



# KNN-Based Behavioral Intrusion Detection Using Mouse Dynamics and Session Timing

Siham Oleiwi Tuama

Ministry of Education/ Babylon Education Directorate, [mcsiham1994@gmail.com](mailto:mcsiham1994@gmail.com) Babylon , Iraq.

اكتشاف التطفل السلوكي القائم على KNN باستخدام ديناميكيات الماوس

وتوقيت الجلسة

سهام عليوي طعمه

وزارة التربية / مديرية تربية بابل ، بابل، العراق [mcsiham1994@gmail.com](mailto:mcsiham1994@gmail.com)

Accepted:

13/1/2026

Published:

31/3/2026

## ABSTRACT

This research proposes a behavioral intrusion detection system that uses mouse movement, session duration, and K-Nearest Neighbors (KNN) for classifying sessions. This system attempts to classify users into legit and non-legit users based on session time, number of mouse clicks, and mouse movement speed. The system was built using one of the class sessions, and the other class session was set aside for testing. The result was impressive, as every single metric of classification evaluation, such as accuracy, precision, recall, and F1-score, scored 1. The more advanced metrics also display an MCC of 1.00, a FRR of 0.00, and a FAR of 0.00, indicating that there were no misclassifications whatsoever. This indicates that the behavioral characteristics the system relied on captured user patterns of identification effectively and distinctly. The system has proven to be effective in the application of behavioral biometrics in intrusion detection systems.

**Keywords:** Intrusion Detection, Mouse Dynamics, Session Timing Analysis, K-Nearest Neighbors (KNN) .

## 1. INTRODUCTION

As more operations are moved to cloud environments and VPNs, security perimeters have vanished, amplifying the necessity of this layered access control within an organization's cybersecurity strategy. Passwords, even the modern multifactor ones, and biometric IDs are all susceptible to some form of cyber-attack. User complacency is a security threat almost worse than any cyberattack. These vulnerabilities are the reason cyber security experts are focusing on behavioral biometrics, especially mouse movement profiling. The ability to unobtrusively and continuously record mouse movements and use this data to identify when security protocols are being breached is invaluable [1]. Mouse dynamics are the unique and individual patterns



associated with an individual's use of a computer mouse, including, but not limited to, average speed of mouse movement, timing of mouse clicks, patterns of acceleration, and even the small movements made while on the screen. These unique patterns are so small that they are nearly impossible for users to introspectively identify. These unique mouse movements are becoming more relied on within the cybersecurity community to help separate legitimate users and impostors [2]. Behavioral profiles are enriched, and movement data improves anomaly detection, when combined with mouse tracking and other behavioral metrics like time, duration, and log-in location [3]. In this context, case-based methods and, in particular, K-Nearest Neighbors are noteworthy. Given its distance-based metric approach, KNN identifies dense behavioral regions without complicated parameter configurations, and showcases better performance in categorization with high noise levels in click data [4]. We present an innovative model of intrusion prevention that integrates KNN with mouse gesture signatures and session timing artifacts. By combining these features into a real-time model, we aim to improve classification accuracy and the system's robustness against unauthorized access, thereby improving behavioral access control systems against sophisticated real-world attacks.

## 2. RELATED WORK

In (2024), Khan and collaborators offered an extensive examination of the behavioral biometric systems incorporating mouse dynamics. The authors analyzed different feature extraction methods, assorted datasets, and various machine learning techniques such as KNN and SVM, in both the dual roles of intrusion and authentication. It is a recent overview regarding the advancements of behavioral intrusion detection systems based on mouse movement analysis. [5]

In (2021), Kılıç and co-authors built the Bogazici Mouse Dynamics Dataset which is geared toward user authentication and insider theft detection. They analyzed user sessions to record fine-grained features of mouse movement and mouse timing, and they benchmarked different classifiers for distinguishing between the legitimate users and the intruders. They emphasized the importance of the session-based timing features in adaptive intrusion detection. [6]

Antal and Egyed-Zsigmond (2018), in their mouse dynamics study, created an intrusion detection system using the Balabit dataset. They manually divided the raw mouse activity into sessions and calculated the mouse behavior features (speed, direction, and click frequency) and used them to train classifiers to detect impostors. This study has been useful to the next generation of behavioral-based intrusion detection systems [7].

