MÔ HÌNH LƯU TRỮ TÀI LIỆU ĐIỆN TỬ DỰA TRÊN CHUỖI KHỐI TRONG NGỮ CẢNH CHUYỂN ĐỔI SỐ TẠI CƠ SỞ GIÁO DỤC[1]

ThS Hồ Thị Ngọc, ThS Bùi Vũ Bảo Khuyên, ThS Nguyễn Tấn Công
Khoa Thư viện-Thông tin học, Trường ĐHKHXH&NV-ĐHQG Tp.Hồ Chí Minh

Tóm tắt: Trong ngữ cảnh chuyển đổi số hiện nay, nhiều nghiên cứu nỗ lực phát triển các giải pháp lưu trữ và trao đổi tài liệu học thuật hiệu quả. Trong đó, hướng tiếp cận phổ biến là triển khai các trung tâm dữ liệu lưu trữ tập trung trên nền tảng đám mây. Tuy nhiên, hướng tiếp cận này tồn tại một số hạn chế về chi phí bảo trì, không gian lưu trữ lớn, nguy cơ về tính nguyên bản và bảo mật khi chia sẻ tài liệu trên môi trường mạng. Để giải quyết các vấn đề trên, chúng tôi đề xuất một mô hình hiệu quả sử dụng kết hợp công nghệ chuỗi khối, hợp đồng thông minh và hệ thống chia sẻ tập tin phân tán. Những thế mạnh của các công nghệ này giúp tạo ra một phương thức điều khiển truy cập thông minh, từ đó hình thành chính sách phân phối và truy cập tài liệu tin cậy cho các bên tham gia như các tác giả, người dùng tin và trung tâm dữ liệu.

Từ khóa: Hệ thống lưu trữ; chuỗi khối; hợp đồng thông minh; chuyển đổi số.

BLOCKCHAIN-BASED ELECTRONIC DOCUMENT STORAGE MODEL IN THE CONTEXT OF DIGITAL TRANSFORMATION IN EDUCATIONAL INSTITUTIONS

Abstract: In the current digital transformation context, many studies make great efforts to develop efficient solutions for storing and exchanging scholarly documents. In which, the popular approach is to deploy centralized data centers on the cloud-based platform. However, this approach has some limitations in terms of maintenance cost, large storage space, originality and security risks when sharing documents on a network environment. To eliminate these problems, we propose an efficient model that uses a combination of technologies such as the blockchain, smart contract, and distributed file sharing system. Strengths of these technologies contribute to create an intelligent access control method from which to form trusted document access and distribution policies for stakeholders such as authors, users, and data centers.

Keywords: Storage system; Blockchain; smart contracts; digital transformation.

GIỚI THIỆU

Với bối cảnh phát triển năng động ngày nay, mọi hoạt động về kinh tế, sản xuất và giáo dục đều được vận hành trên môi trường số. Điều này dẫn đến nhu cầu phát triển những hệ thống lưu trữ và truy xuất dữ liệu hiệu quả khi dữ liệu số trở nên đa dạng trong mọi ngành nghề và lĩnh vực. Đặc biệt trong khía cạnh giáo dục, tài liệu học thuật không những khổng lồ về số lượng mà còn đa dạng về chủng loại. Bên cạnh đó, chuyển đổi số trong vận hành các cơ sở học thuật là nhu cầu cần thiết trong bối cảnh phát triển của xã hội [Gurdish Sandhu, 2018]. Tuy nhiên, những công nghệ lưu trữ tài liệu điện tử hiện tại tồn tại nhiều hạn chế khi phần lớn được triển khai tại các trung tâm dữ liệu tập trung hoặc dựa vào nền tảng đám mây (Cloud-based platform) [Jin Han và cộng sự, 2021]. Các hướng tiếp cận này còn tồn tại nhiều vấn đề về chi phí bảo trì, yêu cầu không gian lưu trữ lớn, ẩn chứa nhiều nguy cơ về an toàn thông tin, quyền riêng tư và khả năng bảo mật [Sun Y và cộng sự, 2014].

Hiện nay, công nghệ chuỗi khối (Blockchain) [Zibin Zheng và cộng sự, 2017] đã trở thành một trong những chủ đề nghiên cứu quan trọng, không chỉ áp dụng trong ngành tài chính mà còn lan rộng sang các lĩnh vực công nghệ khác như giáo dục, lưu trữ, chăm sóc sức khỏe và sản xuất,... Chuỗi khối mang nhiều ưu điểm khi tạo ra các giao dịch giữa các bên liên quan như khi loại bỏ các đối tượng trung gian, tính phi tập trung, minh bạch, bền vững và bảo

[1] Nghiên cứu được tài trợ bởi Trường Đại học KHXH&NV, Đại học Quốc gia Tp. Hồ Chí Minh trong khuôn khổ Đề tài mã số T2022-25.

mật cao [Wubing Chen và cộng sự, 2018]. Ngoài ra, một kỹ thuật tự động hóa giao dịch được gọi là Hợp đồng thông minh (Smart Contract) cũng sớm được phát triển và vận hành trên chuỗi khối [Shafaq Khan và cộng sự, 2021]. Hợp đồng thông minh có thể hiểu là các chương trình có thể tự chủ xác minh và thực hiện các thỏa thuận trong hợp đồng khi các điều khoản định trước được thỏa mãn, giúp giảm chi phí quản lý các giao dịch. Do đó, tích hợp chuỗi khối và hợp đồng thông minh giúp hình thành chính sách phân phối và truy cập tài liệu tin cậy, cũng như khả năng tự động hóa các tác vụ trao đổi tài liệu.

