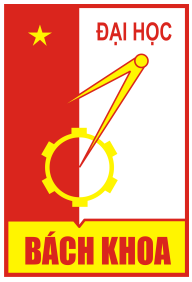
Trường Đại học Bách Khoa Hà Nội

Viện Toán ứng dụng và Tin học



ĐỒ ÁN TỐT NGHIỆP

Ứng dụng Chuỗi khối trong quản lý văn vằng giáo dục

Hướng dẫn:

TS. Đoàn Duy Trung

Thực hiện:

Đỗ Minh Tuấn - 20185419

Hà Nội, 2022

Nhận xét của giảng viên hướng dẫn

Mục đích và nội dung của Đồ án

Kết quả đạt được

Ý thức làm việc của sinh viên thực hiện

Hà Nội, ngày ... tháng ... năm ...

Giảng viên hướng dẫn

(Ký và ghi rõ họ tên)

Lời cảm ơn

Vậy là lễ tốt nghiệp sắp đến, có rất nhiều cảm xúc mà em không thể nói ra bằng lời. Em xin gửi lời cảm ơn đến *thầy Đoàn Duy Trung*, người đã định hướng và giúp đỡ em rất nhiều trong suốt chặng đường vừa qua. Em cảm ơn thầy.

Cùng với đó, em tự hào vì được là sinh viên CTTN Toán-Tin, hạnh phúc khi ở bên có những con người sẵn sàng chia sẻ từng chút "hiểu biết" cho nhau để ghép lên một bức tranh tuyệt đẹp về *tri thức của tình bạn*. Bạn bè giúp ta thấy được thời gian trôi nhanh và quý giá đến nhường nào, để ta cảm nhận được từng khoảnh khắc đáng nhớ trong cuộc sống vốn vội vã, xô bồ.

Tuy một học kỳ là ngắn ngủi, nhưng với định hướng rõ ràng, quả thực, nó mang lại cho em quá là nhiều kiến thức, đưa em đến những chân trời mới về công nghệ. Sau mỗi thành tựu mà ai đó đạt được, sẽ thật thiếu xót khi không nhìn vào sự nỗ lực của chính họ. Bản thân em cũng đã dành ra khá nhiều công sức tìm hiểu, thử nghiệm, triển khai, và mang đến kết quả phù hợp nhất cho đề tài mà mình đã chọn. Em tự tin với chất lượng của bài báo cáo này.

Mục lục

Chương 1

Bài toán quản lý văn bằng giáo dục

Từ lâu, các văn bằng giáo dục là một trong những mục tiêu của quá trình học tập không chỉ ở Việt Nam. Một "tấm bằng" sẽ ghi lại kết quả của một giai đoạn học và tích luỹ kiến thức của cá nhân, thể hiện bằng những điểm số qua các kỳ thi hay các bài kiểm tra cụ thể.

Với văn bằng ở các cấp độ giáo dục từ Trung học trở xuống, các thông tin thường được ghi là kết quả đánh giá tổng thể của một cấp học, đi kèm là một số giấy tờ bổ sung chi tiết về điểm số của học sinh và nhận xét của giáo viên, nhà trường đối với học sinh đó (thường được gọi là *Học bạ*). Đối với các văn bằng ở các cấp độ cao hơn như Trung cấp, Cao đẳng, Đại học, điểm số tổng kết và đánh giá năng lực của cơ sở giáo dục đối với học viên/sinh viên sẽ được thể hiện, cùng với đó là đính kèm bảng điểm chi tiết các học phần trong chương trình đào tạo mà học viên theo học tại cơ sở giáo dục đó (được gọi là *Phụ lục văn bằng*).

Ngoài văn bằng do các cơ sở giáo dục cấp phát, nhiều cơ quan, tổ chức cũng phát hành văn bằng đánh giá kết quả của cá nhân hoặc một nhóm cá nhân theo một lĩnh vực nào đó, thường dưới dạng các *chứng chỉ*, hay *giấy chứng nhận*.

Trong các hoạt động tuyển dụng nhân sự, văn bằng giáo dục (cùng với chứng chỉ) đem đến cho nhà tuyển dụng (các doanh nghiệp, cơ quan, tổ chức) cái nhìn tổng quát đầu tiên về năng lực của ứng viên tương ứng dựa trên đánh giá thể hiện qua thông tin trên văn bằng đó. Những ứng viên với một văn bằng cùng những thông tin tích cực có lợi thế rất lớn trong cuộc đua trở thành "người được chọn", bên cạnh việc thể hiện khả năng của mình trong quá trình làm việc. Việc có được sự đánh giá tốt từ những cơ sở giáo dục chất lượng, có uy tín cũng đem đến sự tin tưởng của nhà tuyển dụng đối với thông tin được ghi trong văn bằng.

Tuy nhiên, quy trình xác thực tính đúng đắn của một văn bằng giáo dục tại Việt Nam hiện nay gặp rất nhiều khó khăn. Phần nhiều là vì, các cơ sở giáo dục tại nước ta không công khai thông tin các văn bằng đã cấp phát bởi nhiều lý do. Thêm nữa, sự thiếu hụt các cổng tra cứu công cộng về văn bằng giáo dục cũng khiến các nhà tuyển dụng phải đặt niềm tin "tạm thời" vào một tờ bìa "có vẻ đáng tin cậy". Tính đến thời điểm bài báo cáo này được viết, *Bộ Lao động, Thương binh và Xã hội* đã cung cấp *Trang thông tin tra cứu văn bằng Giáo dục nghề nghiệp*. Tuy nhiên, đó vẫn là chưa đủ. Một hệ thống tra cứu thông tin về văn bằng với độ tin cậy cao và dễ sử dụng, áp dụng được trên mọi cơ sở giáo dục, tổ chức cấp phát văn bằng trở nên hết sức cần thiết.

Chương 2

Mã hoá bất đối xứng và chữ ký số

2.1 Mã hoá bất đối xứng

2.2 Chữ ký số

Chương 3

Chuỗi khối và mạng Ethereum

3.1 Tổng quan về chuỗi khối

Một chuỗi khối là một sổ cái điện tử *phân tán*[1], *phi tập trung*[2], bao gồm các bản ghi được gọi là *khối* (block) thường được dùng để ghi lại các *giao dịch* (transaction) qua các máy tính.

Một chuỗi khối có thể bao gồm bảy lớp:

1. Cơ sở hạ tầng: Phần cứng
2. Mạng: Khám phá các nút mạng, chuyển tiếp thông tin, và xác thực
3. *Đồng thuận*[3]
4. Dữ liệu: Các khối, các giao dịch
5. Ứng dụng: *Hợp đồng thông minh*[4], *ứng dụng phi tập trung*[5]

[1]Distributed
[2]Decentralized
[3]Consensus
[4]Smart Contract
[5]Decentralized application (Dapp)

3.1.1 Khối

Các *khối* (block) nắm giữ các giao dịch hợp lệ đã được băm, quá trình mã hoá này sử dụng cấu trúc của một *cây Merkle*[6]. Mỗi khối bao gồm mã băm của khối liền trước trong chuỗi khối để liên kết với khối đó. Các khối liên kết với nhau định hình một *chuỗi* (chain). Quá trình tiếp diễn này xác minh tính toàn vẹn của khối trước đó, cho đến khối đầu tiên - được gọi là *khối bắt đầu*[7]. Để đảm bảo dữ liệu bên trong, các khối thường có *chữ ký số*[8].

Đôi khi các khối khác nhau được tạo ra đồng thời, tạo ra sự *phân nhánh tạm thời*[9]. Để bảo mật lịch sử băm, bất kỳ chuỗi khối nào đều có một thuật toán để "tính điểm" các phiên bản khác nhau, phần lịch sử của nhánh nào có "điểm số" lớn hơn sẽ được chọn cho chuỗi khối. Các khối không được chọn để đưa vào chuỗi khối được gọi là *khối mồ côi*[10]. Sự ngang hàng trong mạng khiến cho cơ sở dữ liệu có rất nhiều phiên bản/lịch sử theo thời gian, và phiên bản có "điểm số" cao nhất sẽ được giữ lại. Bất cứ khi nào một thành phần trong mạng nhận được một phiên bản có "điểm số" lớn hơn (thường là phiên bản cũ với một khối mới được thêm vào), nó sẽ mở rộng/ghi đè cơ sở dữ liệu và chuyển tiếp cho các thành phần ngang hàng khác ở trong mạng.