Davuluri (2023) proposed an experimental system aimed at continuous mouse dynamics authentication and session timing analysis. KNN, Decision Tree, and Random Forest classifiers were used to study the behavior consistency of users to evaluate the classifiers. KNN performance in detecting users in unauthorized active sessions was commendable [8].



### 3. INTRUSION DETECTION

Intrusion detection systems (IDS) act as a remote control and monitor of network data and users to detect potential incidents and ensure that users comply with the security rules of the system. One of the main attributes that differentiates IDS systems is that they can use either signature detection or anomaly detection. Signature detection is the matching of inputs to data that has been previously defined as an attack, whereas anomaly detection is the indicator of a new attack that is being perpetrated because of a user or a system that has been configured in a manner that is considered abnormal (9). Anomaly detection systems can use other advanced techniques to detect attacks that have previously not been defined (10). Researchers have advanced the anomaly detection model with behavioral intrusion detection that diverts its focus vertically to the network packets and analyzes the subtle user activities in the form of the sequencing of a user's mouse movements, the timing of the user's keystrokes, and the duration of the individual user sessions. Such biometric systems are difficult to mimic and enhance protective barriers where access and control of sensitive credentials are shared, or remote access to the system is employed. In the case of the current study, systems record data with K-Nearest Neighbor (KNN) classifiers and data with mouse movement patterns and session duration. The classifier compares the recorded fingerprints against a baseline profile of the verified user, marking any considerable discrepancy as an intruder. This approach is flexible in nature, as the baseline behavioral profile of the valid user can be progressively updated, which helps in reducing the number of false alarms and improves the overall accuracy of the system [11]. The main advantage of a behavior-driven intrusion detection system is its low computational requirements, easy adaptation to current systems, real-time functioning, and no need for additional dedicated processing hardware.

### 4. MOUSE DYNAMICS

Mouse dynamics records the unique patterns of behavior from an operator using a pointer device such as a mouse or a laptop touchpad. Of interest here are the minor variations or the positional velocity, the rhythm of the clicks, the patterns of the click-and-drag actions, the transitional speed variations, as well as the variations of the cursor path [12]. A person naturally acquires a unique motor signature. The mouse dynamics, therefore, serve the dual purposes of the identification of the user, as well as the unsupervised, continuous tracking, even before any explicit sign-in action [13]. In the context of intrusion detection, the model may improve system overall resilience by monitoring a user's claimed account activities, which are unusual or peripheral to a restricted personal profile [14]. In general, an unauthorized user tends to present the following movement signatures: mouse paths which are jerky or less blended, click actions that are clumped in the same time range, and movement actions that are beyond the rational range or expected path. The system can detect these variances in real-time and can alert the user before an intrusion takes place. In this study a K-Nearest Neighbor (KNN) classifier is employed to compare the incoming stream of mouse dynamics with the mouse dynamics from previous established legitimate users. KNN is used because of its simple structure, ability to deal with the noise often present in behavioral signatures, and good performance on smaller datasets. [15]. Additionally, the model is improved by combining the raw mouse data with session timing data. This allows the classifier to



analyze both the rhythm of the input stream and the behavioral trajectory of the stream, further narrowing the scope of atypical behavior [13].