Đối lập với cơ chế lưu trữ tập trung thì hướng tiếp cận lưu trữ tập tin phân tán ngày càng cho thấy nhiều ưu điểm vượt trội hơn như tăng tốc độ truyền tải dữ liệu, tăng tính bảo mật, giảm phụ thuộc vào máy chủ. Bên cạnh đó, khả năng lưu trữ tập tin phi tập trung dẫn đến khả năng mở rộng không gian lưu trữ cũng như tiết kiệm chi phí bảo trì hệ thống trong tương lai. Đại diện cho hướng tiếp cận này và được sử dụng rộng rãi hiện nay là hệ thống tệp liên hành tinh (Interplanetary File System-IPFS). IPFS cho phép thiết lập một hệ thống tập tin phân tán ngang hàng, kết nối tất cả các thiết bị máy tính với nhau [Athreya và cộng sự, 2021]. Chính vì thế, tích hợp IPFS giúp loại bỏ những vấn đề cố hữu như vấn đề băng thông, không gian lưu trữ và bảo mật trên những hệ thống lưu trữ tập tin hiện nay.

Từ nhận định trên, chúng tôi đề xuất một mô hình lưu trữ tài liệu học thuật hiệu quả sử dụng kết hợp công nghệ chuỗi khối và hệ thống phân phối tập tin phân tán. Giải pháp có khả năng lưu trữ tài liệu phân tán, truy cập tin cậy và chia sẻ quản lý tại các khoa và bộ môn. Điều này không chỉ giảm gánh nặng công việc cho bộ phận thư viện mà còn hiệu quả khi thu thập tài liệu. Đóng góp chính của chúng tôi bao gồm: Thứ nhất, phân tích quy tắc hoạt động của chuỗi khối, hợp đồng thông minh và IPFS; Thứ hai, thiết kế một mô hình lưu trữ tài liệu điện tử hiệu quả theo mô hình phân tán; Và cuối cùng, định nghĩa tập các giao thức truy cập dữ liệu dựa trên mô hình đề xuất.

1. NHỮNG CÔNG NGHỆ KỸ THUẬT VÀ NGHIÊN CỨU LIÊN QUAN

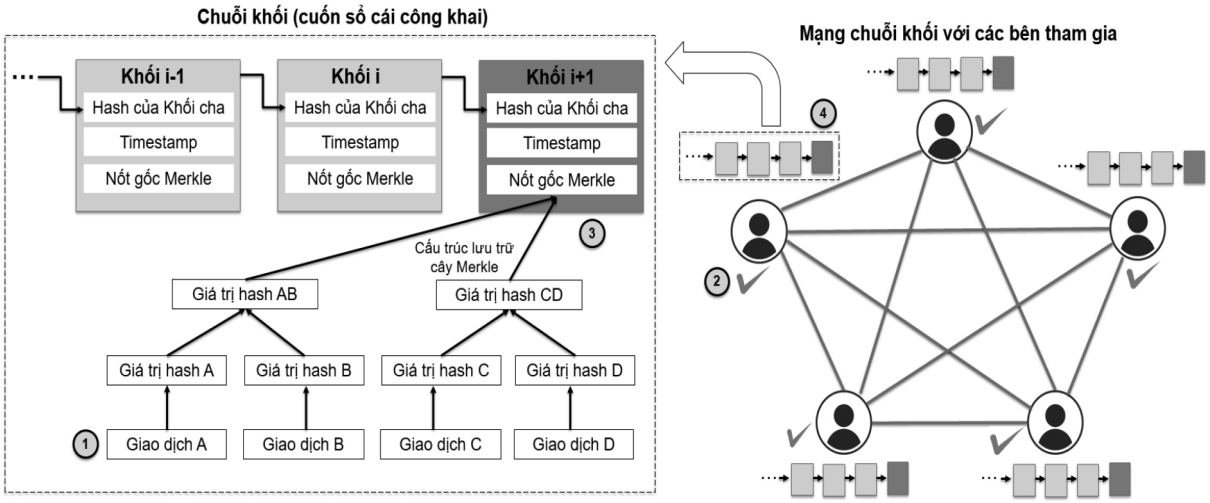
1.1. Chữ ký điện tử (Digital signature)

Cơ chế hoạt động của chuỗi khối dựa trên kỹ thuật chữ ký điện tử với nguyên lý hoạt động được minh họa trong Hình 1, giúp người nhận có thể kiểm tra một tài liệu là nguyên bản và hợp lệ từ người gửi. Đầu tiên, một cặp khóa duy nhất sẽ được tạo ra, trong đó khóa riêng (Private key) được giữ cẩn thận bởi người gửi và khóa công khai (Public key) được công bố cho tất cả người dùng khác. Tài liệu của người gửi đầu tiên sẽ được tính giá trị băm (Hash value) thông qua một thuật toán băm (Hash algorithm). Giá trị băm này sẽ được mã hóa dựa trên khóa riêng của người gửi. Giá trị đã mã hóa được đính kèm cùng với tài liệu và chuyển đến người nhận thông qua mạng. Khi nhận được, người nhận sẽ lấy giá trị mã hóa đính kèm thực hiện giải mã bởi khóa công khai mà người gửi công bố trước đó, từ đó thu được giá trị băm được giải mã. Bên cạnh đó, tài liệu gửi tới sẽ được thực hiện tính giá trị băm một lần nữa. Nếu giá trị băm tính lại từ tài liệu khớp với giá trị băm giải mã thì tài liệu này là nguyên bản từ người gửi.

