

MÔ HÌNH ĐA CHIỀU CHO DỮ LIỆU BLOCKCHAIN DỰA TRÊN QUY TRÌNH ETL

Trần Thị Phương Chi, Nguyễn Quang Hưng

Khoa Công nghệ Thông tin, Trường Đại học Khoa học, Đại học Huế

Email: phuongchi0910@husc.edu.vn , nqhung@husc.edu.vn

Ngày nhận bài: 6/6/2022; ngày hoàn thành phản biện: 12/6/2022; ngày duyệt đăng: 22/6/2022

TÓM TẮT

Sự gia tăng nhanh chóng dữ liệu blockchain thúc đẩy việc triển khai các công cụ phân tích dữ liệu. ETL là quy trình trích xuất, chuyển đổi và tải dữ liệu từ nhiều nguồn dữ liệu khác nhau. Mục tiêu của bài báo này đó là thực hiện một quy trình có thể lưu trữ dữ liệu từ việc chiết xuất dữ liệu blockchain vào mô hình dữ liệu đa chiều tạo sự thuận tiện cho việc phân tích và truy vấn dữ liệu.

Từ khóa: Blockchain, ETL, mô hình đa chiều.

1. MỞ ĐẦU

Với sự gia tăng nhanh chóng dữ liệu từ công nghệ blockchain, việc triển khai các công cụ nhằm chiết xuất và phân tích dữ liệu đóng vai trò ngày càng cấp thiết. Đây chính là động lực để các nhà nghiên cứu hình thành một quy trình khai thác dữ liệu blockchain với mục tiêu tạo ra hệ thống linh hoạt và thuận tiện cho việc trích xuất, phân cụm các địa chỉ và tạo ra dữ liệu hữu dụng cho bất kỳ hình thức phân tích dữ liệu. Từ những nghiên cứu gần đây và với nhu cầu trích xuất dữ liệu một cách hiệu quả, mục tiêu của bài báo này đó là thực hiện một quy trình có thể lưu trữ dữ liệu từ việc chiết xuất dữ liệu blockchain vào cơ sở dữ liệu tạo sự thuận tiện cho việc phân tích và truy vấn dữ liệu.

Để đạt được mục tiêu đó, chúng tôi đã đưa ra ý tưởng xây dựng quy trình để trích xuất dữ liệu blockchain, chuyển đổi và nạp dữ liệu vào trong một cơ sở dữ liệu quan hệ, nhằm tạo điều kiện thuận lợi cho việc phân tích dữ liệu sau này. Dựa theo đó, chúng tôi đã thiết kế và triển khai hệ thống như một quy trình ETL (trích xuất, chuyển đổi, tải dữ liệu).

Quy trình ETL [5] cho phép tự động hóa hoạt động lựa chọn dữ liệu, thu thập dữ liệu và điều hòa dữ liệu từ kho dữ liệu và tạo ra dữ liệu đầu ra ở định dạng tốt nhất cho quá trình xử lý tiếp theo hoặc cho doanh nghiệp. Quy trình ETL trích xuất một loạt

150 khối cùng một lúc (các giao dịch trong một ngày) từ chuỗi khối. Giai đoạn chuyển đổi chuyển đổi dữ liệu blockchain sang cấu trúc dữ liệu có cấu trúc ví dụ như cấu trúc các bảng trong cơ sở dữ liệu SQL. Cuối cùng là giai đoạn tải dữ liệu đã được chuyển đổi vào trong cơ sở dữ liệu đích và truy vấn trên dữ liệu đã được nạp vào cơ sở dữ liệu đích

Trong bài báo này, chúng tôi trình bày kiến thức nền tảng quy trình ETL và cấu trúc dữ liệu của blockchain ở phần 2. Ở phần 3 là một số hướng tiếp cận trong việc ứng dụng ETL trong việc trích xuất, chuyển đổi và nạp dữ liệu từ nguồn dữ liệu blockchain. Trên cơ sở đó chúng tôi đưa ra phương pháp luận với hướng tiếp cận đa chiều trên dữ liệu chuỗi khối Ethereum. Trong phần minh họa cho hướng tiếp cận đa chiều, ở phần 4 chúng tôi sử dụng công cụ đặc tả Ethereum cụ thể là Ethereum ETL và OpenEthereum cho việc đồng bộ hóa dữ liệu được chiết xuất. Cuối cùng là kết luận và hướng phát triển của bài báo được đưa ra.

