International Journal of Research Publication and Reviews, Vol (6), Issue (1), January (2025), Page – 4924-4927

WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An AI based keystroke dynamics for user authentication

N.Tharun[a], N.Sangeetha[b]

[a]First affiliation, Address, City and Postcode, Country
[b]Second affiliation, Address, City and Postcode, Country

A B S T R A C T

Keystroke dynamics, a form of behavioural biometrics, refers to the unique patterns in an individual's typing rhythm, speed, and cadence that can be used to authenticate users or detect anomalies in their behaviour. This project explores the application of keystroke dynamics as a robust method for user identification and security enhancement in digital systems. By analysing the timing and pressure variations between key presses, the project aims to develop a model that can effectively distinguish between different users, even in the absence of traditional biometric identifiers. Through the collection of keystroke data from various participants and the application of machine learning algorithms, the system will be trained to recognize individual typing patterns with a high degree of accuracy. The project investigates the potential of keystroke dynamics in securing online accounts, improving authentication protocols, and detecting fraudulent activities such as identity theft or session hijacking. Additionally, the challenges of accuracy, usability, and privacy concerns will be examined, and potential solutions to address these issues will be proposed. Ultimately, this work aims to contribute to the advancement of biometrics by showcasing the practicality and reliability of keystroke dynamics in real-world applications.

Keywords: Behavioral Biometrics, User Authentication, User Identification, Typing Patterns, Keystroke Analysis, Authentication Protocols

1. Introduction

Our reliance on computers and digital platforms has significantly increased, making our lives more convenient. The integration of automated information systems has led to enhanced performance of networking services, particularly in terms of reliability and reduced computational costs. These advancements have sparked a global interest in accessing online platforms. However, this growing reliance on technology has also brought about an increase in security threats to computer systems. As a result, safeguarding systems from attacks and fraud has become a primary focus for researchers. Consequently, the development of robust measures to prevent unauthorized access is a critical area of research. One effective method for granting access to legitimate users involves detecting their unique behavioral patterns, such as their typing rhythm. Keystroke Dynamics, which analyzes an individual's typing habits, has proven to be a reliable security feature. It is observed that a person's typing style—including finger placement, key pressure, and typing speed—tends to remain consistent. This distinctive pattern can be leveraged to distinguish between legitimate users and intruders, as individuals will naturally type in their usual rhythm. Thus, tracking typing rhythms can provide an efficient way to monitor and secure computer activities. In real time, the classifications of object and knowledge based are merged to fulfil the process of authentications such bank passwords and their PINS. One major disadvantage with regards to this classification-based authentication is the ability to memorize and manage multiple such PINS and recalling them. Therefore, the usage of biometrics authentication is preferred as it overcomes these issues and makes use of automated methods to identify and verify the individual. Also, this form of authentication is gaining worldwide popularity as they provide an extra level of security. In today's world, the Internet serves as the primary platform for communication, facilitating the exchange of vast amounts of sensitive data between computer systems. This makes identity verification more crucial than ever. However, traditional username and password combinations, which have remained largely unchanged for the past decade, are increasingly vulnerable. Attacks such as man-in-the-middle, phishing, and social engineering can easily lead to identity theft, allowing unauthorized access to private information, exploitation of trust relationships, and other criminal activities. M Moreover, with the growing prevalence of single sign-on services—where one set of credentials can authenticate users across multiple websites—the consequences of identity theft are even more severe. While simple password-based authentication may suffice in many cases, environments such as online commerce and banking demand stronger authentication methods to safeguard sensitive and confidential data. Although "strong authentication" doesn't have a universally accepted definition, it generally refers to methods that provide greater security than traditional approaches. Typically, it is implemented through two-factor authentication, relying on knowledge-based or token-based techniques. While these methods improve security, they still primarily validate passwords, tokens, and keys, which can be stolen, lost, shared, or manipulated, ultimately compromising security. To address the limitations of traditional solutions, biometric systems offer a promising alternative by enabling identity verification based on the individual user. However, biometric authentication often requires specialized hardware (such as fingerprint readers), making deployment challenging. Keystroke dynamics, a form of behavioral biometrics, offers a compelling solution for strong authentication on the web. It has the distinct advantage of not requiring specialized hardware while maintaining the security benefits of more common biometric methods. As we will explore in this work, keystroke

* *Corresponding author.* Tel.: +0-000-000-0000; fax: +0-000-000-0000.
E-mail address: author@institute.xxx

dynamics can be easily integrated with existing password-based authentication systems, providing an additional layer of security without the need for additional hardware.