Hình 1. *Nguyên lý hoạt động của chữ ký điện tử*

1.2. Chuỗi khối (Blockchain)

Từ khi được giới thiệu bởi Satoshi Nakamoto [Satoshi Nakamoto, 2008], cảm hứng sử dụng chuỗi khối cho thiết lập các giao dịch an toàn đã dần được áp dụng rộng rãi cho nhiều lĩnh vực trọng yếu [Zibin Zheng và cộng sự, 2017]. Về mặt kiến trúc, chuỗi khối bao gồm hai phần chính là mạng chuỗi khối (Blockchain network) và danh sách liên kết các khối hay còn gọi là sổ cái phân tán (Distributed ledger) như ở Hình 2. Trong mạng chuỗi khối, mỗi đối tượng tham gia hay còn gọi là nốt mạng (Node) liên kết với các nốt khác tạo thành một mạng liên kết và có thể phát sinh giao dịch với nhau. Mỗi nốt bên trong mạng đều lưu một bản sao sổ cái phân tán, là một danh sách liên kết các khối chứa thông tin giao dịch. Khi một khối mới được thêm vào chuỗi, nội dung của khối đó bao gồm thông tin các giao dịch đang thực hiện được lưu trữ hiệu quả dựa trên cấu trúc dữ liệu cây Merkle, cùng với thông tin về mốc thời gian (Timestamp) và giá trị băm được tính của khối trước đó [Qurotul Aini và cộng sự, 2021]. Bởi lưu liên kết dựa vào giá trị băm của khối trước, điều này tạo nên tính chất không thể sửa đổi của giao dịch. Giả sử ta sửa đổi giao dịch hay một khối đã tạo trong quá khứ, điều này sẽ dẫn đến giá trị băm của khối sửa đổi sẽ bị sai khác và không còn khớp với mã băm đã tính và lưu trữ ở các khối đã thêm vào sau khối đó.



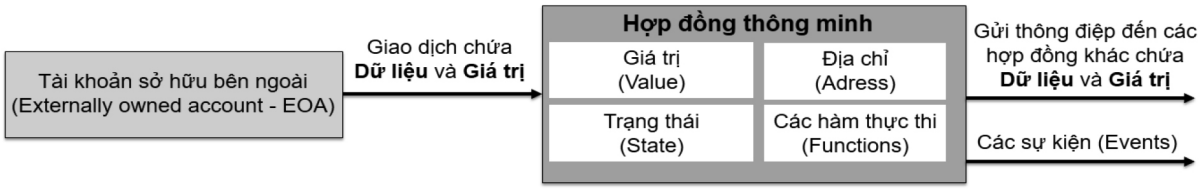
Hình 2. *Mô hình hóa hoạt động của chuỗi khối*

Về hoạt động, cơ bản chuỗi khối trải qua các bước chính như sau [Cao Minh Kiểm, 2019]: (1) các giao dịch khởi tạo bởi người gửi sẽ được tải lên mạng và thông báo đến tất cả các nốt tham gia. Thông điệp giao dịch bao gồm thông tin địa chỉ cũng là khóa công khai của người nhận và giá trị giao dịch. Thông điệp cũng được ký bởi khóa riêng tư của người gửi; (2) tất cả các nút mạng nhận được giao dịch sẽ tiến hành sử dụng khóa ngoại người gửi để xác thực tính chính danh của người gửi; (3) các giao dịch hợp lệ được lưu trữ trên cây Merkle và đóng gói tạo ra khối mới, công việc này được đảm nhận bởi một nốt trong mạng; (4) khối mới tạo được gửi đến tất cả các nốt trên mạng và kích hoạt quá trình xác minh tính hợp lệ của khối mới với thuật toán đồng thuận bởi đa số nốt. Nếu đồng thuận chấp nhận, khối mới sẽ được tất cả các nốt bên trong mạng kết vào bản sao cuốn sổ cái của mình.

