SECURE ILEARNING

Frank Graf (Fraunhofer-IGD, Darmstadt, Germany)

Abstract This paper discusses the specific security requirements of Internet-based learning. Two concrete solutions for improving the degree of security achievable are shown. First a framework for secure testing is introduced. This framework prevents manipulation from the side of the students during learning. Thus allowing a reliable control of learning success. The second solution presented deals with the problem of confidentiality and protection of copyright. By establishing a life-long control of the copyright owner over his documents it prevents illegitimate access and redistribution.

Keywords: Secure learning, Internet and intranet security, watermarking, biometry, encryption

1. ABOUT ILEARNING

In this paper the term iLearning stands for all the various forms of computer assisted learning which are based on Internet technology. The arrival of the WWW marks an important landmark in the history of learning. This new medium was almost immediately adopted for learning purposes and stimulated lot of new research and development activities (Virtual Universities) and has become the de facto standard platform for computer assisted learning.

Even though security is an important issue for iLearning it gained wider interest only recently. Due to the close relation to the WWW many of the security problems are similar to those of other WWW applications, therefore solutions developed there will also apply to iLearning. However, there are some aspects of security which are specific for iLearning.

2. WHY ILEARNING NEEDS SECURITY?

2.1. KNOWLEDGE BECOMES AN IMPORTANT MEANS OF PRODUCTION

New knowledge emerges and must be mastered and incorporated in ever shorter periods of time. As a result of this, continuous, lifelong training is becoming a necessity. From a practical point of view the implications are:

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35413-2_36](https://doi.org/10.1007/978-0-387-35413-2_36)

Figure 1 Contribution of the Copyright Industries to U.S. GDP

Learning must be done more frequently and since it is becoming more and more part of the normal work flow, it must be integrated as seamlessly as possible with the working process. Currently training is often performed in a remote classroom environment, where besides the cost for training there might also be considerable costs for traveling and accommodation. With iLearning training can be provided at the workplace immediately when the demand arises.

2.2. KNOWLEDGE AS PRODUCT

Since knowledge has become a crucial resource for the success of a company, it has also become a product and an object of trade. A 1999 study [IIPA, 1999] by Economist's, Inc for the International Intellectual Property Alliance (IIPA) shows that the copyright industries contribute a substantial part to the U.S. GDP (see figure 1). In 1997 the copyright industries' foreign sales/exports exceed even the combined automobile and automobile parts industries, as well as the agricultural sector.

In a study from May 2000 IDC [Anderson et al., 2000] predicts a 12% annual growth rate in revenues for IT-based training from \$19.4 billion in 1999 to almost \$34 billion in 2004. Given such stakes there is definitely a need to protect the courseware – which is the carrier of knowledge, yet must be exposed to the broad public.

2.3. KNOWLEDGE AS KEY FOR PERSONAL SUCCESS

The knowledge of a person is also the key for his professional success. Since there is no instant test to evaluate the knowledge or the skills of an individual,

Figure 2 Legitimate users as security leaks

certificates such as e.g. university diplomas or professional certification results are a cornerstone of evaluation. Since certificates are essential for the progress of his professional life, achieving such a certificate is of considerable importance for everybody. Certain individuals will go to great length to obtain certificates or claim to own one. Thus the history of education is also the history of cheating. For the traditional forms of learning there are established ways to prevent cheating. iLearning introduces new ways of learning but the students are still the same, and will cheat if given the opportunity. Since iLearning is still in an early phase of development, the mechanisms to prevent cheating have yet to be established. An additional problem lies in validating credentials presented by prospective employees, as a significant portion of these turns out to be either fraudulent or only partially accurate.

3. SECURITY ISSUES IN ILEARNING

3.1. MANIPULATION

Protection against manipulation is also an issue in other WWW applications and several solutions have been developed for this, e.g. encryption, digital signatures, digital watermarks, TLS and firewalls. The problem with these solutions is that they are directed against attacks from the outside while the user is considered trusted. In iLearning the users are the students and one cannot necessarily trust them.

