

# NGHIÊN CỨU TRUYỀN NHẬN DỮ LIỆU GIỮA CÁC THIẾT BỊ DI ĐỘNG VỚI HỆ THỐNG MÁY CHỦ

ĐINH BẢO NGỌC, NGUYỄN TRƯỜNG XUÂN, NGUYỄN HOÀNG LONG,  
TRẦN THỊ HẢI VÂN, NGUYỄN THỊ HẢI YẾN

Khoa Công nghệ thông tin, Trường Đại học Mở - Địa chất

## ***Tóm tắt:***

Ngày nay các ứng dụng Mobile GIS ngày càng được phát triển trợ giúp cho quá trình thu thập, phân tích dữ liệu ngoài thực địa trở nên dễ dàng hơn. Trong đó, việc truyền nhận và lưu trữ dữ liệu không gian giữa các ứng dụng di động và hệ thống máy chủ là hết sức quan trọng. Trong báo cáo này sẽ giới thiệu về công nghệ và chiến lược trao đổi, lưu trữ và phân phối dữ liệu không gian giữa hệ thống server và các ứng dụng GIS trên các thiết bị di động.

## **1. Mở đầu**

Hiện nay, việc phát triển các ứng dụng trên điện thoại thông minh là nhu cầu thực tế trong thời kỳ công nghệ 4.0. Việc phát triển các ứng dụng này được dựa trên 3 kỹ thuật chính, đó là native application, web application và hybrid application trên tất cả các hệ điều hành, phổ biến nhất là IOS và Android. Tùy thuộc vào cách thức lưu trữ, truy xuất và gửi nhận dữ liệu sẽ có các điều kiện ứng dụng khác nhau. Như với các ứng dụng web application, dữ liệu sẽ được nhận về từ hệ thống máy chủ thông qua mạng internet và các ứng dụng dạng này điều kiện bắt buộc phải có kết nối internet để sử dụng. Trong khi đó các ứng dụng native application và hybrid application vừa có thể tương tác với dữ liệu từ máy chủ cũng như từ nội tại được lưu trữ trực tiếp trên thiết bị.

Ngoài lựa chọn kỹ thuật xây dựng phát triển các ứng dụng trên di động, đồng bộ dữ liệu là công việc cần đến trong các hệ thống công nghệ thông tin hiện nay tại Việt Nam, trong đó, chủ yếu là các hệ thống cần thu thập dữ liệu (thiết bị di động) và các hệ thống cập nhật cũng như tác nghiệp trong môi trường kết nối không thường xuyên (mô hình thành phố - quận - phường). Đối với các hệ thống GIS, dữ liệu thường chiếm tỉ lệ rất lớn về chi phí, số người tham gia trong việc hình thành và duy trì hệ thống. Do đó, đồng bộ

dữ liệu không gian là một vấn đề cần thiết cho các ứng dụng GIS. Những hệ thống có xây dựng cơ chế đồng bộ tốt sẽ góp phần tiết kiệm chi phí cho hệ thống. Và hơn nữa, đó cũng là một trong yếu tố góp phần đưa hệ thống gần gũi hơn với người sử dụng và hệ thống được bền vững hơn.

## **2. Giải quyết vấn đề**

### ***2.1. Phân loại các nền tảng xây dựng ứng dụng trên các thiết bị di động***

Native App, Mobile web app, Hybrid App (ứng dụng lai), là những khái niệm cơ bản của lập trình di động. Dựa trên ngôn ngữ lập trình, có thể chia mobile app làm 3 loại, bao gồm: ứng dụng gốc (Native App), ứng dụng web (Web App) và ứng dụng lai (Hybrid App). Trong đó trên 50% nhà phát triển ứng dụng di động chọn làm ứng dụng native [2].

*2.1.1. Ứng dụng gốc (Native app)* là những ứng dụng được viết riêng cho một loại nền tảng như iOS, Android, Windows Phone bằng các ngôn ngữ tương ứng của mỗi nền tảng đó, ví dụ Java trên Android, Object C trên iOS, C# trên winphone.

Mỗi Native App chỉ chạy được trên một nền tảng và không thể mang sang các nền tảng khác. Ví dụ game cho iOS sẽ không thể chạy được trên các máy Android. Đa phần các game mobile hiện nay là Native App.

Ngày nhận bài: 07/01/2020, ngày chuyển phản biện: 12/01/2020, ngày chấp nhận phản biện: 16/01/2020, ngày chấp nhận đăng: 19/01/2020

### Điểm mạnh

Về mặt hiệu năng thì native app chạy nhanh hơn mobile app trong hầu hết trường hợp. Tuy nhiên sự chênh lệch ở các ứng dụng có thể là nhỏ, chấp nhận được và khó nhận biết bởi người dùng. Hiện tại, với game thì native app vượt bậc và thắng tuyệt đối.

Về tính năng native app có khả năng truy cập các phần cứng (như camera, GPS, thiết bị thu âm...) và có thể truy cập đến các quyền local như: sao chép, tạo, ghi, đọc tập tin (trên bộ nhớ của mobile), các thông tin danh bạ, thông tin cá nhân, thực hiện cuộc gọi, nhắn tin... Nói cách khác, native app có thể khai thác được sức mạnh phần cứng nhiều nhất của nền tảng. Trong khi Mobile Web app thì không thể (có trường hợp có thể đọc file, như upload).

Ở chế độ offline: Native app có lợi thế so với mobile web app (luôn phải chạy online). Ở offline-mode, native app sẽ dùng những dữ liệu đã cache trước đó ở lần cuối cùng mà người dùng truy cập online.

### Điểm yếu

Không thể đa nền tảng: một native app chỉ có thể chạy trên một hệ điều hành nhất định.

