UDC 004.855.5(045)
DOI:10.18372/1990-5548.76.17665

[1]O. I. Chumachenko,
[2]A. O. Burkov

DEVELOPMENT OF AN EFFICIENT UKRANIAN KEYBOARD LAYOUT USING A GENETIC ALGORITH

Department of Artificial Intelligence, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mails: [1]eliranvik@gmail.com ORCID 0000-0003-3006-7460, [2]kep.buran@gmail.com.

Abstract*—The aim of this work is to create a more efficient Ukrainian keyboard layout, as the current most popular layout "ЙЦУКЕН" is more adapted to the Russian language than Ukrainian and is therefore not efficient. A mathematical model of a genetic algorithm was developed in this study to achieve this goal. The primary measure of layout efficiency is the total distance traveled by the fingers during typing. Additionally, the study describes the most common typing methods and variables that determine the keyboard's dimensions. This allows for the adaptation of the algorithm to create a keyboard that meets the needs of each individual user. To assess the efficiency of the newly developed Ukrainian keyboard layout, created using a genetic algorithm, experiments and comparisons with existing layouts are conducted. The results of the experiments demonstrate the advantages of the developed keyboard layout for the Ukrainian language.*

Index Terms—Genetic algorithm; keyboard layout; Ukrainian keyboard layout; optimization.

I. INTRODUCTION

The keyboard is an integral part of modern life, as it is one of the primary input devices for computers. Every day, millions of Ukrainian-speaking people worldwide use the keyboard for work, education, and communication. As the number of Ukrainian computer users increases, the importance of productivity and comfort while working with the keyboard for inputting documents, messages, and other textual data in the Ukrainian language grows, along with the need for an efficient Ukrainian keyboard layout.

The relevance of creating a more efficient Ukrainian keyboard layout arises from the necessity of adapting the keyboard to the input of Ukrainian text. Unfortunately, the existing Ukrainian keyboard layout was not specifically designed for working with the Ukrainian language but is merely an adaptation of the most popular Russian keyboard layout. This adaptation reduces its efficiency and usability. Therefore, the goal of this research is to create a more efficient Ukrainian keyboard layout tailored to the requirements of the modern Ukrainian language.

To achieve this objective, this research employs a genetic algorithm, which is a powerful tool for finding optimal solutions to complex problems. Applying a genetic algorithm to create a more efficient Ukrainian keyboard layout is an approach that aims to explore all possible layout variations and identify the one that minimizes finger movement when typing in the Ukrainian language, thereby enhancing comfort and typing speed.

II. EXISTING UKRAINIAN KEYBOARD LAYOUTS

Tracking the history of keyboard layout origins with Cyrillic characters is quite challenging, but the modern and most popular Cyrillic layout is "ЙЦУКЕН" (JCUKEN). It was approved by the Soviet Union's GOST 6431-75 in 1975 for typewriter keyboards [1]. Since this document was applicable to all Soviet republics, "ЙЦУКЕН" quickly spread to Ukraine as well. However, in its original form, this layout includes all the letters of the Russian alphabet. Nowadays, there is a "ЙЦУКЕН" layout for each language that uses the Cyrillic alphabet, including Belarusian, Kazakh, Tajik, and others [2] (Fig. 1).

Of course, "ЙЦУКЕН" also received its adaptation for the Ukrainian language. In 1996, the Ukrainian version of the layout was approved in the DSTU 3470-96 standard [3]. This version is currently used on most devices by Ukrainian-speaking users (Fig. 2).

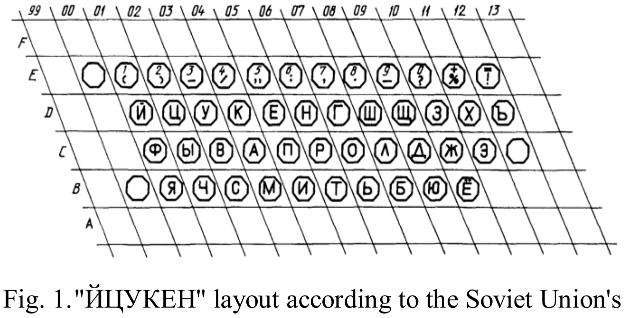

Fig. 1."ЙЦУКЕН" layout according to the Soviet Union's GOST 6431-75 [1]