2. QUY TRÌNH ETL VÀ CẤU TRÚC DỮ LIỆU BLOCKCHAIN

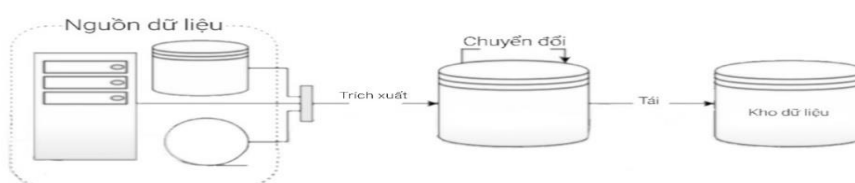
Hệ thống cho việc chiết xuất và phân tích dữ liệu chuỗi khối dựa trên hai giải pháp công nghệ cho phép xử lý dữ liệu hiệu quả và có thể mở rộng hệ thống. Đặc biệt, chúng tôi sử dụng quy trình ETL hướng tiếp cận đa chiều cho dữ liệu blockchain. Để rõ ràng hơn, chúng tôi mô tả các khái niệm cơ bản của hai công nghệ này ở phần dưới đây.

2.1 Kho dữ liệu và ETL

Để hiểu được kho dữ liệu và quy trình ETL trước hết cần hiểu được khái niệm tri thức nghiệp vụ. Tri thức nghiệp vụ BI (Business Intelligence) là quá trình thu thập dữ liệu thô hoặc dữ liệu kinh doanh và chuyển đổi thành thông tin hữu ích và có ý nghĩa hơn. Các dữ liệu thô là các hồ sơ của các giao dịch hàng ngày của một tổ chức như tương tác với khách hàng, quản lý tài chính và quản lý nhân viên,...vv. Những dữ liệu này sẽ được sử dụng cho việc báo cáo, phân tích, khai thác dữ liệu, hay dự đoán. Một kho dữ liệu (Data warehouse) là một cơ sở dữ liệu được thiết kế cho truy vấn và phân tích hơn là cho xử lý giao dịch. Các kho dữ liệu được xây dựng bằng cách tích hợp dữ liệu từ nhiều nguồn. Điều này cho các phép công ty hoặc tổ chức hợp nhất dữ liệu từ nhiều nguồn và tách khối lượng công việc phân tích từ khối lượng công việc giao dịch.

Như vậy kho dữ liệu là cơ sở của tri thức nghiệp vụ và ETL là cơ sở của kho dữ liệu. ETL là từ viết cho Extract - Transform - Load đó là việc trích xuất - chuyển đổi - tải. Đây là quá trình trích xuất dữ liệu từ nhiều nguồn dữ liệu khác nhau, sau đó chuyển đổi các dữ liệu không đồng nhất thành một dạng đồng nhất và nạp vào kho dữ liệu để sử dụng tùy theo mục đích của tổ chức cụ thể:

- Extraction (Trích xuất) Dữ liệu nguồn từ các hệ thống nguồn khác nhau được trích xuất có thể ở nhiều định dạng khác nhau như nhiều loại cơ sở dữ liệu, từ file excel hay từ file thô. Điều quan trọng là phải trích xuất dữ liệu từ các hệ thống nguồn khác nhau và lưu trữ nó vào khu vực dàn dựng trước tiên chứ không phải trực tiếp vào kho dữ liệu vì dữ liệu được trích xuất có nhiều định dạng khác nhau.
- Transformation (Chuyển đổi) Tại đây dữ liệu thô trải qua quá trình xử lý dữ liệu. Dữ liệu được chuyển đổi và hợp nhất cho những trường hợp phân tích sau này.
- Loading (Tải dữ liệu) Đây là quá trình đẩy dữ liệu sau khi được chuyển đổi cuối cùng vào kho dữ liệu. Việc tải dữ liệu vào một nơi lưu trữ tập trung cho phép các nhà phát triển có thể xây dựng ứng dụng và người dùng cuối có thể ra quyết định dựa trên dữ liệu và ứng dụng đó.



