COMPLETENESS, SECURITY AND PRIVACY IN USER MODELLING FOR WEB-BASED LEARNING

Maria Virvou, Nineta Polemi, Katerina Kabassi
Department of Informatics, University of Piraeus, 80 Karaoli &Dimitriou Str., 18534 Piraeus, Greece

Keywords: e-Learning, Public Key Infrastructure, WS-Security.

Abstract: This paper describes a user modelling agent for a web-based learning system that provides personalised and secure interaction with its users (learners). The system involves user modelling in order to provide tutoring that is dynamically tailored to the individual learner's needs. It constantly observes the user silently and processes the observations so that it may generate hypotheses about the user's level of domain knowledge, current goals and possible problems. Personalised Web-based learning requires the gathering of a lot of personal data concerning the user and thus security and privacy issues are raised. Information about the user is maintained centrally on a User modelling Server and concerns each individual learner. Each user model is available to any client application of the system that requests it through Web Services. PKI and WS-security technologies have been used in order to embed security in the F-SMILE system and offer basic security services.

1 INTRODUCTION

There have been educational software technologies that have been particularly effective at personalising tutoring such as Intelligent Learning Environments (ILEs). ILEs base the generation of personalised teaching on their user modelling components. User modelling involves the construction of a detailed representation of the student's cognitive state and behaviour regarding existing background knowledge about a domain. Information about the user ranges from relatively long-term facts such as areas of interest or expertise to quite short term facts such as the problem that the user is currently trying to solve. In view of this, there is a distinction between long-term and short-term user models. A long-term user model consists of information about the user that has been gathered during past interactions. This information may involve the user's level of knowledge of the domain, his/her common errors etc. A short-term user model consists of the user's beliefs at a very specific time and is the output of the reasoning of the system. Ideally, both models should exist in an ILE and should exchange information between them.

It is evident that personalised tutoring requires the gathering of a lot of personal information about the user so that the system may adapt to his/her needs. Moreover, data gathering is mostly performed in an unobtrusive manner and often without users' awareness; this is done to avoid distracting users from their tasks (Kobsa 2002). Thus personalised systems pose privacy and security problems.

Indeed, security can be a very crucial issue in electronic personalised learning systems, since it often involves quite sensitive and critical data (psychological, behavioural learners' characteristics) where any modification, exposal to unauthorised persons or loss may reveal risks. In this respect, it is essential to ensure data integrity and availability, since these data may form the basis for critical decisions on the learners' progress, confidentiality of personal, behavioural, academic and administrative information due to possible social, ethical and psychological impacts, as well as accountability and non-repudiation of data origin, receipt and use.

Despite the above-mentioned strong requirements and needs, data security, is usually underestimated or completely ignored during e-learning information systems and applications design, making relevant later enhancements an almost impossible task.

In view of the above, we have developed Web F-SMILE (Web File-Store Manipulation Intelligent Learning Environment) which is an interoperable,

Virvou M., Polemi N. and Kabassi K. (2006).
COMPLETENESS, SECURITY AND PRIVACY IN USER MODELLING FOR WEB-BASED LEARNING.
In *Proceedings of WEBIST 2006 - Second International Conference on Web Information Systems and Technologies - Society, e-Business and e-Government / e-Learning*, pages 293-298
DOI: 10.5220/0001238102930298
Copyright © SciTePress

personalised learning system operating securely over the Web. In particular, Web F-SMILE is an ILE for novice users of a GUI that manipulates files, such as the Windows Explorer. Our approach concerning the operation of the system over the Web is based on Web Services. The main characteristic of Web Services is that they interact with the applications that invoke them, using web standards such as WSDL (Web Service Definition Language), SOAP (Simple Object Access Protocol) and UDDI (Universal Description, Discovery and Integration). Basing user modelling on web standards has the advantage of enabling the dynamic integration of applications distributed over the Internet, independently of their underlying platforms.

PKI and WS-security technologies have been used in order to embed security in the F-SMILE system and offer basic security services.

2 REQUIREMENTS

Personalised tutoring over the Web poses many functional and non-functional requirements that have to be met by the developed systems. The issue of personalisation over the Web raises many requirements concerning the architecture and operation of the resulting system. Moreover, the need for privacy and security of the personal data about the users to be collected has to be ensured.