Thời gian khối

Thời gian khối[11] là thời gian trung bình để một mạng tạo ra một khối mới trong chuỗi khối. Ngay sau khi khối mới được tạo, dữ liệu trong nó sẽ được xác minh. Đối với tiền điện tử, các giao dịch diễn ra, một thời gian khối ngắn hơn đồng nghĩa với việc giao dịch sẽ nhanh hơn.

[6] Merkle tree
[7] Genesis block
[8] Digital signature
[9] Temporay fork
[10] Orphan block
[11] Block time

Phân nhánh hoàn toàn

Phân nhánh hoàn toàn[12] là việc thay đổi quy tắc trong chuỗi khối khiến các phần mềm xác thực dựa vào quy tắc cũ xác thực các chuỗi mới được tạo dựa trên quy tắc mới không hợp lệ. Trong trường hợp có một đợt phân nhánh hoàn toàn, tất cả các nút cần nâng cấp phần mềm để theo quy tắc mới. Nếu có một nhóm các nút tiếp tục sử dụng phần mềm cũ trong khi các nút còn lại sử dụng phần mềm mới, sự chia tách có thể xảy ra.

3.1.2 Phi tập trung

Bằng cách lưu dữ liệu thông qua *mạng ngang hàng*[13], chuỗi khối giảm được các rủi ro trong lưu trữ dữ liệu tập trung. Chuỗi khối phi tập trung sử dụng *truyền thông điệp tuỳ biến*[14] và *mạng phân tán*[15]. Khi thiếu đi tính phi tập trung, chuỗi khối có thể đối mặt với *tấn công 51%* (51% attack) - một thành phần trong mạng có thể kiểm soát nhiều hơn nửa phần còn lại của mạng và điều khiển các bản ghi trong chuỗi khối như ý muốn, trong đó có *chi tiêu gấp đôi* (double-spending).

Mạng chuỗi khối ngang hàng giảm thiểu khả năng bị khai thác tại các điểm tập trung dễ bị tấn công nào đó. Các phương pháp bảo mật trong chuỗi khối bao gồm việc sử dụng *mã hoá khoá công khai*[16]. Mỗi *khoá công khai* (public key), chuỗi ký tự dài ngẫu nhiên, là một địa chỉ (address) trong chuỗi khối. Các *token* gửi đi khắp mạng sẽ được ghi lại thuộc về một địa chỉ nào đó của chuỗi khối. Mỗi *khoá riêng tư* (private key) giống như một mật khẩu, giúp chủ sở hữu có thể truy cập vào *tài sản số*[17] của họ. Dữ liệu được lưu trong chuỗi khối được cho là không thể bị phá vỡ.

[12]Hard fork
[13]Peer-to-peer network (P2P network)
[14]Ad hoc message passing
[15]Distributed networking
[16]Public-key cryptography
[17]Digial asset

Mỗi *nút* (node) trong hệ thống phi tập trung giữ một bản sao của chuỗi khối. Chất lượng dữ liệu được duy trì bởi sự *nhân rộng cơ sở dữ liệu*[18] và sự tin cậy tính toán. Sẽ không có bản sao "chính" nào, cũng sẽ không có người dùng nào "đáng tin" hơn bất cứ ai trong hệ thống. Các giao dịch được gửi đi khắp/quảng bá (broadcasted) trong mạng qua phần mềm. Các thông điệp được gửi dựa trên *nền tảng nỗ lực tốt nhất*[19]. Các *nút đào*[20] xác thực giao dịch, thêm chúng vào khối mà nút đó đang tạo, và gửi khối đó đi khắp các nút trong mạng khi đã hoàn thành.

3.1.3 Tính mở

Các chuỗi khối mở *dễ dùng*[21] hơn so với các bản ghi truyền thống - mặc dù mở nhưng vẫn cần truy cập vật lý để xem. Tất cả các chuỗi khối trước đây đều *vô quyền* (permissionless). Nhiều tranh cãi đã nổ ra liên quan đến định nghĩa của chuỗi khối: Liệu một hệ thống riêng tư với xác thực viên được uỷ quyền bởi một *trung tâm uỷ quyền*[22] có được coi là một chuỗi khối hay không. Phía ủng hộ các chuỗi phân quyền hoặc riêng tư muốn thuật ngữ "chuỗi khối" có thể áp dụng với bất cứ cấu trúc dữ liệu phân dữ liệu thành các *khối đóng dấu thời gian*[23]. Phía phản đối điều này khẳng định các *hệ thống phân quyền* (permissioned system) giống như cơ sở dữ liệu truyền thống, không hỗ trợ xác thực dữ liệu phi tập trung, không thể chống lại sự giả mạo và sửa đổi.

Sự vô quyền

Một lợi thế của một mạng chuỗi khối mở, vô quyền, hoặc công khai là không cần bảo vệ chống lại các tác nhân xấu, không cần kiểm soát truy cập. Nghĩa là, các ứng dụng có thể được thêm vào mạng mà không cần sự chấp thuận và sự

[18] Database replication
[19] Best-effort basis
[20] Mining node
[21] User-friendly
[22] Central authority
[23] Time-stamped block

tin tưởng của các nút khác trong mạng, sử dụng chuỗi khối như là một *lớp vận chuyển*[24].

Chuỗi khối phân quyền/riêng tư

Các chuỗi khối phân quyền sử dụng một lớp điều khiển truy cập để quản lý những ai truy cập vào mạng. Trái với mạng chuỗi khối công cộng, xác thực viên trong mạng chuỗi khối riêng tư được kiểm tra chủ của mạng. Chủ của mạng không dựa vào các *nút ẩn danh*[25] để xác thực giao dịch hay các quyền lợi từ *hiệu ứng mạng*[26]. Các chuỗi khối phân quyền còn được biết đến với cái tên *chuỗi khối consortium*[27].

3.1.4 Các thuật toán đồng thuận phổ biến

Bằng chứng công việc - PoW

Bằng chứng công việc[28] (PoW) là một dạng của bằng chứng mã hoá, trong đó một bên chứng minh cho các bên khác (bên xác thực) rằng họ đã bỏ ra khối lượng tính toán nào đó. Bên xác thực sẽ tuần tự xác minh tính đúng đắn của bằng chứng này một cách dễ dàng.

PoW được đưa ra lần đầu bởi *Cynthia Dwork* và *Moni Naor* vào năm 1993 như một cách để xác định các cuộc *tấn công từ chối dịch vụ*[29], và các vấn đề liên quan đến lạm dụng dịch vụ như *spam* trong một mạng bằng cách yêu cầu một vài công việc bởi phía yêu cầu dịch vụ, thường là một quá trình tiêu tốn tài nguyên (thời gian, bộ nhớ) của máy tính. Thuật ngữ này được sử dụng lần đầu trong một bài báo của *Markus Jakobsson* và *Ari Juels*. Sau đó, nó được phổ biến bởi *Bitcoin*, như là một thuật toán đồng thuận đầu tiên trong mạng phi tập trung không phân

[24]Transport layer
[25]Anonymous node
[26]Network effect
[27]Consortium blockchain
[28]Proof of Work - PoW
[29]Denial-of-Service attack - DoS attack

quyền.

Trong mạng Bitcoin, các nút đào cần thực hiện giải một bài toán "khó" bằng cách tìm ra một con số được gọi là *nonce* sao cho sau khi kết hợp nó với các thông tin đã có của một khối trong chuỗi khối để thực hiện mã hoá băm, ta được một chuỗi băm bắt đầu với một chuỗi các ký tự "0" liên tiếp nhất định, và dãy số này cũng chính là địa chỉ của khối được tạo ra. Nút đầu tiên giải quyết được bài toán trên sẽ quảng bá khối đó lên toàn bộ mạng, các nút khác sẽ xác thực lại tính đúng đắn bằng cách sử dụng hàm băm để mã hoá lại thông tin của khối, so sánh với địa chỉ của khối: Nếu đúng, khối đó sẽ được thêm vào chuỗi, và nút tạo ra khối đó sẽ được nhận phần thưởng. Phần thưởng ở trong mạng Bitcoin chính là một lượng nhỏ đồng tiền mã hoá Bitcoin, và được gửi tới địa chỉ nút nhận ở khối nhất định sau khối vừa được thêm vào.