However this by no means implies that those security mechanisms are not necessarily suitable for iLearning. In fact, they are needed to prevent manip-

ulations by third parties. But additionally one needs further mechanisms to ward against manipulation by the legitimate users. Maleficent users are also a potential security hazard in normal application scenarios – in fact studies indicate that approximately 80% of attacks are performed by insiders – but the difference in an iLearning scenario is the percentage. In normal application scenarios some criminal energy and extra effort is required to gain personal advantages from manipulation, hence the vast majority of users will stick to the rules. In iLearning students will directly benefit from manipulations and the results of their learning will be of immediate relevance for their future career. Because of this there is a much higher readiness to cheat. But even if no student had the intention to cheat one should nevertheless provide mechanisms to prevent manipulations, otherwise students could cheat themselves (e.g. manipulate time limits for tests or look up solutions while a test is running).

With respect to security the underlying client-server architecture of iLearning is a big asset. Due to this it is possible to isolate the sensitive parts (e.g. test evaluation) of the learning material into entities which are kept on the server. Only learning material without any sensitive components is delivered to the students. During the learning phase this material interacts with the sensitive entities on the server. Since the server is under the control of the training provider, he can take all necessary precautions against manipulations and if the students cannot access the sensitive entities this will dramatically reduce the potential for attacks. This approach requires some discipline on the part of the learning material authors, since they must split up their material into a sensitive and a non-sensitive part and enable them to interact. Currently (client-side) JavaScript is often used for computations of test results, ignoring the fact that this is highly insecure.

3.2. USER AUTHENTICATION

A reliable identification of the user is an elementary feature for an iLearning system because it is the basis for:

Access control. Usually access to the iLearning system is granted only to registered students but even their access might be restricted to a certain subset of the courseware.

Billing. Especially if iLearning is offered on a commercial basis there is a need for remuneration. There are numerous accounting schemes but almost all of them are in some way related to the actions of the individual users.

Maintenance of learning material. Information which students used which material and how well the training worked out, provides the training provider some insight into, what the assets and the deficits of his learning material are and allow him to improve his portfolio.

User profiles. If the learning success of the a user is recorded, this information on his current level of knowledge can be used to optimize a course for his pre-existing knowledge, his preferences and/or his learning goals.

Certification. In case the learning results in some kind of certification (e.g. a diploma), it is essential that the records on which the certificate is based on can, without ambiguity, be attributed to a specific student.

How to handle certification and user profiles in collaborative learning is an interesting topic. There are several good reasons to assign a user a specific ID for learning, opening the possibility of a student to use multiple IDs. Decoupling the working accounts from the iLearning accounts requires that the training provider implements a user administration of his own. Then it is no problem at all if the training provider is not identical with the organization the user belongs to, which is an important premise for real worldwide iLearning. A globally unique naming scheme with a provable correlation between an ID and an individual would be very beneficial; X.501 and X.509 are obvious solutions for this. However since such IDs are quite unwieldy and due to security requirements, they are used best in conjunction with smart cards.

Currently there are three major categories of techniques for user identification.

Passwords are still the standard in user identification. No special hardware is required for their realization. In a networked environment the secure transmission of passwords is a crucial issue.

Smart cards use cryptographic mechanisms, namely public-key encryption, to provide authentication and identification. The security of the smart card does not depend on the computer it is used with, and smart cards can be designed to resist tampering. They provide efficient means of identification and authentication with a two factor (knowledge and possession) mechanism that also ensures non-repudiation properties. In order to use smart cards, all clients must be equipped with special hardware. Currently there is no universal standard for smart cards and smart card readers are not standard equipment on today's computers. Therefore smart cards are not a viable alternative for iLearning except in environments putting emphasis on security and being able to pay for this added security.

Biometric Identification has recently gained much attention. Here unique, invariable biological characteristics of a person (fingerprints, voice, face, handwriting, etc.) are used to authenticate a user. The idea behind this is that if the user itself is the key there is no longer the possibility of stealing or duplicating the key. However, binding the authentication to

personal characteristics raises concerns about privacy. With regard to standards the situation is even worse as with smart cards. All biometric systems need some hardware to read in the biometrical characteristics. Part of them uses rather exotic hardware (e.g. finger or iris scanner) while for others more common hardware like microphones or cameras is sufficient. Since microphones and cameras can be used to realize audio/video conferences with other students or a tutor one would prefer such biometric systems for iLearning which require only these. The development in the area of biometric systems is still in an early state. In a study [Henning Daum, 2000] initiated by the Bundeskriminalamt (BKA, Federal Criminal Investigation Office) and the Bundesamt für Sicherheit in der Informationstechnik (BSI, German Information Security Agency) several of the twelve system tested showed unsatisfactory recognition rates and almost half of the systems could be duped without much effort.

