Bokareva, A. A. (2020). Model of a document protection algorithm for a person based on biometric parameters. *Actual Issues of Modern Science. European Scientific e-Journal*, *6*, 1, 12–19. Hlučín.

DOI: 10.47451/inn2020-09-002

The paper is published in Crossref, ICI Copernicus, BASE, Academic Resource Index ResearchBib, J-Gate, ISI International Scientific Indexing, Zenodo, OpenAIRE, BASE, LORY, ADL, Mendeley, eLibrary, and WebArchive databases.

Alyona A. Bokareva, Researcher in Electronic Documentary, Undergraduate, ITMO University. St. Petersburg, Russia.

Model of a Document Protection Algorithm for a Person Based on Biometric Parameters

Abstract: At the moment, various types of paper documents, which carry various social and legal values relative to the citizen presenting them (passport, driver's license, etc.), are used to identify a person. In the field of electronic document management, electronic digital signatures have been created and implemented to transmit the generated document packages using network technologies. This paper discusses the principle of an algorithm for protecting an identifying document using biometric data without using third-party servers for both storing and processing information. The entire operation of the system is focused on the transition from paper documents to the electronic (cloud) version without loss of uniqueness and security with the possibility of multi-copying and use outside the work of data transmission networks.

Keywords: authentication, electronic document management, verification, personal identification, biometric parameters, personal data.

Алёна А. Бокарева, исследователь в области электронной документации, магистрант, Университет ИТМО. Санкт-Петербург, Россия.

Модель алгоритма защиты персональных документов на основе биометрических параметров

Аннотация: В настоящее время для идентификации личности используются различные виды бумажных документов, которые несут различные социальные и правовые ценности по отношению к гражданину, их предъявляющему (паспорт, водительские права и т.д.). В области электронного документооборота были созданы и внедрены электронные цифровые подписи для передачи сгенерированных пакетов документов с использованием сетевых технологий. В статье рассматривается принцип алгоритма защиты идентифицирующего документа с использованием биометрических данных без использования сторонних серверов как для хранения, так и для обработки информации. Вся работа системы ориентирована на переход от бумажных документов к электронному (облачному) варианту без потери уникальности и безопасности с возможностью многократного копирования и использования передачи данных вне работы сетей.

Ключевые слова: аутентификация, электронный документооборот, верификация, идентификация личности, биометрические параметры, персональные данные.

Introduction

At the moment, various types of paper documents, which carry various social and legal values relative to the citizen presenting them (passport, driver's license, etc.), are used to identify a person. In the field of electronic document management, electronic digital signatures have been created and implemented to transmit the generated document packages using network technologies.

Currently, many public and private services have started to use modern methods of paperless processing and exchange of documents, which can significantly reduce the time spent on processing documents, transactions and their exchange, as well as improve and reduce the cost of preparing, delivering, recording and storing them. Digital signature for individuals is a way to speed up and simplify interaction with government agencies, employers, etc.

At the same time, when switching to electronic document management, important questions, i.e., the question of the authorship of the document, and, no less important, the question of its authenticity and protection from changes, arise. The most convenient means of protecting electronic documents from distortion, while allowing uniquely to identify the sender of the message, is an electronic digital signature. The external expression of an electronic signature has nothing in common with a handwritten signature, but the purpose of both types of signatures is the same: document authentication (*Khachaturova, 2016*).

Based on the described purpose, an electronic digital signature does not allow to identify a person, which poses a new task to switch to an electronic document format that allows uniquely to identify a person with the possibility of universal use with a full set of parameters that reduce the number of documents used by a person to one without losing the functionality and originality of the document.

To solve the problem of switching to a full-fledged electronic document flow in the field of identification and authentication of an individual, it is necessary not only to determine the initial parameters, the use of which will uniquely identify the person, the kind and type of electronic document and the method of its storage but also to determine the method of cryptographic protection of this document using the basic principles and theoretical basis of this document flow.

To authenticate an individual using an electronic document, it is necessary to include an area in the document that stores information about the biometric parameters of the individual. This article considers a distributed point record of data in the entire volume of an electronic document, the storage coordinates of which will be determined using a unique encryption algorithm determined and calculated from scanned biometric data, thereby reducing the probability of errors when reading the document, data falsification, and the possibility of dynamic changes in the document itself and the variability of its use (*Bokareva, 2020*).

Results

Nowadays, the interest of specialists in the field of information security, in particular, protection of electronic documents from forgery, focuses on the development of cryptography methods based on the use of deterministic-chaotic processes more and more often (*Dovgalet al., 2004*; *Dovgal & Zacharov, 2008*). This class of processes is characterized by instability, which results in non-repeatability of the values of elements in a sequence of numbers with a representative bit grid. So-called nonlinear mappings depending on their dimensionality are used as generators of chaotic deterministic processes. For example, the values of all variables (twos, threes, etc.) are not repeated in two-dimensional and three-dimensional maps (*Parker & Zhua, 1987*; *Thompson, 1985*).

