

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ



HOÀNG XUÂN PHƯƠNG

**XÂY DỰNG QUY TRÌNH VÀ CÔNG CỤ BIÊN TẬP DỮ
LIỆU BIÊN GIỚI, ĐỊA GIỚI CHO NỀN TẢNG CUNG
CẤP DỊCH VỤ ĐỊA CHỈ VIỆT NAM**

LUẬN VĂN THẠC SĨ HỆ THỐNG THÔNG TIN

Hà Nội – 10/2020

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

Hoàng Xuân Phương

**XÂY DỰNG QUY TRÌNH VÀ CÔNG CỤ BIÊN TẬP DỮ
LIỆU BIÊN GIỚI, ĐỊA GIỚI CHO NỀN TẢNG CUNG
CẤP DỊCH VỤ ĐỊA CHỈ VIỆT NAM**

Ngành: Hệ thống thông tin

Chuyên ngành: Hệ thống thông tin

Mã Số: 8.80104.01

LUẬN VĂN THẠC SĨ HỆ THỐNG THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC:

TS. Bùi Quang Hưng

Hà Nội – 10/2020

Mục lục

LỜI CẢM ƠN.....	III
LỜI CAM ĐOAN.....	IV
DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT.....	V
DANH MỤC BẢNG BIỂU.....	VI
DANH MỤC HÌNH VẼ.....	VII
 CHƯƠNG 1:GIỚI THIỆU.....	 1
1.1 ĐẶT VẤN ĐỀ, ĐỊNH HƯỚNG NGHIÊN CỨU	1
1.2 MỤC TIÊU CỦA LUẬN VĂN	2
1.3 PHẠM VI NGHIÊN CỨU CỦA LUẬN VĂN.....	2
1.4 NỘI DUNG CỦA LUẬN VĂN.....	2
 CHƯƠNG 2:.. TỔNG QUAN VỀ DỮ LIỆU BIÊN GIỚI ĐỊA GIỚI CHO	
NỀN TẢNG CUNG CẤP DỊCH VỤ ĐỊA CHỈ	4
2.1 NỀN TẢNG CUNG CẤP DỊCH VỤ ĐỊA CHỈ VIỆT NAM	4
2.1.1 Giới thiệu	4
2.1.2 Bản đồ số	7
2.1.3 Loại dữ liệu trên bản đồ số.....	7
2.2 DỮ LIỆU BIÊN GIỚI, ĐỊA GIỚI.....	9
2.2.1 Bản đồ địa hình.....	9
2.2.2 Biên giới.....	12
2.2.3 Địa giới	13
2.2.4 Bản đồ địa hình 1/50.000.....	14
2.2.5 Hệ quy chiếu VN2000	14
2.2.6 OSM	15
 CHƯƠNG 3:.. XÂY DỰNG QUY TRÌNH VÀ CÔNG CỤ BIÊN TẬP DỮ	
LIỆU BIÊN GIỚI ĐỊA GIỚI.....	16
3.1 QUY TRÌNH BIÊN TẬP DỮ LIỆU BIÊN GIỚI ĐỊA GIỚI.....	17
3.2 TRÍCH XUẤT DỮ LIỆU BIÊN GIỚI QUỐC GIA VÀ DỮ LIỆU ĐỊA GIỚI HÀNH	
CHÍNH TỪ TẬP DỮ LIỆU BẢN ĐỒ ĐỊA HÌNH	18
3.3 GỘP CÁC MẢNH BẢN ĐỒ THEO CẤP HÀNH CHÍNH: XÃ, HUYỆN, TỈNH.....	23
3.3.1 Gộp các mảnh rời rạc lại thành một mảnh thống nhất:	23

3.3.2	<i>Loại bỏ đường chia cắt của các mảnh sau khi gộp:</i>	25
3.4	BỔ SUNG THÔNG TIN HÀNH CHÍNH CẤP HUYỆN, TỈNH	26
3.5	SỬA LỖI SAU KHI GỘP	30
3.6	CHUYỂN ĐỔI HỆ QUY CHIỀU TỪ VN2000 SANG WGS84	39
3.6.1	<i>Tham số chuyển đổi và lưới chiếu</i>	39
3.6.2	<i>Thêm phương pháp chuyển đổi hệ quy chiếu:</i>	41
3.6.3	<i>Chuyển đổi hệ quy chiếu của dữ liệu:</i>	43
3.7	XÂY DỰNG CÔNG CỤ CHUYỂN ĐỔI DỮ LIỆU TỪ SHAPEFILE SANG ĐƯỜNG BOUNDARIES TRÊN OSM	44
3.7.1	<i>Cấu trúc của một tập tin tin OSM</i>	45
3.7.2	<i>Các bước tiền xử lý dữ liệu shapefile</i>	47
3.7.3	<i>Quy trình xây dựng chương trình xử lý</i>	51
CHƯƠNG 4: KẾT QUẢ VÀ ĐÁNH GIÁ		53
KẾT LUẬN		61
TÀI LIỆU THAM KHẢO		62