2.1 Requirements for Web-based User Modelling for Educational Systems

Important requirements that arise in web based personalised educational systems:

Availability of the personalisation functionality at any place and at any time. This means that the system should be able to work fully, whether the user accesses the application from a computer at a lab, or a PC from his/her home or elsewhere. Preferably, the personalisation functionality should be available both online (when the user is connected to the Web) and offline so that the student's work on a PC is not disrupted in case a Web connection is temporarily unavailable for some reason. To have the personalisation functionality both online and offline two copies of current user models are needed, one on the server (for the online case) and one on the user's PC (for the offline case).

Accuracy and completeness of user models. The accuracy and completeness of user models ensure that the personalisation addresses the real needs of a particular user. If the user model is not accurate or incomplete then the system will probably generate the wrong hypotheses about the user and it will adapt tutoring and help in the wrong way. In such cases the system's adaptivity may result in the user's frustration or even irritation and it will lose its credibility. To ensure the accuracy and completeness, the user models have to contain all the information about the user collected locally from the PCs-clients and centrally from the server.

Availability of long term information about the user. The student's history record is important while making hypotheses about his/her current cognitive state. To ensure that all past information exists in one place that contains the whole picture about the user, the long term user model should preferably reside on a Web server that is frequently updated by the clients and passes the long term information about the user to the clients that request it.

Guaranteed and timely update of long term information about the user. The long term information about the user is always needed for the system to be able to generate plausible hypotheses about the current state of the user. If the long term user model resides on a server, there has to be a guaranteed way of its update from information on local PCs even in cases when the PC works offline locally for some time.

2.2 Security Requirements

There are various security requirements that arise in web based personalised educational systems:

Authentication of origin ensures that students are really the ones who they claim to be. Authentication of the student engaged in an e-learning action is necessary for the learning environment to uniquely and irrevocably identify the parties involved in any action and particularly for the authentication of the centralised learners' models. This requirement can be addressed by the application of XML digital signatures in combination with tamper resistant cryptographic modules such as smart cards.

Integrity of the content of the learning material and student model ensures that they cannot be altered intentionally or accidentally during transmission or storage. Thus, the involved parties can be confident with respect to the content of the transacting e-documents. A cryptographic hash function (Nash, 2001) provides message integrity checks and can be used either separately or as part of the digital signature process.

Confidentiality and privacy ensures that no one other than the sender and the designated recipients can read the data. XML Encryption as specified in the W3C Recommendation (Eastlake, 2002) and the Web Services Security recommendation for encryption in SOAP messages (Hartman, 2002) provide confidentiality.

Integrity of the sequence of the data assists in avoiding any gaps occurring in the transactions and in strengthening the performance of the system. This requirement is implementation specific and can be fulfilled by enforcing a tight sequence issuance scheme for the reference number embedded in each action.

Availability ensures that the students can use the e-learning service at any time without disruption. On one hand, the system should be robust and protected against intrusion and hacking, which can be ensured by standard network elements such as intrusion detection systems, antivirus and firewalls. On the other hand, some form of public directory usage for publishing the offered services will foster services dissemination.

Electronic Storage of files. The conditions for electronic storage of e-files e-objects and the technical requirements of the electronic storage system are integral components of the security requirements concerning e-learning. Authenticity, integrity and readability should be guaranteed throughout the storage period. A native XML database can ensure that XML files/objects are stored exactly in the original format in which they were received for the correct creation of the students' model or any future audit. Furthermore, the combination of XAdES and such a database can guarantee the secure long-term archiving of e-learning data.

Secure Sections. Secure sections allow temporary replacement of the current security execution context, so that the enclosed code executes on behalf of the new principal. The new context remains in scope within the section and propagates between virtual machines, if necessary. For example, if a remote client with limited rights invokes a simple database query service, it may not have the required privileges to perform the query. If the query was executed on behalf of a second (intermediate) user with sufficient privileges, the client could then retrieve the necessary result set without having been explicitly granted the necessary permissions. These new rights or privileges only remain in effect inside the scope of the secure section.

Privilege Delegation. Privilege delegation occurs when a calling entity authorizes an intermediate entity to perform a task using a set of the rights granted to the calling entity. For delegation, the underlying security mechanisms must be notified that a boundary has been crossed, that the calling entity (principal) is no longer in scope. This occurs only when entering a secure section.

