Zbigniew PIOTROWSKI[1], Krzysztof SAWICKI[1], Arkadiusz ŻELECHOWSKI[2]

Military University of Technology, Faculty of Electronics, Institute of Telecommunications (1),
Military University of Technology, Faculty of Cybernetics (2)

E-learning system for assessing the susceptibility of a subscriber to manipulation of identity and voice in telephony

***Abstract**. In this paper an e-learning system designed to train people in the field of secure communications on a daily basis when using GSM and PSTN phones for business and personal reasons is demonstrated. The system has been implemented using publicly available tools (Apache, Asterisk, PHP). The system performs synchronous playback over a phone and displays questionnaires on a PC, and is completely automatic.*

Streszczenie.** W artykule przedstawiony został system e-learningowy przeznaczony do szkolenia osób z zakresu bezpiecznej komunikacji na co dzień posługujących się telefonami GSM i PSTN w celach służbowych i prywatnych. System zrealizowany został z wykorzystaniem ogólnodostępnych narzędzi (Apache, Asterisk, PHP). System realizuje synchroniczny proces odtwarzania przez telefon i wyświetlania formularzy pytań na komputerze PC i jest całkowicie automatyczny. (**System e-learningowy do oceny podatności abonenta na manipulację tożsamością i wiadomością głosową w telefonii).

Słowa kluczowe: e-learning, telefonia, bezpieczeństwo, podszywanie się pod tożsamość głosową, impersonacja, podatność na atak impersonacji, nieautoryzowana zmiana wiadomości głosowej
Keywords: e-learning, telephony, security, impersonating voice identity (spoofing), impersonation, vulnerability to impersonation, unauthorized change of a voice message

Introduction

E-learning systems are becoming more and more popular [1,2,12,13]. Many universities are conducting studies using special e-learning platforms prepared for their students. Online studies carried out by the Warsaw University of Technology [3] are an example of such a solution.

So far many different types of e-learning systems have been implemented. In [4] a system is presented for learning over the Internet, which focuses on cooperation between participants and exchange of notes. Another approach to the problem has been presented in [5]. An e-learning system introduced at the University of Aveiro in Portugal is presented there, whose primary goal is to support the academic learning process of engineers. An interesting solution is presented in [6]. The described system makes it possible to post multimedia lectures (video, audio, still images), sets of questions for review, as well as enabling sharing of materials and notes.

Many studies focused on the analysis of the effects of studying with the help of e-learning systems have also been performed. Paper [7] proposed a theory of e-learning as a new field of scientific interest. The author of document [8] assesses the quality of distance learning in Greece, Italy and Cyprus among students in the technical and medical fields. Meanwhile, the behaviour of Chinese part-time students - frequency of use of the e-learning system, use of these systems, and even emotional changes among participants of e-learning - has been analysed in [9].

Skills tested

The designed system, called BruteForceVoice [11], allows ten different skills to be explored through testing sixteen tasks. Each task consists of at least one question. The following skills are tested:

Identification [14,15] of a short voice profile - test conducted through four experiments. The test participant listens to recordings of voices that must be recognised. These recordings can be listened to repeatedly, so that the person can memorise the voice profile as well as possible. He or she then hears six speakers. The task of the participant is to identify which of the six speakers was the one from the first recording. The first experiment consists of indicating one voice from among six speakers reading the same text of a few seconds in length. In the second experiment, the task is to recognise one voice from among six speakers reading various texts of a few seconds in length. In the third experiment, two people are presented at the beginning. The test participant is then required to indicate which of the four dialogues featured both of the speakers presented before. All four dialogues are identical in terms of content in this experiment. The fourth experiment differs from the third in that the content of the dialogues differs. Later in this paper the ability to recognise a short voice profile will be marked as G01.

The second skill (G02) is the recognition of artificial (synthesized) voices. This skill is tested through an experiment in which the listener is presented with five different recordings, only one of which is a recording of an actual person. The listener has to indicate which recording features a human voice.

The third examined skill (G03) is the ability to understand information communicated by voice in a variety of listening conditions. The experiment that examines this skill consists of the test participant listening to a few seconds of noisy text. Then the person answers three short questions about the text. During the experiment the listener is presented three increasingly noisy texts.

The fourth tested skill (G04) is the ability to extract additional information from the signal, such as the acoustic background from the recording heard. This skill of the subscriber is tested through an experiment that consists of presenting the listener with recordings with different acoustic backgrounds (e.g. a plane taking off, a moving train, background music). The task of the listener is to recognise the acoustic backgrounds heard and to indicate the order in which they were played.

