Authentication in E-learning systems: Challenges and Solutions

Quang-Huan Luu*, Duy-Minh Nguyen, Hoang-Anh Pham, Nguyen Huynh-Tuong

ABSTRACT
Digitization is gradually penetrating all aspects of modern society. As it changes the way people communicate, technology has revolutionized education and training in the 21st century. With the advantages of reasonable costs and flexible study time, online training is increasingly seen as an attractive alternative to the full-time on-campus training model. To assure quality of distance training and learning, it is crucial for the online learning management system to make sure the person accessing the course resources and performing learning activities is actually enrolled in the course. One of the important factors determining the security of this process is user authentication. In most cases, this role is done with a password, but the evidence shows that this method is easily compromised. While there are many alternatives available such as biometric methods, user-challenging methods, smart card methods, etc. The strong development of technology that requires confidentiality and authentication must be tightly coupled. A qualitative survey of user authentication systems is being used in today’s E-learning systems and a comparative study of various different authentication mechanisms presented in this paper. There are many methods of user authentication for online learning systems, but each method will have different advantages and disadvantages and has not completely solved the challenges of user authentication. The issue of user authentication still has many challenges that need to be solved thoroughly to improve the security of the system as well as the trust of users and society. This paper provides an overview of our approach and recommendations to address the mentioned issues. In addition, we propose a number of feasible approaches to improve user data privacy as well as improve the effectiveness of the authentication process in the online learning system.

Key words: Decentralized Authentication, Privacy, Merkle Tree, Blockchain

INTRODUCTION
Many top universities in the world have launched online courses up to master level such as the Massachusetts Institute of Technology, Harvard University and the University of Pennsylvania. By collaborating with online training platforms such as Coursera and edX, these institutions have opened entirely remote courses via the Internet. The distance learning process is facilitated by an online learning management system (also known as distance learning or e-learning system). This is a set of software applications that manage the teaching and learning process and the examination procedures. With no more than an Internet-connected computer, a student can access lectures, books and other learning materials, ask questions, submit assignments, and take graded tests just like with traditional learning methods. Originally, the e-learning management system was simply a piece of software that enabled a user to do different things online, including playing lecture video clips and participating in discussion forums. With the current needs, however, the online learning management system has grown into an independent educational environment. Students no longer have to go to lecture halls to meet their instructors; instead, they can interact via the Internet. Some online learning platforms even allow the students to remotely take exams or go through the admission procedure without visiting the campus. This online learning method requires learners to be proactive in their work. To assure quality of distance training and learning, it is crucial for the online learning management system to make sure the person accessing the course resources and performing learning activities is actually enrolled in the course. From the point of view of computer science, the point is to identify and reference a person in the real world as a user in the system. The entity in the system or the user identifier is represented by access to a computer location or resource. In an online learning management system, it is the right to access learning materials, interact with instructors and peers, submit assignments, and take exams. The management of user identification and authentication is among the challenges facing security researchers. The remainder of this article is divided into five sections. In the next one, we present some security
challenges facing online learning systems, analyzing the security elements and risks when authenticating based on user attributes. The following section provides an overview of our approach and recommendations to address the mentioned issues. The overall architecture and assessment will follow after that. The final section summarizes our key findings and proposes future research directions.

**CHALLENGES OF USER AUTHENTICATION IN E-LEARNING SYSTEMS**

For online training systems to continue growing and be accepted as an official form of training free of discrimination, security issues must be thoroughly addressed. The system must demonstrate its reliability and win the trust of users and the society regarding its quality of training and transparency, especially in online training. One prominent challenge is how to know if a student’s performance in the system is indeed his or hers. In traditional training, academic records including transcripts and examination results are stored and managed via written documents. Today, both online and offline training systems employ digital records, and digital data seem more likely to be erased or altered than are physical data. Therefore, it is imperative for students’ online learning results to be stored and processed in a clear, objective and transparent manner. Let’s have a closer look at this challenge via two common security issues: identity misuse and integrity of students’ academic results.