Do sự cải tiến về mặt công nghệ, các máy tính ngày nay có sức mạnh tính toán vô cùng lớn, tạo ra thách thức với các cơ chế đồng thuận dựa trên khối lượng tính toán, trong đó có PoW. Chính vì vậy, độ khó của bài toán mà các nút cần giải quyết ngày càng tăng lên. Tính tới thời điểm tháng 11 năm 2021, số lượng ký tự "0" liên tiếp trong phần đầu của địa chỉ khối trong mạng Bitcoin đã lên tới 7 chữ số, và thời gian để một giao dịch được thực thi trên mạng này xấp xỉ 10 phút. Thời gian thực thi giao dịch lâu khiến cho trải nghiệm của người dùng giảm thấp, đôi khi gây tắc nghẽn hệ thống mạng do nhiều giao dịch không được xử lý. Các cơ chế đồng thuận khác được đưa ra để giải quyết hạn chế này, trong đó có *Bằng chứng cổ phần* - PoS.

Bằng chứng cổ phần - PoS

Bằng chứng cổ phần[30] (PoS) là một lớp các cơ chế đồng thuận hoạt động bằng cách chọn ra các xác thực viên dựa trên tỷ lệ tiền mã hoá mà họ nắm giữ.

[30]Proof of Stake - PoS

Cơ chế PoS cho phép chủ sở hữu của các đồng tiền mã hoá "cọc" (stake) một lượng tiền để có thể trở thành một nút xác thực. *Cọc*[31] là khi một nút bỏ ra một số tiền để tham gia quá trình xác thực giao dịch. Số tiền này sẽ bị khoá khi *cọc*, và cần *huỷ cọc* (unstake) để có thể sử dụng giao dịch.

Khi một khối các giao dịch sẵn sàng để thực thi, cơ chế PoS sẽ chọn một nút xác thực (xác thực viên) để đánh giá khối đó. Xác thực viên sẽ kiểm tra thông tin các giao dịch trong khối có chính xác hay không, nếu đúng, khối đó sẽ được thêm vào chuỗi khối. Nút thêm khối mới vào chuỗi, tất nhiên, sẽ nhận được phần thưởng. Tuy nhiên, nếu nút đó thêm một khối với thông tin không chính xác, số tiền cọc của nút đó sẽ mất (một phần hoặc toàn bộ).

Mỗi cơ chế PoS có cách chọn xác thực viên khác nhau. Thông thường, quá trình lựa chọn ngẫu nhiên sẽ diễn ra, các yếu tố ảnh hưởng đến có thể kể ra như số lượng tiền mã hoá nút đó bỏ ra để cọc, nút đó tham gia quá trình cọc bao lâu rồi, vân vân. Mặc dù, ai cũng có thể tham gia cọc, nhưng tỷ lệ được chọn làm xác thực viên là rất thấp nếu số tiền bỏ ra để cọc. Vì lý do này, các thành viên tham gia quá trình cọc gia nhập vào các *bể cọc*[32]. Chủ của mỗi bể cọc sẽ thiết lập một nút tham gia quá trình xác thực, và các thành viên trong bể cọc sẽ "dồn tiền" để nút đó tham gia cọc để gia tăng cơ hội được chọn làm nút xác thực. Phần thưởng khi nút đó thêm khối mới vào chuỗi sẽ được chia cho các thành viên trong bể cọc đó.

Các cơ chế đồng thuận khác

Ngoài hai cơ chế đồng thuận phổ biến PoW và PoS, có rất nhiều các thuật toán đồng thuận khác. *Bằng chứng bộ nhớ*[33] (PoC) là một cơ chế cho phép chia sẻ không gian bộ nhớ của một nút cho mạng chuỗi khối. Nút nào có càng nhiều không gian bộ nhớ, nút đó càng có nhiều quyền duy trì mạng. *Bằng chứng hoạt động*[34] (PoA),

[31] Staking
[32] Staking pool
[33] Proof of Capacity - PoC
[34] Proof of Activity - PoA

được sử dụng trên chuỗi khối *Decred*, là một cơ chế kết hợp giữa PoW và PoS. *Bằng chứng tiêu thụ*[35] (PoB) lại yêu cầu các nút gửi lượng nhỏ tiền của chúng tới một địa chỉ ví không thể truy cập. Một cơ chế khác là *Bằng chứng lịch sử*[36] (PoH), được phát triển bởi *Solana Project*, tương tự như cơ chế *Bằng chứng thời gian còn lại*[37] (PoET), mã hoá thông tin thời gian trôi qua để đạt được sự đồng thuận mà không cần tiêu tốn nhiều tài nguyên.

3.2 Mạng Ethereum

3.2.1 Lịch sử

Ethereum là một nền tảng mã nguồn mở dựa trên chuỗi khối, hỗ trợ *Hợp đồng thông minh*[38]. Ethereum khá nổi với *đồng tiền mã hoá*[39] của nó với tên gọi là *Ether* (ký hiệu: ETH). Dựa vào sự phân tán của công nghệ chuỗi khối, Ethereum khá an toàn, và cũng nhờ bảo mật cao nên giá trị của đồng tiền ETH tích luỹ ngày càng lớn trên thị trường tiện điện tử.

Bắt đầu ý tưởng từ năm 2013 bởi lập trình viên *Vitalik Buterin* và một số cộng sự, công việc phát triển Ethereum được vận hành và kêu gọi vốn từ cộng đồng vào năm sau đó. Mạng Ethereum chính thức "lên sóng" vào ngày 30 tháng 7 năm 2015.

3.2.2 Hợp đồng thông minh

Nền tảng Ethereum còn hỗ trợ mạng lưới các *ứng dụng phi tập trung*[40]. Mạng này vận hành xoay quanh các hợp đồng thông minh. Phần lớn các ứng dụng sử dụng

[35]Proof of Burn - PoB
[36]Proof of History - PoH
[37]Proof of Elapsed Time - PoET
[38]Smart contract
[39]Cryptocurrency
[40]Decentralized applications (DApps)

hợp đồng thông minh để liên kết với công nghệ chuỗi khối. Có thể nói, hợp đồng thông minh chính là nhân tố trung tâm của nền tảng Ethereum.

Hợp đồng thông minh là *hợp đồng tự thực thi*[41] với các điều khoản được viết bởi các dòng lệnh hay các đoạn mã lập trình. Các đoạn mã này tồn tại khắp các nút trong mạng chuỗi khối, điều hành sự thực thi các giao dịch, và không thể thay đổi. Hợp đồng thông minh mang đến các giao dịch đáng tin cậy, sự đồng ý với các điều khoản trong hợp đồng tới các bên "ẩn danh" mà không cần qua một bên trung gian hay một cơ chế thực thi bên ngoài.

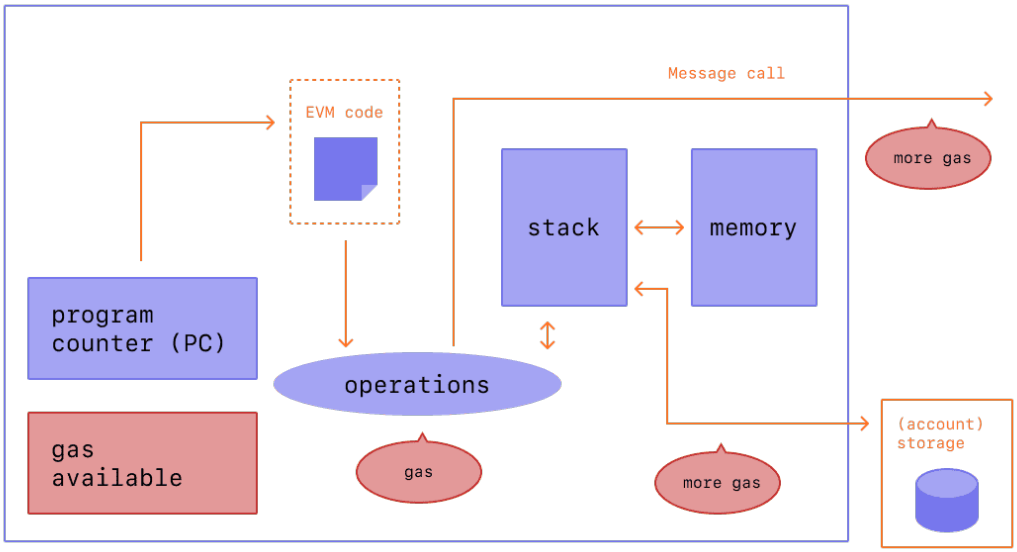
Trên Ethereum, các đoạn mã của hợp đồng thông minh được viết bằng ngôn ngữ lập trình *Solidity* hoặc *Vyper*. Solidity là ngôn ngữ lập trình hướng đối tượng bậc cao dựa theo *C++*, *JavaScript*, *Python*, và được thiết kế để tích hợp được với *Máy ảo Ethereum*[42] (EVM). Vyper là ngôn ngữ đang trong quá trình thử nghiệm.

