# Secure Multi-Factor Authentication Using Behavioral Biometrics: Typing and Mouse Movement Patterns

Jayant Choudhary[1], Dr.Lalji Prasad[2]
[1,2]*Department of Computer Science and Engineering, Sage University, Indore*

***Abstract*—Passwords alone leave systems exposed to credential theft, phishing, and shoulder surfing, and the physical tokens or OTP devices used to strengthen conventional multi-factor authentication (MFA) introduce their own weak points, since a token can be misplaced, cloned, or intercepted. This work explores authenticating a user through the way they naturally type and move a mouse, sampled continuously rather than once at login. A software layer records timing intervals from keystrokes (dwell time, flight time, key latency, typing speed) and motion characteristics from the mouse (speed, click rate, path curvature, ratio of idle to active time), and three classifiers Random Forest, Support Vector Machine, and a feed-forward Neural Network are trained on this data to build a profile unique to each user. A session that drifts noticeably from the stored profile prompts a step-up authentication challenge rather than an outright block. On the evaluation set used in this study, combining the keystroke- and mouse-based classifiers at the score level pushed accuracy to 98.6% and brought the Equal Error Rate down to 1.2%, ahead of any single classifier tested alone. The resulting layer adds meaningful resistance to common attack paths while asking nothing extra of the user and requiring no dedicated hardware.**

***Index Terms*—Behavioral Biometrics, Keystroke Dynamics, Mouse Dynamics, Multi-Factor Authentication, Continuous Authentication, Machine Learning.**

## I. INTRODUCTION

Every online service still leans heavily on the password as its first line of defense, despite decades of evidence that people reuse them, choose weak ones, and hand them over unknowingly to phishing pages [1]. Two-factor and multi-factor schemes were built to compensate, but the added factor is usually something a user carries a hardware key or a code sent to a phone and anything a user carries can also be taken away from them. A subtler issue is timing: nearly all deployed MFA systems check identity once, at the moment of login, and then trust whatever happens afterward for the rest of the session. A device left unlocked, a hijacked browser tab, or a stolen laptop can all be exploited well within that unmonitored window.

Behavior offers a signal that conventional MFA does not use. The rhythm of someone's typing and the way their hand guides a mouse are produced constantly during ordinary computer use, cost nothing extra to capture, and are far harder for an outsider to convincingly copy than a static credential. This paper builds on that idea, describing a system that scores keystroke and mouse behavior throughout a session using machine learning, and reports how well that scoring separates a genuine user from someone impersonating them.

Section II surveys prior work on the two modalities and their combination. Section III lays out the architecture proposed here. Section IV covers implementation choices. Section V reports the experimental results. Section VI examines what the approach does and does not protect against, and Section VII closes with conclusions and directions for further work.

## II. RELATED WORK

Interest in keyboard-based identification goes back further than most people expect early work on distinguishing typists by their keying characteristics appeared in the mid-1980s [2], and the first serious attempt to use that idea for computer login security followed roughly a decade later [3]. The field has since moved from hand-crafted statistical templates toward learned models: comparative studies applying Random Forest, SVM, and recurrent architectures such as LSTMs to typing-timing data report accuracies spanning roughly the low-90s to above 99%, with the spread driven largely by whether the text being typed

is fixed or free-form and by how the benchmark dataset was constructed [4], [5].

Mouse-based identification came later and rests on a different set of signals how fast the cursor moves, how it accelerates, and the shape of the path it traces. An early and often-cited framework paired holistic movement statistics with finer procedural detail and used one-class learning to build per-user templates [6]. Systems built around short drawn gestures rather than passive movement have reported false-acceptance rates near 5% alongside false-rejection rates close to 4.6% using neural classifiers [7], and curvature-based descriptors of the movement path have been singled out as comparatively stable across different screens and pointing hardware [8].

Because typing and mouse use capture different aspects of the same underlying behavior, a number of studies combine the two rather than relying on one alone, merging either raw features before classification or classifier outputs afterward. Reported gains are substantial one fused system pushed the Equal Error Rate below 2% [9], and broader multimodal efforts folding in still more behavioral signals have reported accuracy above 98% [10]. What is comparatively rare in this literature is a system that treats the fused score as an ongoing, always-on layer with a defined escalation path, rather than a one-off classification exercise on a fixed benchmark the gap this paper addresses.