Even if all these methods would work perfectly there is still a conceptual problem which none of them can overcome. In the best case a training provider can ensure that a certain user has been involved in particular learning activity. What he cannot ensure is that recorded actions of the user are original to this specific user and he had not used illegitimate resources. In order to guarantee this learning has to be done in an environment, over which the training provider has full administrative control. However, this contradicts the core principles of iLearning (time- and space independent learning). But traditional learning as well is not done exclusively in environments the teacher has total administrative control over (take for example homework or writing a thesis). As a consequence there is no absolute security nor will there be such a thing, but over the years procedures have been established which give an adequate amount of security. The same will happen for iLearning. Possible approaches are:

Legal Steps. Here the student must submit an affidavit stating that he made no use of illegitimate resources. Any certificate is granted under the condition that this statement is true. Although with this approach cheating cannot be prevented it provides much better leverage for legal actions once illegitimate actions are uncovered. This approach is independent from any technical provision, thus it can and ought to be used in parallel to technical means.

Separation of learning and certification. This is another well known approach also found in traditional learning which can be adapted to iLearning, too. Here one distinguishes between a learning and an explicit testing phase. The iLearning system can be used in both phases, but while the

learning can be done in a time and space independent fashion, the testing is done in a controlled environment. User actions are recorded in both phases, but only the reliable data from the test phase is used for certification. This approach allows a high degree of security, although there are some disadvantages. Testing is additional effort for the student although it ideally should lead to the same results that could be derived from the learning phase. The fact that there is an upper limit for the duration of a test has consequences, too: First the actual performance of the student on day of testing may not be representative of the student's capabilities. Then there is also an upper limit to the complexity of the problems that can be used in the test. So only a subset of the entire course content can be tested. Although the tester will try to select a subset as representative as possible, there are aspects which cannot be tested in an such a setting.

Integrated testing. Instead of distinguishing between a learning and a testing phase iLearning could go into the opposite direction. All the information concerning the learning success gathered during the entire learning process is used for certification. In this way the entire performance of the user can be taken into account. However for this the gathered information must be of higher granularity than just simple test results. Essential for this approach are tools which assist the tutor in analyzing user profiles, evaluate the students learning success and detecting suspicious learning pattern. It cannot be guaranteed that every piece of information describes the situation accurately, either because the student tried to manipulate or because he was distracted. If the interpretation is done with care and over a longer period of time, this should result in an accurate overall evaluation. However there will always be a certain amount of insecurity. In case the tutor has some doubts, he can revert to explicit tests. In this case the information available gives good hints what to test.

Innovative course design. A very fascinating approach, but very challenging for the authors, would be to find course material or forms of learning where the user cannot cheat. As one example the user could be given a concrete task. In order to find the solution he must – effectively as side effect of the course – acquire the desired knowledge and the course is designed in a way that he is guided towards this information. Collaborative elements, in which the student is forced to present his solution might also be an interesting possibility here. Since the tutor evaluating the presentation will obtain a good impression in this way of his actual knowledge the student is obliged to learn in order to perform a good presentation of his solution.

3.3. CONFIDENTIALITY

Due to fact that courseware by its nature must be distributed to the outside, industrial espionage and data theft is typically not a major problem in iLearning. Nevertheless in certain scenarios this is also of relevance. Take for example a company using iLearning for in-house training. In this case the courseware might contain sensitive information so that confidentiality and restricting flow of information to well-defined groups is definitely an issue.

3.4. COPYRIGHT PROTECTION

A rule of thumb is that in order to create courseware of fairly good quality equivalent to one hour of course one has to invest approximately 100 hours of manpower. This effort cannot be automated to a significant extend. So courseware will always represent a significant investment of expertise and work and copyright holders of courseware therefore have a strong interest in protecting their courseware from illicit use and distribution.

