ỨNG DỤNG HỆ THỐNG MÃ HÓA LAI VÀO CÔNG TÁC BẢO MẬT TRONG TRUYỀN TẢI ĐỀ THI

APPLICATION OF THE HYBRID CRYPTOSYSTEM TO SECURITY OF EXAM QUESTIONS TRANSMISSION

Nguyễn Trần Quốc Vinh
Trường Đại học Sư phạm, Đại học Đà Nẵng
Email: ntquocvinh@yahoo.com

Nguyễn Tùng Sinh
HVCH Đại học Đà Nẵng

TÓM TẮT

Bảo mật đề thi có vai trò hết sức quan trọng đối với các kỳ thi. Đặc biệt, với các kỳ thi có quy mô lớn, đề thi thường được truyền tải qua mạng internet. Người ta thường dùng mật mã phi đối xứng trong truyền tải thông tin qua mạng vì nó đảm bảo không cho phép bên thứ ba truy cập dữ liệu đã mã hoá tại các điểm trung gian. Tuy nhiên, trường hợp đề thi có kích thước lớn, như đề thi bao gồm các nội dung đa phương tiện, mật mã phi đối xứng có thể dẫn đến vấn đề yêu cầu tài nguyên lớn hơn trong việc mã hoá, giải mã. Bài viết kết hợp tốc độ của phương pháp mật mã đối xứng với tính an toàn trong việc truyền tải qua mạng của mật mã phi đối xứng trong hệ thống mã hóa lai để xây dựng hệ thống truyền tải đề thi. Ngoài ra, bài viết còn so sánh các phương pháp mật mã khác nhau để chọn giải pháp tối ưu.

Từ khóa: đề thi; nội dung có kích thước lớn; AES; ECC; hệ mật mã lai

ABSTRACT

Security of exam questions is very important for all exams, especially for large-scaled exams. Exam questions are often transmitted through the internet. The asymmetric encryption is often used for the transmission of information over the network because it certainly does not allow the third party to access encrypted data at intermediate points. However, in the case of large-sized exam questions, since the exams questions consist of multimedia contents, asymmetric encryption can lead to the problem that requires greater resources in the encoding. The author combines the speed of symmetric encryption method with the security of the transmission through the network of asymmetric encryption in hybrid encryption system to build an exam questions transmission system.

Key words: Exam questions; large-size contents; AES; ECC; hybrid cryptosystems

1. Đặt vấn đề

Đề thi là một trong những tài liệu mật của quốc gia. Hằng năm, các trường học phải thường xuyên tổ chức các kỳ thi nhằm tuyển chọn học sinh vào trường, kỳ thi đánh giá kết quả học tập của học sinh như: Thi tuyển sinh đầu vào, kiểm tra chất lượng, thi học kỳ, thi tốt nghiệp, thi học sinh giỏi… Trong các kỳ thi đó, có những đợt thi các trường thi chung đề thi của Bộ Giáo dục và Đào tạo, của Sở Giáo dục và Đào tạo (SGD&ĐT).

Hiện nay, SGD&ĐT bảo mật đề thi của các kỳ thi bằng cách niêm phong các túi đề thi. Sau đó, phối hợp cùng với lực lượng an ninh mang đến các trường đặt làm địa điểm thi để bàn giao cho các hội đồng thi. Từng hội đồng thi có trách nhiệm lập phương án bảo vệ đề thi trong suốt quá trình tổ chức kỳ thi. Với việc nhận và chuyển đề thi theo phương thức này có thể gặp nhiều trở ngại cũng như việc đảm bảo an toàn, bí mật cho đề thi chứa đựng nhiều yếu tố rủi ro, kinh phí cho việc giao nhận và bảo vệ đề thi rất tốn kém.

Để góp phần khắc phục một phần những hạn chế trên, việc sử dụng các công cụ của mật mã học ứng dụng vào công tác bảo mật đề thi trong truyền tải đề thi qua mạng là một vấn đề mang tính thời sự và cấp thiết.