3.2.3 Chi phí giao dịch

Gas là đơn vị thể hiện cho khối lượng tính toán để thực hiện một hành động nào đó trên mạng Ethereum. Do mỗi giao dịch trên mạng Ethereum đều cần tài nguyên tính toán để được thực thi, vì thế mà nó phát sinh ra *chi phí giao dịch*. Khi đó, *gas* thể hiện chi phí để thực hiện giao dịch thành công trên mạng.

Gas được trả bằng đồng *ether* (ETH). *Giá gas*[43] có đơn vị là *gwei*, mỗi *gwei* tương ứng với một phần một tỷ của một *ether*: 1 *gwei* $= 10^{-9}$ *ether*. Vì vậy, thay vì nói chi phí giao dịch là $0,000000001$ *ether*, ta có thể nói giao dịch đó tiêu tốn 1 *gwei*. Ngoài ra, 1 *gwei* chính là một tỷ *wei*; *wei* (được đặt tên theo *Wei Dai* - nhà khoa học máy tính nổi tiếng đưa ra lý thuyết về thanh toán bằng tiền mã hoá) là đơn vị nhỏ nhất trên Ethereum.

[41] Self-executed contract
[42] Ethereum Virtual Machine - EVM
[43] Gas price


Message call
EVM code
more gas
stack
memory
program
counter (PC)
operations
(account)
storage
gas
available
gas
more gas

Chương 4

Hệ thống tệp tin phân tán IPFS

4.1 Tổng quan về IPFS

4.2 Cơ chế hoạt động

4.3 Bảo mật và quyền riêng tư

Chương 5

Quản lý văn bằng giáo dục trên mạng Ethereum

5.1 Đề xuất giải pháp

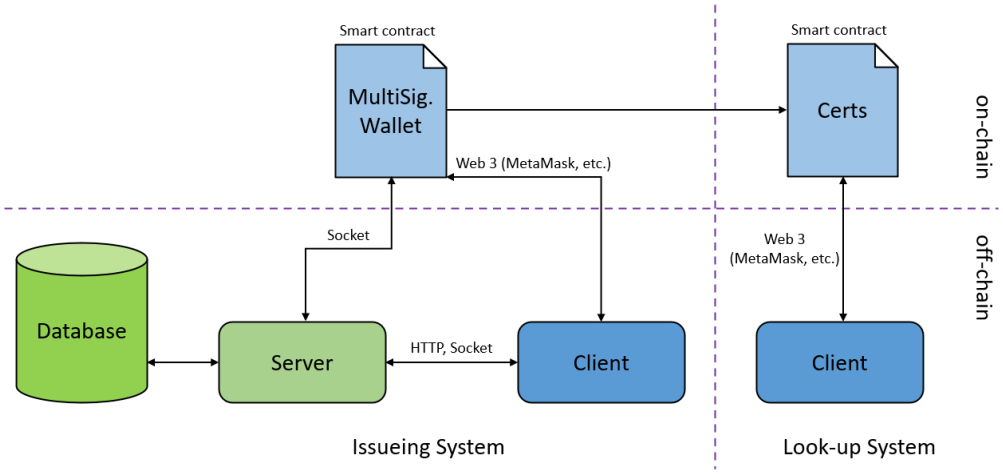
Giải pháp phổ biến hiện nay là xây dựng cơ sở dữ liệu chung về văn bằng giữa các cơ sở giáo dục. Nhà tuyển dụng có thể dễ dàng tra cứu thông tin văn bằng khi truy cập vào hệ thống sử dụng cơ sở dữ liệu này. Tuy nhiên, vấn đề hiện hữu là sử dụng cơ sở dữ liệu chung truyền thống tiềm ẩn rất nhiều rủi ro về dữ liệu. Việc nhiều bên truy cập và chỉnh sửa dữ liệu (các cơ sở giáo dục có quyền như nhau đối với cơ sở dữ liệu này) có thể phát sinh mất mát và ảnh hưởng đến dữ liệu của các bên khác. Ngoài ra, nếu giao quyền cập nhật dữ liệu cho một hoặc một số lượng hạn chế các bên tham gia, quy trình rà soát và sửa sai có thể kéo dài và các yêu cầu thay đổi dữ liệu được gửi từ các bên không có quyền cập nhật sẽ không được xử lý kịp thời.

Không phải nói quá, chuỗi khối giải quyết quá tốt các bài toán về cơ sở dữ liệu chung, đặc biệt là khi tính minh bạch của dữ liệu được ưu tiên. Đặc biệt, tương tác với chuỗi khối ngày càng trở nên đơn giản với *DApp* - các ứng dụng phi tập trung với giao diện thân thiện, dễ dùng, đồng thời tốc độ truy xuất thông tin

chấp nhận được, việc ứng dụng chuỗi khối càng được ưa chuộng. Việc lưu trữ thông tin văn bằng giáo dục trên chuỗi khối đảm bảo được dữ liệu không thể bị chỉnh sửa, mọi người đều dễ dàng truy cập. Nhiệm vụ cấp phát văn bằng được đưa về phía từng cơ sở giáo dục, lưu trữ trên mạng chuỗi khối chung, không thể chỉnh sửa. Đó là ưu điểm khi việc quản lý văn bằng sẽ không quá tập trung vào một số bên nhất định, nhưng cũng là thách thức cho các cơ sở giáo dục trong vấn đề xây dựng hệ thống tích hợp với mạng chuỗi khối một cách hợp lý và tốn không quá nhiều chi phí.

5.2 Xây dựng hệ thống

Với yêu cầu của bài toán, ta cần xây dựng hệ thống quản lý văn bằng cho các cơ sở giáo dục, và hệ thống tra cứu thông tin văn bằng cho phía doanh nghiệp (hoặc người có nhu cầu tra cứu).



Hình 5.1: Sơ đồ tổng quan các hệ thống và mạng chuỗi khối

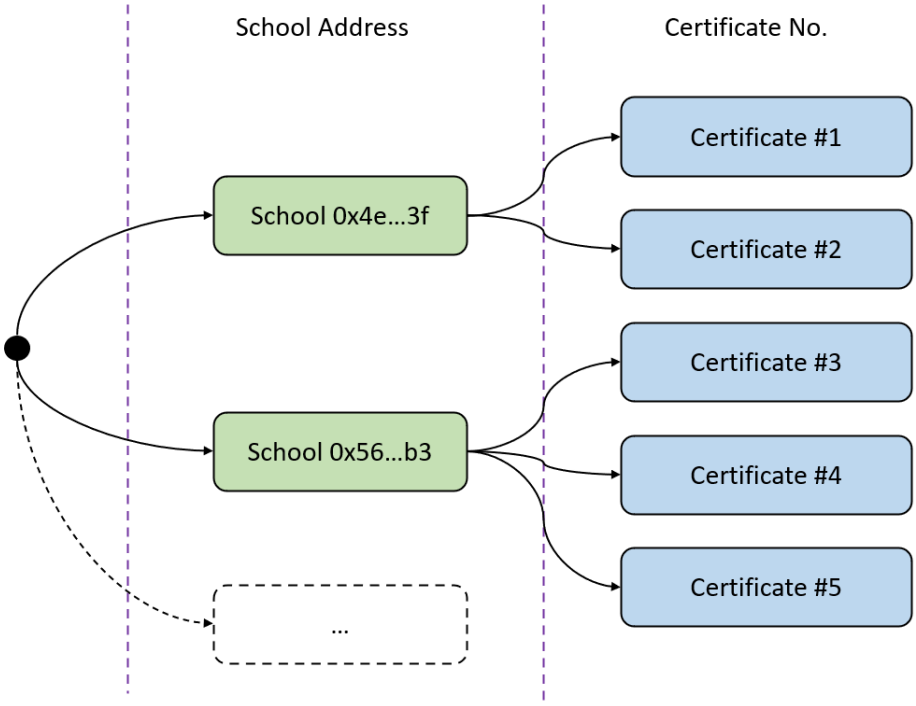
Những hệ thống này sẽ tương tác với các hợp đồng thông minh trên mạng chuỗi khối, bao gồm *Kho văn bằng* (Certs) và *Ví đa chữ ký* (MultiSig. Wallet). Trong bài báo cáo này, em xin phép tập trung trình bày về phần thiết kế các hợp đồng thông minh được triển khai cùng hệ thống.

