INTI JOURNAL | Vol.2026, Issue 2, No.9
eISSN:2600-7320

Secure Authentication Framework based on Adaptive Sparse Features and Enhanced LSTM Networks

A. Kalamani[1*], M. Suganya[2]

[1]Computer Science Department, RVS College of Arts and Science, Sulur, Coimbatore, India.
[2]Department of Information Technology, RVS College of Arts and Science, Sulur, Coimbatore, India.

***Email**: kalaa.mca@gmail.com[1]

Abstract

In the modern world, security has become a critical need across every domain. Among the various approaches, biometric authentication offers stronger protection compared to traditional security mechanisms. Within biometric systems, physical traits such as iris, fingerprint, and palm print, along with behavioural traits like gait, voice are widely used in all areas. Passwords have historically been the standard technique for limiting access to computer systems, however this strategy has a number of fundamental problems. Keystroke dynamics is a relatively recent biometric identification technique that offers a low-profile, reasonably priced way to make the standard login and password process more difficult. Keystroke dynamics enables user verification based on individual typing behaviour like rhythm and speed. In this study, Statistical measures including mean and standard deviation are calculated from keystroke features such as latency, duration, and digraph timing. To improve reliability and identify the optimum features, Adaptive sparse feature selection is applied, which efficiently identifies the most relevant feature subset and enhances the overall performance. For the classification phase, Enhanced LSTM models are employed to classify the users based on their typing patterns. The system's effectiveness is evaluated using standard performance measures such as false positives, false negatives, true positives, and true negatives. Keystroke dynamics supports SDG in the way of energy saving and reduce E-waste. The Accuracy of the system is improved while using the Adaptive Sparse with the proposed Enhanced LSTM models. This system identifies genuine users and intruders with 91% of Accuracy Rate. It is useful for more security areas.

Keywords

Mean, Standard Deviation, Adaptive Sparse Feature Selection, Enhanced LSTM.

Submission: 10 June 2026; **Acceptance:** 14 August 2026; **Available online:** August 2026

Introduction

Soft computing (Blackburn et al., 2007; Okwu & Tartibu, 2020) constitutes a broad category that encompasses various algorithms utilized to derive approximate solutions for intricate issues within the domain of computer science. Biometric research represents a subset of soft computing, employing statistical methodologies to scrutinize biological data pertaining to individuals. This methodology is pivotal for the identification and verification of an individual's identity predicated on distinctive characteristics. Such attributes encompass retinal scans, iris patterns, fingerprints, vocal characteristics, and even typing styles (Monrose & Rubin, 2000). These systems find applications across numerous commercial sectors, encompassing forensic analysis, access control for files, banking security, and attendance monitoring.

Biometrics (Blackburn et al., 2007; AlRousan & Intrigila, 2020) entails the quantitative measurement and statistical examination of individuals' distinctive physiological and behavioural attributes (Tripathi, 2011; Hossain & Chetty, 2011). Each individual is identified by their unique characteristics. Physical identification methodologies are predicated on the examination of a person's physiological traits, such as fingerprints, facial recognition, iris patterns, and the configurations of palm or finger veins. Behavioural identification methodologies (Tripathi, 2011) rely on the analysis of an individual's behavioural traits—attributes intrinsic to each person during the execution of actions. Signature verification, keystroke dynamics, vocal patterns, and gait analysis are grounded in behavioural characteristic identification methodologies.

Methodology

Keystroke dynamics constitutes one of the prominent authentication methodologies employed for the identification of individuals. The unique typing rhythm of each individual is discerned and validated through a variety of advanced techniques. This technology finds application across numerous domains, including attendance monitoring, forensic investigations, banking, and other sectors requiring enhanced security measures (Monrose & Rubin, 2000). The principal aim of this investigation is to diminish both the False Acceptance Rate and the False Rejection Rate (Teh et al., 2007).

Keystroke dynamics facilitates sustainable development by offering an economical, energy-efficient, and hardware-free biometric authentication mechanism that fortifies digital security, minimizes electronic waste, promotes digital inclusion, and enhances secure infrastructures in accordance with Sustainable Development Goals (SDGs) such as SDG 9, SDG 12, SDG 13, and SDG 16.