Trong bài viết này, các tác giả sử dụng hệ mật mã lai với sự kết hợp giữa mật mã đối xứng và mật mã phi đối xứng. Mật mã lai đã được đề cập trong nhiều tài liệu khác nhau. Tuy nhiên, cho đến nay, các tác giả chưa gặp công trình nào công bố kết quả kết hợp mật mã đường cong eliptic (ECC, Elliptic curve cryptography) và chuẩn mã hoá nâng cao (AES, Advanced Encrypttion

Standard), đặc biệt là so sánh các hệ mật mã khác nhau và lựa chọn sự kết hợp tốt nhất.

2. Mật mã đối xứng với AES và phi đối xứng với ECC

Ngày nay, với sự phát triển bùng nổ của internet, nhu cầu trao đổi thông tin qua mạng ngày càng được ứng dụng rộng rãi trên tất cả mọi lĩnh vực với lượng thông tin giao dịch ngày càng lớn và đa dạng, dung lượng tệp trao đổi rất lớn. Tuy nhiên, đây cũng chính là môi trường thuận lợi để tội phạm máy tính ngày càng gia tăng, chúng thực hiện các cuộc tấn công vào các hệ thống nhằm khai thác thông tin, lấy cắp tài khoản để trục lợi, lừa đảo người dùng… Tội phạm máy tính rất đa dạng, ngày càng gia tăng về số lượng, độ tinh vi, mức độ nghiêm trọng và tổn thất.

Với sự xuất hiện của máy tính, các tài liệu văn bản và các thông tin quan trọng đều được số hóa và xử lý trên máy tính, đồng thời được truyền đi trên một môi trường mà mặc định là không an toàn. Do đó, yêu cầu về việc có một cơ chế, giải pháp để bảo vệ sự an toàn và bí mật của các thông tin nhạy cảm, quan trọng ngày càng trở nên cấp thiết. Việc bảo vệ dữ liệu là vấn đề mà tất cả những ai sử dụng máy tính đều phải quan tâm. Mã hóa được xem là mức bảo vệ tối ưu nhất đối với dữ liệu, giúp thông tin không bị lộ và nâng cao độ an toàn trong các giao dịch truyền tải thông tin. Sự ra đời của ngành mật mã học đã giải quyết phần nào mong muốn đó.

Cho đến nay, đã có nhiều phương pháp mã hóa và các thuật toán tương ứng với mỗi phương pháp được ứng dụng để mã hóa thông tin, tiêu biểu là mật mã đối xứng (symmetric cryptography) và phi đối xứng (asymmetric cryptography).

Mật mã đối xứng sử dụng cùng một khóa cho cả hai quá trình mã hóa và giải mã. Ưu điểm của phương pháp này là tốc độ xử lý nhanh. Tuy nhiên, khả năng bảo mật chưa thực sự được an toàn khi cần trao đổi thông tin ở mức xử lý với nhiều bên nhận và gửi dữ liệu. Thuật toán mật mã đối xứng được biết đến rộng rãi là AES. Đây là một chuẩn mật mã cao cấp với khóa bí mật cho phép xử lý các khối dữ liệu đầu vào sử dụng các khóa có độ dài 128, 192, 256 bit.

Mật mã phi đối xứng, hay còn được gọi là mật mã khóa công khai (public key cryptography) có nguyên tắc hoạt động là mỗi bên tham gia truyền tin sẽ có hai khóa. Một khóa được dùng để mã hóa và có thể được công bố công khai để bất kỳ ai cũng có thể sử dụng khóa này để gửi tin cho chủ thể (được gọi là khóa công khai – public key) và một khóa được dùng trong quá trình giải mã và được giữ bí mật (gọi là khóa bí mật – private key). Khóa giải mã không thể tính toán được từ khóa mã hóa.

Ưu điểm của mật mã phi đối xứng là việc quản lý khóa sẽ linh hoạt và hiệu quả hơn. Người sử dụng chỉ cần bảo vệ khóa bí mật của mình. Tuy nhiên, nhược điểm của mật mã khóa công khai nằm ở tốc độ thực hiện, nó chậm hơn rất nhiều so với mã hóa đối xứng.