onluyen.net

LỜI CẢM ƠN

Trước tiên tôi xin dành lời cảm ơn chân thành và sâu sắc đến thầy giáo, TS. Bùi Quang Hưng – người đã hướng dẫn, khuyến khích, chỉ bảo và tạo cho tôi những điều kiện tốt nhất từ khi bắt đầu cho tới khi hoàn thành công việc của mình.

Tôi xin dành lời cảm ơn chân thành tới các thầy cô giáo khoa Công nghệ thông tin, trường Đại học Công nghệ, ĐHQGHN đã tận tình đào tạo, cung cấp cho tôi những kiến thức vô cùng quý giá và đã tạo điều kiện tốt nhất cho tôi trong suốt quá trình học tập, nghiên cứu tại trường.

Đồng thời tôi xin cảm ơn tất cả những người thân trong gia đình tôi cùng toàn thể bạn bè những người đã luôn giúp đỡ, động viên tôi những khi vấp phải những khó khăn, bế tắc.

Cuối cùng, tôi xin chân thành cảm ơn các đồng nghiệp của tôi tại Trung tâm tích hợp liên ngành và giám sát hiện trường - FIMO đã giúp đỡ, tạo điều kiện thuận lợi cho tôi học tập và nghiên cứu chương trình thạc sĩ tại Đại học Công nghệ, ĐHQGHN.

Luận văn này được hỗ trợ bởi đề tài nghiên cứu ứng dụng và phát triển công nghệ cấp quốc gia: "Nghiên cứu xây dựng Nền tảng cung cấp dịch vụ dữ liệu địa chỉ Việt Nam phục vụ phát triển các ứng dụng dân sinh", Mã số: ĐTCT-KC-4.0-03/19/25.

LỜI CAM ĐOAN

Tôi xin cam đoan rằng luận văn thạc sĩ hệ thống thông tin “Xây dựng quy trình và công cụ biên tập dữ liệu biên giới, địa giới cho nền tảng cung cấp dịch vụ địa chỉ Việt Nam” là công trình nghiên cứu của riêng tôi, không sao chép lại của người khác. Trong toàn bộ nội dung của luận văn, những điều đã được trình bày hoặc là của chính cá nhân tôi hoặc là được tổng hợp từ nhiều nguồn tài liệu. Tất cả các nguồn tài liệu tham khảo đều có xuất xứ rõ ràng và hợp pháp.

Tôi xin hoàn toàn chịu trách nhiệm và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan này.

Hà Nội, ngày 01 tháng 10 năm 2020

Hoàng Xuân Phương

DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Thuật ngữ, khái niệm	Định nghĩa	Ghi chú
Các từ viết tắt		
CSDL	Cơ sở dữ liệu.	
OSM	Open Street Map – Nền tảng bản đồ mở.	
NĐL	Nền địa lý.	
GIS	Geographic information system – Hệ thống thông tin địa lý.	
ArcMap	Là một chương trình trong bộ phần mềm ArcGIS. ArcMap cho phép người sử dụng thực hiện các chức năng: hiển thị trực quan, tạo lập bản đồ, trợ giúp quyết định, trình bày, ...	
Merge	Gộp 2 hoặc nhiều lớp bản đồ lại thành một lớp bản đồ duy nhất.	
Dissolve	Chập các đối tượng kề nhau thành một đối tượng duy nhất thông qua một hoặc nhiều thuộc tính chung.	
VNPost	Tổng công ty bưu điện Việt Nam	
APIs	Application Programing Interface – Giao diện lập trình ứng dụng.	
Node	Một điểm trên OSM được xác định bằng kinh độ, vĩ độ và ID.	
Way	Được sử dụng trên OSM xác định bằng các Node để xác định các đối tượng dạng đường và ranh giới khu vực.	
Relation	Được sử dụng trên OSM để xác định mối quan hệ giữa các đối tượng Node, Way.	
admin_level	Được sử dụng trong OSM để xác định cấp độ của đường boundaries.	

