*Grenze International Journal of Engineering and Technology, June Issue*

# Continuous user Authentication through Keystroke Dynamics

Suvika K V[1], Darshan Kumar M[2], Guntur Lohith Sai[3], Jasin Jayachandran[4] and Kapil Adarsh S A[5]
[1]Assistant Professor, Dept. of CSE, School of Engineering, Dayananda Sagar University
Email: kvsuvika@gmail.com
[2-5]Students, Dept. of CSE, School of Engineering, Dayananda Sagar University
Email: {darshandeepa2002, gunturlohithsai0117, jasinjay2000, kapiladarshsa2003}@gmail.com

***Abstract*— Our research aims to tackle the urgent concern of personal computer security, specifically concentrating on enhancing user authentication. Conventional authentication techniques, such as passwords and PINs, have demonstrated susceptibility to diverse cyber threats, encompassing brute-force attacks and phishing exploits. Additionally, the user experience with these methods often falls short, leading to frustration and potential security vulnerabilities arising from weak or reused passwords. To address these issues and propel the field of personal computer security forward, we advocate for a novel solution: User Authentication Based on Keystroke Dynamics. Keystroke dynamics involves the analysis of an individual's distinct typing patterns, encompassing keypress duration and intervals, to authenticate users. This approach capitalizes on the uniqueness of each person's typing rhythm, rendering it a robust and highly secure authentication method.**

***Index Terms*— STM, Keystroke Data, Continuous User Authentication**

## I. INTRODUCTION

In the ever-evolving landscape of the IT industry, where technological integration has propelled productivity and efficiency to unprecedented heights, there comes a concomitant rise in challenges, notably the pervasive threat of impersonation and data breaches infiltrating business systems.

Our endeavour embarks on a critical mission to confront and mitigate this formidable challenge through the strategic deployment of keystroke analysis and continuous authentication. While technology has ushered in unparalleled advancements, it has also exposed vulnerabilities that demand innovative solutions.

At the heart of our project lies the proactive adoption of continuous authentication, a dynamic paradigm that unfolds seamlessly, mirroring and adapting to the user's behavioural nuances over time. This departure from conventional authentication methods, which rely on discrete moments of validation, reflects our commitment to staying one step ahead of potential security threats.

The cornerstone of our approach lies in the distinctive rhythm and pattern of individual keystrokes—a unique cadence that transcends various tasks and applications. This project meticulously delves into the intricacies of keystroke intervals, typing speed, and even the pressure time applied to the keys. The culmination of this meticulous examination is the generation of a comprehensive and unique biometric profile for each user.

In essence, our project not only identifies the challenges posed by contemporary cybersecurity threats but also presents a cutting-edge solution grounded in the granular analysis of user behaviour. By leveraging continuous

*Grenze ID: 01.GIJET.10.2.702*

authentication and keystroke analysis, we strive to fortify the digital perimeters of business systems, ushering in a Our program is made to work stealthily in the background, monitoring a user's precise typing habits in real time. Through the painstaking tracking of keystroke duration, this novel technology aims to identify distinct patterns inherent in the typing style of the user. The gathered information creates an extensive dataset that depicts a variety of situations and typing habits from a range of individuals. The software extracts crucial components including typing speed, interkey timings, and key press lengths using advanced feature extraction techniques. Developing a sophisticated depiction of the user's typing habits is the main goal. The system goes through a rigorous training procedure using sophisticated machine learning models, such as LSTM (Long Short Term Memory). This gives it the potential to detect pre-fed patterns and adjust to the changing subtleties and peculiarities of each individual user's typing style, opening up new possibilities for applications in behavioural analysis and authentication.

## II. Data Architecture For Authentication

A Keystroke analysis data application system should include a four-tier architecture that includes collecting of data, storing of data, analysing of data, and interchanging and sharing of data, according to research and studies on the usage of data thus far.