**Identity misuse**

A student’s identity in the system is used by another person. Possible causes: the student actively sharing the account or the account being attacked. Two testing-related scenarios could take place as follows:

- The online test is conducted in a controlled environment, on a university’s premise for instance. This is common in most of today’s educational institutions. Students study remotely on the e-learning platform, then when the time comes for term-end exams, they come to the institution’s campus to take the test, which is usually hosted online. Before entering the examination room, students present their student identification (ID) card to the examination officer for identity verification. When the number of students is large, this process is laborious and sometimes impractical. It is also open to error as the officer may be unable to determine if the ID card holder is its legitimate owner.

- The online test is conducted in an uncontrolled environment, off campus where educational institutions do not have any control over student identity. This is a typical situation for most online learning platforms. It is then the learning management system’s job to ensure the test-taker is a legitimate registrant on the system.

In the two cases above, the objectivity and reliability of the E-learning system, particularly of online testing, depends on its ability to ensure testing results are free from cheating, involuntary or voluntary tampering, and impersonation. This challenge pertains to authenticating test-takers, online or offline. When applying the right authentication mechanism, the educational institution can rest assured that student identity is in good check both before and during the test.

**Integrity of students’ academic results**

This aspect concerns the storing and handling of students’ academic results. This is particularly vital if the outcomes of distance learning are to be seen as equal to traditional training outcomes. Traditionally, student results are kept and maintained in paper records. In online learning, these records are stored digitally, often in databases. Undesirable data alterations happen when an intruder attacks the system, acquires unauthorized access to the record database, and modifies test results and transcripts. On the other hand, it is expected for users to perceive these (digital) data not as “real” (as written data) and open to modifications and deletion. In these cases, the challenge is to ensure data integrity and guarantee the transparency of students’ learning results.

**REVIEW OF EXISTING AUTHENTICATION METHODS**

**D/Password-Based Authentication**

User ID/Password is one of the most common authentication mechanisms used in online systems. Regardless of user type and user role, each user has a unique identifier to distinguish it from other users. Usually in the authentication process, the user ID is used along with the password. Users must provide both login information correctly to gain access to the system or application. This ID is used to assign permissions, monitor user activity and manage common activities on a specific system, network or application. Like other information systems, E-learning systems often use user ID and password as the main authentication mechanism. Regarding passwords, people often choose a password that is easy and intuitive; Today people have to have different passwords to be
authorized in many different systems. Therefore, these passwords are often similar and not complicated enough. The registration number or date of birth is used as well as the name and they have a habit of writing them on paper or some other place. To create a good password some rules must be followed (avoid personal names, use special characters, use capital letters, etc.).

Passwords generated by following rules are not intuitive and not easy to remember so users can forget their passwords. With the known risks of the authentication system through accounts and passwords such as disclosure, theft or users actively share this account with others to attend school instead. E-learning systems have used other methods to authenticate user identifiers.

### Biometric-Based Authentication

Authentication based on biometrics or characteristics is done by verifying the physical or behavioral characteristics of an individual. Biometrics frees users from having to memorize passwords or carry them because users themselves are locked to identify. Several biometric authentication features have been developed in recent studies and implemented in online learning systems including: fingerprint recognition, iris identification, face recognition, identification audio or combining these features in multimodal biometrics.

### Behavior-Based Authentication

The behavior-based authentication uses devices such as smartphones, smartwatches or other IoT devices. All of these devices offer a wide range of sensors that can detect different kinds of user behavior. The user behavior outcomes are processed and consolidated into a single value called the trust level. This trust level is sent to web services instead of passwords, the web service determines which trust threshold is needed to access their service or what features are available.

### User authentication by challenge questions

Based on the assumption that only the user knows his personal information and his past activities, the user attributes-based authentication model challenges the user with a set of security questions. These questions are generated based on user attributes, behavior, and past activities. Only by passing these questions can a user prove that he is an entity with the corresponding attributes in the system.

