METHODS OF LONG-TERM AND SHORT-TERM ADAPTATION FOR BIOMETRIC IMAGE OF KEYSTROKE DYNAMICS

INTRODUCTION

One of the most pressing problems with biometric identification systems is the change in the user’s image over time. Such changes can occur both over a gradually long period of time due to natural behavioral or physiological changes in the user, and abrupt changes over a short period of time, which, as a rule, are associated with external factors (CONTINUOUS IDENTIFICATION OF THE USER BY KEYBOARD HANDWRITING USING A STATE CONTEXT-BASED VIEW, 2020).

With regard to keyboard handwriting, changes over time can be associated with several factors:

1. continuous formation of keyboard handwriting as you gain experience with the keyboard;
2. stabilization of keyboard handwriting after switching to the other device;
3. various psychophysical conditions of the user during the day that affect the mechanics of typing on the keyboard, such as stress or fatigue. Abrupt changes in the biometric image of the user’s keyboard handwriting can be caused by various external factors, among which are:

   1. injuries that interfere with standard text input;
   2. keyboard defects;
   3. serious deviations of the psychophysical state of the user.

As a rule, large deviations of the biometric image arising for the indicated reasons will cause access denial as an unregistered user. To continue the work of such a user in the system, additional measures will be required, such as the suspension of the employee from work, the intervention of the system administrator, or the use of alternative means of authentication / identification. This issue is beyond the scope of this work (BHATIA, HANMANDLU, 2017).

Tracking gradual changes in the keyboard handwriting and the subsequent self-learning of the user identification system based on the keyboard handwriting is one of the main tasks of this research work.

Depending on the nature of gradual changes in the keyboard handwriting, several groups of time intervals can be distinguished, during which deviations of the biometric image will be monitored (SHEPHERD, 1995):

1. short-term period: from several hours to a day;
2. long-term period: several days or longer.

Comparison of images for a short period allows you to track minor deviations of the psychophysical state. In the process of further self-learning of the system, it is possible to identify such user states. In addition, in this range, it is also permissible to track the stabilization of the keyboard handwriting of an experienced user, caused by a change of device, with the subsequent adaptation of the keyboard handwriting;

Tracking changes over a long period allows for gradual retraining of keyboard handwriting...
modules as the keyboard handwriting gradually transforms due to the accumulation of typing experience on the keyboard.

Based on the foregoing, it is necessary to develop self-learning methods specific to each of the proposed time intervals. These methods should, on the one hand, expand the formed biometric image under conditions of short-term changes in the keyboard handwriting, and on the other hand, transform the formed image under the conditions of the formation or transformation of the keyboard handwriting over time.

**SHORT-TERM ANALYSIS OF CHANGES IN KEYBOARD HANDWRITING**

Numerous studies show that performance is not a stable characteristic. It changes in the process of labor in certain phases. Ultimately, its dynamics is determined by the dynamics of changes in the reflex activity of the human nervous system (DOWLAND; FURNELL, 2004).

The first phase is characterized by increasing efficiency, "accumulation of working potentials". Here, functional restructuring and the establishment of a dynamic stereotype are carried out. In the pre-working period, a person either rests or is engaged in any activity (household, sports, etc.). In both cases, the relationship between mental (and physiological) processes, as well as the characteristics of each of them, differ from those required for work. The initial period of work is characterized by the well-known "initial mismatch" between the new requirements for a person and the present state of his functions. All other things being equal, the magnitude of this mismatch determines the duration of entry into work (the period of operation). The speed, and sometimes the accuracy of human actions in the first phase, is low (GUNETTI, PICARDI, 2005).

The second phase - relative stable performance = “is the period when the establishment of the stereotype is completed, and the body’s activity acquires harmonious unity and integrity, provided by the stereotype without significant losses in the speed and accuracy of reproduction" of actions. This phase is characterized by the "attunement of the rhythms and rates" of the activity of "individual parts of the nervous system." The duration of this phase depends on the nature of the work, as well as on the level of training and condition of the employee.

The third phase - a drop in working capacity - is due to quenching. According to modern concepts, "fatigue is not a direct result of wasting potentials ..., but an expression of a change in the functional state of the central nervous system. It is a natural reaction to work. In the process of fatigue, the harmonious unity of nervous activity is disturbed, the dynamics and interrelation of the main nervous processes change. Changes in braking, which become unstable, vague and superficial, are especially significant. With fatigue, changes in the bioelectric activity of the brain are observed: a decrease in the alpha rhythm due to an increase in the beta rhythm, a decrease in the potential and the recovery period of the alpha rhythm. This indicates the formation of regional stagnant foci in the cerebral cortex (DANIELE, CLAUDIA, 2005).