©National Aviation University, 2023

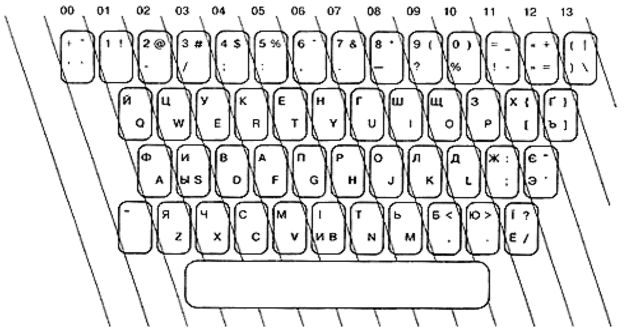

Fig. 2. Ukrainian "ЙЦУКЕН" keyboard layout [3]

Additionally, there are alternative Ukrainian keyboard layouts, but most of them, such as the "ol03a" and "ol03g" layouts [4], are modified versions of "ЙЦУКЕН" that include functional keys that are not relevant to this work. There is also the "Шарапівка" (Sharapivka) Ukrainian keyboard layout, which arranges the most commonly used letters in the Ukrainian language in the middle row to optimize typing using touch typing techniques [5] (Fig. 3).

Fig. 3. "Шарапівка" (Sharapivka) keyboard layout [5]

The alternative "Шарапівка" layout considers the letter frequency in the Ukrainian language but does not take into account the most common letter combinations in words. It is specifically designed for touch typing. Therefore, creating a more efficient Ukrainian keyboard layout for different typing methods remains a relevant task.

III. TYPING METHODS

The efficiency and speed of typing on a keyboard are influenced not only by the key layout but also by the finger placement on the keyboard and finger movement during typing. Of course, the keyboard itself also affects the used technique because while a physical keyboard allows for the use of all ten fingers, it is usually not possible on devices with virtual touch-sensitive keyboards like smartphones since users also use their fingers to hold the device. In this work, we will consider the most popular typing methods for physical and touch-sensitive keyboards.

The most popular and one of the most efficient typing methods for physical keyboards is touch typing or the ten-finger method. This method gained popularity after Frank Edward McGurrin's victory on July 25, 1888, in a typing speed competition where he used touch typing [6]. The idea of this method is to use all ten fingers while typing, with the index, middle, ring, and little fingers of both hands resting on the middle row of the keyboard, and the thumbs on the spacebar key. Figure 4 shows the finger placement scheme on the alphanumeric keys of the keyboard when using touch typing. The initial finger positions highlighted in the figure are used to start typing. When typing with any finger, it immediately returns to its initial position if it does not correspond to the next letter in the word.

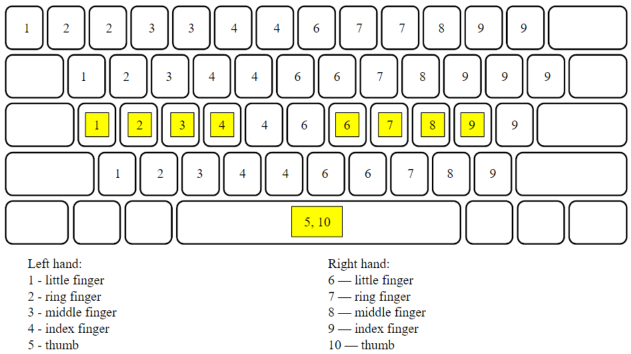

Fig. 4. Touch typing method scheme

However, touch typing is not optimal for touch-sensitive devices since the total width of the fingers is often larger than the width of smartphone screens. Additionally, people usually type on smartphones while holding them in their hands, so their fingers are also engaged in holding the device and cannot be solely used for typing on the keyboard. Therefore, for such cases, there is a method of typing on a touch-sensitive keyboard using two thumbs.

To formalize this method of typing using two thumbs, the touch-sensitive keyboard is divided vertically into two parts, with each part corresponding to one thumb. Figure 5 illustrates the finger placement scheme when typing on a touch-sensitive keyboard using two thumbs.