Hình 1. Quy trình ETL

2. 2 Cấu trúc dữ liệu blockchain

Blockchain về cơ bản là một cơ sở dữ liệu được phân phối của các bản ghi hoặc sổ cái công khai của tất cả các giao dịch hoặc các sự kiện kỹ thuật số đã được thực hiện và chia sẻ giữa các bên tham gia. Mỗi giao dịch trong sổ cái công khai được xác minh bởi sự đồng thuận của đa số những người tham gia trong hệ thống. Ngoài ra, một khi đã nhập thông tin không bao giờ có thể bị xóa bỏ

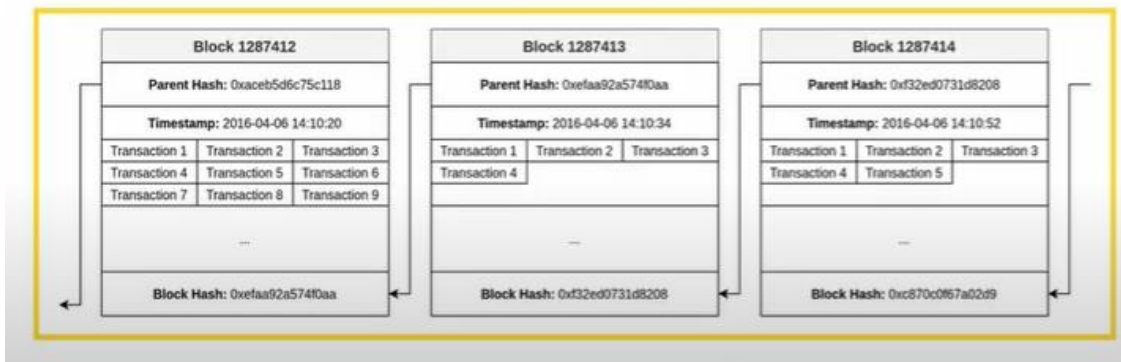
Công nghệ Blockchain sử dụng mã hóa Public key và hàm hash để đảm bảo tính minh bạch, toàn vẹn và riêng tư của dữ liệu; sử dụng mỗi một nút trong mạng như một Client và cũng là một Server để lưu trữ bản sao ứng dụng, và áp dụng các nguyên tắc đối với các nút tham gia vào hệ thống đều phải tuân thủ sự đồng thuận.

Cấu trúc dữ liệu Blockchain bao gồm 2 phần chính:

- Khối (Block): các khối này chứa dữ liệu
- Chuỗi (Chain): bao gồm các khối trên liên kết với nhau tạo thành chuỗi

Mỗi khối bao gồm các thành phần chính: Giao dịch , Mã Hash của khối hiện tại (Mã hàm băm) và Mã Previous Hash (mã Hash khối trước đó) và dấu thời gian (Timestamp).

- Một khối chứa một tập hợp các giao dịch (transactions). Giao dịch là một sự thay đổi dữ liệu đã được thỏa thuận từ trạng thái này sang trạng thái khác. Các giao dịch trong một khối là “nguyên tử” có nghĩa là chúng không thể được chia thành các tập hợp nhỏ hơn
- Data (Dữ liệu): Các bản ghi dữ liệu đã được xác minh được bảo vệ bằng các thuật toán mã hóa phụ thuộc vào mỗi chuỗi khối.
- Mã Hash của khối hiện tại (Mã hàm băm): Đây là một chuỗi ký tự và số được tạo ngẫu nhiên không hoàn toàn giống nhau. Nó đại diện cụ thể cho khối và sử dụng một thuật toán mã hóa để mã hóa nó. Mã này được sử dụng để phát hiện các thay đổi trong khối. Mã là duy nhất, không trùng nhau.
- Mã Previous Hash (mã Hash khối trước đó): Nó được sử dụng để cho các khối liên kế biết khối nào ở phía trước và khối nào ở sau, để liên kết đúng với nhau. Tuy nhiên khối đầu tiên, bởi vì không có khối nào trước nó nên mã Hash của nó là một chuỗi số 0. Khối đầu tiên này được gọi là Genesis block tức là “Khối nguyên thủy” hay khối gốc.
- Dấu thời gian trong chuỗi khối(Timestamp) là giá trị ngày-giờ được lưu trữ trong một khối. Như đã đề cập trước đó, bất kỳ giao dịch nào đều không hoàn chỉnh nếu không có thông tin ngày và giờ. Trong blockchain, dấu thời gian cho biết khối đó được tạo vào thời điểm nào.