3 OPERATION OF THE SYSTEM

Web F-SMILE (File-Store Manipulation Intelligent Learning Environment) is an intelligent learning environment for novice users of a GUI (Graphical User Interface) that manipulates files. It works in a similar way as Windows 98/NT Explorer (Microsoft Corporation, 1998) but additionally it can dynamically adapt its interaction to individual learners for helping and tutoring them. For this purpose, Web F-SMILE silently observes the students while they are actively engaged in their usual activities for their file manipulation. If Web F-SMILE judges that a student has been involved in a problematic situation (as indicated by the user modelling component) it provides individualised advice and tutoring at its own initiative.

The system can work both as a Web-based application and as a standalone application when the learner's computer is not connected to the Internet. The system keeps two copies of user models, one on the Server and one on the user's PC so that the system may work both online and offline. When the system works online, information about the learner is stored on a User modelling Server and is given to any client of the application that requests it. When the system works offline information about the learner is stored on the PC. Web F-SMILE uses Web Services for the interaction of the components of the system with the Web Server.

A simple example of the system's operation taken from a real interaction of a user with Web F-SMILE is presented in Table 1. The learner's initial file store state of the floppy disk is illustrated in figure 1. The learner's final intention is to format the floppy disk A. However, the floppy disk contains a folder with some lecture notes which apparently are useful. Therefore, the learner wants to move this folder to a safe place (the hard disk of his/her computer).

In order to achieve his/her goal the user issues a cut command (action 1) in order to move the folder 'lecture notes'. However, it appears that the learner does not know how to complete this plan because in the second action, s/he falsely uses a 'copy' command instead of a 'paste' command. Web F-

SMILE finds this action suspect because if it was executed it would delete the content of the clipboard before this was used anywhere. Therefore, the system tries to generate alternative actions that the learner may have meant to issue instead. In order to select the most appropriate advice, the system uses the information about the learner that is available on the user model. The alternative action that is considered by Web F-SMILE as more likely to have been intended is the action 'paste(C:\Courses\)' for several reasons. First, it uses effectively the content of the clipboard. Moreover, the commands 'copy' and 'paste' are considered quite similar because they both involve the clipboard. Thus the user may have confused them.

Table 1: An example of a learner's interaction.

1. cut(A:\lecture notes\)
2. copy(C:\courses\)
Web F-SMILE's reasoning: Suspect Action.
Suggestion: paste(C:\courses\)
Additional tutoring themes:
Copying Objects,
Moving Objects.
3. paste(C:\courses\)
4. format(A:\)

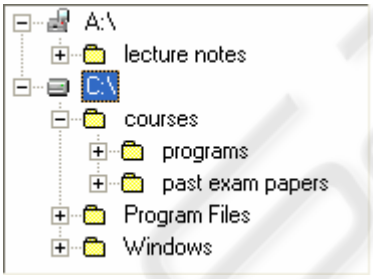

Figure 1: The learner's initial file store state.

The system also produces additional tutoring concerning the topic of copying and moving objects. The information of the user model indicates that the particular user has not sufficient experience in copying and moving objects and that in the past s/he had repeatedly made mistakes due to lack of knowledge on the topic. Indeed, the learner finds the system's advice very helpful and, therefore, adopts its suggestion in action 3. Then, in action 4, the learner formats the floppy disk, which was his/her final goal. In case the learner had used a standard file manipulation program, his/her error in command 2 would not have been recognised and the learner would have formatted the floppy disk and would have lost useful information.

4 SYSTEM'S ARCHITECTURE AND REASONING MECHANISMS

Web F-SMILE's architecture consists of six components, namely, Short Term User modelling (STUM) component, Long Term User modelling (LTUM) component, Advising component, Tutoring component, Domain Representation component and the user interface. The architecture of Web F-SMILE is illustrated in Figure 2. The components cooperate in order to provide individualised advice and tutoring in case this is considered necessary. Advice is provided to learners who have made an error with respect to their hypothesised intentions. All these components work locally on the learner's computer and only the LTUM component is responsible for the interaction with a Web Server for user modelling.