The initial instance of an individual engaging with a biometric system is referred to as enrollment or feature extraction (Karnan et al., 2009). In the Enrollment, the system collects and saves a person's biometric information. In this research, duration, latency and digraph features are extracted from 11 users.

Keystroke Dynamics Working Methodology illustrates the methodology followed for collecting and extracting keystroke-dynamics features (Figure 1). To effectively capture a keystroke, it is requisite for users to input their password a total of 10 times. The system acquires these features through three distinct methodologies, which pertain to the temporal aspects (in milliseconds) that a specific user sustains the key in a pressed state (duration time), the interval that elapses between the release of one key and the subsequent pressing of another (latency time), and the blend of the above-mentioned termed Digraph as shown in Figure 2.

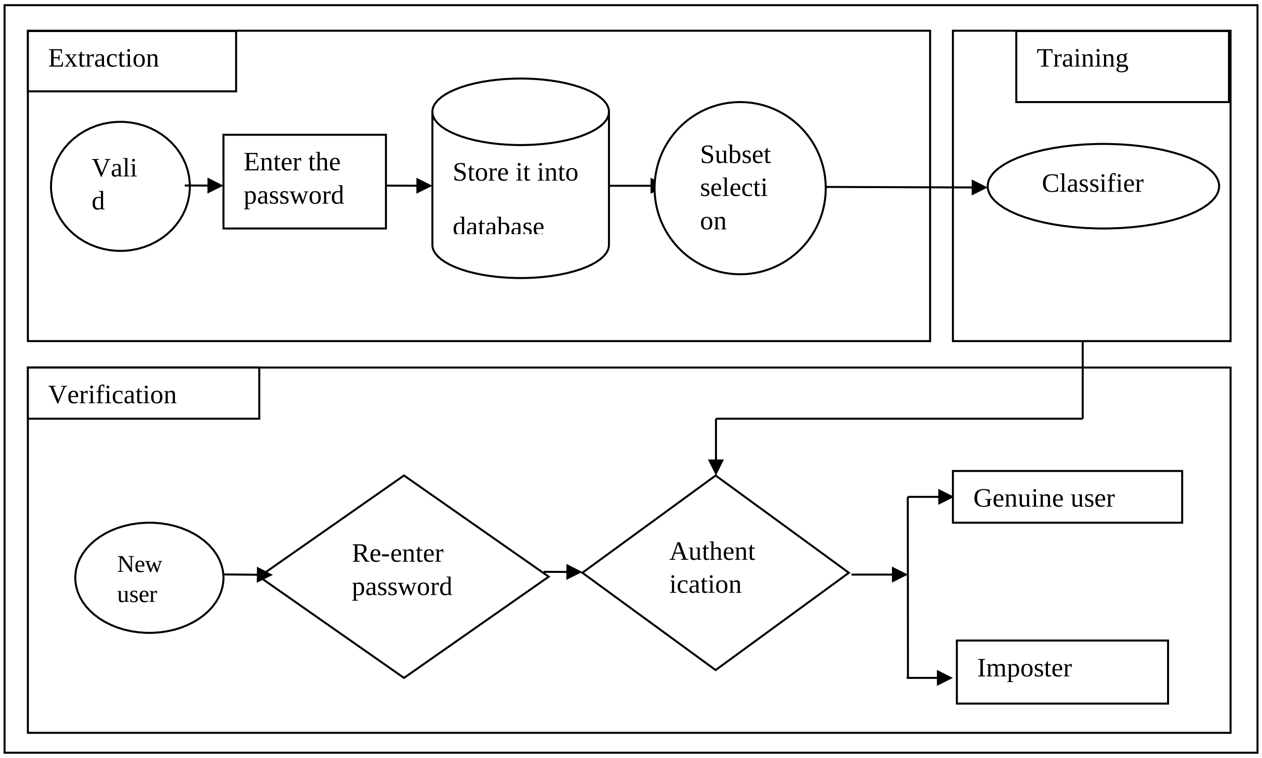

Figure 1. Keystroke Dynamics working Methodology.

The data was collected from 11 participants, each utilizing different passwords. The mean and standard deviation values were subsequently computed.