Hình 2. Cấu trúc dữ liệu của blockchain

3. ỨNG DỤNG QUY TRÌNH ETL CHO DỮ LIỆU BLOCKCHAIN

3.1 Các hướng nghiên cứu liên quan

Ngày nay một trong các mối quan tâm lớn của các nhà nghiên cứu là đề xuất các kỹ thuật phân tích dữ liệu blockchain.

Ali và Wrembel đã nghiên cứu các kỹ thuật trong việc thiết kế mô hình khái niệm và mô hình logic cho quy trình ETL [1]. Họ đã đánh giá các phương pháp tiếp cận hiện có và chỉ ra các hạn chế mà các phương pháp đang gặp phải. Dựa trên việc đánh giá họ đưa ra kết luận cần có mô hình tiêu chuẩn hóa cho quy trình ETL. Họ cũng nhấn mạnh tầm quan trọng của việc xây dựng khung ứng dụng cho các nhà phát triển ETL trong việc mô hình hóa và nghiên cứu các kỹ thuật hỗ trợ một cách hiệu quả.

Pinna và cộng sự đã sử dụng Petri Nets để mô hình hóa các giao dịch và địa chỉ hóa Bitcoin [2]. Họ mô hình hóa địa chỉ dưới dạng địa điểm và giao dịch dưới dạng chuyển tiếp. Các thử nghiệm đã được thực hiện bằng cách sử dụng dữ liệu chuỗi khối Bitcoin trong 1,5 năm đầu tiên. Với mô hình này, Pinna và các cộng sự phân cụm các địa chỉ riêng lẻ thành nhóm các địa chỉ thuộc cùng một thực thể. Petri Nets mô hình hóa dữ liệu Bitcoin và mang lại cơ hội lớn cho việc phân tích dữ liệu blockchain.

Galici và cộng sự [3] đã sử dụng các kỹ thuật kho dữ liệu truyền thống để thực thi một cách hiệu quả quy trình ETL cho chuỗi khối Bitcoin. Mục tiêu là một cơ sở dữ liệu quan hệ cho phép các truy vấn người dùng trên địa chỉ và các giao dịch Bitcoin. Dựa trên kỹ thuật kho dữ liệu truyền thống, Galici và cộng sự có thể phân tích dữ liệu, nghiên cứu các thay đổi trong hoạt động của người dùng trên mạng Bitcoin.

Medvedev và nhóm D5 [4] đã phát triển Ethereum ETL một công cụ mã nguồn mở cho việc trích xuất, chuyển đổi và tải dữ liệu chuỗi khối Ethereum. Việc trích xuất dữ liệu xuất phát từ một phần mềm nút Ethereum sau đó chuyển đổi và tải vào một cơ sở dữ liệu quan hệ.