Another skill (G05) is the ability to decide whether a voice message has been modified. The tested person first listens to the original recording. Then he or she listens to three tracks, of which two are modified (e.g. a single word being cut). The participant is to identify the recording that has not been changed.

The sixth skill checked (G06) during the test is the ability to listen and simultaneously note information. The participant hears recordings where bank account numbers are dictated. These numbers must be noted and then transcribed onto a computer by the test subject. In the first

experiment the recordings feature added noise, while in the second experiment the recordings are accelerated so as to check whether the test participant can quickly and accurately record information.

The seventh skill (G07) is the ability to compare and check messages read and heard. The participant sees an account number on the screen and hears five speakers reading the number. Four of them make a mistake when reading this number. The task of the participant is to identify which speaker read the account number impeccably.

The eight tested skill (G08) consists of memorising and comparing a variety of short pieces of information. In two experiments the participant is given at the beginning a date or name, surname, date of a meeting and its subject. Then the test participant listens to recordings in which information is provided. Only one of them features the correct information. The participant is to identify the correct recording.

The ninth skill (G09) is reading and listening comprehension. In one experiment the participant is presented with a short text that must be read. Then five short recordings are played back. The task of the user is to identify whether the speaker was telling the truth or lying. In the next experiment the participant listens to a longer read text. He or she then has to answer three specific questions about the heard text.

The tenth and final skill (G10) is the ability to indicate whether the voice transmission contains a watermark. The participant is presented with a couple of recordings with an added digital acoustic watermark [10]. One of them does not contain a watermark, but in the remaining ones the watermark is encoded with varying signal strength. The participant is to identify the recording that does not feature the included watermark.

Architecture

The system was implemented in a client-server architecture. The client's role is played by a Web browser and a telephone (GSM/UMTS or PSTN). The following services are installed on the server side: Apache web server, PostgreSQL database server and Asterisk software telephone exchange. The system architecture foresees a separation of the system into two separate servers. This is due to security concerns and the specific conditions of functioning of the network in which the system has been initiated. One of the machines plays only the role of a WWW server that makes it possible to display the user interface in the form of a website. On the second machine the other services are installed and initiated. Communication between the machines is ensured through an internal protocol based on a database. A description of the architecture is presented in Figure 1. System logic and supervision over its operations has been implemented in PHP.

The system is equipped with an administrative panel that makes it possible to control its operations and manage (create, edit and delete) test tasks.

Synchronisation of operations and internal protocol

There is no ready solution that allows direct control over a telephone connection managed through the Asterisk telephone exchange using a web browser. Therefore there was a need to develop such a solution. Since the web server that accepts information from the system's user is located on a different machine than the software telephone exchange, the exchange of messages between the two threads - the control thread of the telephone connection and the web interface thread - is done using a database.

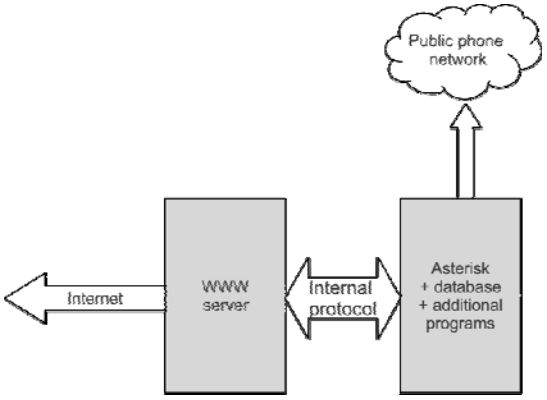

Fig. 1. System architecture

The database also includes a special table, within which two simultaneously running threads record information regarding their operational state. The connection controller process, after a call is taken by the test participant, provides information regarding this by writing a new record in the database. The looped interface process checks whether such a record was written. If it has appeared, it starts to show the content of the task. The connection controller process then waits for information from the Web interface on whether the test participant has already read the instructions for the task. This check is also periodical. Once this information is received, test recordings are played. After they are played back, the interface process is informed of the fact that it is to display the questions for a task. Once an answer is provided the interface process moves to the next task. When all the tasks are completed, the connection controller automatically disconnects the ongoing call.

Method of operation

he system will eventually be available to all interested parties. Three types of users are distinguished in the system: administrator - as the person responsible for operating the system, manager - as the person responsible for the registration of participants and employee - training participant. The manager can independently register within the system. His or her task is to introduce basic data regarding participants who are registered for training. The manager indicates basic information: name, surname, e-mail address and phone numbers (both mobile and land lines) that will be used during the training. Another task of the manager is to pay for the cost of training. Once registered by the manager the participants can proceed to undergo the training.