Every time the learner issues a command, the STUM component, which works on the client side, reasons about the command in terms of the learner's goals and possible problems. The Short Term User modelling (STUM) has two underlying reasoning mechanisms: one performs goal recognition based on the effects of users' commands and the other one performs error diagnosis (Virvou & Kabassi 2002).

The two reasoning mechanisms are independent of each other in the way they function. However, the compatibility of the hypotheses generated from these two mechanisms increases the certainty degree of these hypotheses. When an action is issued by the user, it is first examined by the goal recognition mechanism and in case it is believed that it contradicts the user's goals, the error diagnoser is used to generate similar alternative commands that the user may have intended to issue instead of the one issued which was problematic.

As soon as the alternative actions are generated, they are sent to the Advising component, which is responsible for selecting the alternative action that the learner was more likely to have intended. Furthermore, in case the STUM Component thinks that the learner's misconception was due to the learner's lack of knowledge, it informs the Tutoring Component about it. The Tutoring Component is responsible for forming an adaptive presentation of the lesson to be taught to the learner. The Advising and the Tutoring Component request information about the learner from the STUM Component. This is done so that they may adapt the advice and/or the lesson produced to the needs and the interests of each individual learner. The Advising and the

Tutoring Component, however, do not need to communicate with the Server, directly since their reasoning mechanisms reside on the client.

Both the Advising Component and the Tutoring Component send their results to the user interface, which is also located on the client. The user interface is responsible for the overall communication with the learner. This usually involves the collection of the learner's queries and the presentation of advice and tutoring in case the learner is diagnosed to have been in a problematic situation.

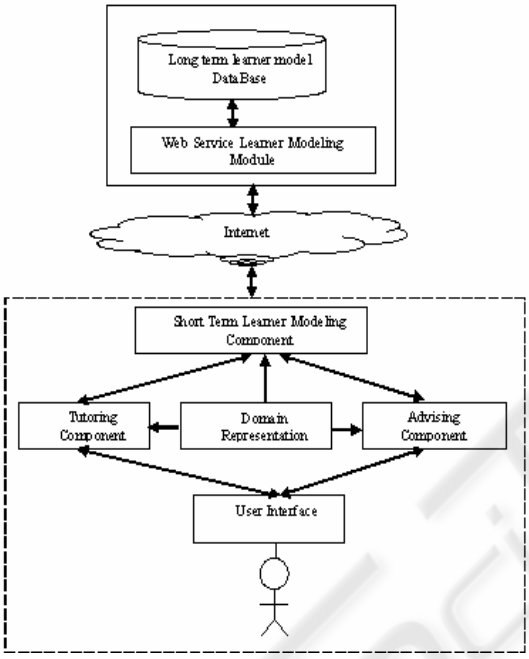

Figure 2: Web F-SMILE's architecture.

Every time the STUM Component acquires new information about the learner that interacts with the system, it sends it to the LTUM Component.

Generally, the LTUM Component, maintains and manages the learner profiles and provides relevant information to the STUM Component whenever this is considered necessary. Furthermore, the LTUM Component is responsible for the interaction with the Web Service User modelling (WS-UM) Server in order to maintain and update the information stored in user models, both on the Web Service Server and the client.

5 SECURITY COMPONENTS FOR SAFETY IN USER MODELLING

F-SMILE design adopts the most advanced and widely adopted standards for secure interoperable service provision. F-SMILE relies on XML and Web Services for security and interoperability, a fact which enables smooth integration with existing accounting software that organisations may use, as well as stand-alone operation of the service. This is achieved by publishing the provided service in UDDI (Universal Description, Discovery and Integration Protocol) based directories from which the service description can be retrieved formulated as specified by WSDL (Web Services Description Language). This enables other Web Services conforming to the appropriate message formats to interact with the F-SMILE Web Service.

The User Interface is a Signed Java Applet running on a standard web browser. The user interacts with the system through this interface to create, manage and send e-files/objects. The interface is able to produce XAdES signatures according to the hosting organization signature policy. It communicates with five other entities:

- The user's smart card for authentication and signing purposes. The communication protocol uses the PKCS#11 standard (Nash 2001).
- The F-SMILE e-learning system to deliver e-learning data. The communication is performed through the use of SOAP over HTTP.
- The CA (Certification Authority) to request certificate status information. The protocol used is OCSP (Online Certificate Status Protocol).
- The TSA (Time Stamping Authority) to request time stamps. This communication uses an implementation of the standard time stamping protocol.
- The XML database to retrieve e-learning specific information (existing files, contact details etc.). The communication protocol is SOAP over HTTP.

