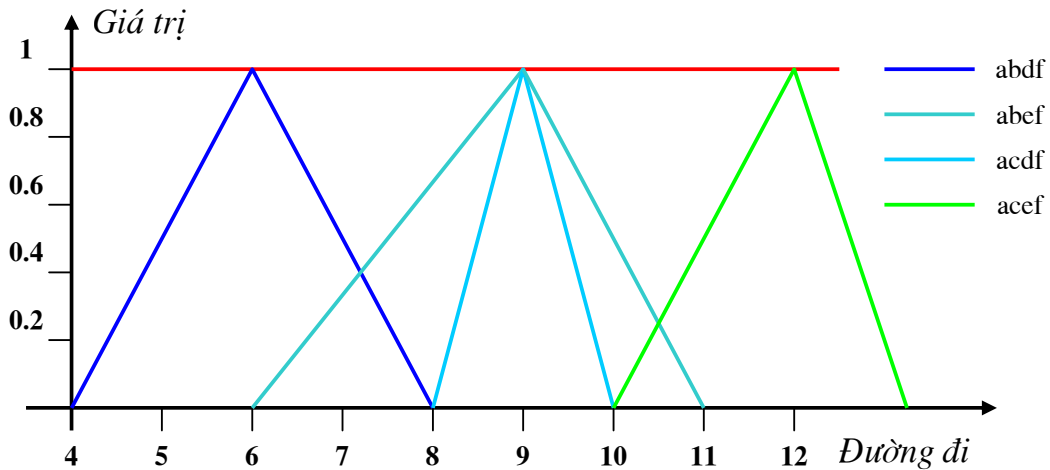


Hình 4.15. Đồ thị G có hướng V- mờ

4.2.3 Kết quả đạt được

Hình trên chỉ ra kiểu trọng số đồ thị mờ V. Đỉnh a là điểm khởi hành và đỉnh f là điểm đến của đường đi. Các trọng số có thể là các số cứng của chúng hoặc các số tam giác mờ.

Các chiều dài mờ đối với 4 đường đi từ đỉnh a tới đỉnh f được liệt kê trong hình trên - từ điều này chúng ta thấy rằng $k=8$ và đường đi abdf có giá trị mờ $\pi S(abdf)=1$, đường đi abef có giá trị mờ $\pi S(abef)=2/5$, và các đường khác có giá trị mờ $\pi S(acdf) = \pi S(acef) = 0$ trong tập mờ các đường đi ngắn nhất. Hình sau đây minh họa đường đi ngắn nhất mờ.



Hình 4.16. Đường đi ngắn nhất mờ của đồ thị mờ G

4.3 Bài toán tìm vị trí xây dựng nhà máy xi măng

4.3.1 Phát biểu bài toán

Tỉnh Quảng Ninh là tỉnh giàu tiềm năng về Công nghiệp khai thác mỏ và Du lịch. Do sự phân bố về mỏ và các loại tài nguyên thiên nhiên khác. Ba huyện Hoành Bồ, Ba chẽ, Yên hưng là các huyện có tỷ trọng về công nghiệp khai thác thấp mà tiềm năng của huyện này đa dạng và phong phú. Lãnh đạo tỉnh muốn phát đầu tư và xây dựng nhà máy xi măng tại cụm 3 huyện trên với mục đích sử dụng các nguồn nguyên liệu tại chỗ như đất sét, than, đá vôi...và nguồn nhân lực tại chỗ; nhưng cũng đặc biệt tới vấn đề bảo vệ môi trường vịnh

Hạ Long di sản thiên nhiên thế giới. Theo ý kiến của chuyên gia tiêu chí để chọn vị trí xây dựng nhà máy gồm:

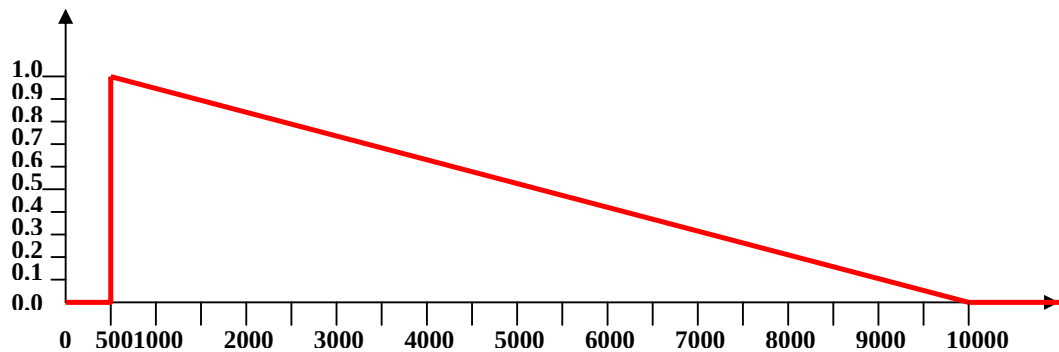
- (1). Gần với các mỏ than để hình thành khu công nghiệp liên hoàn khai thác than và sản xuất xi măng.
- (2). Gần các khu vực mỏ đất sét nguyên liệu để sản xuất xi măng.
- (3). Gần mỏ đá vôi nguyên liệu để sản xuất xi măng.
- (4). Gần cảng biển để thuận lợi cho việc bốc dỡ hàng hóa.
- (5). Cách đường giao thông chính một khoảng nhất định vừa bảo đảm vận chuyển và không ảnh hưởng tới môi trường giao thông.
- (6). Cách Vịnh Hạ Long một khoảng nhất định để không bị ảnh hưởng ô nhiễm tới môi trường vịnh Hạ Long.

4.3.2 Phương pháp tiến hành

Nhận xét bài toán.

Do đặc thù địa hình khu vực Hoàn bì, Ba chẽ, Yên hưng là các huyện chưa phát triển về mặt công nghiệp và đô thị cho nên các tiêu chí ảnh hưởng của vùng đất quy hoạch đô thị hầu như không có. Các bước tiến hành như sau:

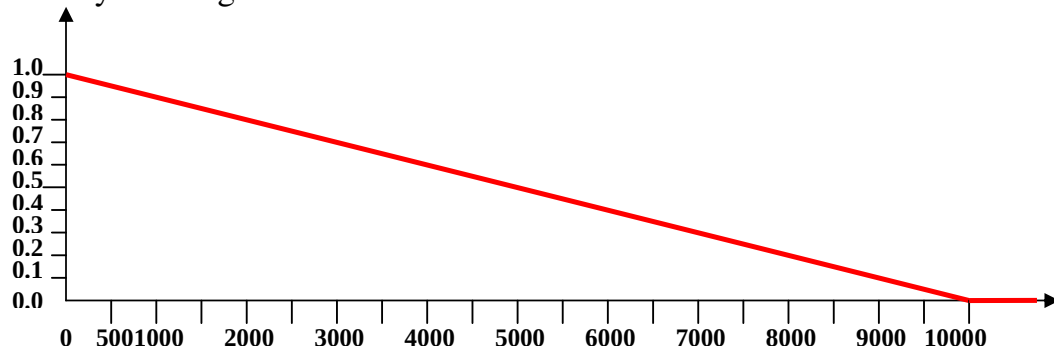
*Lớp thông tin về tính gần các mỏ than được tạo thành các vùng đệm bao quanh các vị trí mỏ than theo các khoảng cách 500 m. Để thuận lợi các nhà máy nên cách xa trên 500 m để thuận lợi cho việc khai thác than; các phạm vi trong vòng 10.000 m thuận lợi cho việc vận chuyển bằng các loại xe vận tải. Hàm mờ được sử dụng để mờ hoá là hàm tuyến tính giảm như sau:



Hình 4.17. Hàm mờ sử dụng lớp thông tin gần mỏ than

$$\text{Hàm mờ sử dụng } \mu_{\text{mỏ than}}(I) = \begin{cases} (10000 - x)/9500 & 500 \leq x \leq 10000 \\ 0 & x > 10000; x < 500 \end{cases}$$