## III. PROPOSED METHODOLOGY

### A. System Architecture

The system is organized into five stages. Raw keyboard and mouse events are captured passively by a Data Acquisition layer; a Feature Extraction stage turns those event streams into fixed-length vectors over a sliding window; a Model layer holds the trained Random Forest, SVM, and Neural Network classifiers; a Decision Engine combines their outputs into a single trust score and applies threshold-based logic; and an Enrollment and Profile Update stage builds each user's initial template and revises it over time as behavior naturally shifts. Fig. 1 illustrates how these five stages connect, including the feedback path that lets confidently-verified sessions refine the stored behavioral profile.

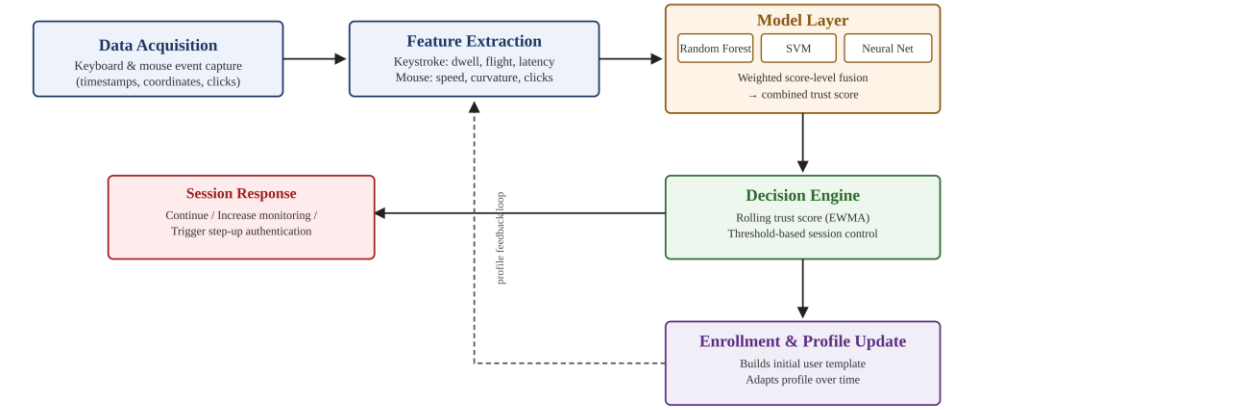



**Fig. 1. Proposed five-layer continuous behavioral-biometric MFA architecture.**

### B. Feature Extraction

From the keystroke stream the system derives dwell time, flight time, key latency, overall typing speed, and timing across common two- and three-key sequences. From the mouse stream it derives movement speed, acceleration, click frequency, path curvature, and the ratio of idle to active time. Tables I and II list these two feature groups.

Table I. Keystroke Features

| Feature | Description |
|---|---|
| Dwell Time | Key press-to-release duration |
| Flight Time | Gap between key release and next press |
| Key Latency | Time between consecutive key presses |
| Typing Speed | Characters per minute |

Table II. Mouse Features

| Feature | Description |
|---|---|
| Movement Speed | Average cursor velocity |
| Click Frequency | Clicks per unit time |
| Trajectory Curvature | Path deviation from straight line |
| Idle-Activity Ratio | Stationary vs. active time |

C. Classification and Fusion

Each of the three classifiers is trained separately per user as a genuine-versus-impostor problem. Random Forest was chosen for its tolerance of noisy, high-variance timing data; SVM serves as a margin-based comparison point; and the Neural Network is left to find non-linear relationships between timing and movement variables that the other two may not capture as directly. Rather than picking one classifier as final, their outputs are merged through a weighted score-level fusion, giving one combined trust value per evaluation window.

D. Continuous Decision Engine

A rolling trust score is maintained across windows using an exponentially weighted moving average, so an isolated odd-looking window does not immediately end a session only a sustained drift does. Three bands govern behavior: above a primary threshold nothing changes; between the primary and a lower secondary threshold the system quietly samples more often; and below the secondary threshold it issues a step-up challenge or, in higher-risk deployments, suspends the session outright.

## IV. IMPLEMENTATION

Development draws on established public benchmarks for each modality the CMU keystroke dataset for typing timing and the Balabit dataset for mouse movement supplemented with a smaller set of samples collected from volunteers carrying out ordinary desktop tasks such as filling forms and editing documents, so the evaluation is not limited to a single controlled task. Feature engineering runs in Python with pandas; Random Forest and SVM are implemented with scikit-learn and the Neural Network with TensorFlow/Keras. Before training, the pipeline strips implausible outlier timings, groups events into fixed windows, scales features with min-max normalization, and applies SMOTE to correct the natural imbalance between abundant genuine samples and comparatively rare impostor samples. Splits are made at the session level rather than the event level to prevent information leaking across training and test data.