Với mỗi hệ điều hành thì người lập trình phải viết riêng native code cho nó, dẫn đến không nhất quán giữa các phiên bản ứng dụng, chi phí phát triển cao và đòi hỏi khả năng thành thạo nhiều ngôn ngữ ở lập trình viên. Đôi khi đòi hỏi lập trình viên phải sử dụng những công cụ độc quyền và chỉ được làm việc trên một môi trường duy nhất như: dùng xcode trên Mac OS (nếu bạn muốn phát triển ứng dụng cho iOS).

Hơn nữa, việc bảo trì hay nâng cấp sẽ làm mất nhiều thời gian.

Tính đa dạng phiên bản của hệ điều hành làm cản trở việc phổ dụng các native app. Các thiết bị chạy phiên bản hệ điều hành cũ hơn sẽ không cài đặt được.

2.1.2. *Ứng dụng web* (web-based app) là một loại chương trình có khả năng làm cho máy tính

thực hiện trực tiếp một công việc nào đó người dùng muốn thực hiện.

Web app có thể được thiết kế chạy trên nền tảng web hoặc các trình duyệt của thiết bị di động. Ví dụ: <https://m.facebook.com> là một Web App cho di động.

### Điểm mạnh

+ Đa nền tảng (Cross platform): Có thể chạy trên tất cả trình duyệt của mobile hỗ trợ phiên bản HTML và javascript.

+ Không cần cài đặt trên máy.

+ Thuận lợi cho các nhà phát triển:

+ Với một phiên bản duy nhất cho tất cả, nên giảm chi phí và thời gian cho phát triển, bảo trì, cũng như nâng cấp sau này.

+ Ngôn ngữ lập trình phổ dụng là HTML và Javascript mà hầu hết các lập trình viên đều biết.

+ Không giới hạn môi trường lập trình, lập trình viên có thể làm việc trên hầu hết các HDH.

+ Ứng dụng không cần phải được build lại qua SDK hoặc một công cụ phát triển độc quyền khác như xCode (chỉ cần deploy lên server và chạy).

+ Việc cập nhật phiên bản sẽ đến ngay lập tức và trong suốt với người dùng (ngay sau khi deploy) mà không cần thông qua thủ tục và chờ đợi approval như các chợ ứng dụng.

+ Có thể được đánh dấu bởi search engines (tốt cho SEO).

+ Dễ dàng quảng bá: người dùng có thể tìm thấy thông qua search engine, share link từ email, từ social networks hoặc từ các dịch vụ quảng cáo trực tuyến như Google Ads.

### Điểm yếu

Mặc dù HTML5 có tính tương thích rộng lớn nhưng cùng với đó là cả điểm yếu. Nó phụ thuộc vào trình duyệt, nền tảng, loại máy. Đặc biệt khi mức độ phân mảnh thị trường của các thiết bị di động lớn như hiện nay lại càng làm các nhà lập trình khó khăn trong việc lựa chọn đúng phần

nào của HTML5 có thể sử dụng.

Về hiệu năng, web-based app không chạy nhanh như native app, và không mang lại trải nghiệm trọn vẹn mà nền tảng đó có thể cung cấp.

Luôn phải chạy online. Khi mạng chập chờn thì ứng dụng web cũng ngay lập tức thể hiện điểm yếu của mình.

2.1.3. *Ứng dụng lai* (Hybrid app) là ứng dụng kết hợp của cả Mobile Web App và Native App. Theo đó, các phần cơ bản của ứng dụng vẫn viết bằng ngôn ngữ web, nhưng được đặt trong native container, nên vẫn có thể đưa lên các App Store.

### Điểm mạnh

Hybrid có thể kết hợp điểm mạnh của cả native app và mobile app và khai thác được trải nghiệm cao nhất của nền tảng.

Ví dụ như: Hiển thị nội dung trên tất cả các thiết bị di động, tận dụng tối đa các tính năng của thiết bị di động như GPS, Camera, cảm biến gia tốc... điều này không thể làm được trên mobile web, sử dụng ứng dụng ngay cả khi không có kết nối internet. Những ứng dụng mobile của một số trang tin tức hiện nay như Vnexpress, Dantri v.v.. đều là những ứng dụng lai.

## 2.2. Lưu trữ cơ sở dữ liệu trên thiết bị di động

### 2.2.1. Lưu trữ dữ liệu ngoại tuyến (offline) sử dụng SQLite và Spatialite

SQLite là một thư viện phần mềm mà triển khai một SQL Database Engine truyền thống, không cần Server, không cần cấu hình và nhỏ gọn. SQLite là một trong những Database Engine phát triển nhanh nhất, nhưng sự phát triển của nó là về mặt tính phổ biến, không phải là về mặt kích cỡ. SQLite là mã nguồn mở.

SQLite là phần mềm quản lý cơ sở dữ liệu (DBMS) tương tự như Mysql, PostgreSQL... Đặc điểm của SQLite là gọn, nhẹ, đơn giản. Chương trình gồm một file duy nhất chưa đến 400kB, không cần cài đặt, không cần cấu hình hay khởi động mà có thể sử dụng ngay. Dữ liệu

Database cũng được lưu ở một file duy nhất. Không có khái niệm user, password hay quyền hạn trong SQLite Database.

Ưu điểm của SQLite:

+ SQLite không yêu cầu một tiến trình Server riêng rẽ để hoạt động.

+ SQLite không cần cấu hình, nghĩa là không cần thiết phải cài đặt.

+ Một SQLite Database đầy đủ được lưu giữ trong một disk file đơn.

+ SQLite là rất nhỏ gọn, nhỏ hơn 400kB đã được cấu hình đầy đủ hoặc nhỏ hơn 250kB khi đã bỏ qua các tính năng tùy ý.

+ SQLite là tự chứa, nghĩa là không có sự phụ thuộc vào ngoại vi.

+ Các giao dịch (Transactions) trong SQLite là tuân theo đầy đủ chuẩn ACID, đảm bảo truy cập an toàn từ nhiều tiến trình hoặc thread.