DANH MỤC BẢNG BIỂU

Bảng 1. Số lượng đơn vị địa giới hành chính tính đến 01/10/2020.....	13
Bảng 3-1. Bảng luật topology trong ArcMap	31
Bảng 3-2. Bảng múi chiếu	40
Bảng 4-1. Dữ liệu biên giới địa giới đã được biên tập từ csdl địa hình 1/50.000	53
Bảng 4-2. Địa danh hành chính việt nam tính tới 01/10/2020	59

onluyen.net

DANH MỤC HÌNH VẼ

Hình ảnh 2.1. Kiến trúc tổng thể của nền tảng cung cấp dịch vụ dữ liệu địa chỉ Việt Nam	5
Hình ảnh.2.2. Giao diện OSM	15
Hình ảnh 3.1. Kiến trúc tổng quan của nền tảng cung cấp dịch vụ địa chỉ	16
Hình ảnh 3.2. Các bước biên tập dữ liệu biên giới địa giới của Việt Nam	18
Hình ảnh 3.3. Cấu trúc dữ liệu của bản đồ địa hình 1/50.000	19
Hình ảnh 3.4. VN2000 múi 3 ⁰ cho FME.....	20
Hình ảnh 3.5. Chuyển đổi dữ liệu sang shapefile	21
Hình ảnh 3.6. Các shapefile thu được.....	22
Hình ảnh 3.7. Các bước gộp dữ liệu	23
Hình ảnh 3.8. Công cụ Merge của ArcMap.....	24
Hình ảnh 3.9. Kết quả sau khi Merge	24
Hình ảnh 3.10. Công cụ Dissolve của ArcMap	25
Hình ảnh 3.11. Kết quả sau khi Dissolve.....	26
Hình ảnh 3.12. Bảng dữ liệu của ranh giới địa giới cấp xã sau khi gộp.....	27
Hình ảnh 3.13. Bảng dữ liệu của ranh giới địa giới cấp huyện sau khi gộp...	27
Hình ảnh 3.14. Bảng dữ liệu của ranh giới địa giới cấp tỉnh sau khi gộp	28
Hình ảnh 3.15. Join sử dụng ArcMap.....	29
Hình ảnh 3.16. Kết quả thu được sau khi join	30
Hình ảnh 3.18. Luật must not have gaps	36
Hình ảnh 3.19 Các vùng bị lỗi topology.....	37
Hình ảnh 3.20. Lỗi chồng đè giữa 2 polygon	38
Hình ảnh 3.21. Lỗi polygon không khép kín	38
Hình ảnh 3.22. Lỗi đường phân chia ranh giới giữa các mảnh	39
Hình ảnh 3.23. Lý lịch bản đồ	40
Hình ảnh 3.24. Công cụ projections and transformation.....	41
Hình ảnh 3.25. Thiết lập tham số cho projections and transformation.....	42

Hình ảnh 3.26. Thiết lập tham số cho projections and transformation.....	43
Hình ảnh 3.27. Chuyển đổi hệ quy chiếu.....	43
Hình ảnh 3.28. Kết quả sau khi chuyển đổi.....	44
Hình ảnh 3.29. Bảng admin boundaries của Việt Nam trên OSM	45
Hình ảnh 3.30 . Cấu trúc của một tập tin OSM	46
Hình ảnh 3.31. Polygon sang lines của QGIS	48
Hình ảnh 3.32. v.clean của QGIS	48
Hình ảnh 3.33. Trước khi sử dụng v.clean	49
Hình ảnh 3.34. Sau khi sử dụng v.clean	50
Hình ảnh 3.35. Dữ liệu gốc.....	50
Hình ảnh 3.36. Dữ liệu sau khi chạy v.clean	51
Hình ảnh 3.37. Tập tin OSM chứa thông tin về các đường biên giới địa giới Việt Nam.....	52
Hình ảnh 4.1. Nội dung trong tập tin OSM.....	54
Hình ảnh 4.2. Bản đồ số viet bản đồ.....	55
Hình ảnh 4.3. Chồng lớp nền bản đồ Việt bản đồ	56
Hình ảnh 4.4. Bản đồ số GIS Chính phủ	57
Hình ảnh 4.5. Chồng lớp nền bản đồ GIS Chính phủ.....	58