2.1. Chuẩn mã hoá nâng cao (AES)

Thuật toán AES còn được biết đến với một tên gọi khác là giải thuật Rijndael bởi giải thuật này do nhà mật mã học người Bỉ, Vincent Rijmen và Joan Daeman phát triển. Thuật toán AES thực hiện trên mảng các byte đầu vào có độ dài 128 bit với các tùy chọn độ dài khóa là 128 bit, 192 bit và 256 bit. Cơ sở của thuật toán là các phép toán cộng ($\oplus$) và nhân ($\otimes$) trên trường Galoris GF. Bắt đầu của một quá trình mã hóa (hoặc giải mã) mảng byte đầu vào được sao chép vào mảng trạng thái S và được biến đổi tuần tự qua các vòng. Kết quả của vòng trước sẽ là đầu vào của vòng biến đổi tiếp theo. Như vậy, mảng trạng thái S chính là phần chứa kết quả trung gian qua mỗi vòng xử lý. Mỗi vòng sẽ bao gồm các phép biến đổi SubByte(), dịch vòng Shiftrows(), phép trộn cột Mixcolumns() và phép bổ sung khóa AddRoundkey() cho mỗi vòng dựa vào bảng khóa mở rộng được sinh ra từ khóa mã ban đầu bởi phép KeyExpansion(). Quá trình giải mã được thực hiện theo thứ tự ngược lại.

2.2. Hệ mật mã đường cong elliptic (ECC)

Cùng với thuật toán nổi tiếng RSA, năm 1985, hai nhà khoa học Neal Koblitz và Victor S. Miller đã độc lập nghiên cứu và đưa ra đề xuất ứng dụng lý thuyết toán học đường cong elliptic trên trường hữu hạn. Thuật toán mã hóa dựa trên đường cong elliptic (ECC) thực hiện

việc mã hóa và giải mã dựa trên tọa độ của các điểm trên đường cong elliptic.

Một đường cong elliptic là một vật thể với tính chất kép. Một là, nó là một đường cong, một vật thể hình học. Hai là, chúng ta có thể thêm vào các điểm trên đường cong như thể nếu chúng là những con số, do đó, nó là một đối tượng đại số.

Một đường cong elliptic là một đường cong được tạo ra bởi phương trình dạng mẫu:

$$y^2 = x^3 + ax + b.$$

Thuật toán mật mã dựa trên đường cong elliptic thực hiện việc mã hóa và giải mã dựa trên tọa độ của các điểm trên đường cong elliptic. Xét đẳng thức Q = kP; với Q, P là các điểm nằm trên đường cong elliptic. Có thể tính Q khá dễ dàng nếu biết k và P, nhưng rất khó xác định k nếu biết Q và P (phép nhân được xác định bằng cách cộng liên tiếp cùng điểm P). Bên cạnh công thức của đường cong elliptic, một thông số quan trọng khác của đường cong elliptic là điểm cơ sở - điểm G. Điểm G đối với mỗi đường cong elliptic là cố định. Trong hệ mật mã ECC thì một số nguyên lớn k đóng vai trò như một khóa riêng, trong khi đó kết quả của phép nhân giữa k với điểm G được coi như là khóa công khai tương ứng. Điểm mấu chốt của thuật toán mật mã sử dụng đường cong elliptic là dựa vào độ phức tạp cao của bài toán logarit rời rạc trong hệ đại số.

3. Giải pháp và xây dựng chương trình

Trên thực tế, mật mã đối xứng và mật mã phi đối xứng, mặc dù có nhiều ưu điểm so với các phương pháp mã hóa trước đó, nhưng với nhược điểm của mình nên vẫn chưa thể được sử dụng một cách rộng rãi trong tất cả mọi lĩnh vực ứng dụng.

Để khắc phục các nhược điểm trên, tác giả đề xuất kết hợp hai phương pháp mã hóa này, cụ thể là kết hợp chuẩn mật mã AES và hệ mật mã đường cong elliptic với nhau và được gọi là hệ thống mã hóa lai (Hybrid Cryptosystems). Với sự kết hợp này, hệ thống đã tận dụng được các điểm mạnh của hai hệ thống ở trên đó là tốc độ của mật mã đối xứng và tính an toàn của mật mã phi đối xứng.