+ SQLite hỗ trợ hầu hết các tính năng của một ngôn ngữ truy vấn trong chuẩn SQL92.

+ SQLite được viết bằng ANSI-C và cung cấp API đơn giản và dễ dàng sử dụng.

+ SQLite là có sẵn trên UNIX (Linux, Mac OS-X, Android, iOS) và Windows (Win32, WinCE, WinRT).

Spatialite là thành phần mở rộng về không gian của SQLite cung cấp chức năng thao tác với vector geodatabase. Bản thân SQLite đã hỗ trợ lưu trữ và các thao tác cơ bản với dữ liệu không gian, do đó Spatialite không cần thiết được dùng để quản lý dữ liệu không gian. Tuy nhiên, Spatialite lại hỗ trợ mạnh mẽ việc truy vấn dữ liệu không gian nâng cao, hỗ trợ nhiều hệ tọa độ bản đồ

Một ưu điểm lớn của Spatialite cho phép nhập và xuất các tệp định dạng shapefile.

### 2.2.2. Lưu trữ cơ sở dữ liệu trực tuyến (online) trên hệ thống Server

2.2.2.1. Hệ quản trị cơ sở dữ liệu mã nguồn mở PostgreSQL/PostGIS

PostgreSQL là một hệ thống quản trị cơ sở dữ liệu quan hệ-đối tượng, là hệ thống cơ sở dữ liệu mã nguồn mở tiên tiến nhất hiện nay.

PostgreSQL sở hữu một hệ tính năng đa dạng giúp hỗ trợ phát triển xây dựng app, bảo vệ toàn vẹn dữ liệu, và tạo ra một môi trường chịu lỗi fault-tolerant giúp quản lý dữ liệu bất kể tập dữ liệu lớn hay nhỏ. PostgreSQL cũng có khả năng mở rộng tuyệt vời. Ví dụ, có thể định nghĩa các kiểu dữ liệu riêng, xây dựng các hàm tùy chỉnh, hay viết mã từ các ngôn ngữ lập trình khác nhau mà không cần biên dịch lại CSDL.

PostGIS là phần mở rộng của hệ quản trị CSDL PostgreSQL được cung cấp miễn phí cho phép CSDL quản lý các đối tượng GIS. Trên thực tế, PostGIS là một hợp phần quản lý thông tin không gian của PostgreSQL, nó cho phép quản trị CSDL không gian dùng trong GIS, giống như SDE của ESRI hoặc Oracle's Spatial extension. PostGIS hỗ trợ cho OpenGIS vì "tính năng kỹ thuật SQL đơn giản" và đã được chứng nhận phù hợp với "Kiểu và Chức năng" của OpenGIS.

PostGIS là một dự án mã nguồn mở về CSDL không gian đang được nghiên cứu và phát triển bởi Refrations Research. PostGIS được phát hành theo GNU General Public License (phát hành miễn phí bản quyền sử dụng và phát triển). Dự án tiếp tục phát triển PostGIS, bổ sung các công cụ giao diện với người dùng, hỗ trợ các topology cơ bản, chuẩn hóa dữ liệu, chuyển đổi tọa độ, lập trình API và nhiều hơn nữa.

### Ưu điểm của PostGIS

+ PostGIS - một CSDL không gian có những ưu điểm để xử lý các thông tin như trả lời các truy vấn về những đối tượng ở gần một vị trí nào đó, những đối tượng nằm trong phạm vi hoặc ở vùng phụ cận của một đối tượng khác, phạm vi của một vùng nơi mà có một hoạt động nào đó đang xảy ra là gì, những đối tượng nào nằm bên trong một đối tượng khác.

+ PostGIS có khả năng lưu trữ và thao tác với dữ liệu rất tốt. Nó cung cấp những khả năng xử lý thông tin địa lý bên trong một môi trường cơ

sở dữ liệu. Những hàm SQL bao gồm buffer, intersection, within, distance... Những hàm này lấy dữ liệu hình học từ các cột trong bảng PostGIS và trả về những hình học mới hoặc những thông tin khác..

+ PostGIS cho phép dễ dàng khi kết nối dữ liệu không gian với dữ liệu phi không gian trong một môi trường dữ liệu không gian và cung cấp đầy đủ sức mạnh của ngôn ngữ truy vấn cấu trúc (SQL) để thực hiện những bài toán phân tích..

+ PostGIS còn hỗ trợ thêm GiST-based and R-Tree spatial indexes, đây là điểm riêng biệt nhất của PostGIS so với các hệ quản trị CSDL khác giúp PostGIS dễ dàng quản lý, lưu trữ, khai thác và tìm kiếm thông tin không gian.

Một số đối tượng không gian theo chuẩn OGC trong PostGIS

+ PostGIS hỗ trợ các đối tượng và các hàm được định nghĩa chi tiết trong OGC "Simple Features for SQL". Với PostGIS ta có thể lưu trữ các điểm, đường, vùng, tập hợp điểm, tập hợp đường, tập hợp vùng và các thông tin hình học khác. PostGIS mở rộng chuẩn với sự hỗ trợ cho các tọa độ dạng 3DZ, 3DM và 4D.

+ OpenGIS định nghĩa hai chuẩn để diễn tả các đối tượng không gian: Well-known Text (WKT) và Well-Known Binary (WKB). Cả Well-known Text và Well-Known Binary đều bao gồm thông tin về kiểu của đối tượng và tọa độ của nó.

+ PostGIS được hiện thực theo bản đặc tả "OGC Simple Feature Specifications" cho chuẩn SQL. Bản đặc tả này định nghĩa những thuật toán và lược đồ SQL để thêm, truy vấn, thao tác và xóa những đối tượng không gian. Những tọa độ của các đối tượng không gian được lưu trữ trong các bảng Feature. Mỗi một bảng Feature có thể chứa một kiểu đối tượng hình học là điểm, đường, đa giác. Những tọa độ của mỗi đối tượng được lưu trữ trong một trường với một kiểu riêng biệt.