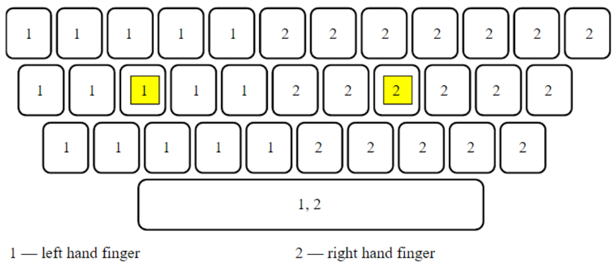

Fig. 5. Two fingers typing method

Additionally, a single finger is used for data input on touch-sensitive devices. In this case, all keyboard keys are associated with that finger, and the starting position is the key located in the middle of the middle row of the keyboard. Figure 6 shows the finger placement scheme when typing on a touch-sensitive keyboard using a single finger.

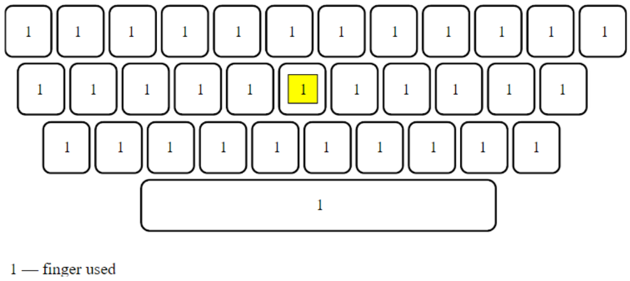

Fig. 6. One-finger typing method

In conclusion, there are three typing methods: one for physical keyboards – touch typing, and two methods for touch-sensitive keyboards – typing with two thumbs and typing with a single finger.

IV. KEYBOARD EFFICIENCY METRICS

To measure the efficiency of a keyboard layout, we will calculate the total distance that a finger travels while typing. To measure the total distance, we will use a text dataset. In this study, we will take journalistic articles in the Ukrainian language on various topics such as politics, sports, culture, etc., as the text dataset. The text dataset is necessary for sequentially traversing each character contained in the text. In each iteration, when a specific character is considered, the distance that the finger must travel from its current position to the key containing that character in the given layout is calculated and added to the sum of distances traveled so far. The finger is selected using a typing method that specifies the coordinates of the keys it can press for each individual finger used in the method.

The initial finger positions are specified by the typing method. The fingers automatically move to their initial positions if a finger was not used in an iteration and the distance it travels while returning is not considered. This is necessary for effectively modeling real typing on a keyboard since returning to the initial position is a secondary action that does not affect the sequential input of text during typing. The finger used in an iteration retains its position on the key that was used.

The distance from a finger to a key is calculated using the specified key width and length, the vertical and horizontal distances between keys, and the row offsets of the keyboard layout from the far-left point. Figure 7 illustrates a scheme depicting all these keyboard metrics.

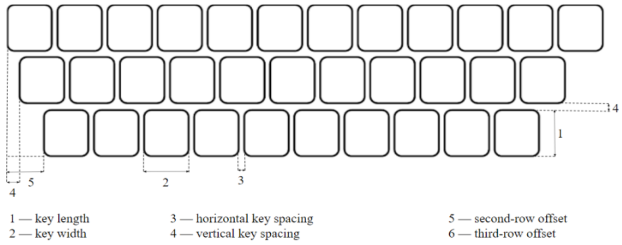

Fig. 7. Keyboard metrics scheme

Taking into account the width and length of keys, as well as the distances between keys, we can calculate the distance between the current position of the finger and the target key, considering that the finger presses the center of the key. Adding the row offset or the distance from the far-left point of the layout allows us to consider the position of the key within the context of the keyboard layout. The distance function between two keys relies on the Pythagorean theorem and can be expressed as:

$$f\left(a_{i,j}, a_{m,n}\right) = \left(|x_i - x_m| + |i - m| \cdot |d_h - w|\right)^2 + \left((j - n)\cdot(d_v - l)\right)^2, \quad (1)$$

where $a_{i,j}$ is the key in row i and column j; $b_{i,j}$ is the key in row m and column n; x_i, x_m is the row offsets of rows i and m; d_h, d_v is the horizontal and vertical distances between keys; w, l is the width and length of the keys.

