An Investigation of Secure Access and Privacy Protection in e-Learning Environment

Radi Romansky, Irina Noninska

Abstract: *Privacy is a fundamental human right of individuals and personal data protection occupies the main part in it. Since the contemporary digital world creates new challenges for information security an e-learning system must follow basic rules ensuring privacy. In this reason the aim of secure access to all resources of an e-learning environment is very important and adequate technological and organizational measures for authentication, authorization and protection of personal data must be applied. Strong security procedures should be proposed to protect user's profiles, designed after successful registration and all personal information collected by educational processes. The goal of this article is to present the authors' point of view for organization of security access and personal data protection in an e-learning environment with internal and external resources based on cloud and social computing. A formal description is proposed and Markov chain model is designed. Analytical investigation based on partial multi-factor experimental plan has been carried out and several statistical assessments delivered by Develve software are discussed.*

Key words: *eLearning Processes; Secure Access; Privacy; Formalization and Modelling.*

INTRODUCTION

The global network Internet proposes many opportunities for collaboration and remote access, making communications easy and fast. At the same time Web applications which usually share personal information, determine a necessity for secure Internet connections, hence network providers must guarantee user's privacy [1, 2]. It is well known that the privacy is a fundamental human right and it very often depends on secure processing of personal data. Different components of the digital word require creation of personal profiles that consist of personal data and they should be protected by improving the legislation [3] and by ensuring adequate level of security [4]. In this reason the European Commission has proposed a new regulation in the field of privacy in the cyber space and has promoted the new paradigm "right to be forgotten / to be erased" [5].

Different models and schemes for digital education are used as basic components of contemporary digital world and all aspects of digital privacy and secure access to the profiles with personal data must find adequate solution. At present day e-Learning environments are extended by opportunities that give social computing [6] and cloud services [7] which outline new challenges for digital privacy [1, 2, 8]. Different techniques for investigation of e-learning approaches, able to validate secure e-learning schemes are used. Some of the most popular methods are graph formalization [9], statistical modelling [10], stochastic modelling [11], etc. The Markov chain theory is defined as one of well applicable apparatus for investigation of processes in e-learning structures [12].

This article discuses basic security measures and privacy rules in e-learning environment and presents an idea for organization of a complex e-servicing system with internal educational resources and remote access to external educational space based on cloud and social computing technologies. The proposed structure unites two sub-systems where functions for security and privacy protection are divided: *Front office* is designed for user's authentication and personal profiles creation; *Back office* is responsible for access control based on authorization, digital rights management and personal data protection. A Markov model for investigation of processes in this heterogenic e-learning environment is designed. This model is used for components' assessment, where special attention on processes of authentication and authorization of the users is paid.

SECURITY AND PRIVACY IN E-LEARNING ENVIRONMENT

The traditional e-learning system consists of sub-systems as a virtual learning environment for collaboration and assessment, library management and sub-systems for grading and content management. At the same time contemporary e-learning processes

usually are realized via the Internet using different Web technologies. In this reason, an e-learning environment could be defined as an integration of educational, information, network and knowledge based technologies and platforms. This complex structure could be extended by social computing (social media, social networks, social bookmarkers, blogs/micro-blogs, etc.) and cloud computing (cloud services, data centers, etc.). These technologies determine the necessity of strong policies for information security and personal data protection (PDP), as well.

An e-learning environment usually implements different procedures for control over users' activities, preventing unallowed access and data modification. In the frame of information and communication security reliable authentication schemes could be applied on the base of specialized hardware, software or biometric tools and cryptographic algorithms. They must be put at the root of a Digital Rights Management System (DRMS) which is able to guarantee successful data protection during input, processing, archiving and transfer via the Internet. In addition, each e-learning environment should enforce strong policy for personal data protection having in mind that data exchanged via the global network are frequently subject of non-authorised using, modification, corruption or loss. Information vulnerability in clouds and social networks is as high as in enterprise applications for computing systems communicated via the Internet. Frequently applied attempts in practice which could successfully overcome proposed security measures and violate privacy are summarized into the following four groups:

♦ Embarrassment with identification of data controller, data processor and data subject in the frame of data processing obligations;

♦ All rights of a user concerning data subject could not be guaranteed, bearing in mind that every user must know the goal of data collection and policy for their protection, he/she should be able to revise and block some data, require data deletion, etc.;

♦ Possibility for multiple data transfers between different locations including to the countries with low level of personal data protection policy;

♦ Low level of organizational and technical measures used for personal data protection, which is main obligation of data controllers.