+ Kiểu mô tả một tập hợp tọa độ không gian của đối tượng là WKT (Well Known Text). Khi

những đối tượng không gian được đưa vào cơ sở dữ liệu thì PostGIS sẽ chuyển chúng từ WKT sang WKB (Well Known Binary) để việc lưu trữ được tốt hơn.

So sánh PostgreSQL với một số hệ quản trị cơ sở dữ liệu khác

+ Hệ điều hành hỗ trợ

|            | Windows | Mac OS X | Linux | BSD   | UNIX | zOS <sup>[16,19]</sup> |
|------------|---------|----------|-------|-------|------|------------------------|
| MySQL      | Có      | Có       | Có    | Có    | Có   | Có thể                 |
| Oracle     | Có      | Có       | Có    | Không | Có   | Có                     |
| PostgreSQL | Có      | Có       | Có    | Có    | Có   | Có thể                 |

+ Các tính năng cơ bản

|            | ACID                         | Referential integrity        | Transactions                 | Unicode                     |
|------------|------------------------------|------------------------------|------------------------------|-----------------------------|
| MySQL      | Phụ thuộc <sup>[16,20]</sup> | Phụ thuộc <sup>[16,20]</sup> | Phụ thuộc <sup>[16,20]</sup> | Có / UTF-8 (3-byte) & UCS-2 |
| Oracle     | Có                           | Có                           | Có                           | Có                          |
| PostgreSQL | Có                           | Có                           | Có                           | Có / UTF-8 (4-byte)         |

+ Hỗ trợ bảng và khung nhìn:

|            | Bảng tạm | Khung nhìn cụ thể               |
|------------|----------|---------------------------------|
| MySQL      | Có       | Tương tự <sup>[17][18,96]</sup> |
| Oracle     | Có       | Có                              |
| PostgreSQL | Có       | Tương tự <sup>[17][18,77]</sup> |

+ Chỉ mục

|            | Cây B-Cây R+                               | Hàm băm          | Biểu thức tập trình | Chỉ mục sóng phân | Chỉ mục đồ             | Bitmap | GST   |
|------------|--|------------------|---------------------|-------------------|------------------------|--------|-------|
| MySQL      | truy SQL, số MyISAM, BDE, hoặc bảng InnoDB | chỉ số bảng HEAP | Không               | Không             | Không                  | Không  | Không |
| Oracle     | chỉ số phân bổ EE                          | biểu gộp         | Có                  | Không             | Có                     | Có     | Không |
| PostgreSQL | Có   | Có               | Có                  | Có                | Có <sup>[16,100]</sup> | Có     | Có    |

+ Các đối tượng khác

|            | Domain | Cursor | Trigger | Hàm <sup>[16,103]</sup> | Thuật ngữ <sup>[16,103]</sup> | External routine <sup>[16,103]</sup> |
|------------|--------|--------|---------|-------------------------|-------------------------------|--------------------------------------|
| MySQL      | Không  | Có     | Có      | Có                      | Có                            | Có                                   |
| Oracle     | Có     | Có     | Có      | Có                      | Có                            | Có                                   |
| PostgreSQL | Có     | Có     | Có      | Có                      | Có                            | Có                                   |

Hình 2: So sánh PostgreSQL với một số hệ quản trị cơ sở dữ liệu khác

Đặc điểm của PostGIS:

+ Các kiểu dữ liệu hình học như Point, Line, Polygon, vv. Các kiểu dữ liệu hình học này được lưu trữ như các đối tượng hình học

+ Các toán tử không gian cho phép xác định các phép đo không gian địa lý như tính diện tích, tính khoảng cách, tính độ dài, tính chu vi. PostGIS hỗ trợ các hàm như: ST\_Area(), ST\_Length(), ST\_Distance(), vv các hàm này thực hiện chức năng các phép đo.

+ Các toán tử không gian cho phép xác định không gian địa lý. Các thao tác như phép hợp, so sánh sự khác nhau giữa các đối tượng hình học. Các toán tử PostGIS hỗ trợ như ST\_Difference() cho biết sự khác nhau giữa các đối tượng hình học.

+ PostGIS cung cấp việc đánh chỉ mục không gian tốc độ cao sử dụng GisT hoặc R-Tree.

2.2.2.2. Hệ quản trị cơ sở dữ liệu thương mại SQL Server

Khái niệm Hệ quản trị cơ sở dữ liệu không gian (SDBMS – Spatial Database Manage System), đây là một hệ quản trị cơ sở dữ liệu, có hỗ trợ mô hình dữ liệu không gian, kiểu dữ liệu không gian (SDT – Spatial Data Types) và ngôn ngữ truy vấn SQL không gian, hỗ trợ kiểu dữ liệu không gian trong lập trình: cung cấp chỉ số không gian (Spatial Indexing) và thuật toán liên kết không gian (Spatial Joins).

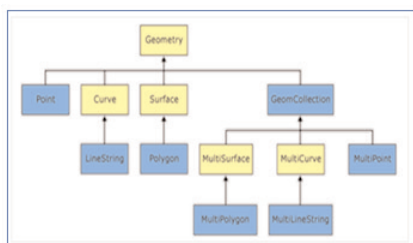
Hiện nay có rất nhiều hệ quản trị CSDL không gian như: DB2 (phần mở rộng IBM DB2 Spatial Extender), Oracle (Oracle Spatial), PostgreSQL (phần mở rộng PostGIS), Sqlite (phần mở rộng SpatialLite), và Microsoft SQL Server phiên bản 2008 trở lên.