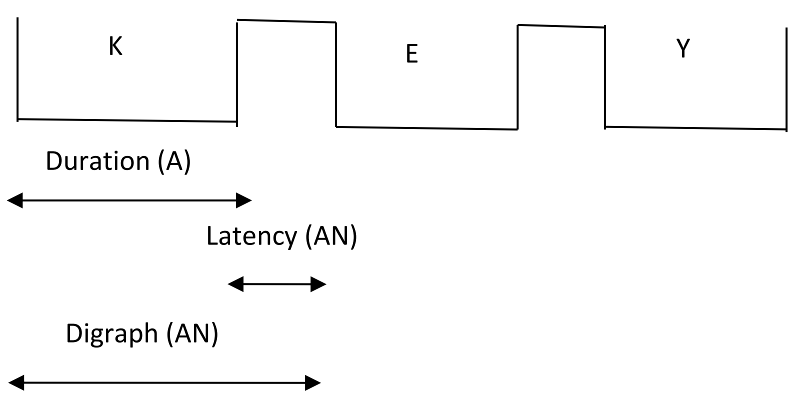

Figure 2. Measuring Method of Features Duration, Latency and Digraph

A keystroke recording application software was conceptualized and implemented utilizing the Python programming language for the purpose of archiving the measurements of duration, latency, and digraph for each user sample. In the process of generating the raw data file, the mean (μ) and standard deviation (σ) (Teh et al., 2007; Karnan et al., 2009) for each feature (i) within the pattern set (x) are computed for N samples.

Selecting the optimum features from the extracted features, a process known as Feature subset selection (Leberknight et al., 2008; Desai et al., 2013), is crucial for effective classification. The extracted features consist of Duration, Latency, and Digraph, and these values are input to the Adaptive Sparse Feature Selection (ASFS) Algorithm (Shradha et al., 2015; Ali et al., 2016), as illustrated in Figure 3, which depicts the three different extracted features processed through the ASFS algorithm.

The algorithm begins with Step 1, where features are extracted in milliseconds and normalized using statistical measures. In Step 2, the Mutual Information (MI) value for the normalized features is calculated using the formula $MI(X,Y) = \sum P(x,y) \log(p(x,y)/p(x)p(y))$, where $x \in X$ and $y \in Y$, with $P(x)$ and $P(y)$ representing the marginal probability distributions of the feature and its corresponding label, respectively; MI quantifies the extent to which one variable (feature) conveys information regarding another variable (label/class). Step 3 involves identifying the features with the highest MI values, where Latency and Digraph provide high MI values compared with Duration, as determined through Python coding; consequently, the lowest MI feature, Duration, is removed, leaving Latency (with the highest MI) as the most significant feature, followed by Digraph. In Step 4, L1-regularized logistic regression (LASSO) is applied to enforce parity, which helps remove redundant features and select those with the highest impact; the optimal weights after LASSO are W(D)=0.5 (low impact), W(L)=0.7 (high impact), and W(DT)=0.6 (moderate impact), with the option to drop the lowest value if it is too low. Finally, in Step 5, the optimal features selected are Latency (L) and Digraph, and subsequent classification is performed using the proposed classification algorithm, Enhanced LSTM networks.