1.3. Hợp đồng thông minh (Smart Contract)

Hợp đồng thông minh có thể được hiểu như là một mẫu chương trình chứa đoạn mã lập trình, nơi cài đặt những xử lý dựa vào các điều khoản định trước. Hợp đồng thông minh được đại diện như một tài khoản với một địa chỉ định danh duy nhất do hệ thống

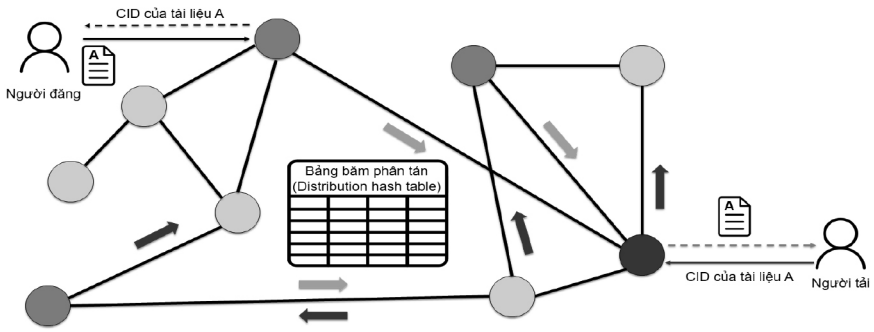
kiểm soát, trong khi đó tài khoản sở hữu bên ngoài (Externally owned account-AOE) thì được kiểm soát bởi người dùng [Arshdeep Bahga và cộng sự, 2016]. Như được trình bày ở Hình 3, cấu trúc của một hợp đồng thông minh cơ bản chứa tập các hàm thực thi (Functions) và các biến lưu trạng thái. Một giao dịch của người dùng gửi đến hợp đồng thông minh chứa dữ liệu và giá trị là các tham số đầu vào yêu cầu bởi hàm thực thi trong hợp đồng đó. Khi các hàm được kích hoạt, các biến trạng thái (Sate) sẽ thay đổi phụ thuộc vào luận lý xử lý bên trong các hàm đó. Hợp đồng thông minh có thể được viết và biên dịch dựa trên ngôn ngữ cấp cao như Python hoặc Solidity. Khi thực thi, hợp đồng thông minh có thể phát ra các sự kiện và ghi vào nhật ký biên nhận giao dịch, cung cấp một cách thức để truy vết thông tin về giao dịch đã thực hiện. Ngoài ra, chúng có khả năng gửi thông điệp để kích hoạt các hợp đồng thông minh khác, cơ chế này giúp người thiết kế linh động khi phân hoạch và định nghĩa các xử lý phức tạp dựa trên dây chuyền xử lý kết hợp các hợp đồng thông minh.



Hình 3. *Cấu trúc của một hợp đồng thông minh*

1.4. Hệ thống tệp liên hành tinh (Interplanetary File System)

Hệ thống tệp liên hành tinh (Interplanetary File System-IPFS) là một giao thức lưu trữ tập tin phân tán cho phép các máy tính lưu trữ và phân phối tập tin như một phần của mạng ngang hàng (Peer-to-peer network) mà không cần bất kỳ một máy chủ nào giữ vai trò làm trung tâm dữ liệu. Bất kỳ máy tính nào đều có thể kết nối vào IPFS và được xem như là một nốt (peer) của mạng thì có thể bắt đầu lưu trữ cũng như phân phối tập tin [Athreya và cộng sự, 2021]. Khi một tập tin được tải lên IPFS, tập tin đó sẽ được chia thành các phần nhỏ để lưu trữ, được tính giá trị băm (Hash value) và giá trị này được xem như là một dấu vân tay duy nhất, được gọi là mã nhận dạng nội dung (Content identifier-CID). Khi một tập tin bị sửa đổi và cố tình tải vào IPFS, hàm giá trị băm mật mã của nó sẽ khác và do đó, nó sẽ nhận được một CID mới. Điều này có nghĩa là các tệp được lưu trữ trên IPFS có khả năng chống giả mạo và kiểm duyệt. Khi người dùng muốn tải về tài liệu, IPFS duy trì một cấu trúc chỉ mục Bảng băm phân tán (Distributed hash table) với mục đích định tuyến và truy tìm những nốt chứa tập tin cần tìm, từ đó phép thiết lập kết nối và tải tập tin như minh họa ở Hình 4.



Hình 4. *Cơ chế hoạt động của hệ thống tệp liên hành tinh IPFS*

2. HỆ THỐNG LƯU TRỮ TÀI LIỆU ĐIỆN TỬ HỌC THUẬT PHÂN TÁN TRÊN NỀN TẢNG CHUỖI KHỐI

2.1. Ngữ cảnh vận dụng cho hệ thống lưu trữ tài liệu điện tử phân tán mở

Cơ chế hoạt động của IPFS dựa vào bảng mã băm phân tán trên mạng lưới phi tập trung ngang hàng, trong khi đó chuỗi khối dùng sổ cái phân tán. Hai công nghệ này có khả năng kết hợp cùng nhau vì cùng có cấu trúc tương đồng. Hiện nay, nhiều hệ thống lưu trữ được đề xuất bởi sự kết hợp từ hai công nghệ này. Shubham Desai [Shubham Desai và cộng sự, 2020] giới thiệu một hệ thống đa người dùng hỗ trợ kiểm soát quyền truy cập sử dụng chuỗi khối và cung cấp cơ chế lưu trữ dữ liệu phân tán. Ian Zhou [Ian Zhou và cộng sự, 2019] đề xuất hệ thống đánh giá bài viết học thuật phục vụ trong môi trường nghiên cứu. Edlira Martiri [Edlira Martiri và cộng sự, 2018] đề xuất mô hình mang tên DMS-XT áp dụng trong lĩnh vực giáo dục để lưu trữ an toàn và tin cậy các văn bằng phục vụ bảo vệ quyền tác giả và chống sao chép ý tưởng.