5.2.1 Kho văn bằng

Kho văn bằng (hay *kho chứng chỉ*) là một hợp đồng thông minh nắm giữ thông tin các văn bằng được lưu trữ trên mạng chuỗi khối. Thông tin được tổ chức theo cấu trúc dạng cây, nắm giữ địa chỉ ví của các cơ sở giáo dục và danh sách các văn bằng cơ sở đó đã cấp phát.

Để lưu trữ văn bằng trên mạng chuỗi khối, các cơ sở giáo dục cần sử dụng một

địa chỉ ví để tương tác với hợp đồng thông minh được triển khai trên mạng đó. Mỗi cơ sở phát hành văn bằng sẽ có một địa chỉ ví xác định, và trên cấu trúc cây được mô tả như trên, địa chỉ này sẽ được lưu trữ ở các *nút bậc 1* (level-1 node), ký hiệu `School Address`; thông tin văn bằng được thể hiện qua các số hiệu tương ứng với các *nút lá* (leaf) với ký hiệu `Certificate No.`.



Hình 5.2: Cấu trúc lưu trữ thông tin trong *kho văn bằng*

Kho văn bằng cung cấp một số chức năng liên quan đến lưu thông tin, tra cứu thông tin văn bằng, phụ lục văn bằng, thông tin của các cơ sở giáo dục, tổ chức cấp phát văn bằng, chứng chỉ. Tuy nhiên, trong phạm vi bài báo cáo này, ta quan tâm đến hai chức năng cơ bản:

- Lưu thông tin văn bằng
- Xem thông tin văn bằng

Đối với việc *lưu thông tin*, địa chỉ ví của cơ sở (hay người) gửi yêu cầu lưu thông tin văn bằng sẽ được lấy làm `School Address`, và thông tin được gửi khi tương tác với hợp đồng thông minh sẽ được lưu tương ứng với số hiệu `Certificate No.`. Như vậy, sẽ không có tình huống một cơ sở giáo dục phát hành văn bằng với địa chỉ của cơ sở khác, trường hợp nhầm lẫn do vô ý hoặc có chủ đích không thể xảy ra.

Để *xem thông tin*, người tra cứu cần cung cấp địa chỉ ví của cơ sở phát hành văn bằng (`School Address`), và số hiệu văn bằng của cơ sở đó (`Certificate No.`).

5.2.2 Ví đa chữ ký

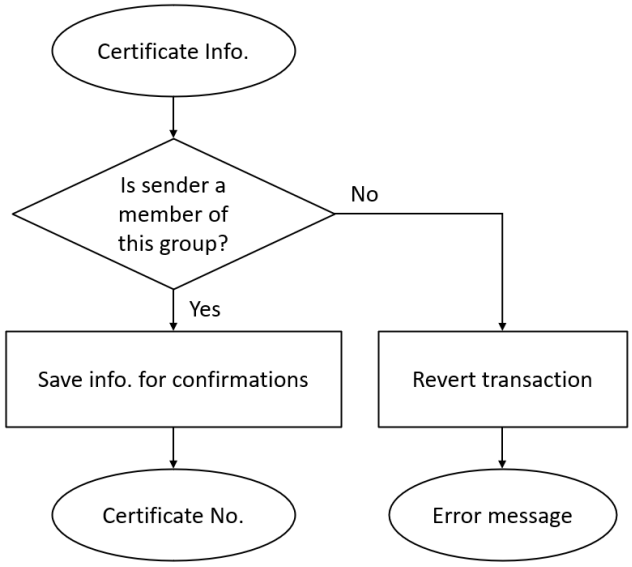
Ví đa chữ ký là một hợp đồng thông minh với mục đích tăng tính bảo mật cho quá trình tương tác thay đổi thông tin văn bằng trên chuỗi khối.

Với việc "đẩy" thông tin văn bằng lên mạng chuỗi khối một cách thông thường, mỗi cơ sở giáo dục sử dụng địa chỉ ví của một cá nhân đại diện để tương tác, hoặc lựa chọn một địa chỉ ví và sử dụng chung cho cá nhân trong cơ sở. Điều này đảm bảo mỗi cơ sở cấp phát chứng chỉ có một địa chỉ `School Address` duy nhất. Tuy nhiên, khi nhiều cá nhân cùng dùng một địa chỉ ví, khả năng mất cắp tài sản liên kết với địa chỉ này càng lớn, đặc biệt khi nó còn được sử dụng trong các giao dịch khác có giá trị về tài chính (như địa chỉ sở hữu tiền mã hoá với giá trị cao trên các *sàn giao dịch*[1], hay liên kết với các *DApp* khác). Do đó, một cơ chế giúp giảm thiểu khả năng nhiều người cùng sở hữu một địa chỉ ví và có thể sử dụng địa chỉ ví để xác thực thông tin cơ sở cấp phát văn bằng là vô cùng cần thiết. *Ví đa chữ ký* ra đời để giải quyết vấn đề này.

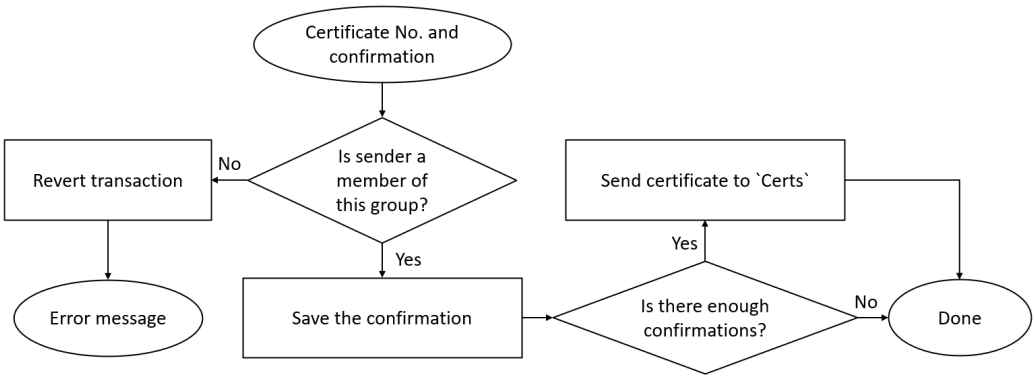
Không giống với các *hệ thống xác thực đa chữ ký*[2] khi ít nhiều phụ thuộc vào các cơ chế xác thực phức tạp, *ví đa chữ ký* sử dụng các tính năng, lợi thế của hợp đồng

[1] Exchange
[2] Multi-signature authentication system

thông minh và mạng chuỗi khối. Ở *ví đa chữ ký*, mỗi hành động cần thực thi (ở đây là việc cấp phát văn bằng) yêu cầu một số lượng nhất định sự đồng ý từ cá nhân. Địa chỉ ví của các cá nhân này đã được thêm vào danh sách "thành viên" ngay từ khi hợp đồng thông minh này được triển khai, và họ được coi như các "cổ đông" của "doanh nghiệp" cấp phát văn bằng khi có "tiếng nói" trong các "hoạt động" ở đây. Mỗi cơ sở cấp phát văn bằng sử dụng một *ví đa chữ ký* duy nhất, và địa chỉ của hợp đồng thông minh này đại diện cho địa chỉ ví của cả cơ sở đó. Các văn bằng cần được đẩy lên *kho văn bằng* sẽ được một cá nhân trong cơ sở gửi lên "ví" này. Các thành viên khác trong cơ sở có thể xem thông tin các văn bằng được gửi lên, và đưa ra biểu quyết "đồng ý" hay "không đồng ý" trên hợp đồng thông minh. Khi số lượng sự đồng ý đạt ngưỡng nhất định (được thiết lập từ đầu), các văn bằng đó được đẩy lên "kho", và thông tin được lưu trữ trên mạng chuỗi khối.