3.1. Giải pháp

Hiện tại, các trường trung học phổ thông (TTHPT) và SGD&ĐT đều đã được trang bị các máy tính có cấu hình đủ mạnh để chạy các phần mềm có yêu cầu cấu hình cao và đều đã được kết nối internet. Do đó, việc tin học hóa hệ thống quản lý đề thi trong các trường phổ thông dưới sự quản lý của SGD&ĐT có thể được triển khai theo mô hình sau:

Tại máy chủ của SGD&ĐT (bên A) và các TTHPT (bên B), chúng ta cài đặt phần mềm, ví dụ có tên "Phần mềm truyền tải đề thi". Mỗi khi tiến hành đợt thi, việc truyền tải đề thi sẽ được thực hiện thông qua phần mềm theo nguyên tắc dùng mật mã phi đối xứng ECC để truyền khoá mật mã đối xứng AES, dữ liệu là nội dung đề thi sẽ được mã hoá theo phương pháp đối xứng AES và truyền tải qua kênh thông thường (Hình 1). Khoá bí mật do SGD&ĐT lưu giữ. Khóa công khai sẽ được gửi cho các TTHPT hoặc được công bố công khai trên website.

❖ **Thực hiện việc trao đổi khóa mật mã đối xứng**

- Bên B sinh khoá mã hoá AES cho phiên làm việc, gọi là SK. SK là một chuỗi ký tự ngẫu nghiên.

- Bên B dùng khoá công khai ECC mã hoá SK và gửi dữ liệu SK đã mã hoá cho bên A.

- Bên A dùng khoá bí mật ECC để giải mã và thu được SK. Cả bên A và bên B đều lưu giữ SK để dùng cho mã hoá dữ liệu đề thi cũng như các dữ liệu trao đổi khác bằng AES.

❖ **Bên A gửi dữ liệu (đề thi) cho các trường**

- Bên A sử dụng SK đã nhận được để mã hóa đề thi;

- Đề thi đã được mã hóa sẽ được chuyển cho bên B qua kênh thông thường;

- Bên B sử dụng SK giải mã và thu được nội dung đề thi.

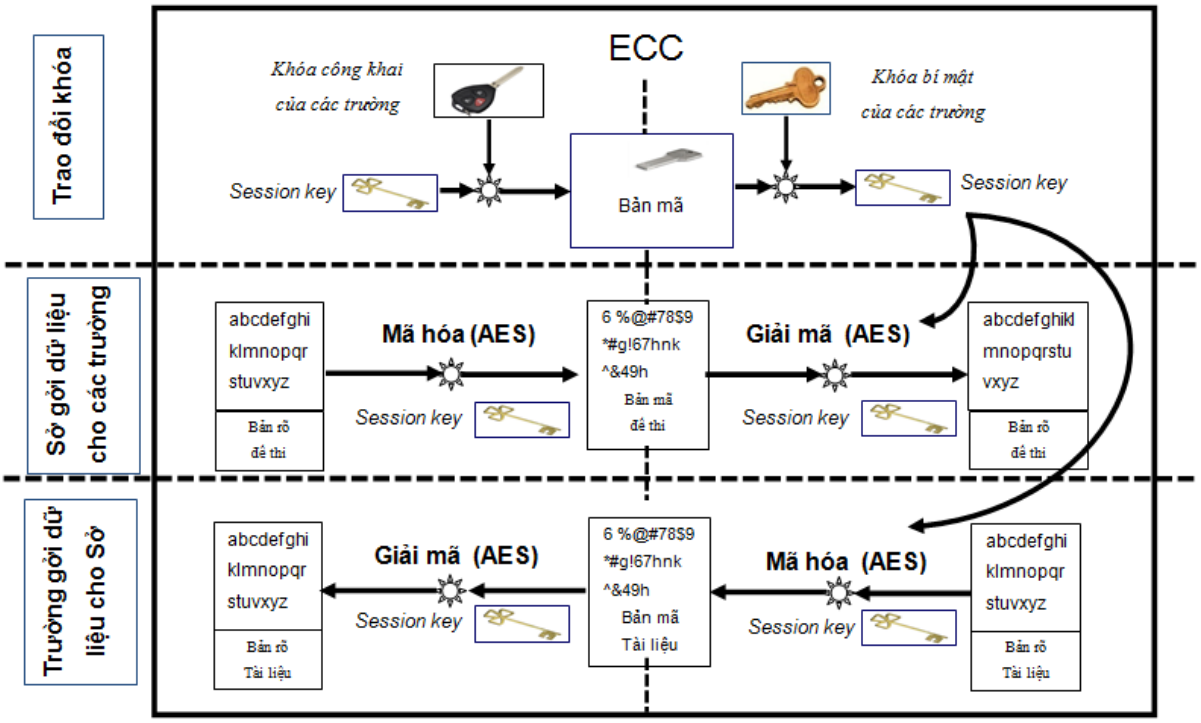
❖ **Trường phổ thông gửi dữ liệu cho SGD&ĐT**