Several basic architectures for e-learning systems development are presented in [10] as service-oriented architecture, distributed architecture, event driven architecture and cloud oriented architecture. An example of e-learning cloud architecture based on cloud services PaaS, SaaS and IaaS is proposed in [9]. All data supported and used in such system are stored in data centers and could be managed by remote access to the resources via the Internet. In this case the main security and privacy affairs of cloud computing will be transferred to the learning processes. These affairs could be assigned to social computing as well, since creation and supporting of user's profiles in the network space permit different personal information to be accessed by other users via the global network. This could cause very undesirable consequence for users' identity and violate their privacy. In this reason the PDP should be important obligation for any operator or stakeholder, who provides distributed services.

FORMAL DESCRIPTION OF E-LEARNING ORGANIZATION

Formal description of an e-learning environment with heterogenic structure is presented in fig. 1. The structural scheme includes four systems which are listed below.

♦ Front office – system for supporting user's access to e-learning services by input point (web portal). The main functions of this system are the user's requests identification, checking the legitimate access, starting the procedure for registration (for a new user) and carried out secure authentication by adopted tools. It must be noted that the registration procedure is connected with creation a personal profile and every user must be informed for this and for his/her rights. Additional procedures that accomplish functionality of these systems are registration of each login (time, IP address and additional attributes),

supporting audit journal (file) for the parameters of each access and collection of statistical information for accessed components. All these additional functions are proposed to be used for analysis of inputs profiles after unallowed access or cyber attacks.

♦ Back office – it performs basic administrative procedures and guidance of requests to a suitable component for processing. The gate of the back office is the module "Request Analysis" which has very important role as a main requests' distributor. If a current request is intended to access any collaborated external resource (cloud services, data centres, social media networks, specialized web sites, etc.) it is directed to the security shell of these resources (it is assumed that each attached external environment accessed via the Internet has individual system for security and data protection – in reason of the rules of the EU legislation). If a request needs any internal resource then the internal security sub-system will start with preliminary authorization on the rules of the realized internal DRMS. An authorization procedure ensures access to two databases – with personal profiles and with administrative information to determine access rights for concrete resources by user's profile (student, tutor or administrator). The first database should be reliable protected by organizational and technical measures according to the rules of the data protection law. The next steps supported by the back office are connected with processing of the request and could be realized after correct authorization only.

♦ Educational space that includes all internal information (learning) resources;

♦ External network resources attached to the e-learning environment and accessed via the Internet on the base of rent service or by providers. Each external system must have its own personal security system and is obliged to support an adequate policy for PDP.