The training begins with a telephone connection being made. The system calls a training participant at a previously indicated GSM or PSTN phone number. The participant, after receiving the call, is presented with the content of the first task (Fig. 2).

Fig. 2. Example of task description

The training participant can read the task instructions and optionally repeatedly listen to the recordings attached to the instructions. After moving forward, he or she hears a test recording. These recordings are played only once and there is no possibility of repeating them. Following playback, the system automatically presents the training participant with a question that he or she has to answer (Fig. 3).

Fig. 3. Sample question

Fig. 4. An example of a certificate issued by the training system

After the user provides an answer, he or she is automatically redirected to the next task. Once the test finishes, the test participant is given a percentage score for correct answers given and a brief summary of the test. Then the test subject may proceed to review all tasks completed in the test and repeatedly listen to all the recordings presented earlier. There is also an option to automatically generate a certificate confirming competencies gained during the test (Fig. 4).

During the study the user sees all the time on the web application's dashboard how many tasks remain to be done and how much time is left to complete the test. These two pieces of information are meant to motivate the user. During the first tests, a maximum test duration of 50 minutes was specified.

Test results

Twenty people were tested during the trials of the system. Each of these people did all the tasks. The average duration of the entire test was 42 minutes (with an assumed maximum of 50 minutes). None of the participants exceeded the maximum allotted time. The average test score was 74%. Figure 5 presents the results of all carried out tests. A value of 100% means flawless execution by the telephony subscriber of all sixteen tasks in the interactive test.

As can be seen in Fig. 5, most of the participants achieved a score within the range of 70 - 80%. Such results seem satisfactory, especially since the people involved in the experiments use phones for mostly private purposes. However, it is worth taking note of a summary of results for individual skills (Fig. 6).

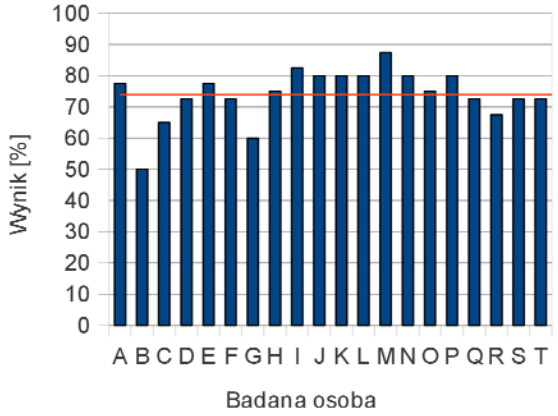

Fig. 5. Results of tests carried out by all test participants with the indicated average value

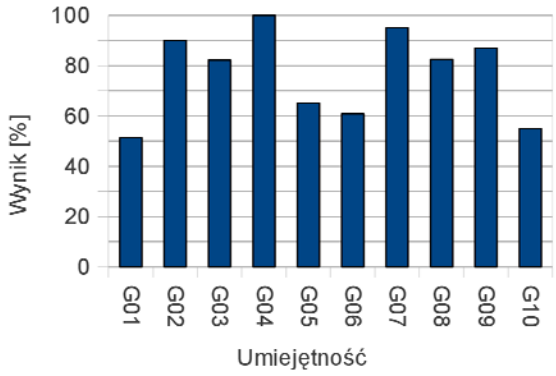

Fig. 6. Summary of results for individual skills (skills marked in accordance with the "Skills tested" section)

Study participants fared almost perfectly in terms of tasks associated with two skills: identifying synthesised voices and recognising background noise. High scores (approximately 90% of correct answers) were also obtained for the following skills: comparing a read and heard message, as well as reading and listening comprehension. Low scores were obtained for the short voice profile recognition skill (G01), which may mean that the participants are at risk of a voice spoofing type attack (impersonating voice identity). A very low score for the ability to recognise modified messages (G05) further supports this assumption. The participants also achieved poor scores in tests consisting of notation of dictated

information and recognition of recordings with an added digital acoustic watermark.