3. 2 Phương pháp luận nghiên cứu

Các hướng nghiên cứu

Các nghiên cứu và thí nghiệm được trình bày trong bài báo này được tiến hành dựa trên phương pháp luận được dựa trên nền tảng Ethereum ETL. Trong bài báo này việc ứng dụng quá trình ETL cho dữ liệu blockchain cụ thể dựa trên nền tảng Ethereum ETL, chúng tôi tiến hành theo các bước như sau:

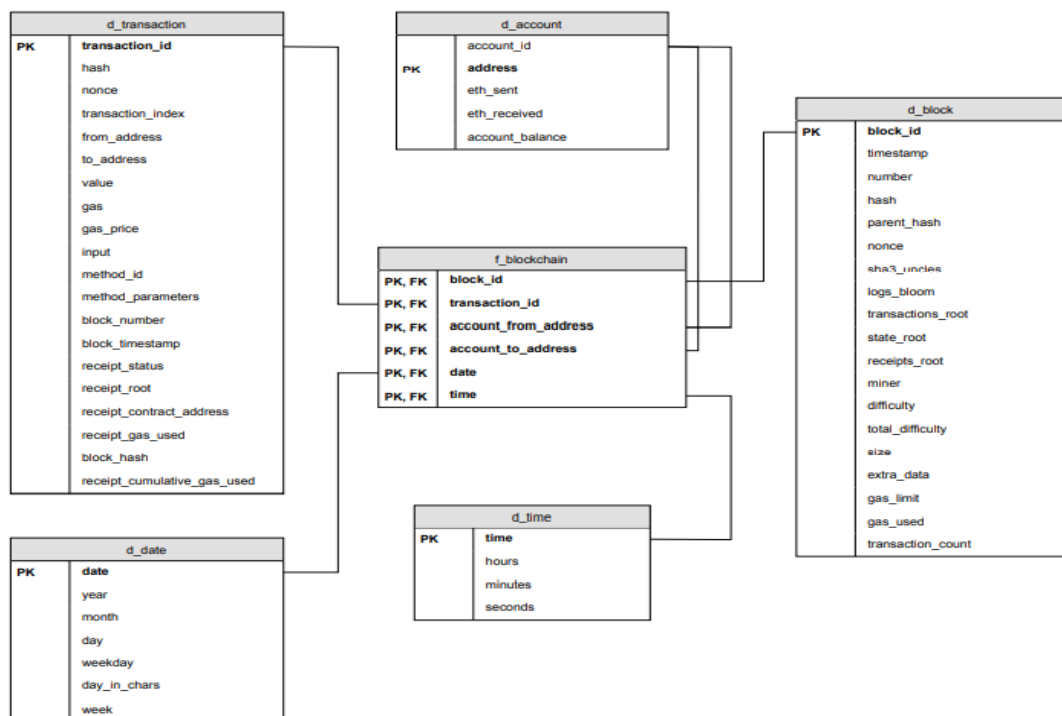
- Bước 1: Trích xuất dữ liệu blockchain từ nhiều nguồn dữ liệu.
- Bước 2: Chuyển đổi dữ liệu sang mô hình dữ liệu đa chiều cụ thể chuyển đổi lược đồ hai chiều: giao tác và bảng khối sang mô hình đa chiều.
- Bước 3: Tải dữ liệu đã được chuyển vào lược đồ hình sao.

- Bước 4: Truy vấn dữ liệu blockchain.

Bước đầu tiên trong việc thiết kế là khái niệm hóa về một lược đồ dữ liệu tương thích với phân tích các thuộc tính đa chiều. Tùy thuộc vào thực thể nào sẽ được trích xuất, các bảng sẽ được thiết lập cho quá trình trích xuất.

Trong phần nghiên cứu, chúng tôi giới hạn đối với thực thể được trích xuất từ các giao dịch và bảng khối, hai bảng tương ứng sẽ được triển khai tương ứng. Việc thiết kế trong giai đoạn này bắt nguồn từ các đặc tả của Ethereum ETL, công cụ được sử dụng để việc trích xuất dữ liệu. Mục tiêu ở đây là tạo một lược đồ cơ sở dữ liệu đa chiều bắt nguồn từ dữ liệu hai chiều. Điều này được thực hiện thông qua một chuỗi các bước biến đổi. Lược đồ cơ sở dữ liệu cho phân tích đa chiều bao gồm các giai đoạn trích xuất, chuyển đổi và tải dữ liệu vào lược đồ hình sao.