Trong SQL Server 2008 giới thiệu các 2 kiểu dữ liệu không gian là:

+ Kiểu dữ liệu hình học: Là kiểu dữ liệu không gian cơ bản được sử dụng để biểu diễn các đối tượng trong hệ tọa độ Euclide, đối tượng không có thông tin về SRID (Spatial Reference Identifiers, đây là một hệ quy chiếu đã được định

nghĩa trước ứng với một số nguyên).

+ Kiểu dữ liệu địa lý: Là kiểu dữ liệu không gian cơ bản được sử dụng để biểu diễn các đối tượng trong hệ tọa độ trái đất (tọa độ Ellipsoid). Đối tượng có chứa thông tin về SRID.



Hình 3: Mô hình thiết kế các đối tượng không gian

Trong SQL Server 2008, các đối tượng hình học trên bản đồ được mô tả bởi Well-Know Text (WKT) hoặc Well-Known Binary (WKB), các định nghĩa đầu tiên được đưa ra bởi Open Geospatial Consortium (OGC) trong đó:

+ Well-Know Text (WKT): ngôn ngữ đánh dấu kiểu văn bản dùng để biểu diễn cho các đối tượng hình học ở định dạng vector trên bản đồ. WKT bao gồm các thông tin về kiểu của đối tượng và các tọa độ dạng đối tượng

Ngoài ra, WKT cũng được sử dụng để mô tả hệ quy chiếu như Datum, Geoid, Hệ tọa độ và Phép chiếu bản đồ của các đối tượng không gian.

+ Well-Known Binary (WKB): ngôn ngữ đánh dấu kiểu nhị phân dùng để biểu diễn cho các đối tượng hình học trên bản đồ (tương tự như WKT) ở hệ thập lục phân.

Với sự hỗ trợ làm việc với các đối tượng không gian, người sử dụng SQL Server 2008 có thể thực hiện việc lưu trữ dữ liệu không gian cũng như thực hiện truy vấn không gian mà các hệ quản trị cơ sở dữ liệu truyền thống (không chứa thành phần không gian không thể thực hiện được. Tính năng này sẽ cung cấp cho các nhà quản lý trong các lĩnh vực liên quan đến yếu tố không gian một hệ quản trị CSDL phù hợp, đặc biệt là những nhà quản lý về tài nguyên-môi trường, nhà quản lý giao thông, các hệ thống điện – cấp thoát nước, người nghiên cứu về địa

lý, cũng như các nhà quy hoạch phát triển kinh tế.

### 2.3. Đồng bộ dữ liệu không gian giữa ứng dụng di động và hệ thống Server

#### 2.3.1. Đồng bộ dữ liệu trong hệ thống quản trị cơ sở dữ liệu

##### Đồng bộ dữ liệu trong hệ thống SQL Server

Với các hệ thống quản trị CSDL như Microsoft SQL Server, chức năng đồng bộ dữ liệu thuộc tính được xây dựng từ các phiên bản. Trong đó, các mô hình truyền thống thường được biết đến là [1]:

- Snapshot replication: mang ý nghĩa tương tự như bản sao lưu đầy đủ, nghĩa là sao chép toàn bộ dữ liệu (có thay đổi và kể cả không thay đổi) từ máy nguồn sang các máy đích và thay thế toàn bộ dữ liệu tồn tại trên các máy đích. Thông thường, dạng đồng bộ này được sử dụng khi có nhu cầu hoặc công việc cập nhật định kỳ, ví dụ theo tuần, tháng, quý hoặc năm.

- Transactional replication: chỉ những dữ liệu có thay đổi được gửi đi. Các giao tác được truyền đi theo thứ tự tuần tự, và chỉ chuyển đi các câu lệnh transaction (thay vì dữ liệu). Dữ liệu ở các site không mâu thuẫn nhau vì ta chỉ lấy dữ liệu có sự thay đổi ở một nơi chép cho các nơi khác.

- Merge replication: dữ liệu ở nhiều nơi thay đổi khác nhau. Định kỳ, những thay đổi sẽ được trộn lẫn nhau và giải quyết các xung đột/mâu thuẫn. Phương pháp này không đảm bảo tính nhất quán dữ liệu.

Mỗi loại đồng bộ được sử dụng tùy thuộc vào ứng dụng. Do đó, SQL Server cung cấp 03 đối tượng: nơi phát hành dữ liệu, nơi phân phối dữ liệu và nơi nhận dữ liệu. Ngoài ra, SQL Server cũng hỗ trợ những mô hình vật lý như sau:

- Central Publisher/Distributor: một máy trung tâm làm cả hai nhiệm vụ phát hành và phân phối; các máy còn lại nhận dữ liệu. Trong mô hình này, dữ liệu được đẩy đi từ một trung tâm, nên người quản trị mạng phải đảm bảo dữ liệu

truyền đến các nơi khác chỉ được phép chọn để xem (SELECT), ngoài ra không được phép sửa chữa.

- Central Subscriber/ Multiple Publishers: một máy trung tâm nhận dữ liệu, các máy còn lại làm nhiệm vụ phát hành và phân phối dữ liệu về trung tâm. Trong mô hình này, điều quan trọng là chúng ta phải đảm bảo tất cả dữ liệu trên từng server là duy nhất (nếu không các Publisher khác sẽ ghi đè dữ liệu lên nhau. Do đó, chúng ta cần có sự lọc dữ liệu theo chiều ngang trong trường hợp này. Cụ thể hơn, chúng ta có thể phân dữ liệu theo thuộc tính như các đơn vị phường/xã trong quận hoặc quận/huyện theo thành phố.