Cũng tương tự trường hợp bên A gửi dữ liệu cho bên B, các bên tham gia tiến hành theo

các bước sau:

- Bên B dùng SK để mã hóa dữ liệu;
- Dữ liệu đã được mã hóa sẽ được chuyển cho bên A thông qua kênh thông thường;

Bên A dùng SK để giải mã và thu được dữ liệu mà bên B cần chuyển.



Hình 1. *Mô hình trao đổi dữ liệu qua internet*

3.2. Kết quả

Kết quả, tác giả đã xây dựng được chương trình cho phép chuyển tải đề thi nói riêng và dữ liệu nói chung qua mạng internet với khả năng mã hoá dữ liệu đảm bảo theo nguyên tắc bảo mật của mật mã phi đối xứng nhưng với hiệu quả của mật mã đối xứng. Hình 2 cho thấy chức năng khả năng SGD&ĐT tiếp nhận kết nối từ nhiều điểm thi – TTHPT khác nhau và lựa chọn tệp để truyền tải đến các TTHPT. Tương tự, chức năng gửi tệp từ TTHPT đến SGD&ĐT cũng được xây dựng. Vì mục đích nghiên cứu và chứng minh sự đúng đắn trong việc lựa chọn kết hợp ECC và AES, các tác giả còn xây dựng công cụ đánh giá hiệu quả khi so sánh ba hệ mật mã ECC, RSA và mật mã lai (kết hợp ECC và AES) khi truyền tải dữ liệu ở hai thông số thời gian thực hiện mã hoá và thời gian giải mã. Kích thước của khoá cho cả ba hệ mật mã có thể được thiết đặt ở giao diện cấu hình trên quan điểm đảm bảo tương đương về độ an toàn dữ liệu. Bảng 1 thể hiện kết quả đo thời gian mã hoá và thời gian giải mã các tệp có kích thước khác nhau với các độ dài khoá khác nhau tương ứng với độ dài khoá của EAS là 128bit và 256bit. Tài liệu [9] cung cấp bảng độ dài khoá cho các hệ mật mã để đảm bảo độ bảo mật tương đương. Chẳng hạn, với độ dài khoá cho EAS 128bit thì cho ECC sẽ là 256bit và RSA là 512bit. Với độ dài khoá cho EAS là 256bit thì cho ECC là 512bit và RSA là 1024bit. Cho trường hợp mật mã lai, độ dài khoá của ECC và EAS được chọn tương ứng. Các tác giả cũng tiến hành đo độ dài tệp đã mã hoá và thấy rằng, kích thước tệp đã mã hoá cho EAS và ECC là bằng nhau. Với trường hợp RSA, kích thước lớn hơn không đáng kể.