If the initial display values are unknown, the attacker must perform a complete search of all irrational numbers and all numbers obtained using the generators of deterministic-chaotic process, which significantly increases the cryptographic strength of the proposed methods (*Gordienko et al., 2018*).

Currently, there is an active introduction and use of electronic digital signatures to transfer documents between various public and private entities. When signing electronic documents, a digital signature ensures authenticity, i.e., makes it possible to check the document for integrity and originality. Integrity, in turn, consists of the fact that the content of information in the document cannot be modified by an unauthorized user. The signature is generated by a cryptographic operation, for which the document and private key are presented. This key should only be known to the person signing the document. To verify the document, a public (test) key, which can be known to everyone, is used. Nobody can get a private key from this public key. A test, in turn, is a cryptographic operation, for which a document and a test key are provided (*Bakhimova et al., 2016*).

These methods are widely used, and the stability of the algorithms has been proven both theoretically and over the entire period of use. However, these algorithms are not applicable for solving the tasks due to some factors: the need for fully autonomous operation of the system, which eliminates the possibility of any outside resources, cloud storage of document and the requirement of unambiguous identification.

Using biometric parameters as both a key and a method of cryptographic protection is the best option to solve this problem. If the read biometric data are presented in the form of a coordinate grid, then the process of document verification and identity authentication can be organized through a predefined function for changing the characteristics (digital code) of a point, defined by the coordinate grid of biometric parameters (hereinafter—the change map). Accordingly, when reading biometric parameters, we get:

- a map of changes (coordinates of points that need to be modified for verification and authentication),
- a key to the point change algorithm,
- a low probability of failures of the 1st and 2nd types.

Due to the absence of the need to request and transfer any data to other devices and servers, the system is completely autonomous and allows to reduce the time and cost of using it, which in some cases is extremely important.

Although the retinal identification algorithm was proposed, developed and proved to be unique for each person, its use in practice was significantly difficult due to the high (until recently) cost of the necessary equipment and special scanning conditions, which could give erroneous data with a reading error. However, with the development of technology, especially in the field of mobile devices in terms of technical parameters and the expansion of mobile applications, retinal scanning has moved to a new level and is available to be used almost everyone, who has a smartphone at their disposal. Naturally, applications and technical capabilities affect the quality and level of detail of the scanned image (data), and the use of the simplest applications, which are available to each user, does not provide the ability to operate with the necessary amount of data to identify a person in the framework of the task but provides prerequisites for the development of less expensive devices and software for full-fledged scanning of biometric parameters of a person to organize subsequent work with the read data in the framework of the task.

The proposed algorithm for data encryption in an electronic identification document can be schematically represented as an interacting system consisting of three matrices and a method (relationship) for data modification (transformation) when writing or reading them from the final file. The first matrix represents the original data of the person that has to be saved in an electronic document in a secure form. The second matrix is a map of changes (read biometric data of a person, converted into the necessary form for work). The third matrix is the file itself, in which the changed points (data) from the first matrix will be "appended" using data from the second matrix (namely, the coordinates of the change, the change itself, and the encryption method).

When decrypting data from a file containing an electronic document, the use of biometric data other than the original data of the identified person will lead to erroneous (not readable by humans) data, which in turn will make it impossible to identify the person under other documents, since significant data is read from the map of changes (agreed with it)—the formation of the original document file will be impossible without distortion and loss.

To create the first matrix, data from the scanned retina of the eye is used. This article considers reading the intersections of blood vessels located on the retina. Before starting the scan, a specialist needs to determine whether the eye is alive. When tracking the movement of the eyeball, it is suggested to take two characteristics of the eye. The first is fixing the eye on a specific point on the display. The second is the moment when the eyeball moves when moving the view from one point to another. The program evaluates the data obtained and determines unique characteristics for each case, i.e., for each person including the work of the eyeball muscles. The second characteristic is used only in industrial versions for special purposes and uses an expensive solution—the so-called Solvers, which help to determine the coincidence of characteristics in an acceptable time of 3–5 seconds. Using a smartphone with a camera attachment and the *Peek* software application, the procedure for removing the retinal picture is performed. Through the *Wiegand* data transfer interface and *API* functions of the application, information from the retinal image is transferred to a database or temporary file, where data is read and processed (*Givoino & Rostovtsev, 2016*).

To date, there are two newly developed algorithms to segment blood vessels:

- a method based on the use of a median filter,
- a method based on the use of a series of *Gabor* filters.

The results of testing these algorithms on two retinal databases demonstrate the possibility to use the first one for biometric authentication systems. The second method requires large computational costs, so it is not suitable for biometric authentication systems in the framework of the problem under consideration (*Nafikov, 2016*).