The major drawback for copyright protection in iLearning is that material must be made available to the students, which are also potential security leaks. A training provider can and will restrict access to courseware only to registered students and the courseware can be delivered in encrypted form to the students to avoid that illicit users can get hold of the data. But once the data has reached the student, the training providers depend on the goodwill of its students. With digital data giving a user read access to data always means the opportunity to redistribute the data at will. License management schemes also offer no solutions since they are easily broken.

A good way to reduce the costs for courseware is to reuse courseware created by others. Besides avoiding that the same content is re-implemented over and over again a training provider can so get access to material for content outside its own domain. If training providers exchange courseware, this will normally be done on the basis of some kind of agreement and typically there will be some kind of reimbursement. For this reason training providers require a way to prevent other training providers from using material without approval. In practice this requires being able to prove that he is the legitimate holder of the copyright.

4. PROVIDING SECURE ILEARNING

Although in principle the same technology is used in Internet scenarios as in intranet scenarios there are significant practical differences between those both scenarios. The term intranet is mostly used in the sense 'internal network of an organization' and very fuzzy. In this paper the term intranet is used for

Figure 3 System architecture of the test framework

networks, with distinct organizational structure and where administration is done exclusively by trusted personnel.

4.1. SECURE ILEARNING IN THE INTERNET

Lacking access and control over the computers of the students the training providers have only very limited means for enforcing security. They are restricted to precautions they can implement entirely on the server side. A truly reliable user authentication cannot be done and about the only countermeasure against copyright violations they can take is marking their material with digital watermarks. Typically confidentiality in an Internet scenario is less of an issue because the target group is the general public.

The most pressing issue is the protection against manipulations. Although the WWW base technologies provide sufficient functionality for delivering and displaying learning material they do lack mechanisms to prevent manipulations and also do not provide the degree of interactivity, exactness, reliability and security which is needed for effective monitoring the learning success of students. To overcome these deficiencies an open framework [Hausding, 1999] has been developed at Darmstadt University of Technology which provides the functionality the WWW is lacking.

This test framework uses a client/server approach and has been designed to be used in conjunction with an iLearning system. The basic security mechanisms and general testing functionality have been implemented as an integral

part of the test framework. Therefore when creating test material the authors need not worry about security since it is implemented transparently. Table 1 shows the basic test services provided by the framework and figure 3 shows the system architecture. More details can be found in [Graf and Hausding, 1999] and [Graf and Schnaider, 1999].

Being implemented in Java, this framework can easily be extended to support further types of tests and due to its flexible underlying test concept this framework is especially suited as a test-bed to explore new forms of learning with integrated testing or innovative course designs.

4.2. SECURE ILEARNING IN INTRANETS

The existence of a trusted administration is a big asset for security. In an Internet scenario a training provider must tailor his hardware and software requirements on what he can expect to be available with average users. This will normally be quite a heterogeneous mixture of operating systems, browsers and browser extensions (e.g. plugins), all of which are also available in multiple versions with different bugs and capabilities, so the training provider must usually restrict himself to a subset of features. In an intranet scenario however, the administrator can freely decide what level of security is required or preferred and set up the environment accordingly. In such an environment the use of smart cards or other security systems becomes feasible. It is beneficial too that being part of the same organization structure gives the administrator a better leverage for sanctioning transgressions.

Using smart cards improves the security of user identification and authentication. However neither of them can guarantee that the actions of the user are created using illegitimate resources. The test framework introduced in the previous section can also be used in the intranet to prevent manipulations. Typically, the need for confidentiality is considerable higher in an intranet scenario than in an Internet scenario. While mechanisms to prevent attackers from the outside accessing internal data are well established, there is little available to prevent confidential data being leaked to the outside from an insider. One system to ensure confidentiality is CIPRESS¹ [Busch et al., 2000]. This system allows protection of intellectual property from illegal usage and redistribution by combining patented re-encryption and watermarking technology [Busch et al., 1999].