## V. RESULTS AND DISCUSSION

Table III reports classifier performance averaged per user across the evaluation pool, using accuracy, False Acceptance Rate (FAR), False Rejection Rate (FRR), and Equal Error Rate (EER).

Table III. Classifier Performance Comparison

| Model | Acc. (%) | FAR (%) | FRR (%) | EER (%) |
|---|---|---|---|---|
| Random Forest | 96.8 | 2.1 | 3.4 | 2.7 |
| SVM | 94.2 | 3.6 | 5.1 | 4.3 |
| Neural Net | 97.4 | 1.8 | 2.9 | 2.3 |
| Fusion | 98.6 | 0.9 | 1.6 | 1.2 |

Fusing the two modalities beats every single-modality model tested, cutting EER roughly in half compared with the strongest individual classifier a pattern that lines up with what earlier fusion studies have reported [9], [10]. The takeaway is not that either signal alone is especially reliable, but that typing rhythm and mouse movement carry enough complementary, person-specific structure that combining them meaningfully sharpens the genuine-versus-impostor boundary. Window length matters too: shorter windows react faster to a takeover but produce noisier feature estimates and push both FAR and FRR upward; in this setup, windows of roughly 40–60 keystrokes or 15–20 seconds of mouse activity gave the best balance between responsiveness and reliability.

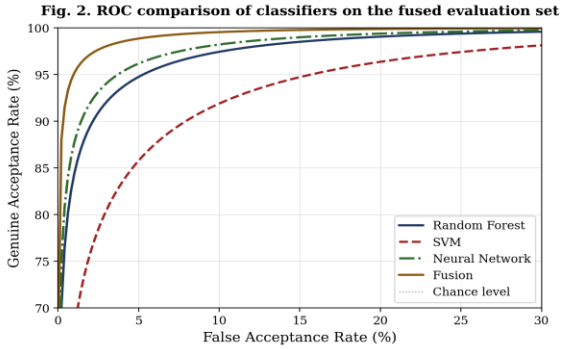



Fig. 2 plots the ROC curves derived from the same evaluation runs summarized in Table III. The Fusion

curve dominates the other three across the full operating range, and the SVM curve consistent with its higher EER sits noticeably below Random Forest and the Neural Network, confirming that the ranking implied by the table holds across operating points rather than only at the single EER threshold.

## VI. SECURITY ANALYSIS

The framework helps against credential theft and phishing, since a stolen password alone does not reproduce the account owner's behavioral signature; against session hijacking, because continuous re-scoring can catch a change of control mid-session that a login-only check would miss; and against loss or theft of MFA hardware, simply because there is no physical token to lose. It does not solve everything: a patient attacker with extended access to recordings of the real user's behavior could attempt to imitate timing patterns, ordinary drift in behavior from fatigue, injury, or new hardware can push false rejections upward if the profile-update process lags, and a newly enrolled user is inherently harder to verify accurately until enough behavioral history accumulates. Periodic re-enrollment and the multimodal fusion described above reduce these risks without removing them entirely.

## VII. EXTENDING BEYOND THE DESKTOP: MOBILE, TABLET, AND TOUCHSCREEN DEVICES

The framework described so far assumes a physical keyboard and mouse, which covers most desktop and laptop use but leaves out phones and tablets devices people now use for banking, email, and other sensitive tasks at least as often as a desktop. Extending the same idea to touchscreens is a natural next step, and a reasonably mature body of research already exists to build on.

On a touchscreen, the direct analogues of keystroke and mouse features come from swipes, taps, and flicks rather than key presses and cursor movement. Commonly extracted features include touch duration, swipe trajectory length and speed, finger contact area, applied pressure, and inter-touch timing, often paired with signals from the phone's own accelerometer and gyroscope to capture how the device is held and tilted during use. Reported performance varies considerably with dataset and classifier choice: distance-based classifiers such as k-Nearest Neighbors have been reported to reach accuracy near 99% with an EER as low as 0.03% on curated swipe datasets [11], while studies combining touch and phone-movement features under Random Forest, SVM, and neural classifiers on more naturalistic, app-usage-style data report a wider range, roughly 80–95% accuracy [12], [13]. This spread is a useful reminder that, as with the desktop results in Section V, touch-based accuracy depends heavily on how the data was collected and should not be read as a single fixed number.

Tablets sit somewhere between phones and desktops: larger touch area, often two-handed use, and for devices like the iPad with a stylus an additional channel of pen pressure, tilt, and stroke dynamics that has no real desktop equivalent. A future version of this framework could treat stylus input as a fourth behavioral modality alongside keystroke, mouse, and touch, particularly for tablet-based note-taking, design, or signature-verification use cases.