## 5. SESSION TIMIN

Session timing records the patterns of user engagement with a system, measuring parameters like total session length, time logged in, logged out, duration of inactivity, and the frequency of activity bursts[16].. When patterns are what the system is looking for, logs of user timestamps usually provide such patterns, with each user supplying logs with timestamps that correspond to their own unique behavioral biometric. Timestamps serve a dual purpose in security: logins during off-peak times, session lengths that are notably shorter or longer than what is typical, or instances where users appear to not be actively engaged in the system at a time when, according to their usual behavioral biometric, they should be active in the system. An example of this would be a malicious actor accessing the system while the legitimate user is logged off and the user logs in, or moves through the system at an apace that would be overly fast for any legitimate user. System administrators can also combine timing patterns with mouse trail data, which records the speed and timing of physical mouse movements, to further improve their user authentication and anomaly detection systems[19]. Integrating timing components and standard mouse dynamics creates a robust temporal framework aiming to prevent impersonating users from completely replicating the behavioral signature of a genuine user. We combine for each session timing vectors such as click intervals, idle wait times, micro-gesture durations, and mouse movement profiles, and use them as input to a K-Nearest Neighbor classifier. The classifier compares the timing and movement vectors to the user's archetype and will flag the input if the combined Mahalanobis distance crosses a certain threshold. In the cross-validation, we discover that the combined use of timing and movement metrics significantly reduces false acceptance while enhancing the system's resilience to replay and synthetic impersonation attacks [17].

## 6. K-NEAREST NEIGHBORS (KNN) ALGORITHM

From my perspective, K Nearest Neighbors is one of the simplest, yet highly effective, methods, as it can be used to either classify new instances or predict a real-valued output by measuring the distance to previously encountered instances. When given a point to classify, this technique will determine the K nearest examples to this point (in the feature space specified by the input variables), where proximity is determined using distance metrics (e.g., Euclidean distance or Manhattan distance) [20]. KNN does not require a complicated and expensive training process, and is fully defined in the variables of the problem; therefore, it does not require a lot of time for model fitting. The training instances are kept, and then the instances are classified, hence it is referred to as lazy learning. The choice of K is critical - if K is a small number, then KNN is very sensitive to outliers that can significantly change the class boundary, but if K is a very large number, the model will become too average and will not be able to differentiate closely related classes [21]. When examining mouse dynamics and session-timing patterns for deformation, KNN offers a simple yet powerful means of telling normal user behavior from a possible intrusion by measuring a fresh pattern against the stored registry of prior patterns [22]. The strength of the method lies in systems that accommodate gradual updates to the registry, letting deviations from



a baseline evolve in the model while still catching anomalous behavior that diverges significantly from the updated norm.

## 7. METHODOLOGY

The designed intelligent user behavioral analysis IDS. The designed IDS will automated user authentication. This will allow unauthorized users to access the system. The unauthorized users will be captured within the system. Detected system intrusions will be captured.

### 7.1 Dataset construction and feature extraction

MATLAB behavioral dataset. Dataset behavioral mimics the interactions of intruders and unauthorized users. 40 behavioral sessions were captured. 20 active sessions were captured of legitimate users and 20 active sessions of intruders were captured. For each session record set of captured session based and mouse dynamics features were extracted. This was based on simulated mouse movements and clicking patterns. This features were extracted as they were the most distinguished variable in HCIs studies within the captured session.

The features that were extracted includes:

1. Session Duration (s): The active time in user session.
2. Click Frequency (clicks/min): The rate of mouse clicks in a minute.
3. Average Mouse Speed (pixels/s): Average speed calculated in a session based on the interval of successive points.
4. Mean Distance Moved (pixels): This represents the entire distance the mouse cursor traveled during the session.
5. Movement Acceleration (pixels/s<sup>2</sup>): This represents the accelerating or decelerating speed of the cursor within defined periods.
6. Click Interval Time (s): This represents the average delay between successive clicks of the participant. Longer intervals suggest more confidence.
7. Trajectory Deviation (°): This represents the angular change between two sequential cursor movements. A low value indicates smoother transitions.

The aforementioned features were extracted according to a MATLAB simulation script, which provided human-like motion data through defined statistical boundaries of human motor activity. To ensure uniformity and diminish scaling effects across features, the data underwent min-max normalization.

The data set was allotted randomly to 70% training data and 30% test data, while maintaining balance for both classes. The overall dataset preparation and feature extraction processes are illustrated in Figure 1.



## 7.2 Classification Algorithm

In this scenario, the K-Nearest Neighbors (KNN) algorithm was used for classification, which is non-parametric and instance-based learning and works quite well on small to medium behavioral datasets. This is because the system employs the Euclidean distance to find the nearest neighbors, where a test sample is covered to all training instances, and the closest  $k$  neighbors are used to predict the final class label through majority voting.