In CIPRESS, an operating system extension will ensure that all data is stored and transmitted in encrypted form only. Encryption and decryption is done transparently and mandatorily. The keys used are specific for each document and user and are obtained from a remote server, the Key Center which is operated in a secured facility. They are requested from the Key Center every time the document is accessed (see figure 4), so it is impossible to view or

<p>Secure delivery of material relevant to testing</p> <ul style="list-style-type: none"> ■ There must be a well-defined point of time for the start of the test. ■ From this point of time on all test-relevant material must be available to the student. ■ Up to this point of time none of the test-relevant material must be accessible.
<p>Tamper-resistant evaluation of tests</p> <ul style="list-style-type: none"> ■ The solution must be kept secret before and during the test. ■ The user must not be able to access or manipulate the evaluation mechanism. ■ Secure transfer of the test result from the evaluation to the user profile.
<p>Reliable timing control</p> <ul style="list-style-type: none"> ■ Exact association of user actions and corresponding points of time. ■ Accuracy to single-seconds precision. ■ Secure against manipulation. ■ Reliable termination of tests at expiration time.
<p>Reliable recognition and evaluation of user actions</p> <ul style="list-style-type: none"> ■ Monitoring of user actions and filtering out actions significant for the test status on the client side. ■ Interpretation and recording of significant user actions on server side.
<p>Creation and delivery of feedback</p> <ul style="list-style-type: none"> ■ Fast creation of feedback for the case of interactive test material. ■ Secure feedback creation, if feedback contains test relevant information. ■ Recording of feedback for consideration at evaluation time.

Table 1 Basic functionality provided by the test framework

Figure 4 Overview CIPRESS

redistribute a document without the Key Center becoming aware of this and approving it by delivering a key. Access control is realized on top of an operating system independent user and group management. The Key Center logs all accesses, thus the entire usage trail of a document can be traced. Since a normal user can neither manipulate the operating system extension nor the Key Center, there is no way for him to circumvent this mandatory encryption.

In addition CIPRESS monitors all incoming and outgoing network traffic. When receiving encrypted data the data will transparently be decrypted too. In case the communication takes place over an insecure connection CIPRESS will encrypt the outgoing data using SSL. To guarantee confidentiality CIPRESS will allow communication only between certified hosts and deny traffic to and from unauthorized hosts.

The combination of these two mechanisms secures all distribution channels involving a computer. However, they cannot come into effect when the data leaves the computer, e.g. when a user prints out the material. For such situations CIPRESS embeds hierarchical digital watermarks holding identifiers of the most recent users into the document content itself. These watermarks can be traced back to a perpetrator and that the copyright owner can prove that a given document is in fact his property. Further on CIPRESS provides a mechanism to correctly handle secondary copyright.

Using CIPRESS in conjunction with an iLearning system results in the following benefits [Graf et al., 1999]:

Access control. Only legitimate users can access the courseware. Within CIPRESS access control does not end once the courseware is delivered to

the student. Instead, it is maintained over the entire life span of the courseware. Thus sensitive or confidential information can be included into courseware since CIPRESS prevents illicit usage and distribution even for courseware which is located on the students' PCs.

User profiling. Since CIPRESS detects and records access to material at any time and not only when the user downloads courseware, one can reach a much higher level of accuracy of the user profiles. Furthermore students can no longer access courseware outside a learning session without the iLearning system becoming aware of it.

Accounting and billing. Since the Key Center keeps record of every key delivery all the basic data needed for accounting and billing is readily available at the Key Center.

One unique feature of our approach is the fact that, due to the re-encryption technology, the owner retains permanent control over his courseware. With CIPRESS, the student can use the courseware only if he receives the keys for it from the Key Center. Besides preventing the illicit re-distribution of courseware, this lifelong control in conjunction with the CIPRESS user administration allows some beneficial features, which are not possible in standard iLearning environments.

Time restricted use of courseware. Normally a student can – once the material is delivered to him – access the courseware as long as he keeps the data. The owner of the courseware can, by setting the CIPRESS permissions, at any time cause the Key Center to no longer distribute any keys for a specific piece of courseware for one or even all students. With this feature a training provider can restrict the student's usage of the courseware to exactly the time span the student is entitled to.