- Cốt lõi của quy trình này là dựa trên lược đồ hình sao. Lược đồ dữ liệu hình sao bao gồm sáu bảng được mô tả ở hình 4. Ở giữa là bảng dữ kiện bao gồm tất cả các khóa chính của các bảng đa chiều. Ở đây chúng ta có các chiều: d_transaction, d_block, d_date, d_time và d_account tương ứng với các bảng đa chiều. Trong đó chiều d_account có hai tham chiếu, một tham chiếu từ account_to và một tham chiếu từ account_from.



Hình 3. Lược đồ sao: bao gồm bảng dữ kiện và năm chiều lưu trữ dữ liệu blockchain

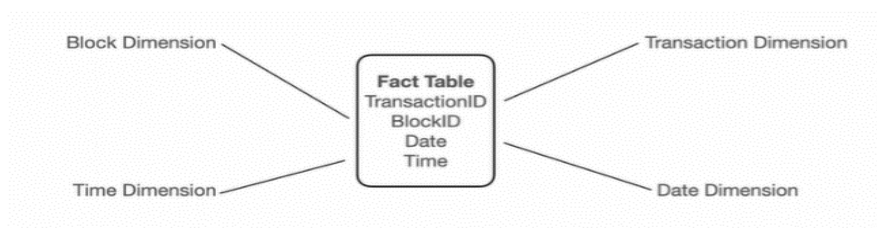
3.3. Quy trình ETL cho dữ liệu blockchain

3.3.1 Trích xuất dữ liệu

Giai đoạn trích xuất là bước đầu tiên của quy trình ETL liên quan đến việc thu thập dữ liệu thô từ một hoặc nhiều nguồn dữ liệu. Hướng tiếp cận của bài báo dựa trên nền tảng Ethereum ETL do đó quá trình trích xuất dữ liệu yêu cầu quyền truy cập vào máy khách Ethereum. Chúng tôi sử dụng ETL, một công cụ mã nguồn mở và công cụ này bao gồm các tập lệnh Python trích xuất dữ liệu từ các giao dịch và khối từ một nút Ethereum.

3.3.2 Chuyển đổi

Dữ liệu sau khi được trích xuất từ dữ liệu blockchain, dựa vào cấu trúc dữ liệu của blockchain là một lược đồ hai chiều trong đó các giao dịch và các bảng khối được chuyển đổi thành một lược đồ đa chiều. Mục tiêu ở đây là tạo dữ liệu mới từ dữ liệu đã có sẵn. Điều này được thực hiện thông qua một loạt các bước chuyển đổi cú pháp và ngữ nghĩa.



Hình 4. Mô hình dữ liệu đa chiều

3.3.3 Tải dữ liệu

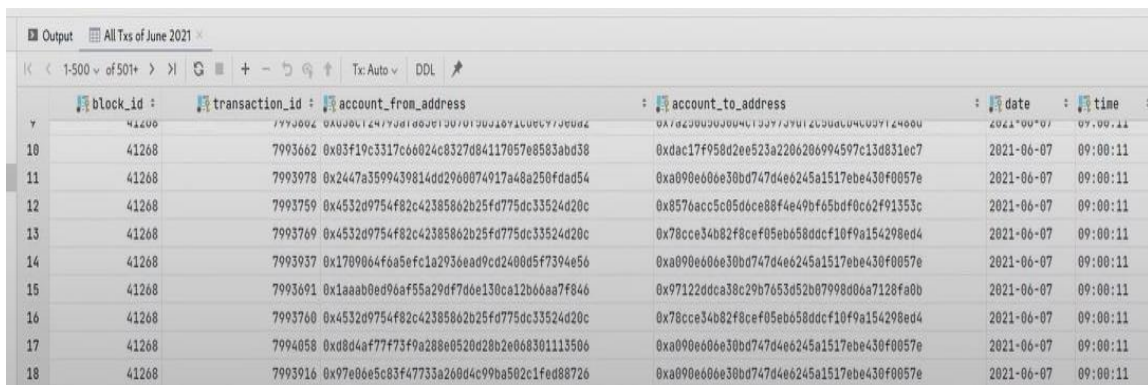
Giai đoạn cuối cùng là giai đoạn tải dữ liệu đã biến đổi và phân tích dữ liệu theo các chiều đã xác định trước đó. Quá trình tải dữ liệu vào lược đồ hình sao được tự động hóa và diễn ra khi tất cả các phép biến đổi đã được thực hiện. Mỗi bảng dữ liệu được chèn vào một chiều tương ứng. Khi dữ liệu đã được chèn vào các chiều tương ứng, các thuộc tính bảng dữ liệu được tạo bởi các thuộc tính khóa của các chiều ngoại trừ chiều account được tham chiếu kép – và một khóa tổng hợp được tạo ra bởi 2 khóa chính. Sau khi lược đồ hình sao được hình thành, việc truy vấn trên cơ sở dữ liệu với lược đồ hình sao có thể được tiến hành.