Trong ngữ cảnh chuyển đổi số tại các cơ sở giáo dục ngày nay, lưu trữ và sử dụng tài liệu nội sinh giữ vai trò quan trọng giúp thúc đẩy tiến trình nghiên cứu mang tính chất lặp bao gồm kế thừa, phát triển và chia sẻ. Tuy nhiên, phần lớn những tài liệu và công trình nghiên cứu được chủ trì quản lý bởi thư viện. Điều này loại bỏ vai trò của các bộ phận đào tạo chuyên môn khác như các khoa và bộ môn, là nơi tạo ra những tài liệu có giá trị chuyên sâu thông qua các hoạt động giảng dạy như đồ án môn học, tiểu luận và báo cáo môn học,… Những tài liệu này chưa có hình thức lưu trữ và chia sẻ hiệu quả. Bên cạnh đó, nếu tăng cường lưu trữ tập trung các tài liệu kể trên tại thư viện sẽ dẫn đến sự quá tải của hệ thống lưu trữ trung tâm và công việc của thủ thư. Mặt khác, đảm bảo tính nguyên bản của tài liệu và độ tin cậy khi chia sẻ tài liệu là một vấn đề quan trọng khác cần đề cập. Từ thế mạnh của chuỗi khối và hệ thống tập tin liên hành tinh, chúng tôi đề xuất một mô hình thiết kế cho hệ thống lưu trữ tài liệu điện tử phân tán mở (Open distributed e-document storage system-ODESS). Trong đó, cơ sở dữ liệu học thuật sẽ được lưu trữ phân tán không chỉ tại thư viện mà còn tại các khoa và bộ môn với các bên tham gia bao gồm:

- *Tác giả* (Author): là chủ sở hữu của tài liệu điện tử mà họ tạo ra. Tác giả có thể tạo hợp đồng thông minh và lưu trữ trong chuỗi khối và có trách nhiệm xác định quyền truy cập vào tài liệu của họ trong mạng IPFS. Việc xác định này được thực hiện trong hợp đồng thông minh mà họ tạo ra.

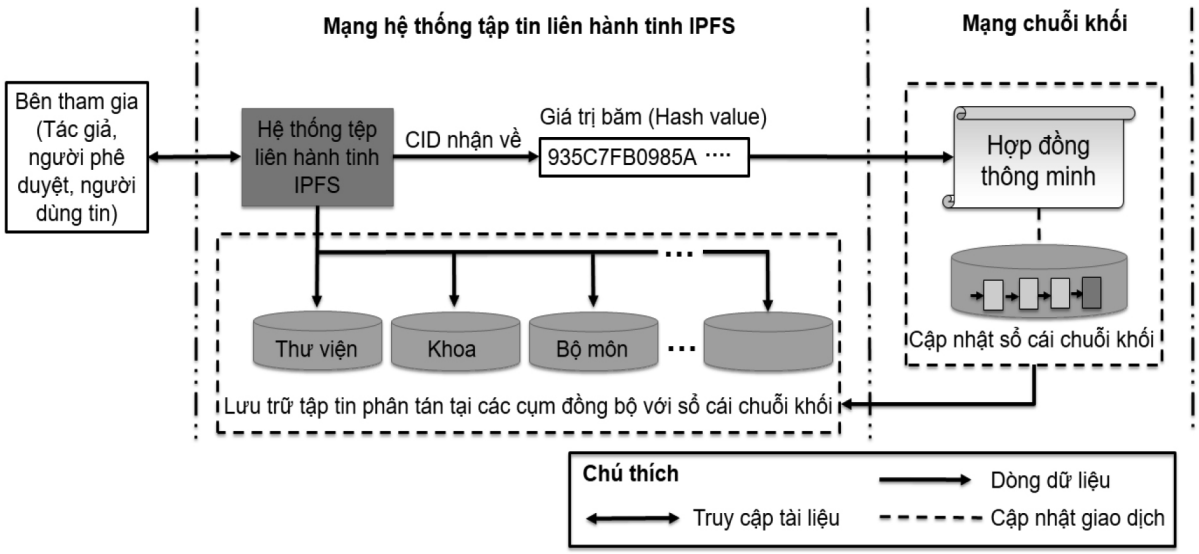
- *Người phê duyệt* (Approver): là người đại diện cụm lưu trữ (Thư viện, Khoa và Bộ môn) mà tác giả gửi tài liệu đến. Trách nhiệm chính của người phê duyệt là tải tài liệu điện tử được mã hóa của tác giả lên IPFS và xác minh giao dịch ban đầu của tác giả trên chuỗi khối.

- *Người dùng tin* (Document Requestor- DR): Các sinh viên, giảng viên và nhóm nghiên cứu quan tâm đến việc truy cập tài liệu của tác giả. Tác giả có thể cấp quyền truy cập cho bất kỳ người yêu cầu nào dựa trên chính sách ủy quyền được xác định trong hợp đồng thông minh.