Hình 5.3: Tạo yêu cầu cấp phát văn bằng lên *ví đa chữ ký*



Hình 5.4: Đưa ra biểu quyết để đưa văn bằng trên "ví" lên "kho"

5.2.3 Các yêu cầu liên quan

Với *hệ thống tra cứu*, không có có quá nhiều yêu cầu cần thực hiện. Hiện nay, các tiện ích, phần mềm được tạo ra, đáp ứng nhu cầu kết nối đến mạng chuỗi khối, trong số đó có thể kể đến *Web3.js*. *Web3.js* là một dự án mã nguồn mở của *ChainSafe*, được viết chủ yếu bởi ngôn ngữ lập trình *JavaScript*, cung cấp các công cụ giúp người dùng tương tác với mạng Ethereum và các hợp đồng thông minh trên mạng này. *Web3.js* có thể được tích hợp cho các *DApp* dạng *web*[3] (webapp). Việc tra cứu thông tin văn bằng trở nên đơn giản và hết sức thân thiện, khi người dùng có thể trực tiếp thao tác qua giao diện trên trình duyệt.

Đối với *hệ thống quản lý*, yêu cầu quan trọng nhất của hệ thống là ghi lại thông tin các văn bằng được tạo ra. Các văn bằng này sau khi được cấp phát cần được cập nhật trạng thái để tránh việc đẩy nhiều lần lên chuỗi khối. Mặc dù việc gửi một văn bằng nhiều lần lên chuỗi khối không gây ảnh hưởng xấu gì về mặt thông tin do các thiết kế các hợp đồng thông minh không cho phép hai văn bằng cùng `School Address` và `Certificate No.` cùng tồn tại; tuy nhiên, mỗi khi tương tác với hợp đồng thông minh, chi phí giao dịch sẽ phát sinh. Như đã trình

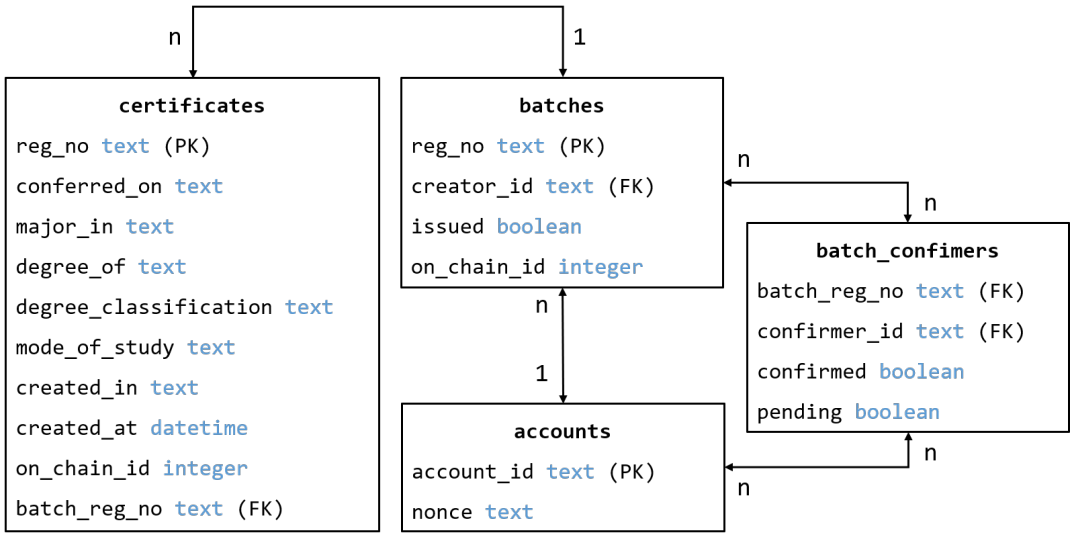
[3]Website

bày ở *Hình 5.1*, phía *máy chủ* (server) sẽ lắng nghe các *sự kiện* (event) khi trạng thái của hợp đồng thông minh thay đổi, nhanh chóng cập nhật thông tin tương ứng trong *cơ sở dữ liệu*, và thông báo lên giao diện người dùng. Ở hệ thống này, để tương tác được với hợp đồng thông minh, các thành viên trong cơ sở cấp phát văn bằng cần sử dụng thêm một *ví mã hoá*[4] lưu trữ các địa chỉ ví của họ. Một số ví mã hoá phổ biến có thể kể đến như *MetaMask*, *Coinbase*.

Việc thiết kế cơ sở dữ liệu cũng cần đảm bảo đáp ứng các yêu cầu của một hệ thống tương tác với chuỗi khối. Đối với những thực thể cần được lưu trữ trên mạng chuỗi khối, ta cần thêm một số thuộc tính định danh trên chuỗi khối. Cụ thể, trong thiết kế cơ sở dữ liệu, bảng `certificates` lưu trữ thông tin cho các văn bằng mà cơ sở giáo dục cấp phát. Ngoài các trường cơ bản - như *số hiệu văn bằng* (cột `reg_no` trong bảng), *tên người được cấp phát văn bằng* (`conferred_on`), *chuyên ngành* (`major_in`), *tên loại văn bằng* (`degree_of`), *đánh giá văn bằng* hay *xếp loại* (`degree_classification`), *hình thức đào tạo* (`mode_of_study`), *nơi cấp* (`created_in`) và *ngày cấp* (`created_at`) - trường *định danh trên chuỗi* (tương ứng với cột `on_chain_id` trong bảng) giúp lưu trữ liên kết giữa dữ liệu trong hệ thống quản lý và mạng chuỗi khối. Để giảm chi phí giao dịch khi tương tác với hợp đồng thông minh, các văn bằng được gom lại thành những *lô* (batch), và được đẩy lên chuỗi khối đồng thời. Chính vì vậy, bảng `certificates` cần thêm một cột lưu thông tin định danh cho các lô được cấp phát, ký hiệu là `batch_reg_no`. Thông tin các lô văn bằng này được thể hiện tương ứng trong bảng `batches`, với các trường thông tin liên quan tới *định danh* `reg_no`, *người tạo* (`creator_id`), và tất nhiên là cả *định danh trên chuỗi* (`on_chain_id`).

Với các thành viên tham gia cấp phát văn bằng, thông tin được lưu trong bảng `accounts`. Để truy cập vào hệ thống quản lý này, như thông thường, các thành viên cần xác thực tài khoản. Nhằm tận dụng các tính năng bảo mật của *mã hoá*