A. *Collecting Data and Storing Data:* The purpose of the data gathering procedure is to create a comprehensive dataset from the complex keystrokes that users within our program contribute. Typing activities are captured as users interact with the keyboard and are automatically saved into an Excel sheet. The Data Collection is done by the four features they are press time, release time, hold time and flight time. We used the Python library Keyboard to record the keystroke data, which includes the key name, press time and release time. The hold time and flight time are calculated from the press time and release time. The difference between release time and press time of the same key is equal to Hold time. The difference between press time of the present key and release time of the previous key is equal to Flight time. For every key, there is press time, release time, hold time and flight time. Our technology attempts to build a comprehensive picture of user typing behaviours by carefully recording these temporal dynamics.

B. *Cleaning Data:* The collected data are Hold Time (HT), Press Time (PT), Release Time (RT), and Flight Time (FT) of the user, which are in milliseconds. To improve the accuracy of the model, we increased the decimals of the timings. To get more decimals, we used a Python library called decimal, which allows printing more decimals of the timing. The collected data must be cleaned with assumptions or with averages. Here, the noise data occurs when a key is not released. If you pressed two keys simultaneously, then the key not released condition occurs. The program records the timings of the last released key, and the other key undergoes a key not released condition. In a key not released condition, we have only the press time. To remove the noise data, we need to take the averages of release time, hold time, and flight time. The average of the previous list of hold times, added to the current press time, is equal to the average release time. The average of the previous list of hold times is equal to the average hold time. The average of the previous list of flight times is equal to the average flight time. We are removing the noise data in the execution of the code.

C. *Exchanging and Sharing Data*: Data sharing and exchange should make it possible for a distributed SOA architecture and larger data interchange using document, Web Service, database, and other means in addition to easing data sharing integration and centralizing data collecting, sorting, and dissemination. This supports multiple interface and standard specifications, allows for the integration, data interchange, and sharing of key crucial application systems and service platforms, and provides tight integration within companies and loose integration between services.

## III. Discussion Of Previous Work

Strong authentication procedures are vital for securing data from outside threats, as recent developments in cybersecurity have shown. Passwords and other traditional static user authentication (SUA) techniques have been found to be particularly insecure due to its static nature [4] and vulnerability to different types of cyberattacks. By providing continuous user authentication (CUA), this literature review delves into the field of keystroke dynamics as a potentially effective means of addressing the drawbacks of traditional authentication systems.

As per our research the weaknesses of classic SUA systems—most notably, their incapacity to offer continuous user identity verification—have been extensively studied. Because of this gap, researchers are looking into CUA systems that use behavioural biometrics—like keystroke dynamics—to guarantee continuous, real-time authentication. With the goal of providing a continuous, non-intrusive verification procedure that improves

security without interfering with user activities, these systems seek to authenticate users based on their distinctive typing patterns [1].
In order to capture keystrokes we needed a model which will capture keystrokes from users and we found the model which has a registration phase where the user will train the model by giving their typing cadence and rhythm of a password to the model this model is then trained such that when the user is logged in to the the application it will check the users keystrokes pattern and cadence in order to find if the user is a imposter or authentic user[3] with this model it helped us to get the model on how to take in keystrokes from users.
We saw that many statistical models like K-nearest neighbours classifiers, Bayesian classifiers and support vector machines (SVM) to perform the keystroke classification. In this SVM showed better results but still had lot of drawbacks and the use of deep learning techniques was the best method for handling intra-class data with noise.[5] Thus the use of deep learning models was required for keystroke authentication.
As we need to continuously authenticate the users we are using keystroke dynamics which are used to individually identify individuals based on their typing behaviour—rhythm, speed, and pressure—it has been a focus of scientific interest. With the use of cutting-edge machine learning algorithms and keyloggers, keystroke dynamics analysis offers a minimally intrusive way to improve user authentication procedures. This stands in stark contrast to more invasive biometric methods, providing an approachable substitute that works well with current workflows [2].
In order to improve accuracy and usability, we found a model combines a Convolutional Neural Network (CNN) and a Recursive Neural Network (RNN) to assess high-quality aspects of free-text keystrokes. This model, in contrast to earlier research on fixed texts, converts keyboard sequences into feature sequences while taking the chronological order into account for better representation. The suggested method adds an intrusion decision condition and uses sliding windows and packet formats for continuous authentication. This all-inclusive approach advances keystroke dynamic authentication research by offering a strong solution for continuous identity authentication based on free-text keystrokes.[6]
A paper focusing on Keystroke Dynamics Recognition (KDR) as a behavioural biometrics method. The proposed solution introduces a robust recurrent confidence model (R- RCM) utilizing a recurrent neural network (RNN) to enhance the sequential analysis of keystroke data. The system aims to improve CUA efficiency by promptly identifying impostor users and enhancing overall performance. By using this paper we need a RNN for our model. We found an RNN model that is LSTM (Long Short-Term Memory) as one of the RNN model which helps in authenticating keystrokes by using its typing rhythm and cadence, by using this method they were able to get an good accuracy of authentication, they also showed how well this model worked for them thus we are going to implement this model for this project in order to authenticate users continuously.[7]