2.2. Hệ thống lưu trữ tài liệu điện tử phân tán mở ODESS

Trong hệ thống đề xuất, tài liệu điện tử sẽ không lưu trữ trong chuỗi khối để tránh mở rộng kích thước dẫn đến sự phình to của chuỗi khối. Vì thế, chúng tôi sử dụng chuỗi khối để quản lý định danh tập tin và kiểm soát truy cập của những người tham gia. Trong khi đó, tập tin dữ liệu thực sự sẽ được mã hóa và tải lên mạng IPFS để ngăn chặn truy cập trái phép. Những người tham gia có thể xem tài liệu an toàn bằng cách trao đổi khóa mã hóa. Điều này đảm bảo tính nguyên bản tài liệu, bảo mật và ngăn rò rỉ dữ liệu đến những người dùng không liên quan.

Kiến trúc tổng thể đề xuất của ODESS được trình bày trong Hình 5, bao gồm hai thành phần chính là mạng chuỗi khối và mạng IPFS. Người dùng tham gia bao gồm tác giả, người phê duyệt và người dùng tin có thể tương tác với hệ thống cho các yêu cầu truy cập dữ liệu. Các tài liệu học thuật được mã hóa và được lưu trữ trong mạng IPFS theo phương thức phân tán ở các cụm lưu trữ tại thư viện, khoa và bộ môn. Mỗi tài liệu được định danh bởi CID là giá trị băm sinh ra bởi IPFS. Thông tin định danh tài liệu và kiểm soát truy cập của các bên tham gia được quản lý bởi mạng chuỗi khối. Trong đó, hợp đồng thông minh chịu trách nhiệm kiểm soát truy cập, thực thi các giao dịch trao đổi tài liệu và lưu trữ thông tin giao dịch bền vững vào sổ cái chuỗi khối. Có thể hiểu khái quát, IPFS giữ vai trò lưu trữ dữ liệu thực tế của tập tin, trong khi đó chuỗi khối đảm nhận kiểm soát truy cập người dùng và siêu dữ liệu của tập tin.



Hình 5. *Kiến trúc tổng quan của hệ thống đề xuất ODESS*

2.3. Giao thức truy cập của hệ thống ODESS

Hợp đồng thông minh được sử dụng tạo ra các giao thức để thực thi chính sách kiểm soát truy cập nội dung tập tin trên chuỗi. Sau đây là tập các xử lý cơ sở, từ đó giúp xây dựng tập giao thức truy cập và tương tác với hệ thống, bao gồm:

- *Tạo hợp đồng* (create_contract): chỉ được tạo và thực thi bởi tác giả để cấp các vai trò tương ứng cho người dùng tin và thông tin liên quan để truy cập tài liệu.

- *Yêu cầu truy cập* (request_access): được thực thi bởi người dùng tin để có được quyền truy cập tài liệu từ tác giả.

- *Chấp nhận truy cập* (approve_document_request): được thực thi bởi tác giả, cấp hoặc từ chối quyền truy cập với yêu cầu từ người dùng tin.

- *Thu hồi quyền truy cập* (remove_document_request): được sử dụng bởi tác giả sở hữu tài liệu. Quyền truy cập chấp thuận trước đó của người dùng tin được truy hồi và loại bỏ ra khỏi hợp đồng thông minh. Khi đó, người dùng tin đó không còn đặc quyền truy cập nội dung tập tin của tác giả.

- *Theo dõi ủy quyền truy cập* (trace_authorization): được sử dụng bởi người dùng tin và tác giả để theo dõi lịch sử ủy

quyền nơi người dùng tin được chấp thuận hoặc bị từ chối truy cập trong chuỗi khối.

Từ tập xử lý cơ sở trên, các giao thức tương tác hệ thống có thể được thiết lập. Ví dụ, minh họa trong Hình 6, giao thức đăng tải tài liệu được tiến hành theo các bước như sau:

- Bước 1: Là bước tương tác ngoại tuyến giữa tác giả và người phê duyệt. Đầu tiên, tác giả yêu cầu người phê duyệt lưu trữ tài liệu. Khi đó, người phê duyệt yêu cầu tác giả cung cấp khóa để mã hóa cho tập tin. Tiếp theo, tác giả tạo một cặp khóa gồm khóa công khai và khóa riêng. Tác giả gửi khóa công khai cho người phê duyệt để xác thực và mã hóa tập tin gốc, trong khi đó khóa riêng được giữ an toàn bởi tác giả.