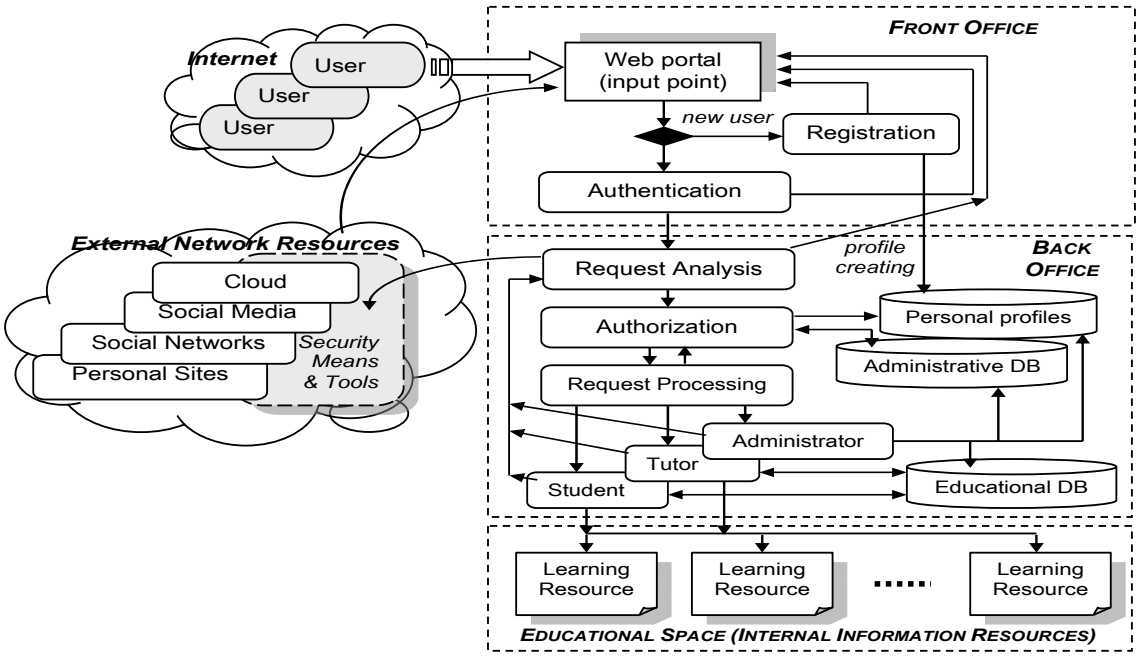


Figure 1. Structural Organization of a Heterogenic e-Learning Environment

PROCESSES INVESTIGATION BY MARKOV MODEL

A Markov chain (MC) with 8 discrete states is used for description of the processes. The model definition is presented in figure 2 with transactions: a – probability for unregistered (new) user; b – probability for correct authorization of registered user; c – probability for a request to access and using external educational resources; d – probability for an input in the back office and authorization (determining the right) for using an internal resource; e – probability for an internal educational resource using (after correct

authorization); f – probability for using system resources and personal data processing. The vector of initial probabilities is $\mathbf{P_0}$ = {1, 0, 0, 0, 0, 0, 0}.

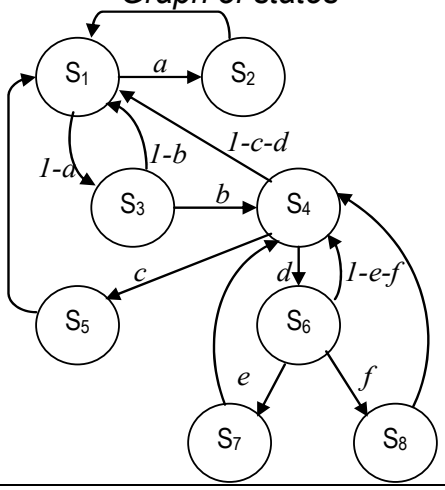
States: **S** = {S_1, ... , S_8}

S_1 – Access to the front office;
S_2 – Registration procedure;
S_3 – Authentication procedure;
S_4 – Request analysis;
S_5 – Access to external resources;
S_6 – Authorization procedure for access to internal resource;
S_7 – Using educational resource;
S_8 – Using system resource & personal data processing

Matrix of transitive probabilities **P** = {p_{ij}}

	1	2	3	4	5	6	7	8
1	0	*a*	1-*a*	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	1-*b*	0	0	*b*	0	0	0	0
4	1-*c*-*d*	0	0	0	*c*	*d*	0	0
5	1	0	0	0	0	0	0	0
6	0	0	0	1-*e*-*f*	0	0	*e*	*f*
7	0	0	0	1	0	0	0	0
8	0	0	0	1	0	0	0	0

Graph of states



Analytical model definition:

$$1)\ p_1 = p_2 + (1-b).p3 + (1-c-d).p_4 + p_5$$
$$2)\ p_2 = a.p_1$$
$$3)\ p_3 = (1-a).p_2$$
$$4)\ p_4 = b.p_3 + (1-e-f).p_6 + p_7 + p_8$$
$$5)\ p_5 = c.p_4$$
$$6)\ p_6 = d.p_4$$
$$7)\ p_7 = e.p_6$$
$$8)\ p_8 = f.p_6$$
$$9)\ p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 = 1$$

Figure 2. Definition of MC-model

A preliminary analysis of procedures is made and some assumptions are accepted to simplify the analytical investigation. For example, it is assumed that the probability of access to the system by new unregistered user is no more than 0,3 ($a \le 0{,}3$) and the unauthorized access to the resources (including attacks from external nodes) is in the frame [10%, 30%], i.e. $0{,}1 \le b \le 0{,}3$. The next assumption is that the refusal of service (in the states "Request Analysis" and "Authorization") is no more than 10% of all cases and this permits to determine the values: $(1\text{-}c\text{-}d) \le 0{,}1 \Rightarrow (c+d) \le 0{,}9$ and $(1\text{-}e\text{-}f) \le 0{,}1 \Rightarrow (e+f) \le 0{,}9$.