Challenge questions are created by extracting personal information such as social security number, day of birth, place of birth, student ID number. This information is managed based on the authentication system. A user profile includes user-specific information that is sensitive. This record is typically stored at the verifier and then used to verify their verification request.

Based on these conventional authentication methods, various instant fixes for solving the authentication challenges have been studied and proposed. These approaches can be divided into three different categories corresponding to what you know (knowledge-based), what you have (ownership-based), and what you are (inherent-based). Table 1 summarizes our investigation on the existing authentication methods.

The first drawback of knowledge-based is to memorize many passwords and passwords that are complex and difficult to remember, which can lead to confusion between passwords. The second is shoulder surfing, in which an outsider can track the user’s keyboard. Passwords are easily attacked by dictionary-based and exhausted methods. It is worth noting that some graphic passwords are also unavoidable with screen capture methods.

In contrast, an inherent-based model is more difficult to break down than a knowledge-based model. However, the lack of this model such as high implementation costs, scars, sunglasses and surgery can cause problems and affect the accuracy of the system. Replay attacks and some fake methods can easily overcome biometric authentication methods. Finally, the ownership model requires users to bring additional physical devices such as security codes, smart cards, and so on. Accordingly, if the user loses his physical device, it will generate some security concerns because anyone who finds it can log into the system.

Further intermediate attacks are threats that can cause problems by collecting data sent by users and servers. Each authentication model has a number of threats and drawbacks that must be considered during the design process, which is summarized in the Table 2.

Since the inception of authentication, a number of methods have emerged. Given the scope of the article, we hereby briefly review the advantages and disadvantages of some of them in Table 3.

### THE PROPOSED APPROACH

**Secure Method to Store Authentication Data**

A hash table is an abstract data structure commonly used to map key and value pairs. A hash function that computes an index into an array in which an element will be inserted or searched. To compute an index,
also known as a hash code, into an array of groups or positions, the desired value can be found. A good hash function that will compute the computational complexity for finding an element in the hash table is $O(1)$.

Hash trees can be used to ensure data integrity for storage, processing, and transmission between computers. The main use of a hash tree is to ensure that blocks of data received from different nodes in the same peer network are received undamaged and undamaged.

Encoding is a method for turning information from a normal format into information that cannot be understood without the means of decoding it. Encryption is essential to secure sensitive information that is passed through two nodes on the network. It is the method of providing data security and end-to-end protection of the data. Encryption is often used to ensure that users' personal data is transmitted, stored securely, and free from malicious attacks or hacks. This encryption keeps the data protected and can only be read by the person holding the secret key.
A linear dimension reducing transform that projects the profile and the verification data to a lower dimension space, while preserving relative distances of the vectors and so correctness of authentication.

**Ensure the Integrity of User Authentication Data**

User authentication data needs to be absolutely secure. In particular, this data needs to be guaranteed to not be changed to pass the authentication step of the system. There have been many attacks on user databases to steal and modify user information for many nefarious purposes. This leads to the need for storage methods to ensure the transparency and integrity of the data. With these strict requirements, blockchain becomes a potential candidate with its preeminent characteristics.

Blockchain technology is commonly known for its applications in the monetary and banking sectors, but it works a little differently from the typical banking system. Instead of relying on centralized regulators, it guarantees the functionality of the blockchain through a set of nodes. This technology ensures immutability, blockchain keeps the information in the best security, not lost, modified and stolen. Transparency and makes it anti-corruption where every node on the system has a copy of the digital ledger. Same rules of consensus so that every node needs to check the validity of a transaction. One feature of blockchain is that once transaction blocks are added to the ledger, no one can go back and change it. Another potential approach is IPFS, which works by storing data on the network in the form of a file structure. This file structure is Merkle DAG, which combines a Merkle tree (which is a form of hash tree to ensure immutability) and Guided Ring Graph (used in Git version control, which also allows users to see content version on IPFS).