- Multiple Publishers/Multiple Subscribers: các máy ngang hàng với nhau, vừa nhận dữ liệu, vừa phân phối dữ liệu. Chúng ta cần phải lưu ý và phân tích tính nhất quán của dữ liệu khi sử dụng mô hình. Mặt khác, dữ liệu phải được lọc hoặc việc phân chia dữ liệu hợp lý. Ví dụ: máy chủ Phòng Kinh tế tại quận/huyện chỉ được cập nhật dữ liệu thông tin trường dữ liệu hộ có đăng ký kinh doanh trên lớp thửa còn các máy chủ tại phường/xã chỉ được cập nhật thông tin trường dữ liệu đăng kí thu gom rác cho mỗi hộ; và Phòng Tài nguyên môi trường quận/huyện chỉ được cập nhật các thông tin thuộc trường dữ liệu hộ bị xử phạt môi trường.

Đồng bộ dữ liệu trong hệ thống mã nguồn mở Postgres

Trên hệ thống Postgres, đến nay, từ phiên bản PostgreSQL 9.0 được trang bị chức năng đồng bộ (sao lập) dữ liệu trên gói Streaming. Hệ thống đồng bộ/sao lập streaming hoạt động dựa trên cơ chế WAL. Và hệ thống Streaming là hệ thống sao lập bất đồng bộ.

Bên cạnh đó, hệ thống Slony cũng là một hệ thống sao lập bất đồng bộ. Dưới đây là một số so sánh giữa hệ thống Streaming và hệ thống Slony.

2.3.2. Đồng bộ dữ liệu trong một số hệ thống Geodatabase

2.3.2.1. Mô hình đồng bộ dữ liệu theo quản lý phiên bản (versioning checkout/check in)

| TT | Tiêu chí so sánh  | Hệ thống Streaming | Hệ thống Slony |
|----|---|--------------------|----------------|
| 1  | Nền tảng cơ chế   | Theo WAL           | Theo trigger   |
| 2  | Có thể sao lập đến các máy con khác                         | Có                 | Có             |
| 3  | Đồng bộ toàn bộ dữ liệu (đầy đủ các đối tượng)              | Có                 | Có             |
| 4  | Chỉ đồng bộ một lần   | không              | không          |
| 5  | Cấu hình hệ thống: máy chủ và máy khách phải giống như nhau | Có                 | Không          |

Trong các phiên bản ArcGIS 8.3 đến ArcGIS 9.1, mô hình quản lý theo phiên bản được giới thiệu. Khi đó, một phiên bản cha sẽ tạo ra các phiên bản con. Và người sử dụng được phân quyền truy cập đến từng phiên bản. Khi hoàn thành các thao tác cập nhật, người sử dụng sẽ sử dụng các thao tác kiểm tra để đưa dữ liệu đồng bộ với phiên bản cha. Và người sử dụng có thể check out dữ liệu thành personal geodatabase nhằm phục vụ việc cập nhật không kết nối với máy chủ. Tuy nhiên, giới hạn về công nghệ của phương pháp này là người quản trị chỉ kiểm tra được một lần duy nhất. Do đó, nếu dữ liệu được hiệu chỉnh nhiều lần, chúng ta phải tạo nhiều bộ dữ liệu đồng bộ checkout.

2.3.2.2. Mô hình đồng bộ dữ liệu hai chiều

Trong mô hình này, người sử dụng có thể đồng bộ nhiều lần mà không cần tạo dữ liệu đồng bộ checkout nhiều lần. Những thay đổi sẽ được chuyển từ máy cha sang máy con và ngược lại. Cơ chế này phù hợp với mô hình hai tổ chức hoặc hai phòng ban (ví dụ: một ở Hà Nội và một ở TP.Hồ Chí Minh) cần chia sẻ chung dữ liệu.

2.3.2.3. Mô hình đồng bộ dữ liệu một chiều

Mô hình đồng bộ một chiều tương tự mô hình đồng bộ hai chiều là chúng ta có thể gửi dữ liệu nhiều lần sau khi đồng bộ. Tuy nhiên, dữ liệu chỉ gửi được một chiều (chiều từ máy chủ chứa dữ liệu “cha” đến máy chủ chứa dữ liệu “con”). Mô hình phù hợp với các nơi cung cấp dữ liệu chuẩn cho các đơn vị khác.

2.3.2.4. Một số vấn đề trong hệ thống Geodatabase khi đồng bộ

Khi sử dụng công nghệ ArcGIS, các cơ chế đồng bộ dữ liệu có thể thực thi trên môi trường mạng cục bộ, webservices, phát hành an toàn

trên mạng Internet và cũng có thể thông qua việc sử dụng tập tin XML khi hai CSDL không được kết nối trực tiếp.

Cụ thể hơn, các mô hình đồng bộ một chiều và hai chiều phải được thực thi trên hệ thống CSDL ArcSDE (bao gồm các phiên bản geodatabase enterprise, new workgroup và cả personal ArcSDE cài đặt trên hệ thống SQL Server Express). Và do cơ chế đồng bộ được ArcSDE quản lý, nên việc đồng bộ không quan tâm đến các nền phần mềm hệ quản trị CSDL khác nhau khi tham gia đồng bộ. Đối với đồng bộ check-out/checkin, CSDL đồng bộ “cha” phải là một CSDL geodatabase ArcSDE và CSDL “con” hoặc sẽ là ArcSDE geodatabase hoặc sẽ là personal geodatabase.

Điều kiện sử dụng đồng bộ dữ liệu là:

- Người sử dụng phải có quyền ghi dữ liệu.
- Dữ liệu phải được đăng kí phiên bản (không được chọn tham số: move edits to base).
- Đối với đồng bộ một chiều và hai chiều: mỗi dữ liệu phải có trường GlobalID và tất cả dữ liệu phải lưu trữ cùng một hệ tọa độ.

- Về bản quyền: sử dụng bản quyền ArcInfo desktop cho các máy khách và ArcSDE, ArcGIS Server Advanced để cung cấp chức năng Web ADF trên máy chủ. Và các phần mềm để hỗ trợ về hạ tầng mạng, như quản trị về bức tường lửa.