The transition from the second phase to the third is characterized by an increase in the variability of actions (primarily in terms of the time of their implementation). If, by the nature of labor operations, a person performs stereotypical, regularly repeated actions, then there is a violation of regularity. At the same time, the overall performance may remain at the level of the second phase for some time.

The latent period of reactions in the phase of fatigue increases significantly. Decreases the accuracy of movements. Complicated skills are particularly affected by fatigue. As F. Bartlett showed, in the process of fatigue, the characteristics of individual movements may not change, but their coordination in time is upset (FURNELL et al., 2001).

**ALGORITHM FOR DETERMINING THE PHASE OF PERFORMANCE ACCORDING TO THE CHARACTERISTICS OF THE KEYBOARD HANDWRITING**

In the problem of user identification by keyboard handwriting, it is proposed to analyze the dynamics of the user’s typing speed as an indicator of the phase of his working capacity and to expand the biometric image of the keyboard handwriting taking into account this feature. This approach will allow not only making an assumption about the current phase, but also taking it into account in the process of dynamic user identification in order to improve the recognition...
Methods of long-term and short-term adaptation for biometric image of keystroke dynamics

Typically, print speed is calculated as the number of characters per unit of time, most often per minute. However, in the task of continuous analysis of keyboard handwriting, the traditional method of calculating the typing speed is not suitable. In the process of working on a personal computer, the periods of keyboard activity can vary greatly, which makes it extremely difficult to accurately calculate the current typing speed by the number of characters in a fixed period of time. Based on this, it is proposed to conduct the concept of instantaneous print speed (DMITRY et al., 220).

Suppose that during the assessment of the typing speed, the subject makes N clicks in the allotted time \( \Delta t \) at intervals \( [FT_{2,1}, FT_{3,2} \ldots FT_{N-1,N}] \), which are recorded by the system during the test. In this case, the print speed for the interval \( \Delta t \) will be calculated as \( v = \frac{N}{\Delta t} \)

Suppose that \( \sum_{i=2}^{N} FT_{i-1,i} = \Delta t \), then the printing speed can be represented as:

\[
v = \frac{N}{\sum_{i=2}^{N} FT_{i-1,i}}
\]

Unlike the traditional method of calculating print speed, this formula does not depend on a fixed amount of time during which the evaluation takes place. By varying the number of characters, it is possible to estimate the current typing speed, all depending on the periods of user activity and the duration of continuous work at the keyboard. Let’s designate this feature as instantaneous print speed. When \( N = 1 \), this indicator is reduced to a sequence of values of the intervals between clicks. In this case, the instantaneous speed indicator will be too noisy and uninformative, as shown in the figure 1.

**Figure 1.** Instantaneous speed indicator

Source: Search data.

To estimate the instantaneous speed, it is recommended to use several successive keystrokes, the interval between which does not exceed a certain threshold, which is also a filter when forming state contexts. It was experimentally obtained that at least 15 consecutive clicks are suitable for assessing the speed. This approach allows you to significantly smooth the values of the indicator, which makes it suitable for further analysis as shown in figure 2.
Figure 2. Smoothed instantaneous speed indicator

Source: Search data.

Since the indicators of the change in the instantaneous speed over time are a time series, the analysis methods used in this field, such as the estimation of frequency and trend, can be applied to it.

Let’s assume that typing speed is directly related to uptime and fatigue. We propose the following model of the dependence of speed indicators on the current phase of performance.

Let’s divide the user's typing speed into three fuzzy sets: high, medium and slow. For each user, the parameters of these sets will be calculated individually in the process of building a keyboard handwriting model.

Since the second phase is a relatively stable phase, during this period, the print speed will be closest to the user’s normal print speed and will be relatively stable throughout the phase. The speed indicator will be high or medium, and the graph of the instantaneous speed will not have a trend, as shown in the figure 3. (The trend is currently calculated using the least squares method for a linear or quadratic function).
In the phase of increasing efficiency, in which the user is in the process of switching from another type of activity, the printing speed will be medium or slow, the graph of the instantaneous speed will have a positive trend.

In the phase of decline in efficiency under the influence of fatigue, the printing speed will also be medium or low, the graph of the instantaneous speed will have a negative trend.