Summary

The proposed e-learning system can be used to identify potential risks associated with the level of awareness of dangers present in telephone networks on the part of people who every day come into contact with customers through mobile or land line PSTN telephony services, for example in emergency number dispatch centres, call-center offices, brokerage houses, banks and departments of state institutions charged with contacts with the public. The described system also reveals a clear gap in training provided in the field of counteracting spoofing of a subscriber's vocal identity and unauthorised modifications of voice phone calls. The system makes it possible to verify the skills of GSM/UMTS and PSTN subscribers within ten test categories, including short-term memory for a voice profile, reading and listening comprehension, memorising as many details in played back voice messages, identification of background noise, distinguishing a synthesized voice from natural articulation, identifying hidden transmission in the background signal.

ACKNOWLEDGMENT

Scientific work funded from assets for science for the years 2010-2012 by the Ministry of National Defense DNiSW/U/5/BN/1.14/2010/779.

REFERENCRES

[1] Clark R. C., Mayer R. E., E-learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning, Pfeifer corp., San Francisco, 2003

[2] Garrison D.R., Anderson T., E-learning in the 21st century: a framework for research and practice, Routledge Falmer, London, 2011

[3] Ośrodek Kształcenia Na Odległość Politechniki Warszawskiej: http://www.okno.pw.edu.pl/

[4] Ping-Lin F., Su-ju L., Han-Jang W., High Interactive Web Annotation Based e-Learning Platform, Proc. of International Conference on Digital Game and Intelligent Toy Enhanced Learning, 2010, pp. 198-201

[5] Sklyarov V., Skliarova I., E-learning tools and remote reconfigurable systems for engineering education , Proc. Of International Conference on Networking and Information Technology 2010, 11-12.06.2010, Aveiro, Portugal, pp. 11-15

[6] Lin P., Sugawara T., Hayasaka N., Miyanaga Y., A New e-Learning System Based on Cooperative Methods - Next-generation Extra University-education System. Proc. Of International Symposium on Communications and Information Technologies, 2006. ISCIT '06. 18.10.2006, pp. 727-730

[7] Huandong C., Research on the Learning Theory of E-learning, Fifth International Joint Conference on INC, IMS and IDC, 2009. NCM '09, 25-27.08.2009, pp. 1185-1187

[8] Brancaleone C., Evaluation of e-learning teaching activities in Mediterranean countries, Proc. Of 9th International Conference on Information Technology and Applications in Biomedicine, 2009. ITAB 2009, 4-7.11.2009, Larnaka, Cypr, pp. 1-4

[9] Wenhui P., Investigation and Analysis on E-learning Behavior of Spare-time Students, International Conference on Internet Computing & Information Services (ICICIS), 2011, pp. 381-384

[10] Z. Piotrowski, P. Gajewski, Novel method for watermarking system operating on the HF and VHF radio links, Computional Methods and Experimental Measurements XIII, CMEM XIII, WIT Press 2007, Southampton, Boston, pp. 791-800

[11] BruteForceVoice: http://abonent.wel.wat.edu.pl

[12] Sławomir Wiak, Maciej Krasuski, Dominik Jeske, Rafał Stryjek, System do prowadzenia nowoczesnego egzaminowania na odległość, Przegląd Elektrotechniczny, 07b/2012, pp.148-151

[13] Remigiusz Rak, Bogdan Galwas, Zaoczne Studia Inżynierskie przez Internet - nowy model kształcenia w Politechnice Warszawskiej. Przegląd Elektrotechniczny, 05/2002, pp.144

[14] D. Cetnarowicz, Sz. Drgas, A. Dąbrowski, Experimental speaker recognition investigations using head/torso simulator and telephone tranmission, Elektronika - konstrukcje, technologie, zastosowania, miesięcznik naukowo techniczny, 2011, nr 5/2011, pp. 94-97

[15] A. Dąbrowski, A. Meyer, A. Chmielewska, R. Weychan Segmentacja mówców w rozmowach telefonicznych na podstawie znaku wodnego, Elektronika - konstrukcje, technologie, zastosowania, miesięcznik naukowo-techniczny, 2011, nr 5/2011, pp. 98-102

***Authors**: ppłk dr inż. Zbigniew Piotrowski, Wojskowa Akademia Techniczna, Wydział Elektroniki, Instytut Telekomunikacji, ul. gen. S. Kaliskiego 2, 00-908 Warszawa, E-mail: zpiotrowski@wat.edu.pl; mgr inż. Krzysztof Sawicki, Wojskowa Akademia Techniczna, Wydział Elektroniki; Instytut Telekomunikacji, E-mail: ksawicki@wat.edu.pl; Arkadiusz Żelechowski, Wojskowa Akademia Techniczna, Wydział Cybernetyki, E-mail: arek@zelechowski.eu.*