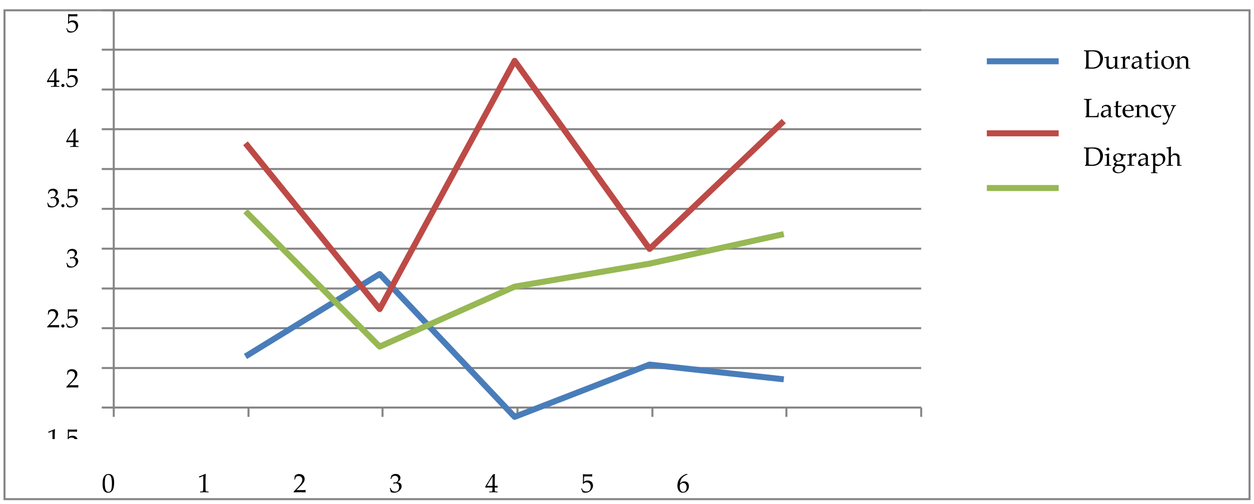

Figure.3 Three Different Extracted features with ASFS Algorithm

Results and Discussion

Following the feature selection process, the finalized optimal features—Latency and Digraph—are fed into the classification phase. During this phase, the system is trained to establish and identify each individual's unique threshold value, which serves as the decision boundary for distinguishing whether two samples originate from the same user or from different users. In practical deployment, when a user re-accesses the system, the newly captured biometric data is compared against the template stored during enrolment; this one-to-one matching process is formally defined as verification (Tripathi, 2011; Mhenni et al., 2021).

For the classification task, an Enhanced Long Short-Term Memory (LSTM) network is employed. While conventional LSTM models are constrained to process sequential data in a single forward direction (Mhenni et al., 2021), the proposed Enhanced-LSTM architecture extends this capability by incorporating both forward and backward passes. This bidirectional processing enables the model to capture more comprehensive temporal dependencies and sequential patterns inherent in keystroke dynamics—a critical advantage for accurately identifying users based on their unique typing behaviour. Specifically, the Enhanced LSTM architecture comprises two distinct LSTM layers that process the input sequence in opposite directions, allowing the network to learn from past and future contexts simultaneously, thereby improving the robustness and discriminative power of the user authentication system.

Table 2. Combined state results

Time Step (t)	Key	Forward ht	Backward ht	Combined Hidden State (ht)
t=1	'K'	[0.15, 0.40]	[0.85, 0.10]	h_1 = [0.15, 0.40, 0.85, 0.10]
t=2	'E'	[0.55, 0.80]	[0.65, 0.35]	h_2 = [0.55, 0.80, 0.65, 0.35]
t=3	'Y'	[0.90, 0.75]	[0.30, 0.55]	h_3 = [0.90, 0.75, 0.30, 0.55]

$$P = [0.03, 0.08, 0.65, 0.07, 0.02]$$

The genuine user is identified based on the highest probability in the Softmax output vector P. As observed in Table 2, the third value of the probability and the Z value of 3 rd output also maximum value. It also satisfies the confidence threshold value of above 60% .so the third user is identified as a genuine user. In keystroke dynamics, Enhanced-LSTMs often achieve better performance than unidirectional LSTMs because they have a more complete view of the sequence.

Accuracy calculation metrics such as FAR, FRR are identified and System performance is measured (Kumar & Arun, 2025; Karnan et al., 2009). FAR = FA / N *100. FAR stands for False Acceptance Rate(Unauthorized accepted), FA for False Acceptance Incidence, and N for Total imposter Attempt. FRR = FR / N * 100, where N is the total number of Genuine user attempt, FRR(Authorized Rejected) is the false rejection rate, and FR is the number of instances of false rejection of Authorized user. System attains a low FAR and Low FRR values such as 0.126, 0.186

An Enhanced-LSTM enhanced using the standard LSTM by allowing the model to look both forwards and backwards in a sequence, providing a more comprehensive contextual understanding which is crucial for tasks where future information is as important as past information for accurate prediction or classification. Adaptive sparse Feature Selection with Enhanced LSTM models provide a more accuracy compared with LSTM Model. The system accept the Genuine user and reject the intruders with the threshold value. It gives 91% accuracy for user Authentication. The classification Error Rate is also reduced to 0.198%.