The total sum of distances is the sum of distances for typing each individual character in the text dataset. Based on the computed total distances for each layout in the population, the layout efficiency value is calculated. It is defined as the number of characters in the text dataset divided by the total distance and by the average distance of the average key on the keyboard to its vertical and horizontal neighbors. By dividing by the average distance, it is possible to normalize the efficiency evaluation and obtain an objective assessment of the layout that minimizes the keyboard size factor.

V. GENETIC ALGORITHM AND RESULTS

To define the genetic algorithm, let's describe its main stages: initialization, selection, crossover, and mutation. The idea of using a genetic algorithm to create an efficient keyboard layout is not new. There are many articles where an efficient layout for an English keyboard was created using a genetic algorithm. Therefore, the crossover and mutation operators were taken from the article [7]. To initialize the algorithm, we will create a population of randomly generated Ukrainian keyboard layouts. However, for the initial positions of the selected typing method, we will place one of the most frequently to used letters in the Ukrainian language. According to the frequency analysis in the article [8], these letters are "о", "а", "н", "и", "і", "в", "т", "е", "р", "с", "л". For selection, the top quarter of individuals based on their efficiency evaluation, as described in the previous section, will be chosen.

The algorithm generated keyboard layouts for each typing method, which are depicted in Fig. 8, and results – Fig. 9. In this case only the positions of letters were changed to not change the position of symbols, which are the same for every language, like ".". Also letter "ґ" is placed in the same place as "г" like at "ЙЦУКЕН" layout because of the rare usage of it. Every method is getting optimized in different amount of population, the more fingers are used, the fewer generations is needed to optimize the layout for this method. As it is shown on histograms at the Fig. 10, generated layout is at least 25% more efficient than others.

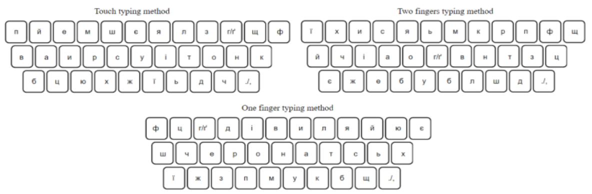

Fig. 8. Algorithm generated keyboard layouts

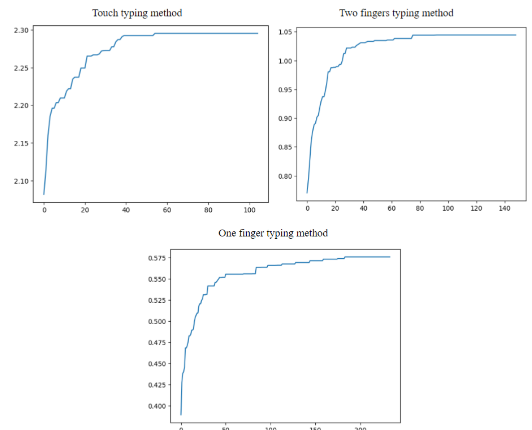

Fig. 9. Algorithm generation best

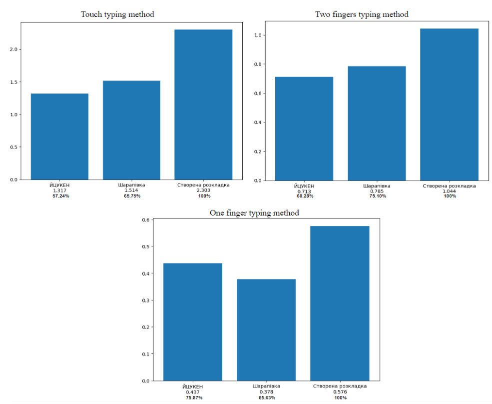

Fig. 10. Keyboard layout comparison histogram

Touch typing was optimized the best, outperforming the next best layout by 34.25%. Other methods that utilize fewer fingers are more challenging to optimize and tend to be less effective, so the general efficiency of layout for touch typing is bigger for all layouts.

VI. CONCLUSION

As a result of this work, a genetic algorithm for creating an efficient Ukrainian keyboard layout has been successfully implemented. Factors such as typing method and keyboard sizes were taken into account, allowing users to create a keyboard layout according to their own needs. Additionally, users have the option to replace the text dataset, for example, inputting their own texts that they type daily. As a result, users obtain the desired keyboard layout. As demonstrated, the algorithm generates a keyboard layout that is at least 25% more efficient than other keyboard layouts.