CHƯƠNG 1: GIỚI THIỆU

1.1 Đặt vấn đề, định hướng nghiên cứu

Đề án Phát triển Hệ tri thức Việt số hóa: Phát triển Hệ tri thức Việt số hóa, là đề án của Chính phủ Việt Nam được phê duyệt theo Quyết định số 677/QĐ-TTg[1] của Thủ tướng Chính phủ ngày 18/05/2017. Đề án có mục tiêu xây dựng một hệ tri thức tổng hợp trong mọi lĩnh vực, góp phần thúc đẩy, tạo điều kiện để mọi người dân học tập suốt đời, làm chủ tri thức, tăng cường nghiên cứu sáng tạo, ứng dụng tiến bộ khoa học công nghệ, thúc đẩy phát triển đất nước. Đề án được xây dựng và cập nhật theo hình thức xã hội hóa, thu hút và khuyến khích mọi người dân và doanh nghiệp tham gia, với vai trò vừa khai thác vừa đóng góp để làm giàu các tài nguyên tri thức số hóa của Việt Nam. Góp phần khơi dậy, lan toả niềm đam mê khoa học và công nghệ, khát vọng sáng tạo, cống hiến của mọi người trong việc tạo lập và phổ biến tri thức. Đề án hướng tới phổ cập thông tin khoa học và công nghệ cho mọi tầng lớp người dân trong xã hội, nhất là các vùng nông thôn, vùng sâu, vùng xa; học sinh, sinh viên, người lao động.

Một trong những nội dung quan trọng của Đề án là Nghiên cứu xây dựng Nền tảng cung cấp dịch vụ dữ liệu địa chỉ của Việt Nam phục vụ phát triển các ứng dụng dân sinh do Đại học Quốc gia Hà Nội chủ trì. Nền tảng cung cấp dịch vụ dữ liệu địa chỉ Việt Nam sẽ là nền tảng cơ bản để trên đó các nhà phát triển xây dựng các ứng dụng dân sinh phục vụ các nhu cầu khác nhau của cộng đồng như tìm kiếm vị trí, địa điểm, địa chỉ, du lịch, văn hóa, giáo dục, y tế, ...

Trong các công việc đó, việc xây dựng lớp dữ liệu biên giới địa giới là nhiệm vụ đầu tiên cần phải thực hiện và cũng là dữ liệu đóng vai trò cơ sở cho các lớp dữ liệu địa chỉ. Việc xây dựng dữ liệu dữ liệu bản đồ cho biên giới địa giới Việt Nam có thể được tổng hợp và biên tập từ nhiều mảnh bản đồ có các độ phân giải khác nhau, dẫn đến dữ liệu khi được tổng hợp thành một thể thống nhất sẽ có nhiều vùng bị khuyết dữ liệu hoặc các vùng không được liền mạch. Bài toán đặt ra cần phải chỉnh sửa hoặc cập nhật lại các vùng dữ liệu bị khuyết đó.

1.2 Mục tiêu của luận văn

Đứng trước các nhu cầu cấp thiết đó tôi đã quyết định chọn đề tài “*Xây dựng quy trình và công cụ biên tập dữ liệu biên giới, địa giới cho Nền tảng cung cấp dịch vụ địa chỉ Việt Nam*” làm đề tài nghiên cứu của mình.

Đề tài này nhằm mục đích nghiên cứu và Xây dựng quy trình và công cụ biên tập dữ liệu biên giới, địa giới cho Nền tảng cung cấp dịch vụ địa chỉ Việt Nam.

1.3 Phạm vi nghiên cứu của luận văn

Công việc thực hiện xây dựng quy trình và biên tập dữ liệu biên giới, địa giới tích hợp vào nền tảng cung cấp dịch vụ địa chỉ, là một công việc phức tạp yêu cầu nhiều kiến thức chuyên môn khác nhau và cần nhiều thời gian để thực hiện. Vì vậy trong luận văn này tôi chỉ tập chung chính vào một số công việc như sau:

- Xây dựng quy trình gộp các mảnh bản đồ 1:50.000.
- Lỗi và phương pháp xử lý lỗi khi gộp các mảnh bản đồ lại với nhau.
- Xây dựng công cụ chuyển đổi dữ liệu từ dữ liệu shapefile sang định dạng tập tin OSM.
- Chỉnh sửa dữ liệu tập tin OSM sau khi chuyển đổi.

1.4 Nội dung của luận văn

Luận văn thực hiện xuyên suốt trong quá trình từ khi hình thành các khái niệm, ý tưởng nghiên cứu, cho đến khi xây dựng được quy trình và công cụ biên tập dữ liệu biên giới, địa giới cho Nền tảng cung cấp dịch vụ địa chỉ Việt Nam. Nội dung chính bao gồm các phần sau:

- Chương 1: Mở Đầu, đặt ra vấn đề, mục tiêu và giải pháp cho bài toán “*Xây dựng quy trình và công cụ biên tập dữ liệu biên giới, địa giới cho Nền tảng cung cấp dịch vụ địa chỉ Việt Nam*”.
- Chương 2: trình bày các khái niệm cơ bản phục vụ cho nghiên cứu của đề tài.
- Chương 3: Quy trình và phương pháp thực hiện sẽ được trình bày trong chương này.