2. Existing System

Distortion in keystroke dynamics is not a well-researched field, and we did not find many results when doing literature search. However, it measured two typing samples of keystroke dynamics data and used two different measures to compare the samples. However, one method that could be used for detecting distorted timing information is Benford's Law and ZIPF's Law. Benford's Law, or the first-digit law, is an observation in a set of numerical data where the first digit, or leading digit, is more likely to be small. In a balanced distribution of numbers between 1 and 9 there would be exactly 11% for each number to be the leading digit. However, if Benford's Law is obeyed then the change of the leading bits to be small increases. One potential approach to identify distorted or manipulated timing information is through the use of Benford's Law and Zipf's Law. Benford's Law, also referred to as the first-digit law, is an empirical observation in numerical datasets, where the leading digit (the first digit) is more likely to be a smaller number. Specifically, in a dataset with a uniform distribution of numbers between 1 and 9, each digit would appear as the first digit approximately 11% of the time. However, when Benford's Law holds, the occurrence of smaller leading digits (such as 1, 2, and 3) is disproportionately higher than larger digits (such as 8 or 9). This phenomenon is commonly seen in naturally occurring datasets like financial figures, population statistics, and even the frequency of words in languages. The application of Benford's Law to keystroke dynamics timing information could provide valuable insights into whether the data follows expected patterns or exhibits unusual distortions. If timing data adheres to the law, it could indicate that the keystroke dynamics is naturally occurring, while deviations from the expected distribution could point to intentional distortion or errors in the data collection process. In the context of keystroke dynamics, particularly when analysing timing patterns like latency (the time between key presses) and duration (the time a key is held down), our findings suggest that only latency values followed Benford's Law. This indicates that the time intervals between successive key presses in the dataset adhered to a pattern that was consistent with the expected distribution. However, the duration values, which represent the time a key is pressed, did not follow the law's expected distribution. This divergence from the expected pattern could point to several potential issues, including human error, system latency, or manipulation of data.

DISADVANTAGES OF EXISTING SYSTEM

- However, the results showed that only latency values from keystroke dynamics timing information followed the law.
- While duration values did not follow the law.
- Inconsistent Data Patterns
- Resource-Intensive Re-Keying Operations
- Limited Generalization to Different Users
- Vulnerability to Manipulation or Mimicry
- Hardware and Environmental Dependencies

In conclusion, while keystroke dynamics holds promise as a behavioural biometric for enhancing user authentication and security, the distortion of timing information remains a significant challenge. Through techniques like Benford's Law and further investigation into the causes and impacts of timing irregularities, we can develop more accurate, efficient, and secure systems. Additionally, improving system efficiency by addressing the high resource demand during re-keying operations will be essential to making keystroke dynamics a viable solution for a wide range of applications, from online banking to secure enterprise environments.

3. Proposed System

To authenticate a user using keystroke dynamics, the first step is to establish a reference template that accurately reflects the user's unique typing behavior. This template is based on various features of their typing rhythm, such as the speed and timing between key presses (referred to as latency) and how long they press each key (referred to as duration). The accuracy of this template is critical to the success of the authentication process, as it serves as the baseline against which future typing patterns are compared. The process begins with the enrollment phase, where the user is asked to type their password multiple times. During this phase, the system collects data on how the user types, capturing important features such as the time intervals between consecutive key presses (latency) and the duration for which each key is held down. This data is then processed to calculate an average typing rhythm, which serves as the reference template. The idea is that a user's typing behavior remains relatively consistent over time, so this reference template can be used to reliably identify the user in future interactions. There are two main approaches to keystroke dynamics authentication: static and continuous. In static authentication, the user types their password once during login, and the system compares this typing behavior (the probe) against the reference template created during enrollment. In continuous authentication, the system monitors the user's typing behavior throughout their session, continuously verifying that the typing rhythm matches the reference template, thus providing an ongoing check against potential unauthorized access. The key difference between these two methods lies in the scope and frequency of comparison—static authentication typically involves a one-time check, whereas continuous authentication provides a real-time assessment. When a user attempts to authenticate, their typing rhythm is compared to the reference template. This comparison is done by analyzing the features (duration and latency) of their typing, and the system determines how closely the current typing pattern (the probe) matches the reference template. A decision threshold is used to make the final determination. The threshold is established by setting a range that includes the minimum and maximum average typing times (in milliseconds) for the user's typing rhythm. If the user's typing rhythm falls within this range, the system will accept the authentication attempt. If the rhythm deviates beyond the threshold, the system will reject the attempt and flag it as suspicious.