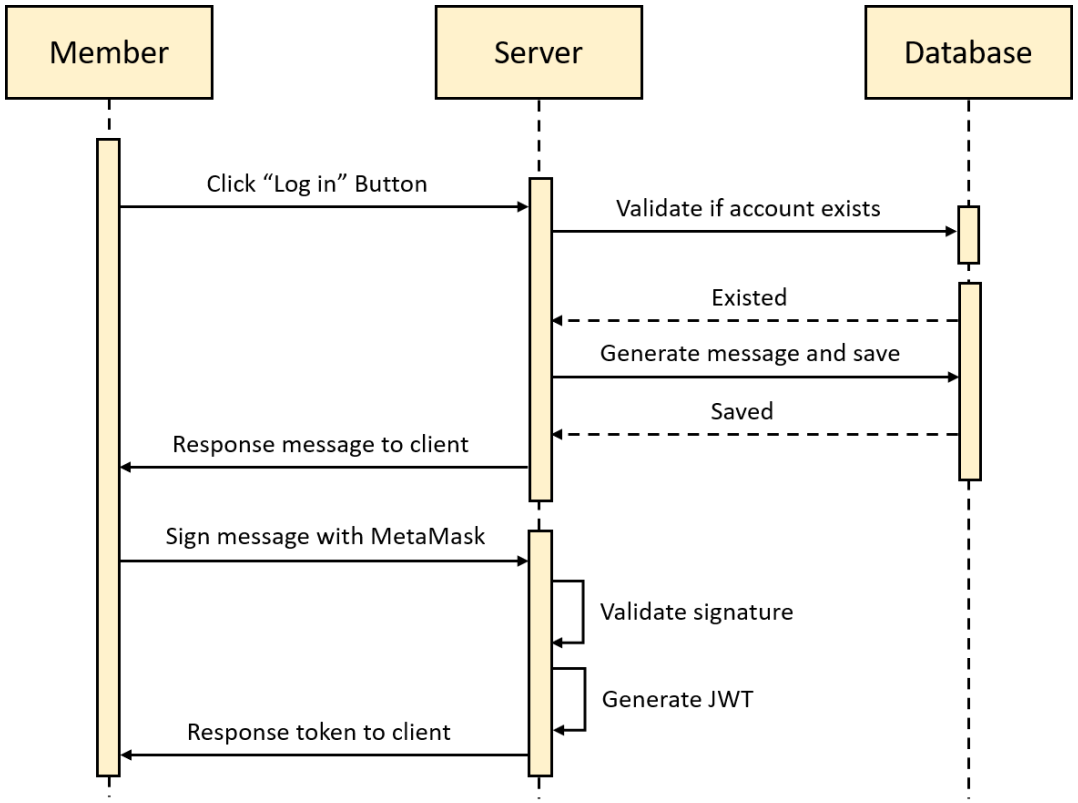
[4]Crypto wallet



Hình 5.5: Biểu đồ cơ sở dữ liệu cho hệ thống quản lý.

bất đối xứng[5] (cụ thể là *chữ ký số*[6]), em xin mang đến một cách xác thực nhanh gọn và an toàn qua ví MetaMask, tạm gọi là "xác thực một nút". Với cách xác thực này, mỗi khi một thành viên muốn truy cập vào hệ thống, một thông điệp ngẫu nhiên sẽ được gửi tới và họ cần ký thông điệp này và gửi lại cho máy chủ hệ thống. Việc "ký" thông điệp hoàn toàn được hỗ trợ bởi ví MetaMask với chỉ một thao tác bấm nút đơn giản. Đồng thời, tính hợp lệ của chữ ký cũng được kiểm tra một cách nhanh chóng dựa trên các trường thông tin như địa chỉ ví và thông điệp. Sau đó, một *mã xác thực có thời hạn*[7] được gửi lại cho phía thành viên, giúp họ tương tác với hệ thống trong một khoảng thời gian nhất định mà không cần phải đăng nhập lại. Với cách xác thực này, các thành viên tham gia cấp phát văn bằng không cần ghi nhớ mật khẩu để đăng nhập, mức độ bảo mật cho hệ thống cũng được đảm bảo. Và quay trở lại bảng `accounts`, định danh cho thành viên trong hệ thống sẽ tương ứng với *địa chỉ ví*, được lưu trong cột `account_id`, và thông điệp (message) dùng được ghi lại tương ứng với cột `nonce` (cách đặt tên dựa theo BitCoin) trong bảng.

[5] Asymmetric cryptography
[6] Digital signature
[7] Expirable token



Hình 5.6: Biểu đồ trình tự đăng nhập với "xác thực một nút".

Với việc sử dụng đa xác thực bằng ví MetaMask, mỗi khi một lô văn bằng cần được đẩy lên mạng chuỗi khối, mỗi thành viên cần biểu quyết đồng ý hoặc từ chối. Thông tin về các biểu quyết này được lưu trong bảng `batch_confirmers`, với hai *khoá ngoại lai*[8] (FK) tương ứng với các *khoá chính*[9] (PK) của bảng `accounts` và bảng `batches`.

[8]Foreign key - FK
[9]Primary key - PK

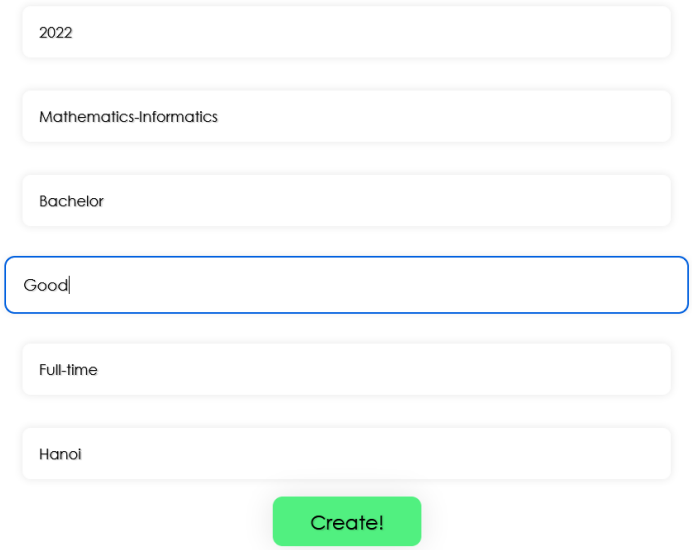
5.3 Triển khai và kết quả

Giao diện *hệ thống quản lý* được xây dựng với thư viện *React* (được phát triển bởi đội ngũ *Facebook* với sự đóng góp của cộng đồng) sử dụng ngôn ngữ lập trình *TypeScript*. Đồng thời, người dùng cần đồng ý kết nối ví mã hoá *MetaMask* với hệ thống để sử dụng các tính năng tương tác với hợp đồng thông minh. Phía máy chủ, *Node.js* được lựa chọn, ta sử dụng *Express* để tạo các *API*[10]. Thông tin cần thiết được lấy từ cơ sở dữ liệu, và việc trao đổi giữa máy chủ và giao diện được thực hiện qua kết nối *HTTP* và *Web Socket*, các cập nhật từ một người sẽ được thông báo ngay lập tức tới các cá nhân khác trong cùng cơ sở cấp phát văn bằng.

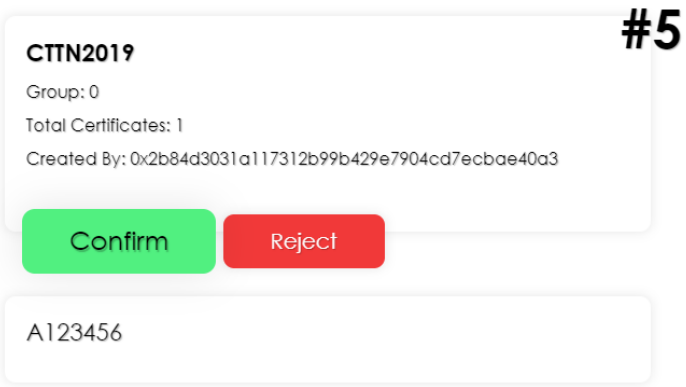
Hệ thống tra cứu sử dụng cấu trúc *trang tĩnh*[11], triển khai trên *GitHub Pages* với mã nguồn công khai, cung cấp cho người dùng một công cụ tra cứu thông tin văn bằng nhanh chóng, đáng tin cậy. Các doanh nghiệp có thể lấy danh sách *địa chỉ ví* của các cơ sở giáo dục tại trang thông tin (website) của cơ sở giáo dục đó, hoặc lấy từ một cơ quan có độ tin cậy lớn (như *Bộ Giáo dục và Đào tạo* chẳng hạn). Thông tin số hiệu văn bằng sẽ được ứng viên cung cấp.

Dưới đây là một số hình ảnh khi người dùng trải nghiệm hệ thống, và em xin lưu ý *đây chưa phải là hình ảnh của hệ thống hoàn thiện*.