## IV. Keystroke Dynamics For Continuous User Authentication Architecture

The inception of this project stemmed from the necessity to fortify personal computer protection systems. However, its utility transcends individual use cases, emerging as a pivotal asset for work-from-home companies, particularly in light of the prevailing trend towards remote work setups. Notably, many companies opt for RSA encryption to safeguard their laptops containing sensitive corporate data. Nonetheless, challenges arise when users fail to respond promptly after receiving their assigned laptops. In such instances, administrators can leverage our keystroke dynamics software to remotely shut down and secure the system, ensuring data integrity and confidentiality.
The proposed model architecture delineates the operational framework of a keystroke dynamics software engineered for user authentication, leveraging previously gathered user data for validation purposes:

A. *Data Collection and Storage:* At the onset, the system initiates data collection by monitoring the user's keystroke dynamics. Subsequently, the captured data is meticulously stored in a structured format within a CSV file to facilitate seamless implementation in subsequent stages.
B. *LSTM Model Implementation*: The core of the system lies in its utilization of Long Short-Term Memory (LSTM) neural networks. This model adeptly manages data streams in fixed-size batches, thereby circumventing overload, data loss, and inefficiencies. The LSTM model operates by analyzing the temporal sequence of the user's keystroke data, juxtaposing it against previously recorded patterns. In cases where the incoming data fails to meet predefined acceptance criteria, the system proceeds to the subsequent stage.
C. *Alerting Mechanism*: Upon surpassing a predefined rejection threshold, an alerting mechanism is promptly activated. In this scenario, if the system suspects unauthorized access, it triggers the activation of the user's screen, simultaneously activating the camera for facial recognition. The facial biometric data captured undergoes rigorous comparison with previously authenticated profiles. If the system successfully verifies the

user's identity, normal system operations resume seamlessly. Conversely, in instances of failed authentication, the system promptly initiates a shutdown sequence, simultaneously dispatching an alert notification to the user's mobile device via SMS

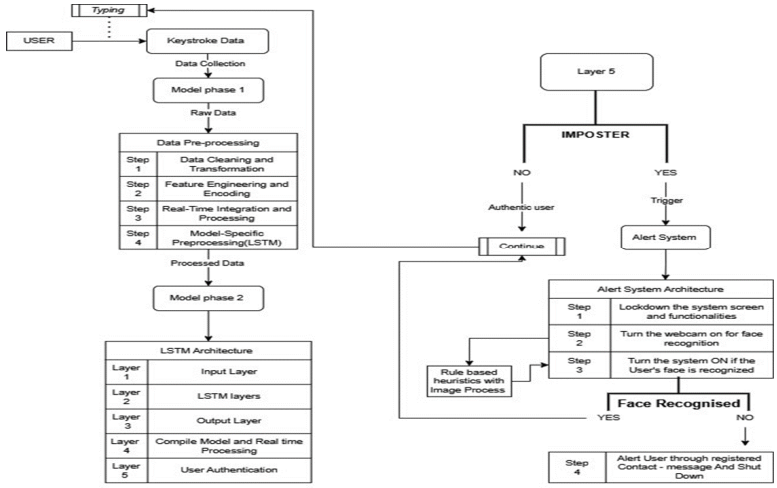


Figure 1: The Complete Design

## V. Open Discussion Of Advantage And Challenges In The Proposed Idea

Unlike RSA encryption, our software presents a cost effective alternative devoid of additional expenses typically associated with hardware-based authentication mechanisms. Furthermore, while RSA encryption relies on static, one time authentication, our model offers continuous authentication, thereby bolstering security measures against potential breaches.