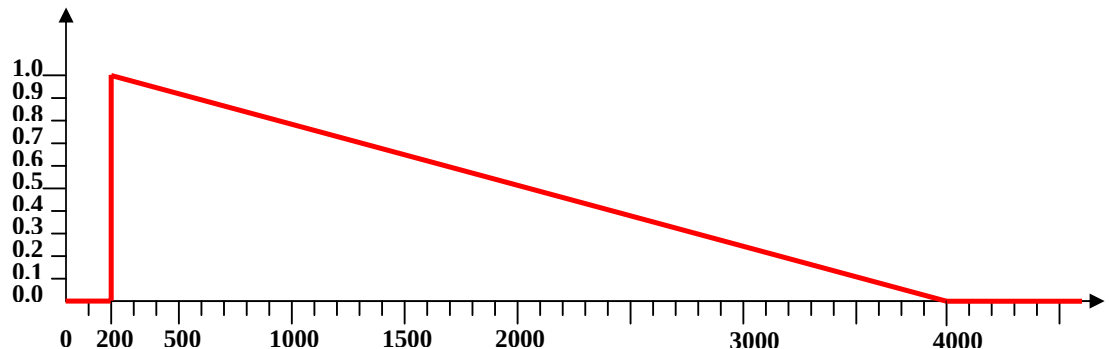
*Lớp thông tin về tính gần các mỏ đất sét được tạo thành các vùng đệm bao quanh các vị trí mỏ đất sét theo các khoảng cách 500 m. Do việc khai thác đất sét có thể khai thác nguyên liệu tại chỗ, và cũng có thể vận chuyển trong vòng bán kính 10,000 m bằng xe vận tải. Hàm mờ được sử dụng để mờ hoá là hàm tuyến tính giảm như sau:



Hình 4.18. Hàm mờ sử dụng lớp thông tin gần mỏ đất sét

$$\text{Hàm mờ sử dụng } \mu_{\text{mỏ đất sét}}(I) = \begin{cases} (10000 - x)/10000 & 0 \leq x \leq 10000 \\ 0 & x > 10000 \end{cases}$$

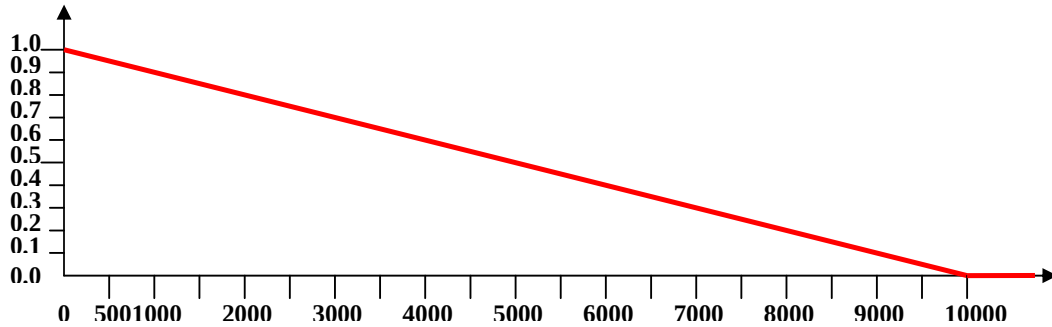
*Lớp thông tin về giao thông theo quy định phải có hành lang giao thông của các tuyến đường (hành lang 200 m). Đối với công nghiệp sản xuất chủ yếu là cơ giới hoá ranh giới phân định tính tới 500 m. Hàm mờ tuyến tính giảm được xác định như sau:



Hình 4.19. Hàm mờ sử dụng cho lớp thông tin giao thông

$$\text{Hàm mờ sử dụng } \mu_{\text{giao thông}}(l) = \begin{cases} (4000 - x)/3800 & 200 \leq x \leq 4000 \\ 0 & x > 4000; 0 \leq x < 200 \end{cases}$$

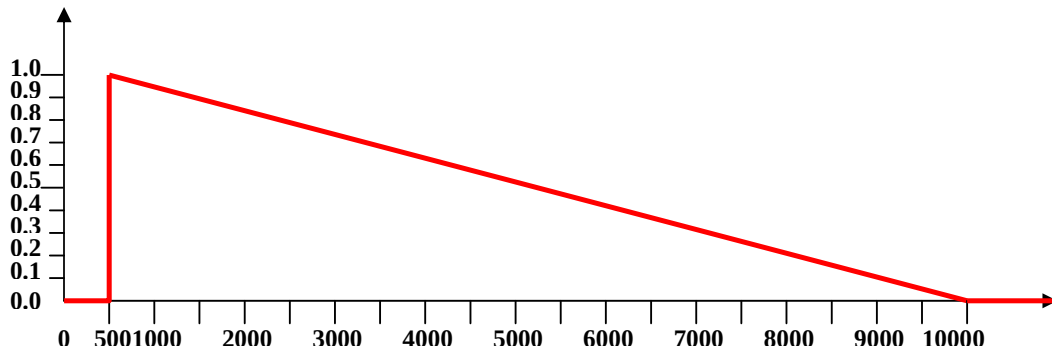
*Lớp thông tin về tính gần các mỏ đá vôi được tạo thành các vùng đệm bao quanh các vị trí mỏ đá vôi theo các khoảng cách 500 m. Do việc khai thác đá vôi có thể khai thác nguyên liệu tại chỗ, và cũng có thể vận chuyển trong vòng bán kính 10,000 m bằng xe vận tải. Hàm mờ được sử dụng để mờ hoá là hàm tuyến tính giảm như sau:



Hình 4.20. Hàm mờ sử dụng lớp thông tin gần mỏ đá vôi

$$\text{Hàm mờ sử dụng } \mu_{\text{mỏ đá vôi}}(l) = \begin{cases} (10000 - x)/10000 & 0 \leq x \leq 10000 \\ 0 & x > 10000; \end{cases}$$

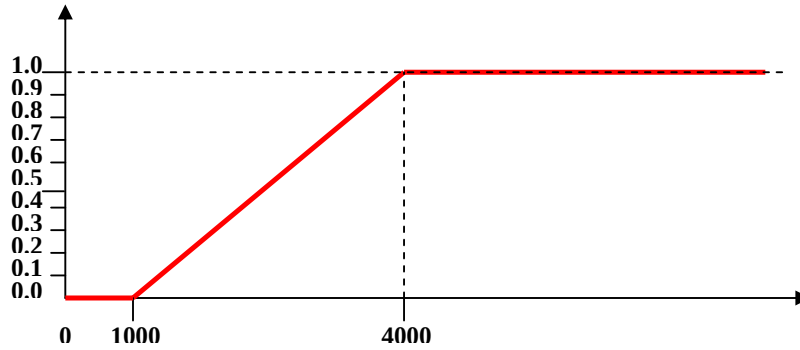
*Lớp thông tin về tính gần các cảng được tạo thành các vùng đệm bao quanh các vị trí cảng theo các khoảng cách 500 m. Các vị trí cảng gần cảng càng tốt. Tuy nhiên để thuận lợi cho việc bốc dỡ hàng hóa các nhà máy nên cách xa trên 500 m để thuận lợi cho việc tiêu thụ hàng hóa và khai thác khu cảng. Hàm mờ được sử dụng để mờ hoá là hàm tuyến tính giảm như sau:



Hình 4.21. Hàm mờ sử dụng lớp thông tin gần cảng

$$\text{Hàm mờ sử dụng } \mu_{\text{gần_cảng}}(l) = \begin{cases} (10000 - x)/9500 & 500 \leq x \leq 10000 \\ 0 & x > 10000; x < 500 \end{cases}$$

* Lớp thông tin về ô nhiễm có thể gây ra do nhà máy đối với vịnh Hạ Long. Hàm mờ tuyến tính giảm sử dụng có dạng sau.



Hình 4.22. Hàm mờ sử dụng cho lớp thông tin ô nhiễm

$$\text{Hàm mờ sử dụng } \mu_{\text{ô_nhiễm}}(l) = \begin{cases} (x - 1000)/3000 & 1000 \leq x \leq 4000 \\ 1 & x > 4000 \\ 0 & x < 1000 \end{cases}$$

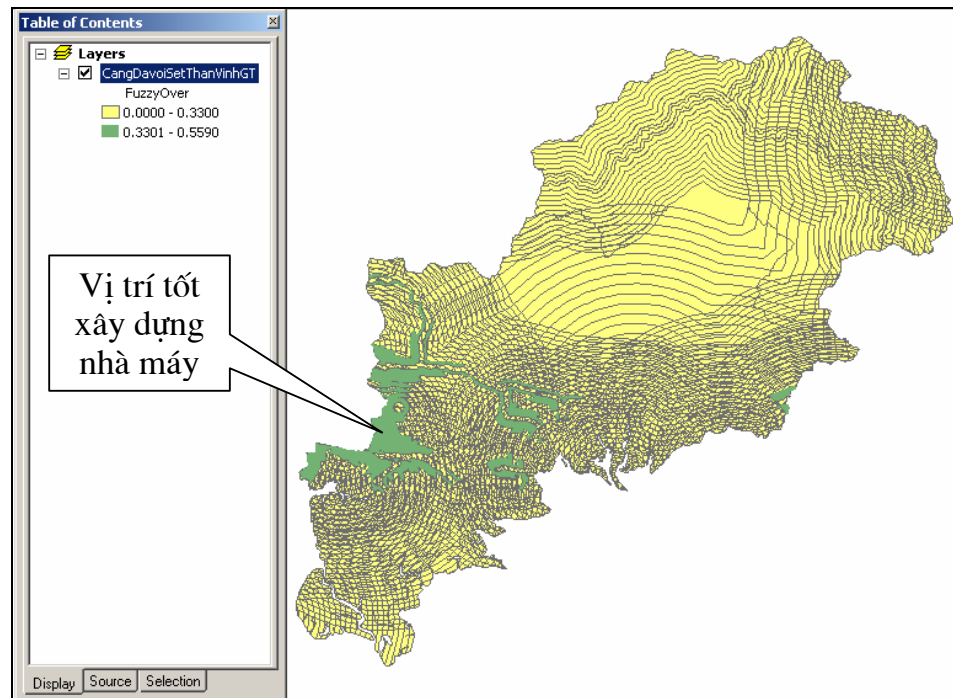
Thực hiện chồng xếp 6 lớp thông tin trên sử dụng công thức:

$$\mu_E(l) = \sum_{i=1}^k [\mu_{A_i}(l)]^q. \text{ Lấy } q = 2 \text{ Ta có: } \mu_E(l) = \{[\mu_{\text{mỏ_than}}(l)]^2 + [\mu_{\text{mỏ_sét}}(l)]^2 + [\mu_{\text{mỏ_đá}}$$

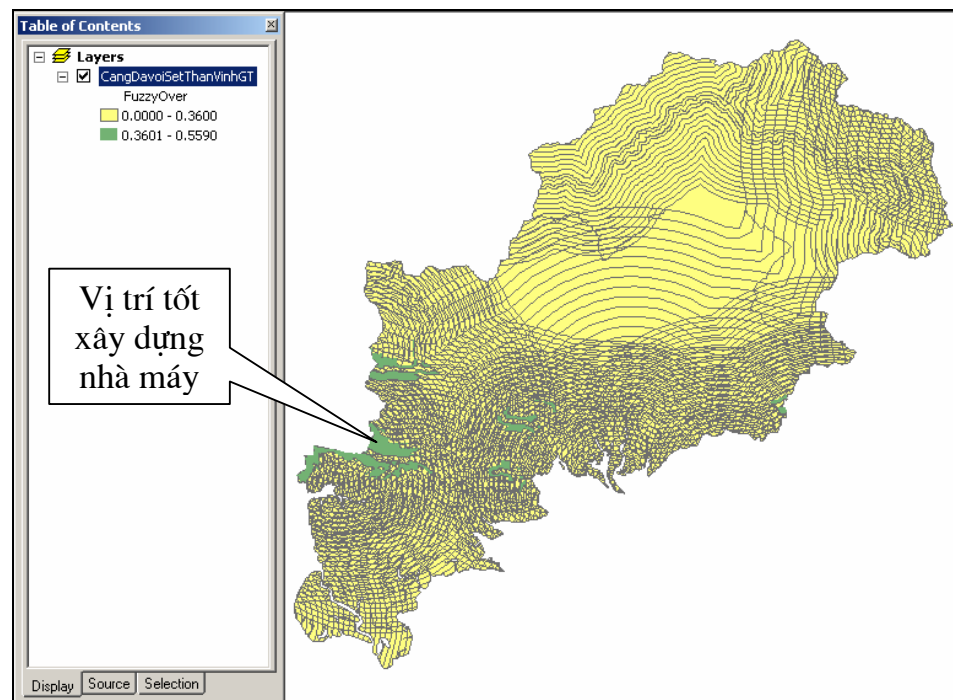
$\text{vôi}(l)]^2 + [\mu_{\text{giao_thông}}(l)]^2 + [\mu_{\text{gần_cảng}}(l)]^2 + [\mu_{\text{ô_nhiễm}}(l)]^2\}/6$; ở đây ta có 6 lớp tham gia chồng xếp để bảo đảm giá trị sau khi tính toán vẫn nằm trong khoảng $[0,1]$ ta nhân với $1/6$ như là trọng số ngang bằng cho sáu lớp nêu trên. μ_E được gọi là độ đo tổng thể từ 6 lớp tham gia trong quá trình chồng xếp.

4.3.3 Kết quả đạt được

Sau khi thực hiện chồng xếp và thực hiện giải mờ với lát cắt $\alpha = 0.33$ và $\alpha = 0.36$ ta nhận được các vùng xanh đậm có thể là vị trí để xây dựng nhà máy xi măng. Tuy nhiên dựa trên kết quả nhận được các vị trí chỉ ra dưới đây có thể sử dụng để xây dựng nhà máy.