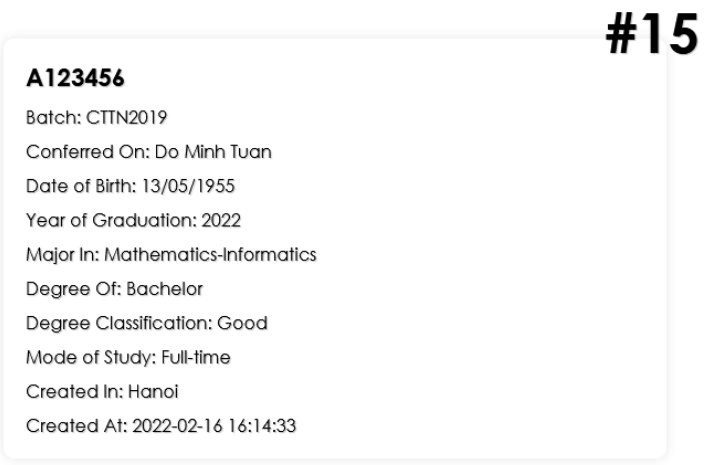
[10] Application Programming Interface
[11] Static web



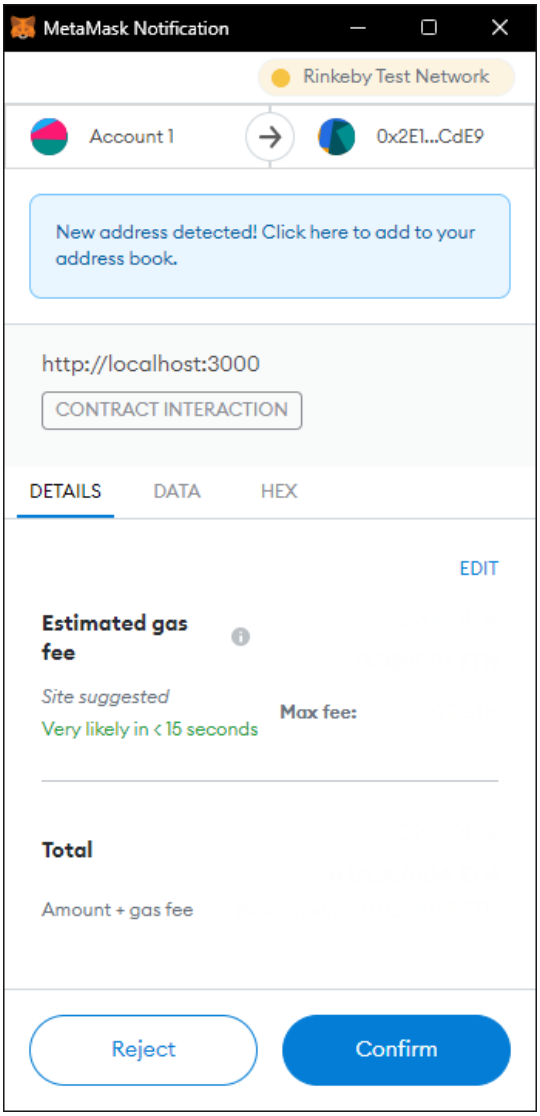
Hình 5.7: Một phần giao diện thêm văn bằng



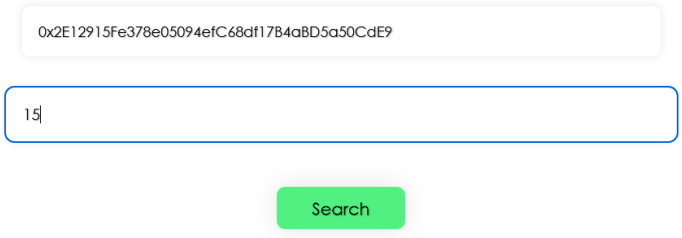
Hình 5.8: Một phần giao diện biểu quyết thêm văn bằng



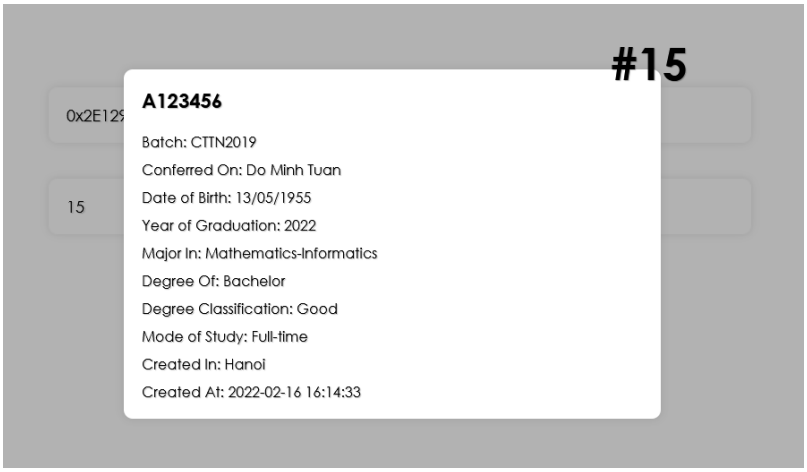
Hình 5.9: Một phần giao diện thông tin văn bằng đã được thêm



Hình 5.10: Giao diện MetaMask tương tác với hợp đồng thông minh



Hình 5.11: Một phần giao diện tra cứu văn bằng



Hình 5.12: Một phần giao diện thông tin văn bằng khi tra cứu

5.4 Đánh giá

Nhìn chung, các hệ thống ta đã trình bày ở trên đáp ứng tốt các yêu cầu của bài toán đã nêu. Tuy nhiên, một số hạn chế vẫn có thể chỉ ra, như:

1. Hiện tại, các hợp đồng thông minh đang được triển khai trên *mạng kiểm thử*[12] (testnet) nên chi phí giao dịch chưa được đề cập. Nếu triển khai trên *mạng chính*[13] của Ethereum, phí giao dịch khá là cao.

2. Hệ thống chưa hỗ trợ lưu và tra cứu phụ lục văn bằng, hay các thông tin kèm theo của một văn bằng/chứng chỉ giáo dục.

Đối với *hạn chế về chi phí giao dịch*, đây là một bài toán khá đau đầu với những *DApp* triển khai trên mạng Ethereum. Tuy nhiên, trong những năm gần đây, rất nhiều giải pháp giảm chi phí và tăng tốc độ xác thực giao dịch trên mạng chuỗi khối đã được đưa ra, trong số đó có thể kể đến như *Plasma*, *Matic*.

Các thông tin bổ sung cho văn bằng trên chuỗi khối sẽ được cân nhắc thêm vào thiết kế, phụ thuộc vào việc giảm thiểu chi phí giao dịch khi tương tác với hợp đồng thông minh.

[12]Testnet
[13]Mainnet

Kết luận

Tính tới thời điểm này, chuỗi khối không còn gì là mới mẻ, nhưng còn rất nhiều vấn đề ta cần giải quyết để khai thác được tiềm năng tối đa của nó. BitCoin với mang đến sự thịnh hành của tiền mã hoá, Ethereum lại phổ biến hợp đồng thông minh và *DApp*, đem tới nhiều cái nhìn tích cực hơn về công nghệ này. Và rồi nay mai đây thôi, những giải pháp cải tiến mạng chuỗi khối tiếp tục ra đời, đồng thời sự phổ cập kiến thức về nó cũng sẽ được đẩy mạnh. Em tin rằng, không lâu nữa, rất nhiều bài toán sẽ có thêm lời giải hợp lý khi đi cùng chuỗi khối.

Cảm ơn thầy, cô, và mọi người đã theo dõi những gì em trình bày trên đây. Em rất mong nhận được sự góp ý, và cũng mong, những gì em viết ra, tạo ra sẽ sớm góp phần nhỏ vào sự phát triển chung, ít nhất là tại đất nước Việt Nam thân yêu này.

Như thường lệ, mọi mã nguồn được công khai và lưu trữ tại GitHub:

`https://github.com/DCerts`

Tài liệu tham khảo

[1] Diniel Drescher, *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, 2017

[2] Investopedia, *Blockchain Definition: What You Need To Know*

[3] Investopedia, *Consensus Mechanism (Cryptocurrency) Definition*

[4] 101 Blockchains, *Blockchain Technology Explained: A Decentralized Ecosystem*

[5] R. C. Hansdah, *A Multisignature Scheme for Implementing Safe Delivery Rule in Group Communication Systems*

[6] Ethereum, *Ethereum Developement Documentation*

[7] ConsenSys Academy, *MultiSig. Wallet Excercise*

[8] ChainSafe, *Web3.js Documentation*

[9] Internet, *các thông tin về Chuỗi khối*