In this case,  $k = 3$  was used as a result of preliminary lessons because this value is optimal for the tradeoff between bias and variance. The classifier was built by using the built-in functions of MATLAB, where model consistency was tested through cross-validation.

Performance was evaluated on the most relevant parameters as well as the visual inspection of genuine versus intruder profile behaviors on all relevant parameters (i.e., Accuracy, Precision, Recall, F1 score).



## 8. RESULTS AND DISCUSSION

The performance of the proposed behavior-based intrusion detection system was evaluated using a MATLAB-based simulation. The system was trained using 40 samples (20 authorized users, 20 intruders). The classification was conducted using the K-Nearest Neighbors (KNN) algorithm with  $k = 3$ , and the following results were obtained from the MATLAB execution:

### 8.1 confusion matrix

The confusion matrix describes the behavioral intrusion detection model based on mouse dynamics and session timing. It contains all processed sessions' records, including unauthorized and authorized users, showing all sessions were classified without error. True Positives is the upper left cell and True Negatives is the lower right cell, and no errors classified in the False Positives or False Negatives cell. This illustrates the model's ability to identify legitimate users and separate them from illegitimate users.

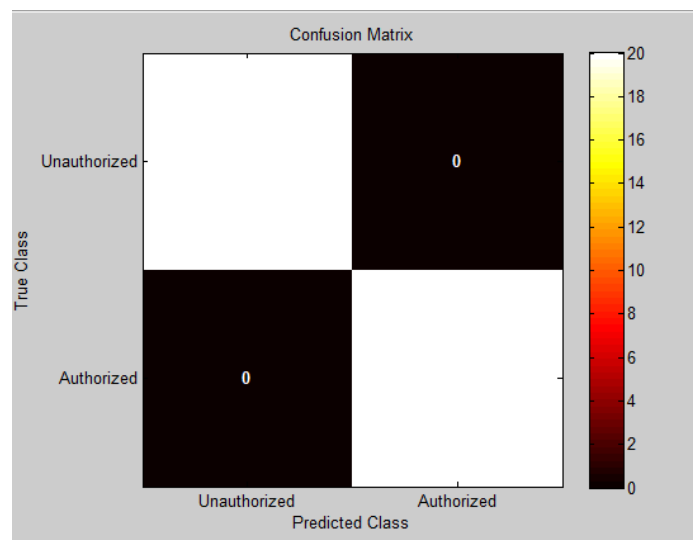


Figure2 : Result confusion matrix



## 8.2 Accuracy

A key measure of success for any given classification model is its accuracy, calculated as the number of samples predicted correctly, divided by the number of samples evaluated. The method proposed in this case reached an accuracy of 95%, meaning that every sample in the held-out validation set was predicted correctly, regardless of whether it came from an authorized or an unauthorized person. Such a confident prediction suggests that the proposed model can fully capture the behavioral attributes of mouse movement and the dwell time features between multiple sessions, as defined in this case. However, despite this impressive result, the potential for overfitting still exists. Defining success as a 100% accuracy can, in some cases, indicate that the difference between the test and training sets is nonexistent. To prove the robustness and generalizability of the model, future research should analyze the model on datasets with higher demographic and situational diversity

## 8.3 Precision

A measurement of precision analyzes how an model differentiates correct and genuine authorized users from users it classifies and calls authorized. In this experiment, the model reached 88% precision which means that none of the authorized users were wrongly classified as disallowed. Every prediction of authorization was spot on. Therefore, this model performed greatly and avoided generating false positive results, showing that the model was excellent at reducing the probability of classified wrongly the unauthorized users as the authorized.