Thus, in order to improve the accuracy of user recognition, taking into account short-term daily fluctuations in the time parameters of keystrokes, it is proposed to highlight the current phase of performance by analyzing the instantaneous typing speed and take this factor into account as an additional attribute of the vector of features of the biometric image of the keystroke dynamics (PAPADAKI et al., 2002).

METHODS OF ADAPTING THE BIOMETRIC IMAGE FOR A LONG-TERM PERIOD

In the problem of dynamic user identification by keyboard handwriting, it is proposed to single out the described phases of performance. At the stage of constructing a biometric image of the keyboard handwriting, this will expand the biometric image, and at the identification stage, take this parameter into account in order to increase the identification accuracy by reducing the probability of a false refusal error. (Even if the user is tired and typing slower, we recognize him).

As mentioned earlier, in addition to short-term fluctuations in the characteristics of the keyboard handwriting depending on the phases of working capacity, there may be changes in the keyboard handwriting over a long period, which are not tracked in the process of short-term analysis and are permanent. These changes may be related to, for example, gaining experience with the keyboard or mastering the blind ten-finger dialing (STEVEN et al., 2004).

Obviously, due to the gradual change in the keyboard handwriting, the resulting image must be modified, and the user recognition model must be retrained.

There are two basic approaches to solving the problem of adapting an image to gradual changes: the growing window method and the sliding window method.

The growing window method assumes the accumulation of samples and additional training of the classifier based on a new training sample. Thus, training is carried out on a growing base of input samples, which, on the one hand, leads to the formation of a more generalized biometric image.
image, but on the other hand, leads to a gradual increase in the biometric image of the keyboard handwriting itself and an increase in the system training time (BOURS; MONDAL, 2020).

The sliding window method assumes that there is a fixed amount of data in the sample database. As new samples are added to the training set, the oldest samples are removed from it, and the analysis system is retrained. Thus, there is a fixed window that contains a sample of data entry, which correspond to the most current state of the biometric image of the keyboard handwriting.

However, both of these approaches, since they involve regular retraining, are extremely resource-intensive, especially in a highly loaded system. Therefore, a more flexible and much less resource intensive approach is proposed (PASHCHENKO et al., 2018).

To solve the problem of adapting an image to gradual changes, it is proposed to perform a cluster analysis of the registered contexts of states and to analyze the displacement of their centroids.

If during the analysis the user is identified as his own, but the centroid of a certain context of states has shifted further than a certain limit distance, the model for this context is retrained.

To increase the speed of the retraining process, it is also proposed to reduce the number of monitored state contexts according to one or several criteria:

1. Stable contexts of states: contexts with the least variance.
2. Common state contexts.

The advantage of this approach is, on the one hand, that the model is rebuilt only if the centroid value goes beyond a certain interval, and not after each successful entry. In addition, the model is retrained only for a particular context, and not for the entire biometric image. Taken together, these benefits help reduce the number of image modifications over an extended period of time and keep image modifications to a minimum.

RESULTS

In the course of experimental verification of the described approaches, the following conclusions were obtained:

- During dynamic analysis of the user’s keyboard handwriting over a long period of time, reduces the probability of false positives by up to 15%. Thus, the system is much less likely to perceive the fatigue recruitment phase of an authorized user as a substitution of an operator.
- Without the use of adaptation mechanisms, recognition accuracy.
- Using the growing window as a biometric image adaptation can reduce the false-value error rate to 5%. However, over time, the time for retraining the model increases, which can be a critical factor in highly loaded systems. How old data entry patterns become new up-to-date biometric images of the user.
- The sliding window also reduces the share of errors of both the first and the second kind by equal, comparable to the growing window, but compared to the previous approach, the model retraining did not increase and the current window of samples in the training set always corresponds to the biometric image of the user. However, the performance problem associated with regular updating of the biometric image and retraining of the model is still relevant.
- Periodic retraining of the model based on the movement of centroids to maintain the identification accuracy that was achieved using the sliding window method and at the same time reduce the resource costs associated with updating the biometric image of the keyboard handwriting and retraining the model by up to 80%.
CONCLUSIONS
The necessity of development of algorithms for adaptation and modification of the biometric image of the keyboard handwriting is substantiated in connection with changes in the parameters of the keyboard handwriting under the influence of internal and external factors.

ACKNOWLEDGMENTS
The reported study was funded by RFBR, project number 19-31-27001

REFERENCES
Pashchenko, D.V. et al. Imitation model for keystroke dynamics base on state contexts representation. IOP Conf. Ser.: Mater. Sci. Eng. 2020, 994 012001. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/994/1/012001. Access: August 10, 2021.