### 2.3.2.2. So sánh về đồng bộ dữ liệu giữa hệ thống ArcSDE và các hệ quản trị CSDL

ArcSDE sử dụng các cơ chế đồng bộ dữ liệu không phải của các hệ quản trị CSDL. Theo đó, nếu sử dụng cơ chế đồng bộ của hệ quản trị hai cơ sở dữ liệu trên hệ thống CSDL geodatabase của ESRI thì người sử dụng sẽ gặp những khó khăn sau:

- Khó khăn về yêu cầu cần hiểu rõ từng bảng trong hệ thống dữ liệu geodatabase.
- Khó khăn về cơ chế chỉ lọc dữ liệu cần đồng bộ và cấu trúc dữ liệu trong CSDL.
- Không thể đồng bộ dữ liệu lịch sử.

- Không thể sử dụng đồng bộ giữa hai hệ quản trị CSDL khác nhau.

- ArcGIS hoàn toàn không hỗ trợ việc đồng bộ này.

Tuy nhiên, việc đồng bộ CSDL trên các hệ quản trị sẽ có những ưu điểm và cũng là những hạn chế khi đồng bộ theo cơ chế ArcSDE không thực hiện, đó là:

- Có thể đồng bộ các bảng dữ liệu không được versioned.
- Có thể đồng bộ toàn bộ CSDL.
- Có thể cấu hình để đồng bộ hoàn toàn.

Ngoài ra, hiện tại, việc đồng bộ dữ liệu trên ArcSDE vẫn còn bị hạn chế khi thay đổi về topology, network và feature dataset, cũng như việc thêm các lớp mô tả quan hệ. Mặt khác, đối với các dự án nhỏ, việc sử dụng đồng bộ dữ liệu theo cơ chế geodatabase ArcSDE sẽ cần một khoản đầu tư ban đầu rất lớn về chi phí bản quyền.

### 2.3.3. Cơ chế đồng bộ dữ liệu không gian trên các thiết bị di động

#### 2.3.3.1. Sơ lược về ngôn ngữ truy vấn không gian

Ngôn ngữ truy vấn không gian (Spatial SQL) được sử dụng nhiều trong các hệ quản trị CSDL có hỗ trợ không gian như Postgres/PostGIS, Oracle, SQL Server 2008. Hơn thế nữa, spatial SQL hiện nay đã trở thành chuẩn và được nhiều gói phần mềm GIS hỗ trợ, như MapXtreme, Spatialite. Với tập lệnh trang bị để xử lý đối tượng hình học và đối tượng địa lý, chúng ta hoàn toàn có thể thêm, sửa các đối tượng. Hơn thế nữa, chúng ta được trang bị các câu lệnh để thực hiện việc phân tích không gian, chồng lớp, buffer, ...

#### 2.3.3.2. Lập trình tương tác với hệ quản trị CSDL không gian Postgres/PostGIS

Với hệ thống Postgres/PostGIS hoàn toàn mở, nhà phát triển được trang bị thêm các thư viện lập trình tương tác. Đó là các thư viện mã



nguồn mở Npgsql. Bằng cách sử dụng các thư viện Npgsql, chúng ta có thể kết nối đến hệ quản trị CSDL Postgres/PostGIS. Dưới đây là một số minh họa như sau trên môi trường phát triển ứng dụng VB.NET 2005:

\* Lệnh khai báo thư viện:

```
Imports Npgsql
```

\* Các lệnh khai báo đối tượng kết nối:

```
Public NpgConn As Npgsql.NpgsqlConnection
```

```
Public NpgCmd As Npgsql.NpgsqlCommand
```

\* Lệnh tạo đối tượng kết nối đến máy chủ:

```
conNectStr = "data source=" & oServer &
";initial catalog=CSDLQuanHuyen;persist security info=False;user id=sa; password = ;packet size=4096"
```

```
dbCon = New NpgsqlConnection(conNectStr)
```

```
dbCon.Open()
```

```
Dim cmd_SqlCommand As New NpgsqlCommand(strSql, dbCon)
```

2.3.3.3. Lập trình tương tác với hệ quản trị CSDL không gian SQLite/Spatialite

Hệ quản trị CSDL theo mô hình tập tin SQLite được phát triển và sử dụng trên các hệ thống di động. Trong đó, hiện nay, hầu hết các thiết bị di động tiên tiến đều sử dụng như: iPhone, iPad và các hệ thống sử dụng hệ điều hành Android. Spatialite là một phần mở rộng của hệ SQLite có khả năng lưu trữ các đối tượng không gian và thực hiện các câu lệnh truy vấn không gian. Một cách đơn giản, vai trò của Spatialite được xem như tương tự với vai trò của thành phần mở rộng PostGIS trong hệ quản trị Postgres.

Tương tự như việc tương tác với hệ quản trị CSDL không gian Postgres/PostGIS, chúng ta cũng có thể tương tác với hệ quản trị CSDL không gian theo mô hình tập tin SQLite/Spatialite. Bằng việc sử dụng các gói thư

viện động

Chúng ta cũng có các lệnh khai báo để việc truy xuất hệ SQLite/Spatialite như sau:

\* Lệnh khai báo thư viện:

```
Imports System.Data.SQLite
```

\* Các lệnh khai báo đối tượng kết nối:

```
Public sqliteCmd As SQLiteCommand
```

```
Public sqliteConn As SQLiteConnection
```

\* Lệnh tạo đối tượng kết nối đến máy chủ:

```
sqliteConn = New SQLiteConnection("data source=" & QHuyen.db & ";Version=3;")
```

```
sqliteConn.Open()
```

```
sqliteCmd = New SQLiteCommand("SELECT load_extension('libspatialite-1.dll');", sqliteConn)
```

\* Và thực hiện một câu truy vấn đơn giản như sau:

```
sqliteStr = "Select ten_phuong from RanhPhuong"
```

```
sqliteCmd = New SQLiteCommand(sqlStr, sqliteConn)
```

2.3.3.4. Đề xuất mô hình ứng dụng đồng bộ CSDL giữa các hệ Postgres/PostGIS thông qua trung gian SQLite/Spatialite

Mô hình ứng dụng đồng bộ CSDL giữa các hệ Postgres/PostGIS thông qua trung gian SQLite/Spatialite được thực hiện trên các nguyên tắc sau:

- Dữ liệu thêm mới và cập nhật trên mỗi máy chủ Postgres/PostGIS được đánh dấu bằng một trường cờ.