REFERENCES

[1] "Typewriters. Arrangement of Keys and Characters on the Keyboard," GOST USSR 6431-75, 1975.

[2] "Windows Keyboard Layouts," IEEE. 20 March 2023. Microsoft. Available: https://learn.microsoft.com/en-us/globalization/windows-keyboard-layouts

[3] "DSTU 3470-96," 1996, State Committee for Technical Regulation and Consumer Policy of Ukraine.

[4] Oleksandr Liutyi, "ol03a and ol03g." February 27, 2017, Wayback Machine. Available: http://web.archive.org/web/20170227000000/http://liutyi.info/ol03a

[5] O. Sharapov,"Sharapovka Keyboard Layout," April 10, 2016, Wayback Machine. Available: https://web.archive.org/web/20160401093256/http://layout.asharapov.com/about

[6] H. Yamada, "A Historical Study of Typewriters and Typing Methods: from the Position of Planning Japanese Parallels," *Journal of Information Processing*, 2(4), 175–202, 1980.

[7] K. Nivasch, & A. Azaria, (Year unknown). A Deep Genetic Method for Keyboard Layout Optimization. Computer Science Department, Ariel University, Ariel, Israel. Available: http://azariaa.com/Content/Publications/Keyboard.pdf

[8] O. O. Arkhipov, & V. M. Zhuravlyov, "Frequency analysis of the usage of letters in the Ukrainian language," *Radioelectronics. Informatics. Control*, 2009. ISSN 1607-3274

Received March 20, 2023

Chumachenko Olena. ORCID 0000-0003-3006-7460.
Doctor of Engineering Science. Professor. Head of the Department.
Department of Artificial Intelligence, Faculty of Informatics and Computer Science, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.
Research area: artificial neural networks, artificial intelligence.
Publications: more than 120 papers.
E-mail: eliranvik@gmail.com

Burkov Anton. Bachelor.
Department of Artificial Intelligence, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» Kyiv, Ukraine
Education: National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» Kyiv, Ukraine, (2023).
Research area: artificial intelligence.
Publications: 1.
E-mail: kep.buran@gmail.com

О. І. Чумаченко, А. О. Бурков. Створення ефективної україномовної розкладки клавіатури за допомогою штучного інтелекту

Метою цієї роботи є створення більш ефективної україномовної розкладки клавіатури, оскільки сучасна найпопулярніша розкладка "ЙЦУКЕН" є більш адаптованою під російську мову, ніж українську, тому не є ефективною. В роботі була розроблена математична модель генетичного алгоритму для досягнення цієї мети. Основною оцінкою ефективності розкладки є загальна дистанція, яку проходять пальці під час друку. Крім того, у роботі описано найпоширеніші методи друку та змінні, які визначають розмірність клавіатури. Це дозволяє адаптувати алгоритм для створення клавіатури, що задовольняє потреби кожного окремого користувача. Для оцінювання ефективності нової україномовної розкладки клавіатури, розробленої з використанням генетичного алгоритму, проводяться експерименти та порівняння з існуючими розкладками. Результати експериментів демонструють переваги розробленої розкладки клавіатури для української мови.

Ключові слова: генетичний алгоритм; розкладка клавіатура; україномовна розкладка клавіатури; оптимізація..

Чумаченко Олена Іллівна. ORCID 0000-0003-3006-7460. Доктор технічних наук. Професор. Завідувачка кафедри.
Кафедра штучного інтелекту, Факультет інформатики та обчислювальної техніки, Національний технічний університет Украіні «Київський Політехнічний Інститут імені Ігоря Сікорського», Київ, Україна.
Кількість публікацій: більше 120 наукових робіт.
Напрям наукової діяльності: штучні нейронні мережі, штучний інтелект.
E-mail: eliranvik@gmail.com

Бурков Антон Олексійович. Бакалавр.
Кафедра штучного інтелекту, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.
Освіта: Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, (2023).
Напрям досліджень: штучний інтелект
Публікації: 1.
E-mail: kep.buran@gmail.com