The threshold is crucial in ensuring that the system doesn't reject legitimate users due to small, natural variations in their typing behavior. For example, the speed at which a user types may vary slightly depending on factors such as keyboard type, device, or emotional state. A well-calibrated threshold allows the system to accommodate these small variations while still ensuring security by rejecting attempts from impostors whose typing rhythms significantly differ from the reference. In summary, the authentication process using keystroke dynamics involves creating a reference template based on the user's typing behavior during enrollment. This template is used as the benchmark for future authentication attempts, and the system compares the user's current typing rhythm against the reference to either accept or reject the authentication request. The decision is based on a threshold that defines an acceptable range of variation, ensuring that the system can distinguish between the legitimate user and potential intruders while accounting for minor fluctuations in typing behavior.

ADVANTAGES:

Non-Intrusive Authentication
Continuous Authentication
Difficult to Mimic
Improved Security
No Need for Physical Biometrics
Adaptive to User Behavior
Low Risk of False Positives
Enhanced User Privacy.

OBJECTIVES:

In input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus, the objective of input design is to create an input layout that is easy to follow.
In Output design, a quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

4. Results and Conclusion

In this project, we implemented a keystroke dynamics-based authentication system, analyzing typing patterns such as key press latency and duration to create a reference template for each user. The system was evaluated using a set of enrollment and probe samples to determine its effectiveness in identifying legitimate users and detecting intruders. During the enrollment phase, users typed their passwords multiple times to generate an average typing rhythm, which was stored as their reference template. For authentication, the system compared the probe (current typing rhythm) with the reference template to make an accept or reject decision. The results indicated that the system was able to distinguish between legitimate users and intruders with a reasonable degree of accuracy. The accuracy of the system was evaluated by setting a decision threshold based on the average typing speed and timing variations for each user. The threshold proved to be effective in accommodating small variations in typing behavior while still rejecting mismatched patterns. However, the system did face challenges in cases where the user's typing rhythm varied significantly due to external factors such as stress or keyboard type, resulting in occasional false rejections. In terms of performance, the system was resource-efficient, with minimal computational requirements during the authentication process. While it was found that latency values generally followed expected distributions (such as Benford's Law), some inconsistencies were noted in the duration values, suggesting that further optimization may be necessary to account for typing behavior variations more accurately. In conclusion, the keystroke dynamics-based authentication system developed in this project offers a promising alternative to traditional password-based systems by adding a layer of behavioral biometric security. It successfully captures and analyses the unique typing patterns of individual users, allowing for effective authentication without the need for specialized hardware. The system demonstrated a high level of security by differentiating between legitimate users and intruders, with reasonable accuracy and minimal false positives. Overall, keystroke dynamics proves to be a valuable tool in strengthening authentication mechanisms and could be seamlessly integrated with existing security frameworks to provide an additional layer of protection against unauthorized access. With further refinement, this technology has the potential to offer a secure, non-intrusive, and scalable solution for user authentication across a wide range of applications.

References

[1] J. V. Monaco and C. C. Tappert, "The partially observable hidden Markov model and its application to keystroke dynamics," Pattern Recognition, 2018.

[2] S. Roy, U. Roy, and D. D. Sinha, “Security enhancement of knowledge-based user authentication through keystroke dynamics,” in Proc. MATEC Web Conf., 2016, vol. 57.

[3] K. S. Killourhy. (2012). A scientific understanding of keystroke dynamics. [Online]. Available: http://reports-archive.adm.cs.cmu.edu/anon/2012/CMU-CS-12-100.pd f

[4] A. Messerman, T. Mustafić, S. A. Camtepe, and S. Albayrak, “Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics,” in Proc. Int. Jt. Conf. Biometrics, 2011.

[5] P. S. Teh, B. J. Andrew Teoh, T. S. Ong, and H. F. Neo, “Statistical fusion approach on keystroke dynamics,” in Proc. Int. Conf. Signal Image Technol. Internet Based Syst., January 2007, pp. 918–923.