4. THỰC NGHIỆM

Để tiến hành thực thi ứng dụng quy trình ETL cho dữ liệu chuỗi khối blockchain, chúng tôi dựa trên nền tảng Ethereum ETL. Ethereum ETL [4] là một công cụ mã nguồn mở được phát triển bởi Evgeny Medvedev và D5 và bao gồm một tập các đoạn mã lệnh Python để chiết xuất và nạp dữ liệu vào cơ sở dữ liệu PostgreSQL hoặc

xuất ra dưới dạng file CSV. Quy trình Ethereum ETL bao gồm các bước sau: (1) thiết lập OpenEthereum, (2) chiết xuất các dữ liệu khối và các giao dịch vào cơ sở dữ liệu PostgreSQL, (3) chuyển đổi và nạp dữ liệu vào lược đồ hình sao và cuối cùng (4) là truy vấn dữ liệu.

```
SELECT date, ROUND(AVG(tx_count), 2) AS average
FROM (
    SELECT date, COUNT(*) AS tx_count
    from f_blockchain
    GROUP BY date, block_id) AS sub
GROUP BY date;
```



block_id	transaction_id	account_from_address	account_to_address	date	time
41268	7993662	0x03f19c3317c66024c8327d084117057e0583abd38	0xdac17f958d2e523a2206206994597c13d831ec7	2021-06-07	09:00:11
41268	7993978	0x2447a3599439814dd2960074917a48a250fdad54	0xa090e606e30bd747d4e6245a1517ebe430f0057e	2021-06-07	09:00:11
41268	7993759	0x4532d9754f82c42385862b25fd775dc33524d20c	0x8576acc5c05d6ce88f4e49bf65bdf0c62f91353c	2021-06-07	09:00:11
41268	7993769	0x4532d9754f82c42385862b25fd775dc33524d20c	0x78cce34b82f8cef05eb658ddcf10f9a154298ed4	2021-06-07	09:00:11
41268	7993937	0x1709064f6a5efc1a2936ead9cd2400d5f7394e56	0xa090e606e30bd747d4e6245a1517ebe430f0057e	2021-06-07	09:00:11
41268	7993691	0x1aaab0ed96af55a29df7d6e130ca12b66aa7f846	0x97122ddca38c29b7653d52b07998d06a7128fa0b	2021-06-07	09:00:11
41268	7993760	0x4532d9754f82c42385862b25fd775dc33524d20c	0x78cce34b82f8cef05eb658ddcf10f9a154298ed4	2021-06-07	09:00:11
41268	7994058	0xd8d4af77f73f9a288e0520d28b2e068301113506	0xa090e606e30bd747d4e6245a1517ebe430f0057e	2021-06-07	09:00:11
41268	7993916	0x97e0e5c83f47733a260d4c99ba502c1fe088726	0xa090e606e30bd747d4e6245a1517ebe430f0057e	2021-06-07	09:00:11