- Dữ liệu bị xóa sẽ được lưu vết trong một bảng dữ liệu xác định.

- Dữ liệu đã đồng bộ sẽ xóa trường cờ đánh dấu và xóa những dữ liệu được lưu vết xóa.

- Dữ liệu trung gian SQLite/Spatialite sẽ đóng vai trò so sánh các tập dữ liệu giữa các

server. Một server Postgres/PostGIS sẽ chuyển dữ liệu trung gian sang tập tin SQLite/Spatialite. Và chúng ta sẽ tạo các công cụ để thực hiện các so sánh như công thức

### 3. Kết quả nghiên cứu và đánh giá

Nền tảng kỹ thuật xây dựng các ứng dụng trên di động rất đa dạng, tuy nhiên chúng ta cần căn cứ vào nhân lực, chi phí xây dựng ứng dụng. Đặc biệt các ứng dụng GIS hiện nay thường được phát triển trên cả hai nền tảng WebGIS và Mobile GIS. Nếu ứng dụng chỉ có mục đích hiển thị dữ liệu, hiển thị thông kê và các lệnh truy vấn đơn giản như tìm kiếm các đối tượng bản đồ và đặc biệt khu vực xây dựng dự án có hệ thống viễn thông và internet ổn định phục vụ việc kết nối dữ liệu từ máy chủ tới thiết bị di động thì nên chọn phát triển theo công nghệ web application vì có chung ngôn ngữ lập trình với webgis là HTML, JavaScript và CSS đó đó thời gian phát triển nhanh chóng, tốn ít chi phí. Nếu ứng dụng cần can thiệp nhiều vào các thiết bị trong điện thoại như máy ảnh, gửi nhận tin nhắn, hệ thống định vị và muốn sử dụng dữ liệu ngoại tuyến thì nên chọn nền tảng native application. Tuy nhiên, nền tảng này có nhược điểm phát phát triển ứng dụng riêng rẽ trên từng nền tảng hệ điều hành khác nhau như IOS và Android dẫn tới chi phí cao, tốn thời gian, nhưng hiệu suất và tốc độ ứng dụng lại cao. Ứng dụng phát triển trên công nghệ hybrid khắc phục nhược điểm của cả hai công nghệ trên, cho phép truy cập vào các thiết bị trong điện thoại, lưu trữ dữ liệu ngoại tuyến, đa nền tảng.

Với đề xuất mô hình ứng dụng đồng bộ CSDL giữa các hệ Postgres/PostGIS thông qua trung gian SQLite/Spatialite, chúng ta hoàn toàn có thể xây dựng các ứng dụng hỗ trợ cập nhật thông tin trên thiết bị di động, đặc biệt các thiết bị có nền tảng Android. Thật vậy, sử dụng các gói trợ giúp hiển thị như gói gvSIG mini (cho hệ điều hành Android), chúng ta dễ dàng lập trình tạo ra các kết nối đến hệ thống CSDL. Khi đó, chúng ta có thể thêm tính năng cập nhật thông tin các điểm và lưu vào tập tin CSDL SQLite. Sau đó, tập tin SQLite sẽ được đồng bộ với máy chủ.

### 4. Kết luận

Đồng bộ dữ liệu là một công việc quan trọng. Trong báo cáo này, các tác giả chứng minh khả năng hình thành một hệ thống đồng bộ dữ liệu không gian đơn giản. Qua đó, chúng ta có thể ứng dụng trong các hệ thống GIS cấp tỉnh thành và quận huyện. Hệ thống đồng bộ dữ liệu còn nên phát triển thêm và thích ứng với nhiều môi trường như: hệ thống thiết kế đạt chuẩn dữ liệu cao, môi trường dữ liệu truy xuất nhiều, môi trường dữ liệu phân tán...

Bài báo là một phần kết quả nghiên cứu của đề tài TN18/CO5 thuộc chương trình Khoa học và Công nghệ phục vụ phát triển kinh tế - xã hội vùng Tây Nguyên trong liên kết vùng hội nhập quốc tế 2018-2020.○

### Tài liệu tham khảo

[1]. M. N. Duc, “Trắc Địa Ảnh, Viễn Thám Và GIS,” 9 9 2013. [Online]. Available: <https://sites.google.com/site/tdgisvaavgk55/2-tin-tuc-cong-nghe-gis/5motsovandhevedhong-bodulieukhonggian>.

[2]. B. T. Thoa, “So sánh Native App, Web App, Hybrid App và Website,” 17 8 2019. [Online]. Available: <https://www.techblog.vn/so-sanh-native-app-web-app-hybrid-app-va-website>.

[3]. M1kma, “Map Services and Web GIS architecture,” 5 1 2018. [Online]. Available: <https://dev.solita.fi/2018/01/05/gis-services.html>.

[4]. B. T. t. v. t. t. Cục tin học hóa, “Tiêu chuẩn kỹ thuật về UDCNTT trong CQNN: Tiêu chuẩn WFS - Giao thức dịch vụ đặc tính web (39),” 29 9 2014. [Online]. Available: <http://www.aita.gov.vn/tieu-chuan-k%E1%BB%B9-thuat-ve-udcntt-trong-cqnn-tieu-chuan-wfs-giao-thuc-dich-vu-dac-tinh-web-39>.○

(Xem tiếp trang 64)