All entities are depicted in Figure 3.

The actors that take part are:

a) *Educational Organisation*. This organization hosts the F-SMILE infrastructure. It takes the appropriate steps to deploy the service and publish it in the Registry, so that other organizations may find it. It also communicates with the TTP to get the proper security credentials.

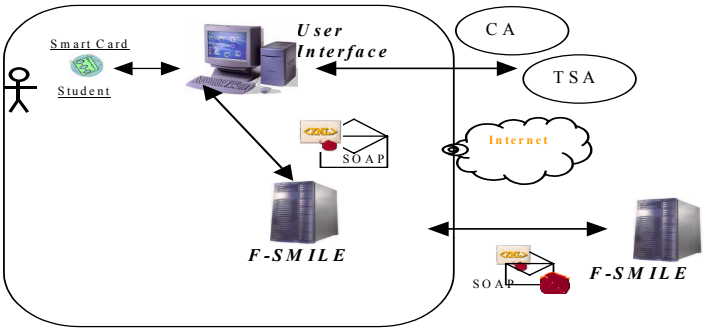

Figure 3: Security components of F-SMILE.

b) *The Learner*. The Receiver organization (or individual) may be hosting the same F-SMILE service or another implementation of a service, which understands the SOAP messages with WS security extensions. In the latter case, the Receiver will have to search for the description of the F-SMILE Web Service in the UDDI and be configured to understand its messages. The Receiver organization will also have to communicate with the TTP to get its proper security credentials.

c) *The TTP*. Before any secure messaging can take place, all participants need to have established an adequate security framework with Trusted Third Parties (TTPs)(Adams 1999). The required TTPs in our solution are at a minimum a Certification Authority (CA) and a Registration Authority (RA) offering the PKI services of registration, certification and revocation status information with OCSP, as well as a Time Stamping Authority (TSA) offering standard based time stamping services.

d) *UDDI directory operator*. This operator hosts a public UDDI directory where Web Services can be published and become publicly available.

6 CONCLUSIONS

In this paper, we have described a personalised learning environment that helps users learn how to operate their file store. The personalisation and adaptivity of learning depends on information about users such as the learners' prior knowledge, abilities and needs which are kept in the long-term and short-term user models. For this reason complete and accurate user models are needed for each user.

This problem is addressed in Web F-SMILE with the incorporation of Web Services for user modelling and the use of smart cards (for authentication purposes) with which the learners can use any PC. Web services are used in Web F-SMILE for the interaction of the agents of the system with a User modelling Server (WS-UM). WS-UM maintains a central database of all user models. In addition, Web F-SMILE keeps for every learner one user model centrally on WS-UM and one user model in each computer that the user uses to interact with Web F-SMILE. In this way, Web F-SMILE overcomes possible problems that may arise due to possible communication failures between a learner's PC and the Server.

The proposed Web Service architecture ensures better accuracy, completeness, security and interoperability of the user models as compared to other traditional architectures that have been used for the deployment of ILEs over the Web

REFERENCES

Adams, C., Lloyd. (1999) Understanding Public-Key Infrastructure – Concepts, Standards and Deployment Considerations, 1st Edition, Macmillan Technical Publishing.

Eastlake,D. , Reagle, J. (editors). (2002). "XML Encryption Syntax and Processing", W3C Recommendation, www.w3.org/TR/xmlenc-core

Hartman, B., Flinn, D., Beszosov, K., & Kawamoto, S. (2003). Mastering Web Services Security, Wiley Publishing.

Kobsa, A. (2002). Personalized hypermedia and international privacy. Communications of the ACM 45(5), 64-67.

Microsoft Corporation, Microsoft® Windows® 98 Resource Kit, Microsoft Press, 1998.

Nash, A., Duane, B., Brink, B., Joseph, C. (2001). PKI: Implementing & Managing E-Security, McGraw-Hill Osborn Media Publishing.

Virvou, M. & Kabassi, K. (2002). Reasoning about Users' Actions in a Graphical User Interface. Human-Computer Interaction, 17(4), 369-399.