Hình 5. Mô hình dữ liệu đa chiều và truy vấn dữ liệu

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Mục tiêu chính của bài báo là xây dựng quy trình ETL để trích xuất dữ liệu blockchain, chuyển đổi và nạp dữ liệu vào trong một cơ sở dữ liệu quan hệ, tạo điều kiện thuận lợi cho việc phân tích dữ liệu sau này. Để triển khai được quy trình này chúng tôi tập trung vào dữ liệu chuỗi khối Ethereum cho nghiên cứu này. Bên cạnh đó, chúng tôi đã thiết kế một mô hình đa chiều sử dụng lược đồ hình sao là cơ sở cho việc truy vấn, phân tích dữ liệu. Cuối cùng chúng tôi đánh giá hướng tiếp cận trên thực nghiệm Research Coin. Tất cả mã lệnh là mã nguồn mở để khuyến khích phát triển thêm quy trình luồng công việc và cung cấp đóng góp vào lĩnh vực nghiên cứu phân tích dữ liệu blockchain. Các công nghệ hỗ trợ bổ sung cho các công nghệ chuỗi khối hơn nữa sẽ là chủ đề của nghiên cứu và thử nghiệm trong tương lai. Tuy nhiên, việc được đề xuất quy trình ETL cho blockchain đặt nền tảng quan trọng để xây dựng và hy vọng sẽ thúc đẩy việc nghiên cứu và phát triển thêm các quy trình ETL và các mô hình đa chiều trong lĩnh vực phân tích dữ liệu blockchain.

TÀI LIỆU THAM KHẢO

- [1]. Ali, S., Wrembel, R. (2017). From conceptual design to performance optimization of etl workflows: current state of research and open problems. *The VLDB Journal* 26. [Online], Website: <https://doi.org/10.1007/s00778-017-0477-2>
- [2]. Pinna, A., Tonelli, R., Orrú, M., Marchesi, M (2018). A petri nets model for blockchain analysis. *The Computer Journal* 61(9), pp. 1374–1388
- [3]. Galici, R., Ordile, L., Marchesi, M., Pinna, A., Tonelli, R. (2020). Applying the etl process to blockchain data. *Prospect and findings. Information* 11, pp.204.
- [4]. Medvedev, E., the D5 team (2018). Ethereum etl. [Online], Website: <https://github.com/blockchain-etl/ethereum-etl>
- [5]. IBM cloud Education (2020). *ETL (Extract, Transformation, Load)*. Website: <https://www.ibm.com/cloud/learn/etl>
- [6]. Roberta Galici, Laura Ordile, Michele Marchesi, Andrea Pinna, Roberto Tonelli [2020]. Applying the ETL Process to Blockchain Data. *Prospect and Findings. Information* 2020. Website: <https://doi.org/10.3390/info11040204>

MULTIDIMENSIONAL MODEL FOR BLOCKCHAIN DATA BASED ON ETL

Tran Thi Phuong Chi, Nguyen Quang Hung

Faculty of Information Technology, University of Sciences, Hue University

Email: phuongchi0910@husc.edu.vn , nghung@husc.edu.vn

ABSTRACT

Abstract: The rapid increase of blockchain data promotes the implementation of data analytics tools. ETL is the process of extracting, transforming, and loading data from a variety of data sources. The goal of this paper is to implement a process that can store data from extracting blockchain data into a specific database with a multi-dimensional data model to facilitate data analysis and queries.

Keywords: Blockchain, ETL, multidimensional model



Trần Thị Phương Chi sinh ngày 09/10/1983 tại Thừa Thiên Huế. Bà tốt nghiệp cử nhân ngành Công nghệ thông tin năm 2005 và thạc sĩ chuyên ngành Khoa học máy tính năm 2009 tại trường Đại học Khoa học, Đại học Huế. Từ năm 2013 đến nay, bà công tác tại khoa Công nghệ Thông tin, Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: Cơ sở dữ liệu.



Nguyễn Quang Hưng sinh ngày 3/11/1979. Ông tốt nghiệp cử nhân ngành Công nghệ Thông tin năm 2001 và thạc sĩ chuyên ngành Khoa học Máy tính tại trường Đại học Khoa Học, Đại học Huế vào năm 2005. Hiện ông công tác tại Khoa Công nghệ Thông tin, trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: mạng máy tính, truyền thông dữ liệu.