Ease of maintenance. This feature is very beneficial for keeping the courseware up to date. Suppose the training provider decides that specific piece of courseware should no longer be used, because the content has become obsolete, is faulty, or for some other reason. With traditional systems the training provider had no means to make sure that the students will not continue to use the now obsolete material delivered to them. With CIPRESS all he has to do is set the access rights accordingly.

Delivery in advance. Although courseware will normally be delivered on demand via the WWW, there are certain scenarios where using a different distribution channel makes sense. It might be desirable to send the courseware in advance on a CD-ROM to avoid significant delays in the flow of learning caused by loading large amounts of data via a slow

network connection. Doing this, it must be guaranteed that the student cannot access the material (exercises, tests) before he is supposed to do so. With CIPRESS this can be achieved by setting the access rights so that when the material arrives the student cannot access the material. Immediately before the student ought to access a piece of courseware, the iLearning system grants the rights and the student can access the material.

Personalization. Students have different backgrounds, learning goals or simply have paid for a different level of support. So ideally, there should be a personalized set of courseware for each of the students' needs. With CIPRESS the training provider can send the same CD-ROM, containing all the available courseware in encrypted form to all students. But students can access only those pieces of courseware for which they received permission; the permissions can be changed any time on an individual basis at the Key Center.

5. CONCLUSION

iLearning has specific security requirements. The crucial drawback is the inherent limitations on trust that can be placed in a student; due to the lack of administrative control over the students learning environment, absolute security cannot be achieved. The common security mechanisms of the WWW are not sufficient. The WWW lacks essential functionality for secure testing, copyright protection and preventing illegitimate redistribution of data by legitimate users. Two additional countermeasures a training provider can take to improve the level of security achievable have been illustrated in this paper. While by using smart cards or biometric systems the level of security for user authentication can be improved, the conceptual problem with user authentication in iLearning – that even if it can be ensured that a user has been involved in some learning activity there is no way to find out or prevent that he is using illegitimate resources – cannot be solved by technical means. The amount of security is directly proportional to the amount of administrative control of the training provider. Thus the amount of security achievable is much higher in an intranet scenario than in an Internet scenario.

Acknowledgments

The author would like to thank the Mitsubishi Corporation, Tokyo, Japan for its generous support and close cooperation in supporting this research.

Notes

1. CIPRESS is the internal code name for a development by the Fraunhofer Institute for Computer Graphics on behalf of the Mitsubishi Corporation.

References

- Anderson, C., Dankens, A.-S., and Julian, E. H. (2000). Worldwide and U.S. Corporate IT Education and Training Services: Forecast and Analysis, 1999-2004. IDC Report #W22154.
- Busch, C., Funk, W., and Wolthusen, S. (1999). Digital Watermarking: From Concepts to Real-Time Video Applications. *IEEE Computer Graphics and Applications*, 19(1):25-35.
- Busch, C., Graf, F., Wolthusen, S., and Zeidler, A. (2000). A System for Intellectual Property Protection. In *Proceedings of the World Multiconference on Systemics, Cybernetics, and Informatics (SCI 2000) /Int'l Conf. on Information Systems Analysis and Synthesis (ISAS 2000)*, Orlando, FL, pages 225-230.
- Graf, F., Busch, C., and Wolthusen, S. (1999). Courseware needs Security. In *Proceedings of ICCE'99, Advanced Research in Computers and Communication in Education*.
- Graf, F. and Hausding, M. (1999). An Open Framework for Monitoring Learning Success in WWW-based Training. In *Proceedings of EAEEIE 99*.
- Graf, F. and Schnaider, M. (1999). Using Patterns to Efficiently Create Tests for Web-based Training. In *Proceedings of 11th International Conference on System Research, Informatics and Cybernetics*.
- Hausding, M. (1999). Ein offenes Framework für Lernkontrolle in WWW-basiertem Training. Master's thesis, Darmstadt University of Technology.
- Henning Daum (2000). BIOIS Study - Comparative Study of Biometric Identification Systems. Technical report, Bundesamt für Sicherheit in der Informationstechnik (BSI). Public Final Report.
- IIPA (1999). Copyright Industries in the U.S. Economy: The 1999 Report. International Intellectual Property Alliance (IIPA).