Hình 2. *SGD&ĐT truyền đề thi đến các trường*

Bảng 1. *So sánh thời gian thực thi giữa thuật toán ECC, RSA và mật mã lai (ECC kết hợp AES)*

#	Kích thước tệp (KB)	Thời gian mã hóa (ms)						Thời gian giải mã (ms)					
		ECC	AES	Hybrid	RSA	Tỉ lệ		ECC	AES	Hybrid	RSA	Tỉ lệ	
						ECC /Hibrid	RSA /Hibrid					ECC /Hibrid	RSA /Hibrid
	Độ dài khoá AES – 128 (ECC – 256, RSA – 512)												
1	2 209	105	90	92	9 877	**107**	**1.15**	116	109	109	69 503	**639**	**1.07**
2	5 378	141	123	125	27 205	**218**	**1.13**	266	207	207	107 681	**520**	**1.28**
3	9 617	274	234	235	35 503	**151**	**1.17**	333	294	294	198 594	**674**	**1.13**
	Độ dài khoá EAS – 256 (ECC – 512, RSA – 1024)												
4	2 209	227	122	123	39 879	**324**	**1.85**	241	136	136	276 957	**2036**	**1.77**
5	5 378	299	158	160	112 381	**702**	**1.87**	542	236	236	437 562	**1854**	**2.30**
6	9 617	563	278	280	149 325	**533**	**2.01**	678	337	337	798 534	**2370**	**2.01**

Các thử nghiệm cho kết quả như trên Bảng 1 được tiến hành trên máy tính có cấu hình: CPU Intel Core i3 2.4GHz, RAM DDR3 2GB, HDD 5400rpm và sử dụng hệ điều hành Windows 7. Các tệp được chọn ngẫu nghiên và có kích thước như ở cột kích thước cột trên bảng 1. Việc đo đạt được tiến hành 5 lần cho mỗi tệp với mỗi độ dài khoá và tính thời gian trung bình.

4. Kết luận

Tác giả đã nghiên cứu mật mã đối xứng và phi đối xứng, cụ thể là AES và ECC, để xây dựng được chương trình cho phép chuyển tải đề thi nói riêng và dữ liệu nói chung qua mạng internet với khả năng mã hoá dữ liệu đảm bảo theo nguyên tắc bảo mật của mật mã phi đối xứng nhưng với hiệu quả của mật mã đối xứng. Các tác giả cũng đã xây dựng công cụ để thử nghiệm các phương án ứng dụng hệ mật mã khác nhau. Kết quả so sánh thời gian mã hoá và thời gian giải mã thông qua tỉ lệ thời gian mã hoá giữa mã hoá lai và ECC, mã hoá lai và RSA và thời gian giải mã giữa mã hoá lai

và ECC, mã hoá lai và RSA trong Bảng 1 cho thấy, mã hoá lai luôn nhanh hơn ECC và nhanh hơn rất nhiều lần so với RSA. Nghiên cứu cần được hoàn thiện thông qua việc áp dụng truyền tải các dữ liệu mẫu theo quy định để đánh giá hiệu quả của giải pháp.

TÀI LIỆU THAM KHẢO

[1] Menezes A., Oorschot P., Vanstone S., *Handbook of Applied Cryptography*, CRC Press, 1997.

[2] New Comparative Study Between DES, 3DES and AES within Nine Factors, *Journal of computing*, Vol 2, ISSUE 3, 2010, ISSN 2151-9617.

[3] Zaidan A.A., Majeed A., "High Securing Cover-File of Hidden Data Using Statistical Technique and AES Encryption Algorithm", *World Academy of Science Engineering and Technology (WASET)*, Vol.54, ISSN: 2070-3724, P.P 468-479.

[4] Rhul M., *Advanced Cryptography*, 2004.

[5] Subasree S., Sakthivel N. K., *Design of anew security protocol using Hybrid Cryptography Algorithms*, School of Computing, Sastra University, Thanjavur – 613401, Tamil Nadu, INDIA, February 2010.

[6] Silverman J., *The Arithmetic of Elliptic Curves*, Springer-Verlag, 1985.

[7] Seroussi Blake I.G., Smart N., *Elliptic Curves in Cryptography*, Cambridge University Press, 1999.

[8] Hofheinz D., Kiltz Eike, "Secure Hybrid Encryption from Weakened Key Encapsulation", *CRYPTO 2007*, Springer.

[9] US National Security Agency, *The Case for Elliptic Curve Cryptography*, http://www.nsa.gov/business/programs/elliptic_curve.shtml (truy cập ngày 15/6/2013).

(BBT nhận bài: 04/10/2013, phản biện xong: 27/12/2013)