However, at the moment, the most stable algorithm with minimal errors of the first and second level (based on testing and practical experiments) is the algorithm based on the search for branching points. This algorithm searches for branching points in the blood vessel system. At the same time, it is more specialized in finding bifurcation and intersection points and much more resistant to noise, but it can only work on binary images. To search for points, segmented vessels are compressed to the thickness of one pixel. Thus, each point of the vessels can be classified by the number of neighbours. This algorithm requires much less computational resources compared to the algorithm based on the phase correlation method. In addition, there are opportunities for its complication to minimize the probability of hacking the system (Authentication via the retina of the eye).

Genetic factors do not actually determine the composition of the blood vessel's structure (the intersection of which creates the desired second matrix) on the human retina. In other words, the structure of the retina (the location of blood vessels) is not reflected in the human DNA and, accordingly, is not an inherited trait, which significantly affects the stability of the algorithm to erroneous access. Up to 400 unique features, from which a map of changes is created according to a given algorithm (or a second matrix for transforming the data of the first matrix into a third one), can be obtained from the retina. The size of the stored information about these unique features is only 96 bytes and is considered the smallest biometric template, which, in turn, allows to minimize the cost of resources (including computational and time), but is not suitable to create and implement the data transformation algorithm of the first matrix due to the critically small size and inconsistency with the selected algorithm for crossing blood vessels, but makes it possible to assess the degree of value of the processed information.

The small size of the primary change map to implement the original transformation of the first matrix into the third allows strengthening the algorithm without significant loss of resources. There are two ways to work with the change map: using the change map recursively or with an additional transformation algorithm. Both methods internally have the format of changing the original second matrix in one way or another and can be used simultaneously to strengthen the protection of an electronic document and reduce the probability of erroneous access and authentication of a person.

When developing and implementing an enhanced algorithm to read the intersections of blood vessels in the retina, it is necessary to take into account both the speed of the system to process the algorithm as a whole and use autonomous data exclusively, i.e., to use keys or encryption methods on a third-party resource in this task is not possible and contradicts the very principle of creating an autonomous electronic document to identify a person.

The use of this algorithm for recording identification and information data about a person is optimal to use an additional function, namely, modifying an existing file (e.g., an existing graphic image) with the introduction of data from the third (already transformed) matrix using the second matrix as a map of changes. This leads to additional requirements for the storage files used, such as document size, readability and use on various platforms and operating systems without loss of quality, the ability to cloud storage and storage in the memory of any technical device.

This algorithm allows not only to create a secure electronic document for identifying and authenticating a person but also to modify the form of storing any document that requires data concealment. For example, if a file is needed to encrypt, it can be placed inside any user-friendly image as described above and stored in the public domain, provided that only the user knows, which graphic image contains the desired file, and can decrypt it unambiguously way according to their biometric parameters.

Conclusion

In conclusion, it should be noted that the proposed concept, in contrast to existing solutions, allows creating a fundamentally new way to both use and store identity documents based on biometric parameters, which, in turn, significantly reduces the risks of erroneous verification, errors in access rights and identification, economic and other resource costs to use identification documents, it makes possible a full-fledged transition from plural of different types certifying and confirming social and legal rights of a citizen to a single electronic document.

References:

Authentication via the retina of the eye. (2019). Modern Electronic Library. Moscow. (In Russ.)

Bakhimova, L. A., Latypova, L. A., & Miftakhova, L. H. (2016). Methods of protection against falsification of electronic signature. *Bulletin of the Kazan Technological University*, *14*, 123–125. (In Russ.)

Bokareva, A. A. (2020). Personal authentication using verified electronic document. *European Scientific e-Journal*, *6*, 1, 4–10.

Dovgal, V. M., Gordienko, V. V., & Elagin V. V. (2004). Methods of a computer based on technology for production of counterfeit-protected electronic documents. *Telecomunication and Radio Engineering*, *8*, 745–753.

Dovgal, V. M. & Zacharov, I. S. (2008). Preventive counteraction to information attack on a text electronic document by the method of chaotic transpositions of symbol code fragments. *Telecomunication and Radio Engineering*, *4*, 995–1005.

Gordienko, V. V., Dovgal, V. M., & Lukina, A. V. (2018). The methods of cryptography and steganography to protect electronic documents against forgery based on the display of the Lorentz. *Auditorium*, *2*(18), 48–53. (In Russ.)

Givoino, A. A. & Rostovtsev, V. N. (2016). Protection of medical data of patients. *Reports of BSUIR (БГУИР)*, *7*(101), 79–83. (In Russ.)

Khachaturova, S. S. (2016). A digital signature is an authentication of the document. *Science, Technology and Education*, *9*(27). (In Russ.)

Nafikov, M. A. (2016). Algorithms for segmentation of blood vessels in the retina. *Applied Informatics*, *3*(63), 39–45. (In Russ.)

Parker, T. S. & Zhua, L. O. (1987). Introduction to the theory of chaotic systems for engineers. *WIEER (ТИИЭР)*, *75*, 8, 6–40. (In Russ.)

Thompson, J. M. T. (1985). *Instabilities and catastrophes in science and technology*. Moscow: Mir. (In Russ.)