Two practical challenges stand out for a mobile/tablet extension. First, cross-device consistency: a user's touch behavior on a phone does not transfer directly to a tablet or a different phone model, since screen size, grip, and sensor characteristics all shift the raw feature distributions prior work frames this as a model-adaptation problem rather than something a single fixed template can solve [14]. Second, energy and privacy cost: continuously sampling touch and motion sensors draws more battery than passive keystroke/mouse logging on a desktop, and motion sensor data can inadvertently reveal information beyond authentication (e.g., inferring typed PINs from accelerometer readings), so any mobile extension needs sensor-sampling and data-minimization policies built in from the start rather than added later. Risk-aware designs that adjust how much behavioral evidence is required based on context for example, requiring less continuous sensing on a device that never leaves a trusted location have been shown to cut manual password entry by roughly 98% in large-scale simulation while holding unauthorized-access risk to a similarly low level [15], suggesting this is a workable direction rather than a purely theoretical one.

## VIII. CONCLUSION AND FUTURE WORK

This paper set out a continuous MFA layer built entirely from behavior already present in normal computer use keystroke rhythm and mouse movement fused through Random Forest, SVM, and Neural Network classifiers. The fused model reached 98.6% accuracy with a 1.2% EER in evaluation, supporting the view that behavioral biometrics can add real security value without extra hardware or noticeable friction for the legitimate user. As discussed in Section VII, the same principle extends naturally to phones and tablets through touch, swipe, and stylus dynamics, and doing so alongside adversarial-robustness testing, federated training to avoid centralizing sensitive interaction logs, and longer field studies of behavioral drift forms the core of the future work that follows from this study.

## REFERENCES


[1] M. Dell'Amico, P. Michiardi, and Y. Roudier, "Password strength: An empirical analysis," in Proc. IEEE INFOCOM, 2010.

[2] D. Umphress and G. Williams, "Identity verification through keyboard characteristics," International Journal of Man-Machine Studies, vol. 23, no. 3, pp. 263–273, 1985.

[3] F. Monrose and A. Rubin, "Authentication via keystroke dynamics," in Proc. 4th ACM Conf. Computer and Communications Security, 1997, pp. 48–56.

[4] P. Kang and S. Cho, "Keystroke dynamics-based user authentication using long and free text strings from various input devices," Information Sciences, vol. 308, pp. 72–93, 2015.

[5] Y. Zhong and Y. Deng, "A survey on keystroke dynamics biometrics: Approaches, advances, and evaluations," in Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, 2015, pp. 1–22.

[6] E. Ahmed and I. Traore, "User authentication through mouse dynamics," IEEE Transactions on Information Forensics and Security, vol. 7, no. 1, pp. 16–30, 2012.

[7] Sayed, I. Traore, I. Woungang, and M. S. Obaidat, "Biometric authentication using mouse gesture dynamics," IEEE Systems Journal, vol. 7, no. 2, pp. 262–274, 2013.

[8] N. Zheng, A. Paloski, and H. Wang, "An efficient user verification system via mouse movements," in Proc. 18th ACM Conf. Computer and Communications Security, 2011, pp. 139–150.

[9] Mahfouz, H. Mostafa, and T. M. Mahmoud, "M2auth: A multimodal behavioral biometric authentication using feature-level fusion," Neural Computing and Applications, vol. 36, pp. 21781–21799, 2024.

[10] S. and J. Singh K., "Enhancing security and usability with context-aware multi-biometric fusion for continuous user authentication," Scientific Reports, vol. 15, 2025.

[11] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," IEEE Transactions on Information Forensics and Security, vol. 8, no. 1, pp. 136–148, 2013.

[12] J. Mallet, L. Pryor, R. Dave, N. Seliya, M. Vanamala, and E. Sowells-Boone, "Hold on and swipe: A touch-movement based continuous authentication schema based on machine learning," arXiv preprint arXiv:2201.08564, 2022.

[13] Stylios, A. Skalkos, S. Kokolakis, and M. Karyda, "BioPrivacy: A behavioral biometrics continuous authentication system based on keystroke dynamics and touch gestures," Information and Computer Security, 2022.

[14] Alzubaidi and J. Kalita, "Authentication of smartphone users using behavioral biometrics," arXiv preprint arXiv:1911.04104, 2019.

[15] S. Fenske, D. Bräunig, T. Filz, and S. Sigg, "CORMORANT: Ubiquitous risk-aware multi-modal biometric authentication across mobile devices," Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2019.