The analytical solution of the system of equations is made by presentation the probabilities by using only the probability p_1 and replacement in the last equation (9):

$$p_1 + ap_1 + (1-a)p_1 + \frac{b(1-a)}{(1-d)}p_1 + \frac{bc(1-c)}{(1-d)}p_1 + \frac{bd(1-a)}{(1-d)}p_1 + \frac{bde(1-a)}{(1-d)}p_1 + \frac{bdf(1-a)}{(1-d)}p_1 = 1$$

After solution of this equation and replacing $\pi = [2\text{-}2d+b(1\text{-}a)(1+c+d+de+df)]$ the final results for the probabilities are determined as follows:

$$p_1 = \frac{(1-d)}{\pi}; p_2 = \frac{a(1-d)}{\pi}; p_3 = \frac{(1-a)(1-d)}{\pi}; p_4 = \frac{b(1-a)}{\pi};$$
$$p_5 = \frac{bc(1-a)}{\pi}; p_6 = \frac{bd(1-a)}{\pi}; p_7 = \frac{bde(1-a)}{\pi}; p_8 = \frac{bdf(1-a)}{\pi}$$

ANALYTICAL INVESTIGATION AND EXPERIMENTAL RESULTS

The assumptions made in the previous section permit to define the working frame for analytical investigation of the proposed Markov model by determining concrete values for the probabilities: $a \in \{0,2;\ 0,25;\ 0,3\}$; $b \in \{0,7;\ 0,75;\ 0,8;\ 0,85;\ 0,9\}$; $c \in \{0,35;\ 0,4;\ 0,45\}$; $d \in \{0,45;\ 0,5;\ 0,55;\ 0,6\}$ and (1-*e*-*f*)=0,1 $\Rightarrow e = f =$ 0,45 (equal access probabilities to educational or system resource after authorization). This factor environment permits to construct partial multi-factor experimental plan based on all combinations of accepted values for the probabilities. Statistical processing of this plan is made over Develve software and several experimental results are presented in fig. 3 and table 1.

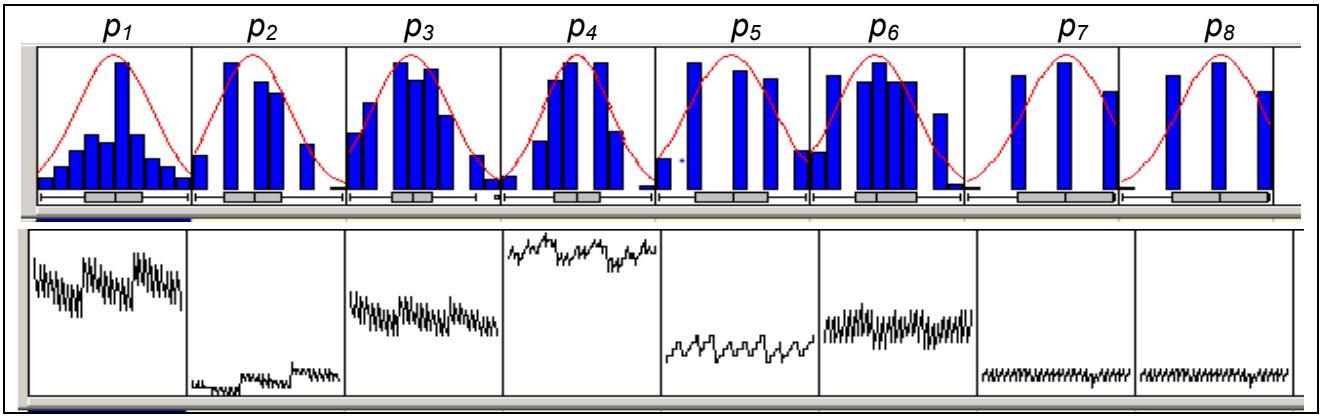


Figure 3. Histograms and time series for the probabilities based on the obtained data sets

Table 1. Generalization of the statistical assessments for the probabilities

	p1	p2	p3	p4	p5	p6	p7	p8
n	108	108	108	108	108	108	108	108
Mean	0,1995	0,0501	0,1493	0,2505	0,1001	0,1318	0,0593	0,0593
Median	0,1992	0,0492	0,1490	0,2512	0,1003	0,1313	0,0591	0,0591
MIN	0,1501	0,0300	0,1151	0,2198	0,0787	0,0989	0,0445	0,0445
MAX	0,2522	0,0757	0,1885	0,2777	0,1216	0,1666	0,0750	0,0750
Δ=[max-min]	0,1021	0,0457	0,0733	0,0579	0,0429	0,0677	0,0305	0,0305
Variance	0,0006	0,0001	0,0003	0,0002	0,0001	0,0003	0,0001	0,0001
St.Dev.	0,02391	0,01139	0,01757	0,01293	0,01051	0,01778	0,008	0,008
Conf.Int.	0,005	0,002	0,003	0,002	0,002	0,003	0,002	0,002

The experimental results show that the probabilities have very small assessments for the characteristics variance, standard deviation (St.Dev.) and confidence interval (Conf.Int.). Equal values for Student's T-distribution and normal distribution are obtained. The difference Δ=MAX-MIN is the largest one for the probability p_1 which presents the loading front office portal by access of different remote users.