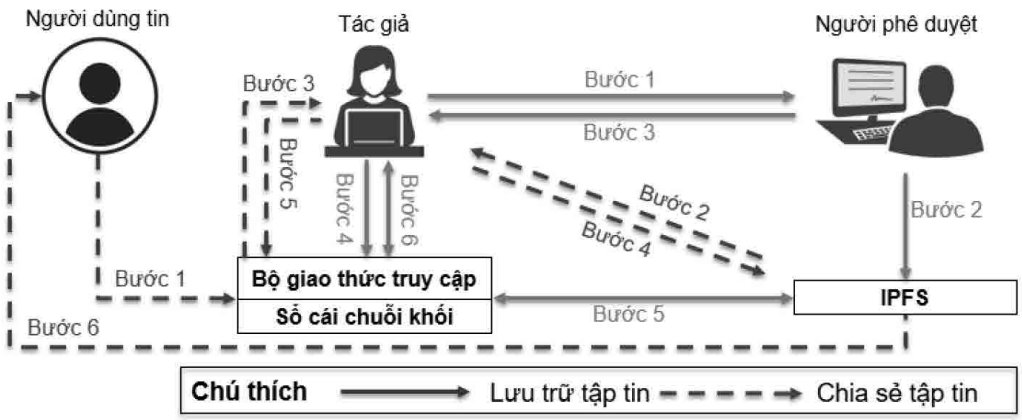
- Bước 2: Người phê duyệt mã hóa tài liệu bằng khóa công khai của tác giả, sau đó tải lên mạng IPFS và nhận về CID của tài liệu đó.

- Bước 3: Sau khi xác minh người gửi và nội dung, người phê duyệt chia sẻ CID của tài liệu về cho tác giả.

- Bước 4: Tác giả tạo một hợp đồng bằng giao thức truy cập *create_contract*.

- Bước 5: Hợp đồng tạo ra ký một giao dịch trên chuỗi khối cùng với khóa công khai tác giả, CID, mốc thời gian, mô tả thuộc tính cho tài liệu. Giao dịch này được xác minh bởi người phê duyệt và được đưa vào chuỗi khối.

- Bước 6: Tác giả sở hữu các tập tin trong hệ thống có thể truy cập, kiểm tra, chứng minh quyền sở hữu và cho phép bất kỳ người dùng tin sử dụng tài liệu của họ bởi sử dụng *trace_authorization*.



Hình 6. *Lưu đồ tương tác giữa các bên tham gia vào ODESS*

Xét về chia sẻ tài liệu, khi người dùng tin muốn truy cập tài liệu cho mục đích nghiên cứu. Khi đó, giao thức các bước để đạt được quyền truy cập diễn ra như sau:

- Bước 1: Người dùng tin thực hiện chia sẻ khóa công khai của mình bằng cách sử dụng *request_access*.

- Bước 2: Tác giả tải tập tin yêu cầu được mã hóa từ mạng IPFS sử dụng CID sinh ra ở giai đoạn lưu trữ tài liệu. Tác giả giải mã tài liệu bằng khóa riêng của mình.

- Bước 3: Tác giả có được khóa công khai của người dùng tin bởi tiến trình xác minh yêu cầu tại Bước 1.

- Bước 4: Tác giả mã hóa tài liệu gốc bằng khóa công khai của người dùng tin và đưa lên mạng IPFS nhận về CID của tài liệu đang chia sẻ.

- Bước 5: Tác giả ký một giao dịch trên chuỗi khối cùng với khóa công khai của người dùng tin, khóa công khai của tác giả và CID có được ở Bước 4 bằng cách sử dụng *approve_document_request.*

- Bước 6 : Người dùng tin truy xuất tập tin chia sẻ sử dụng giá trị CID được gửi tới và giải mã bằng khóa riêng của họ. Bằng cách này, tài liệu được chia sẻ diễn ra an toàn và bảo mật.

Trên đây là hai giao thức chính cho lưu trữ và chia sẻ tài liệu học thuật. Trong đó, các bước trao đổi khóa và mã hóa được mô tả chi tiết ở trên là cho mục đích hiểu rõ giao thức xử lý. Trong thực tế, những thao tác này được xử lý tự động bởi hệ thống.

3. THẢO LUẬN

Tận dụng ưu điểm của chuỗi khối và lưu trữ phân tán IPFS, giải pháp đề xuất có khả năng bảo vệ mật mã khi chia sẻ tài liệu và cho hiệu quả cao hơn về chi phí lưu trữ, mang nhiều điểm mới được thống kê ở Bảng 1 so với hướng tiếp cận thông thường.

- Khả năng mở rộng cơ sở dữ liệu: do dựa trên cơ chế phân tán, nên dễ dàng tích hợp thêm cụm lưu trữ trong tương lai mà không cần sao lưu, phục hồi cũng như ảnh hưởng hệ thống đang vận hành.

- Tránh bị tấn công máy chủ và chia sẻ dữ liệu nhanh: Vì loại bỏ vai trò máy chủ trung tâm nên mô hình đề xuất kháng được những kỹ thuật tấn công máy chủ hiện nay. Nhờ kỹ thuật lưu bộ đệm và khả năng tải dữ liệu đa nguồn, giúp tăng tốc độ tải dữ liệu.

- Tính nguyên bản và an toàn dữ liệu: Tất cả tập tin cũng như giao dịch hoàn toàn được mã hóa. Do đó, loại bỏ khả năng thay đổi dữ liệu, làm giả giao dịch và cung cấp phương thức truy cập dữ liệu tin cậy hơn so với các hướng tiếp cận khác.

- Hiệu quả về quản lý: loại bỏ mô hình quản lý tập trung và mở rộng tham gia quản lý bởi nhiều bên như các Khoa và Bộ môn. Giúp giảm thiểu công việc tập trung ở một bộ phận cụ thể, từ đó hệ thống sẽ linh động và vận hành nhanh hơn.