Usually a website requires centralized data storage for its files in the server to be able to do so. Operating with great advantages over http, IPFS is immune to DDoS attacks, which cause a lot of internet resources concentrated today. Another advantage of IPFS is its ability to connect to IoT devices.

**Efficient Authentication Process**

Most authentication processes require users to provide personal information for authentication, which leads to users having to provide too much sensitive information, and obviously this is a matter of concern. Therefore, the authentication process should only require a small amount of data or even part of a user information field. While providing little data, the authentication process must prove that this data is part of the entire user authentication data.

Besides, we can use an Ethereum address as an identifier (no username or password is required) and the authentication process will be done through smart contracts. This process is described simply by the following steps:

1. User requests access to the service.
2. Service provider sends some challenging questions to the user.
3. User use the private key of their Ethereum account to sign the answer then submit the signed answer.
4. Service provider call the smart contract to verify answer, signature and user address.

With this approach, service providers will not have to store user data, so user data privacy is guaranteed. Besides the authentication process is done by smart contract, and this contract is immutable so the authentication result is transparent and reliable.

**CONCLUSION**

There are many methods of user authentication for online learning systems, but each method will have different advantages and disadvantages and has not completely solved the challenges of user authentication. The issue of user authentication still has many challenges that need to be solved thoroughly to improve the security of the system as well as the trust of users and society. In addition, we propose a number of feasible approaches to improve user data privacy as well as improve the effectiveness of the authentication process in the online learning system.

In the future we will study and propose an effective authentication and identity management solution for online learning systems that not only ensures security but also enhances the privacy of users’ data.

**ACKNOWLEDGEMENT**

This research was supported by Infinity Blockchain Labs (IBL) and Vietnam Blockchain Corporation (VBC).

**CONFLICT OF INTEREST**

The authors declare no conflict of interest in this article.
AUTHOR’S CONTRIBUTION
Quang-Huan Luu and Duy-Minh Nguyen verified the analytical methods. Nguyen Huynh-Tuong and Hoang-Anh Pham supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