The average values obtained for the final probabilities show that the highest value has implementation of the state "Request analysis". On the other hand the assessments for security procedures, defined in the model – authentication (p_3) and authorization (p_6) have a quite small difference (about 0,017) which is due to external resources' access. The difference between cases of process initialization by access to the Front Office (input point – p_1) and using of the Back Office procedure "Request Analysis" (p_4) could be explained with the possibility to process many internal requests by other states as authorization, educational resources, system resources or personal data processing.

CONCLUSIONS AND FUTURE WORK

A preliminary investigation of processes in a heterogenic e-learning environment is proposed in this article. The calculated assessments for the final probabilities permit to analyse the frame of components' implementation for security and privacy protection. Relatively high values obtained after analysis outline their importance in this environment. The analytical investigation could be extended by realization of a full multi-factor experimental plan which is the goal of a future work in order to make a comparison of probabilities in more large borders. For this purpose authors are intending to accomplish investigation applying additional apparatus for modelling and simulation.

REFERENCES

[1] Fischer, A. E. Improving User Protection and Security in Cyberspace, Report of Committee on Culture, Science, Education and Media, Council of Europe, 12 March 2014, (www.statewatch.org/news/2014/mar/coe-parl-ass-cyberspace-security.pdf)

[2] Kinast & Partner. Social Media and Data Protection, 2014 (www.kinast-partner.com/data-protection-law/social-media-and-data-protection/)

[3] Shear, B. S. European Data Protection Authorities Lead the Fight to Protect Digital Privacy, July 2013 (http://www.shearsocialmedia.com/2013/07/eu-data-protection-authorities-are.html)

[4] Symantec Corporation. 2014 Internet Security Threat Report, Volume 19, April 2014 (http://www.symantec.com/security_response/publications/threatreport.jsp)

[5] European Commission. Progress on EU Data Protection Reform Now Irreversible Following European Parliament Vote, MEMO, Strasbourg, 12 March, 2014 (http://europa.eu/rapid/press-release_MEMO-14-186_en.htm)

[6] Neville, K., C. Heavin. Using social media to support the learning needs of future IS security professionals. Electronic Journal of e-Learning, 11(1), 2013, pp.29-38.

[7] Velicanu, A., Longu, I., Diaconita, V. Cloud e-Learning. 9th Int' Scientific Conf. eLearning and Software for Education, 25-26 April 2013, Bucharest, pp.380-385.

[8] Social Media and Data Protection: the Employment Law issues, TaylorWessing (Global Data Hub), February 2013 (http://www.taylorwessing.com/globaldatahub/article_employment_law_issues.html)

[9] Sun X., Z. Li, S. Hu. Directed-Hypergraph Based E-Learning Process Modeling Supporting Dynamic-Personalized-Combined Resource Optimization, Proc. of 5th Int'l Conference of Digital Home (ISBN 978-1-4799-4285-5), 28-30 November 2014, pp.324-330.

[10] Nouri, A. et all. Faster Statistical Model Checking by Means of Abstraction and Learning. RV, September, Toronto, Canada, 2014 (www-verimag.imag.fr/~jcombaz/misc/rv14/pdv)

[11] Abraham, S., S. Nair. Cyber Security Analytics: A Stochastic Model for Security Quantification using Absorbing Markov Chains. Journal of Communications, vol.9, No.12, 2014, pp.899-907.

[12] Taraghi, B. et al. On using Markov Chain to Evidence the Learning Structures and Difficulty Levels of One Digit Multiplication. Proc. of the 4th Int'l Conf. on Learning Analytics and Knowledge (LAK'14), Indianapolis, USA, 24-28 March, 2014, pp.68-72

ABOUT THE AUTHORS

Prof. Radi Romansky, D.Sc., Department of Electronics, Computer Systems and Technologies, College of Energy and Electronics at Technical University of Sofia, Phone: +359 2 965 3295, E-mail: rrom@tu-sofia.bg.

Assoc. Prof. Irina Noninska, PhD, Department of Computer Systems, Technical University of Sofia, Phone: +359 2 965 3471, E-mail: irno@tu-sofia.bg.

The paper has been reviewed.