However, it is imperative to acknowledge the limitations and potential drawbacks inherent in our proposed solution. Chief among these concerns is the propensity for false positives within the model. False positives pose a significant challenge, as they may inadvertently block legitimate users from accessing their systems, leading to operational disruptions and user dissatisfaction. Additionally, there exists a latent risk of harm to users during the authentication process, particularly in scenarios where the system fails to recognize authorized individuals expeditiously. Balancing the imperatives of security with user safety remains a paramount consideration throughout the developmental stages of our model.

## VI. Open Discussion Of Research

The discussion surrounding our research endeavours encompasses multifaceted considerations, encapsulating both the merits and potential pitfalls of our proposed model. Central to this discourse is the imperative to strike a delicate equilibrium between heightened security protocols and user convenience. While our software offers a robust defense mechanism against unauthorized access and data breaches, it also engenders a degree of operational complexity and user inconvenience, particularly in instances of false positives.

Moreover, the advent of remote work culture underscores the exigency for adaptable security solutions capable of accommodating diverse user environments and operational modalities. As such, future iterations of our model must prioritize scalability and user-friendliness, without compromising the integrity of security protocols.

Notwithstanding these challenges, our research holds promise for catalysing paradigm shifts in authentication methodologies, fostering a culture of continuous vigilance and adaptability in the face of evolving cyber threats. By fostering interdisciplinary collaborations and leveraging advancements in machine learning and biometric authentication technologies, we endeavour to surmount existing challenges and forge novel pathways towards enhanced cybersecurity resilience.

## VII. Conclusion

In conclusion, our keystroke dynamics software represents a pioneering endeavour in the realm of cybersecurity, offering pragmatic solutions to contemporary challenges confronting remote work environments. While the journey towards realizing its full potential is fraught with obstacles and uncertainties, the underlying principles of

innovation and adaptability underpin our collective aspirations for a more secure and interconnected digital landscape.

This theoretical framework not only delineates the operational intricacies of our proposed system but also underscores its potential significance in forthcoming research and development endeavours. By prioritizing user security and employing advanced authentication mechanisms, this model lays a robust foundation for future advancements in keystroke dynamics authentication technology.

Moving forward, sustained efforts towards refining and optimizing our model are imperative, with a steadfast commitment to mitigating risks and maximizing user utility. By embracing the ethos of continuous improvement and collaboration, we stand poised to usher in a new era of cybersecurity excellence, wherein the safeguarding of digital assets and user privacy reigns paramount.

## REFERENCES


[1] T. Kiyani, A. Lasebae, and K. Ali, "Continuous User Authentication Based on Deep Neural Networks," Aug. 2020, doi: 10.1109/ucet51115.2020.9205446.

[2] Darabseh and A. S. Namin, "On Accuracy of Classification-Based Keystroke Dynamics for Continuous User Authentication," Oct. 2015, doi: 10.1109/cw.2015.21.

[3] Md. Asraful Haque, Namra Zia Khan, Gulnar Khatoon (2002) Authentication through Keystrokes: What You Type and How You Type

[4] M., Durairaj and Ramasamy, Nandhakumar. (2013). Authentication through Keystrokes: What You Type and How You Type

[5] User Authentication through Keystroke Dynamics by means of Model Checking: A Proposal Fabio Di Tommaso, Michele Guerra, Fabio Martinelli†, Francesco Mercaldo, Massimo Piedimonte, Giovanni Rosa, Antonella Santone

[6] Continuous authentication by free-text keystroke based on CNN and RNN Xiaofeng Lua , Shengfei Zhanga , Pan Hui b , Pietro Lio

[7] TypeNet: Deep Learning Keystroke Biometrics Alejandro Acien, Aythami Morales, John V. Monaco, Ruben Vera-Rodriguez, Julian Fierrez, Member, IEEE