BHATIA, A.; HANMANDLU, M. Keystroke Dynamics Based Authentication Using Information Sets. Journal of Modern Physics, 2017, 08, p. 1557-1583.

BOURS, P.; MONDAL, S. Continuous Authentication with Keystroke Dynamics. IEEE Access, vol. 8, p. 156177-156189, 2020,.

GUNETTI, D.; PICARDI, C. Keystroke analysis of free text. ACM Trans. Inf. Syst. Secur., 2005, 8(3), P. 312-347.

DOWLAND, P.S.; FURNELL, S.M. A long-term trial of keystroke profiling using digraph, trigraph and keyword latencies. In Security and Protection in Information Processing Systems, 147 of IFIP The International Federation for Information Processing, Springer US, 2004, p. 275-289.

FURNELL, S.M.; DOWLAND, P.S.; SINGH, H. A preliminary investigation of user authentication using continuous keystroke analysis. In Proceedings of the IFIP 8th Annual Working Conference on Information Security Management & Small Systems Security, 2001, September.

GUNETTI, D.; PICARDI, C. Keystroke analysis of free text. ACM Transactions on Information and System Security, 2005, 8(3), p. 312-347.

PAPADAKI, M.; DOWLAND, P.S.; FURNELL, S.M. Keystroke analysis as a method of advanced user authentication and response. In Security in the Information Society: Visions and Perspectives, IFIP TC11 17th International Conference on Information Security (SEC2002), 2002, p. 215-226.

PASHCHENKO, D.V.; BALZANNIKOVA, E.A.; SERGINA, I.G. User identification method by means of biometric image of keystroke dynamics with double-chained representation” Issues of radioelectronics. Issues of radioelectronics, 2018, 12, p. 83-89. DOI 10.21778/2218-5453-2018-12-83-89.

SHEPHERD, S.J. Continuous authentication by analysis of keyboard typing characteristics. In European Convention on Security and Detection, 1995, p. 111-114.

STEVEN, M.; FURNELL PAUL, S.; DOWLAND. A long-term trial of keystroke profiling using digraph, trigraph and keyword latencies. In Security and Protection in Information Processing Systems, IFIP 18th WorldComputer Congress, TC11 19th International Information Security Conference, 2004, p. 275-290.

CONTINUOUS IDENTIFICATION OF THE USER BY KEYBOARD HANDWRITING USING A STATE CONTEXT-BASED VIEW. XXI CENTURY: results of the past and problems of the present plus, Vol. 9, 2020, Number 3 (51), C, p. 74-79.
Methods of long-term and short-term adaptation for biometric image of keystroke dynamics

Méthodes de adaptation de longo e de curto prazo para imagem biométrica da dinâmica do teclado

Métodos de adaptación a largo y corto plazo para la imagen biométrica de la dinámica del teclado

Resumo
Este artigo discute um dos principais problemas de identificação do usuário pela escrita à mão no teclado - mudanças de curto prazo na dinâmica de pressionamento de tecla dos usuários em relação ao seu estado psicofísico, bem como mudanças ao longo do tempo associadas à formação da dinâmica de pressionamento de tecla por um novo usuário ou ao trocar para um novo dispositivo. É proposto um método para determinar a fase da capacidade de trabalho pelas características temporais da dinâmica de pressionamento de tecla.

Palavras-chave: Dinâmica de teclas. Biometria. Biometria comportamental. Identificação dinâmica. Máquina de aprendizagem.

Abstract
This article discusses one of the main problems of user identification by keyboard handwriting - short-term changes in the keystroke dynamics of users in connection with its psychophysical state, as well as changes over a long time associated with the formation of keystroke dynamics by a new user or when switching to a new device. A method for determining the phase of working capacity by the time characteristics of the keystroke dynamics is proposed.

Keywords: Keystroke dynamics. Biometry. Behavioral biometry. Dynamics identification. Machine learning.

Resumen
Este artículo analiza uno de los principales problemas de la identificación del usuario mediante la escritura a mano del teclado: los cambios a corto plazo en la dinámica de pulsación de teclas de los usuarios en relación con su estado psicofísico, así como los cambios durante mucho tiempo asociados con la formación de la dinámica de pulsaciones de teclas por un nuevo usuario o al cambiar a un nuevo dispositivo. Se propone un método para determinar la fase de la capacidad de trabajo por las características temporales de la dinámica de pulsaciones de teclas.

Palabras-clave: Dinámica de pulsaciones de teclas. Biometría. Biometría conductual. Identificación dinámica. Máquina de aprendizaje.