REFERENCES
1. Ellis RK. Learning Management Systems. ASTD learning circuits, pages. 2009.p. 1–7.
2. Alix HM, Fan JP. Information security management in e-learning. (July 2015). 2014.p 1–6.
3. Kiennert C, Rocher PO, et al. Security challenges in e-assessment and technical solutions To cite this version. HAL Id : hal-01699388. 2018.Available from: https://doi.org/10.1109/ V.2017.70.
4. Miguel J, Caballe S, Xhafa F. Security for e-Learning. Intelligent Data Analysis for e-Learning. 2016.p  7–23. PMID: 28778063. Available from: https://doi.org/10.1109/B978-0-12-804535-0. 0002-2.
5. Adamski M, Saeed K. Online signature classification and its verification system. Proceedings of the 7th Computer Information Systems and Industrial Management Applications, CISIM. 2008;8(1):189–194. Available from: https://doi.org/10.1109/CISIM.2008.38.
6. Asha S, Chellappan C. Authentication of e-learners using multi-modal biometric technology. IEEE-International Symposium on Biometrics and Security Technologies, ISBAST. 2008; Available from: https://doi.org/10.1109/ISBAST.2008.4547640.
7. Gil C, Castro M, Wyne M. Identification in web evaluation in learning management system by fingerprint identification system. Proceedings - Frontiers in Education Conference, FIE. 2010.p. 1–6. Available from: https://doi.org/10.1109/FIE.2010. 5673638.
8. Aggarwal G, Ratha NK, Jea TY, Bolle RM. Gradient based textural characterization of fingerprints. BTAS 2008 - IEEE 2nd International Conference on Biometrics: Theory, Applications and Systems. 2008.Available from: https://doi.org/10.1109/ BTAS.2008.4699383.
9. Zhao Q, Ye M. The application and implementation of face recognition in authentication system for distance education. 2010 International Conference on Networking and Digital Society, ICNDS 2010. 2010:p.487–489. Available from: https://doi.org/10.1109/ICNDS.2010.5479246.
10. Agulla EG, Rifoi nLA, Castro JLA, Mateo CG. Is my student at the other side? applying bio-metric web authentication to elearning environments. Proceedings - The 8th IEEE International Conference on Advanced Learning Technologies, ICALT 2008. 2008.p. 551–553.
11. Toor AS, Wechsler H, Choo KKR. Visual question authentication protocol (vqap). Computers Security. 2017.p. 76:6-12. Available from: https://doi.org/10.1101/j.ccone.2017.11.017.
12. Gusev PD, Borzunov GI. The analysis of modern methods for video authentication. Procedia Computer Science. 2018;123:161–164. Available from: https://doi.org/10.1016/j. procsc.2018.01.026.
13. Hayes B, Ringwood J. Authenticating student work in an e-learning programme via speaker recognition. 3rd International Conference on Signals, Circuits and Systems, SCS 2009. 2009.p. 1–6. Available from: https://doi.org/10.1109/ICSCSC. 2009.5412484.
14. Shaver CD, Acken JM. Effects of equipment variation on speaker recognition error rates. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2010.p 1814–1817. Available from: https://doi.org/10.1109/ICASSP.2010.5495401.
15. Eno N, Besacier L. Co-inertia analysis for “liveness” test in audio-visual biometrics. 2008.p. 257–261.
16. Meshoul S. Combining Fisher Discriminant Analysis And Probabilistic Neural Network for Effective On-Line Signature Recognition. 2010.p 658–661. Available from: https://doi.org/10.1109/ICBISRC.2010.6505586.
17. Jazahanim KS, Ibrahim Z, Mohamed A. Online zones’ identification using signature baseline. 2nd International Conference on the Applications of Digital Information and Web Technologies, ICADWIT 2009. 2009.p. 363–368. Available from: https://doi.org/10.1109/ICADWIT.2009.5273916.
18. Bressan R, Neve AL, Ruggiero WV. A continuous authentication system based on user behavior analysis. ARES 2010 - 5th International Conference on Availability, Reliability, and Security. 2010.p 380–385. Available from: https://doi.org/10.1109/ARES.2010.63.
19. Muthumanickam K, Ilavarasan E. Behavior based authentication mechanism to prevent malicious code attacks in windows. 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems. 2015.p. 0–4. Available from: https://doi.org/10.1109/ ICESS.2015.7241289.
20. Ullah A, Xiao H, Lilley M. Using Challenge Questions for Student Authentication in Online Examination. International Journal for Infonomics. 2016;5(3/4):631–639. Available from: https://doi.org/10.20533/ij17.4712.2012.0072.
21. Ullah A, Xiao H, Barker T, Lilley M. Graphical and text-based challenge questions for secure and usable authentication in online examinations. 2014 9th International Conference for Internet Technology and Secure Transactions, ICTST 2014, pages 302-308, 2014.Available from: https://doi.org/10.1145/ 1473167.1473182.
22. Ullah A, Xiao H, Lilley M. Profile Based Student Authentication in Online Examination. International Conference on Information Society (i-Society). 2012.p. 109–113.
23. Jiang H. Strong password authentication protocols. ICDE 2010 - 2010 4th International Conference on Distance Learning and Education, Proceedings. 2010.p. 50–52.
24. Kumar SA, Ashish K, Tanu G. Survey of Remote User Password Authentication Scheme Using Smart Cards. International Journal of Advanced Research. 2015;3(4).
25. Elwahab AA, Eldin AMB, et al. A security layer for smart card applications authentication. The 2009 International Conference on Computer Engineering and Systems, ICCES’09. 2009.p. 514–517. Available from: https://doi.org/10.1109/ ICCES.2009.5838211.
26. Lim SF, Fotsing PT, et al. Blockchain Technology the Identity Management and Authentication Service Disruptor: A Survey. International Journal on Advanced Research, Engineering and Information Technology, 2018;8(4-2):1735. Available from: https://doi.org/10.18517/ijareit.8.4.2.6838.
27. Toor AS, Wechsler H, Nappi M, Choo KKR. Visual Question Authentication Protocol (VQAP). Computers Security. 2018;76:285–294. Available from: https://doi.org/10.1016/j. csee.2017.11.017.
28. Ometov A, BezZateev S. Multi-factor authentication: A survey and challenges in V2X applications. International Congress on Ultra Modern Telecommunications and Control Systems and Workshops 2017. 2018.p. 129–136. Available from: https://doi. org/10.1109/ICUMT.2017.8352306.
29. Benet J. iFPS - Content Addressed, Versioned, P2P File System. (Draft 3). 2014.
30. Bouterin V. A next generation Smart Contract & Decentralized Application platform. 2009.p. 1–36.
Xác thực người dùng trong hệ thống học tập trực tuyến: Thách thức và giải pháp