Hình 4.23. Giải mờ với lát cắt $\alpha = 0.33$



Hình 4.24. Giải mờ lấy lát cắt $\alpha = 0.36$

KẾT LUẬN

Lý thuyết tập mờ được xem như là phương tiện thiết kế các công cụ một cách hiệu quả để hỗ trợ các xử lý ra quyết định đối với các bài toán không gian mà đặc thù của nó là dữ liệu không rõ ràng. Trong luận văn này đã nghiên cứu sự hợp nhất của lý thuyết tập mờ trong hệ thống cơ sở dữ liệu quan hệ GIS và ứng dụng thành quả nghiên cứu vào thực tiễn mà điển hình là bài toán mở rộng Thành phố Thái Bình.

Nó chỉ ra sự hiệu quả như thế nào của lý thuyết tập mờ để có thể thực hiện các diễn tả và phân tích dữ liệu địa lý. Ở đó các đặc trưng không rõ ràng là khái niệm cần xử lý. Sự đóng góp của luận văn có thể được tóm tắt như sau: Thứ nhất giới thiệu ngắn gọn về hệ thống thông tin địa lý các tiến bộ và lịch sử phát triển của nó, các khuynh hướng phát triển của các hệ thống thông tin địa lý, trong đó logic mờ là một hướng phát triển có triển vọng trong tương lai. Thứ hai phân tích tính không rõ ràng, không chắc chắn và mập mờ của dữ liệu trong các hệ thống thông tin địa lý và các giới hạn khi thực hiện với lý thuyết tập hợp kinh điển trong cả diễn tả và phân tích dữ liệu địa lý, thay thế nó bằng lý thuyết tập mờ. Để có thể tăng cường lý thuyết tập mờ vào trong các hệ thống thông tin địa lý cần thiết phải mở rộng mô hình dữ liệu không gian tổng thể để thích hợp với sự không rõ ràng, không chắc chắn của các thực thể địa lý. Sau khi đã mở rộng mô hình dữ liệu không gian, các phép toán trong nó cũng được mở rộng để hỗ trợ các lập luận không gian mờ. Trong phần thực nghiệm tác giả giải bài toán quy hoạch mở rộng thành phố Thái Bình. Đây là một ứng dụng rất có ý nghĩa trong tiến trình công nghiệp hóa, hiện đại hóa đất nước. Bài toán mở rộng Thành phố Thái Bình là mô hình ứng dụng tiêu biểu có thể áp dụng cho các thành phố tương tự khác. Điều đó khẳng định rằng việc mở rộng và tăng cường lý thuyết tập mờ trong GIS là hướng đi đúng và thực tế, nó trang bị cho các nhà quy hoạch các công cụ mềm dẻo để giải quyết các vấn đề không gian phức tạp khi dữ liệu và thông tin trong chúng là mờ.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Nguyễn Cát Hồ, Lý thuyết tập mờ và công nghệ tính toán mềm , hệ mờ, mạng nơron và ứng dụng, nhà xuất bản khoa học kỹ thuật.
2. Trần Đình Khang, Xây dựng hàm đo trên đại số gia tử và ứng dụng trong lập luận ngôn ngữ, tạp chí Tin học và điều khiển học (1997).
3. Trần Đình Khang, Tích hợp các đại số gia tử cho suy luận ngôn ngữ, tạp chí Tin học và Điều khiển học (1997).
4. Nguyễn Thanh Thủy, Hồ Cẩm Hà, Đại số quan hệ và nguyên lý xử lý câu hỏi trên một mô hình cơ sở dữ liệu mờ, Hội nghị khoa học 19 Trường Đại học Bách khoa Hà Nội.(2001).

Tiếng Anh

5. Robert Steiner, Fuzzy Logic in GIS.
6. Wolfgang Kainz, Introduction to FuzzyLogic and Applications in GIS.
7. Graeme F.Bonham - Carter, Geographic Infomation systems for Geoscientists, Modeling with GIS.
8. Altman, D. Fuzzy set theoretic approaches for handling imprecision in spatial analysis.
9. Emmanuel Stefanakis and Timos Sellis. Enhancing a Database Management System for GIS with Fuzzy Set Methodologies.
10. Michael F.Goodchild and Karen K.Kemp(1990), Technical Issues In GIS.