Conclusion

The study demonstrated that keystroke dynamics can be used to extract characteristics that can be utilized to identify computer users since they are rich in unique mannerisms and behaviours. Ten samples of each of the eleven users' features are extracted. Subsets of features are selected using Adaptive Sparse Feature Selection Algorithm. Enhanced LSTM models used for classification. Adaptive Sparse optimization techniques with Enhanced LSTM classification perform well with Accuracy Rate of 91%. Objective of this research FAR, FRR rate are reduced as 0.126 and 0.183. SGD of E-waste reduce and Energy Saving also improved using this system. Dynamic Authentication with touch screen monitoring is the future enhancement of this research.

Acknowledgements

The authors would like to thank RVS College of Arts and Science and the Head of the Department of Information Technology and Computer Science for their support in preparing this paper. Special thanks are extended to Dr. Shamina, Head of the Department of Biochemistry, for her valuable support and encouragement.

References

Ali, M. M., Mahale, V. H., Yannawar, P., & Gaikwad, A. T. (2016). Overview of fingerprint recognition system. In *2016 International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT)* (pp. 1334–1338). https://doi.org/10.1109/ICEEOT.2016.7754900

AlRousan, M., & Intrigila, B. (2020). A comparative analysis of biometrics types: Literature review. *Journal of Computer Science, 16*(12), 1778–1788. https://doi.org/10.3844/jcssp.2020.1778.1788

Blackburn, D., Miles, C., Wing, B., & Shepard, K. (2007). *Biometrics overview*. National Science and Technology Council, Subcommittee on Biometrics. https://www.hsdl.org/c/view?docid=464488

Desai, N., Dhameliya, K., & Desai, V. (2013). Feature extraction and classification techniques for speech recognition: A review. *International Journal of Emerging Technology and Advanced Engineering, 3*(2), 5–9. https://www.academia.edu/73258850/Feature_Extraction_and_Classification_Techniques_for_Speech_Recognition_A_Review

Hossain, S. M. E., & Chetty, G. (2011). Human identity verification by using behavioural biometric traits. *International Journal of Bioscience, Biochemistry and Bioinformatics, 1*(3), 199–205. https://doi.org/10.7763/IJBBB.2011.V1.36

Karnan, M., Akila, M., & Kalamani, A. (2009). Feature subset selection in keystroke dynamics using ant colony optimization. *International Journal of Engineering and Technology, 1*(5), 72–80. https://academicjournals.org/journal/JETR/article-abstract/B0738825442

Kumar, R., & Arun, A. (2025). Keystroke dynamics based user authentication system. *Recent Trends in Artificial Intelligence, 4*(2), 70–82. https://doi.org/10.46610/RTAIA.2025.v04i02.009

Leberknight, S., Widmeyer, G. R., & Recce, M. L. (2008). An investigation into the efficacy of keystroke analysis for perimeter defense and facility access. In *Proceedings of the IEEE Conference on Technologies for Homeland Security* (pp. 345–350). https://doi.org/10.1109/THS.2008.4534475

Mhenni, A., Rosenberger, C., & Ben Amara, N. E. (2021). Keystroke dynamics classification based on LSTM and BLSTM models. In *2021 IEEE International Conference on Design & Test of Integrated Micro & Nano-Systems (D&T)* (pp. 295–298). https://doi.org/10.1109/CW52790.2021.00056

Monrose, F., & Rubin, A. D. (2000). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems, 16*(4), 351–359. https://doi.org/10.1016/S0167-739X(99)00059-X

Okwu, M. O., & Tartibu, L. K. (2020). *Metaheuristic optimization: Nature-inspired algorithms and applications*. Springer. https://doi.org/10.1007/978-3-030-61111-8

Shradha, T., Chourasia, J. N., & Chourasia, V. S. (2015). A review of advancement in biometric systems. *International Journal of Innovative Research in Advanced Engineering, 2*(1), 22–26. https://www.ijirae.com/volumes/Vol2/iss1/30.JAEC10091.pdf

Teh, P. S., Teoh, A. B. J., Ong, T. S., & Neo, H. F. (2007). Statistical fusion approach on keystroke dynamics. In *Proceedings of the Third International IEEE Conference on Signal-Image Technologies and Internet-Based Systems*. pp. 918–923. https://doi.org/10.1109/SITIS.2007.46

Tripathi, K. P. (2011). A comparative study of biometric technologies with reference to human interface. *International Journal of Computer Applications, 14*(5), 10–15. https://doi.org/10.5120/1842-2493