Lưu Quang Huân*, Nguyễn Duy Minh, Phạm Hoàng Anh, Huỳnh Tường Nguyên

Tóm tắt
Ngày nay, học tập trực tuyến đang ngày càng phổ biến và được xem như là một phương pháp thay thế hàng xóm với học tập truyền thống. Một hệ thống học tập trực tuyến là sự kết hợp của các công cụ giảng dạy, thảo luận cũng như các công cụ hỗ trợ kiểm tra nhằm đánh giá kỹ năng của người học. Một trong những yếu tố quan trọng quyết định tính bảo mật và quan trọng ngày càng là xác thực người dùng. Cũng như các hệ thống thông tin khác, vấn đề xác thực danh tính người dùng trong hệ thống học tập trực tuyến cũng là một vấn đề còn chưa được giải quyết thỏa đáng. Trong phần lớn các trường hợp, vai trò này được thực hiện bằng một mật khẩu, nhưng bằng chứng cho thấy phương pháp này dễ bị xâm phạm. Trong khi đó có nhiều sự lựa chọn thay thế tồn tại như phương pháp sinh trắc học, phương pháp thử thách người dùng và nhiều các phương pháp khác. Sự phát triển mạnh mẽ của công nghệ đòi hỏi tính bảo mật và xác thực phải được kết hợp chặt chẽ với nhau. Một cuộc khảo sát tình hình việc các hệ thống xác thực người dùng đang được sử dụng trong các hệ thống e-learning ngày nay và một nghiên cứu so sánh về các cơ chế xác thực khác nhau được trình bày ở phần bài báo này. Có nhiều phương pháp xác thực người dùng cho các hệ thống học tập trực tuyến, tuy nhiên mỗi phương pháp sẽ có những ưu điểm khác nhau và chưa giải quyết được những thách thức về xác thực người dùng. Văn đề xác thực người dùng vẫn còn nhiều thách thức cần được giải quyết triệt để để nâng cao tính bảo mật của hệ thống cũng như sự tin tưởng của người dùng và xã hội. Ngoài ra, chúng tôi đề xuất một số phương pháp khả thi để cải thiện quyền riêng tư dữ liệu người dùng cũng như cải thiện hiệu quả của quá trình xác thực trong hệ thống học tập trực tuyến.

Từ khóa: Xác thực phân tán, Tính riêng tư, Cây Merkle, Blockchain

Trích dẫn bài báo này: Huân L Q, Minh N D, Anh P H, Nguyên H T. Xác thực người dùng trong hệ thống học tập trực tuyến: Thách thức và giải pháp. Sci. Tech. Dev. J. - Eng. Tech.; 3(SI1):SI95-SI101.