Bảng 1. *So sánh đặc trưng giữa ODESS và hướng tiếp cận truyền thống*

Đặc trưng	Tập trung trên máy chủ	Tập trung trên đám mây	ODESS đề xuất
Cơ sở dữ liệu	Tập trung	Tập trung	Phân tán
Cơ chế chia sẻ tập tin	HTTP	HTTP	IPFS
Kháng tấn công máy chủ	Không	Không	Có
Khả năng linh động mở rộng lưu trữ	Thấp	Trung bình	Cao
Mã hóa dữ liệu	Không	Không	Có
Hiệu quả khả năng quản lý	Thấp	Thấp	Cao

KẾT LUẬN

Sản phẩm học thuật là tài sản quý giá của bất kỳ cơ sở giáo dục nào. Những tài liệu được sản sinh trong hoạt động giáo dục là nguồn tài liệu nội sinh có giá trị và cần khai thác hiệu quả. Tuy nhiên, hướng tiếp cận lưu trữ tập trung hiện nay tồn tại những hạn chế về chi phí quản lý, khả năng mở rộng và truy cập tin cậy. Tương phản với hướng tiếp cận cũ, nghiên cứu này đề xuất một giải pháp lưu trữ tài liệu điện tử phân tán, với sự kết hợp thế mạnh quản lý giao dịch bằng chuỗi khối và nền tảng lưu trữ phân tán của IPFS, cung cấp

khả năng lưu trữ và chia sẻ tài liệu điện tử hiệu quả phù hợp với ngữ cảnh chuyển đổi số hiện nay. Mở rộng khả năng lưu trữ phân tán và chia sẻ quản lý tại các Khoa và Bộ môn, là nơi phần lớn tài liệu nội sinh được tạo ra. Điều này không chỉ giảm bớt gánh nặng công việc cho bộ phận Thư viện mà còn hiệu quả cho công tác thu thập tài liệu.

TÀI LIỆU THAM KHẢO

1. A.Athreya, Ashwin Kumar, S. Nagarajath, Gururaj H L, Ravi Kumar V, Sachin D N, and Rakesh K R (2021). Peer-to-Peer Distributed Storage Using InterPlanetary File System. 711-721.

2. Arshdeep Bahga and Vijay Madisetti (2016). Blockchain Platform for Industrial Internet of Things. Journal of Software Engineering and Applications 09 (01 2016), 533-546.

3. Cao Minh Kiểm (2019). Công nghệ Blockchain và tiềm năng ứng dụng vào lĩnh vực TT-TV. Tạp chí Thông tin và Tư liệu, Tr 3-11.

4. Edlira Martiri and Gentjana Mua (2018). DMS-XT: A Blockchain-based Document Management System for Secure and Intelligent Archival. In Proceedings of the 3rd International Conference on Recent Trends and Applications in Computer Science and Information Technology (RTACSIT).

5. Gurdish Sandhu (2018). The Role of Academic Libraries in the Digital Transformation of the Universities. In 2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS). 292-296.

6. Ian Zhou, Imran Makhdoom, Mehran Abolhasan, Justin Lipman, and Negin Shariati (2019). A Blockchain-based File-sharing System for Academic Paper Review. In 13th International Conference on Signal Processing and Communication Systems. 1-10.

7. Jin Han, Cheng Wang, Jie Miao, Mingxin Lu, Yingchun Wang, and Shi Jin (2021). Research on Electronic Document Management System Based on Cloud Computing. Computers, Materials Continua 66 (01 2021), 2645-2654.

8. Qurotul Aini, Ninda Lutfiani, Nuke Puji Lestari Santoso, Sulistiawati Sulistiawati, and Erna Astriyani (2021). Blockchain For Education Purpose: Essential Topology. APTISI Transactions on Management (ATM) 5, 2 (May 2021), 112-120.

9. Satoshi Nakamoto. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 23 Sep 2022)

10. Shafaq Khan, Faiza Loukil, Chirine Ghedira, Elhadj Benkhelifa, and Anoud Bani-Hani (2021). Blockchain Smart Contracts: Applications, Challenges, and Future Trends. Peer-to-Peer Networking and Applications 14 (09 2021).

11. Shubham Desai, Rahul Shelke, Onkar Deshmukh, Harish Choudhary, and Santosh S Sambare (2020). Blockchain based secure data storage and access control system using IPFS. In Journal of critical reviews JCR. 7 (2020), 1254-1260.

12. Sun Y, Zhang J, Xiong Y, Zhu G. Data Security and Privacy in Cloud Computing. International Journal of Distributed Sensor Networks. 2014;10 (7).

13. Wubing Chen, Zhiying Xu, Shuyu Shi, Yang Zhao, and Jun Zhao (2018). A Survey of Blockchain Applications in Different Domains. In Proceedings of the 2018 International Conference on Blockchain Technology and Application (ICBTA 2018). Association for Computing Machinery, New York, NY, USA, 17-21.

14. Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. In 2017 IEEE International Congress on Big Data (BigData Congress). 557-564.

(Ngày Tòa soạn nhận được bài: 5-5-2023; Ngày phản biện đánh giá: 6-6-2023; Ngày chấp nhận đăng: 15-7-2023).