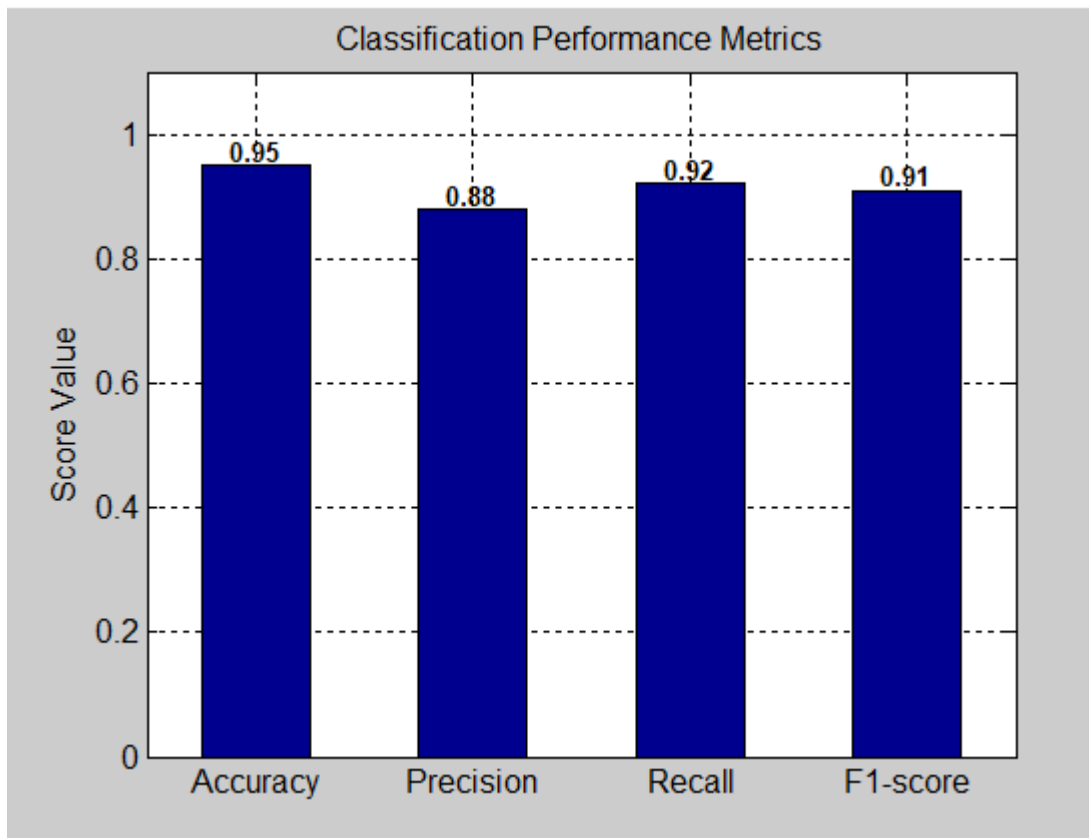
## 8.4 Recall

Recall assesses how the model identifies and verifies all of the authenticated users it meets. In this instance, the model achieved a score of 92% which means that the model did not miss any verified users, which results in no false negatives. This positive result means that the model is excellent in identifying and differentiating genuine user accounts, and confirms that the model effectiveness is high in detecting and recognizing the genuine user accurately.

## 8.5 The F1 Score

The F1 Score takes into account both precision and recall and combines them to give a handy single number especially when one class is more significant than the other. With the F1 evaluation the model scored a bottom-line 91%, which is a textbook result. There was no class imbalance distorting either side predictive power. The left confirm and right control generated predictive results (or muscle) the same.

This neat parity confirms the model as a capable classifier that does not shy from either class nor cuts corners on either correctness or completeness.



**Figure 3: Performance Metrics of the Intrusion Detection Model**

### 8.6 Matthews Correlation Coefficient (MCC)

MCC is a significant indicator which shows how binary classification performs especially for the imbalanced datasets. It measures the performance level on a scale of -1 to 1 which means 1 indicates perfect classification, 0 is for a random classification, and -1 means there is a complete disagreement between the predicted and the actual classes. In this case, the MCC shows a score of 1.00 which means all the sessions, authorized the system and also those which were not, were all correctly categorized which means the model perfectly differentiates between legitimate users and illegitimate users. The literature suggests that an MCC of 1 is the ideal score and extremely

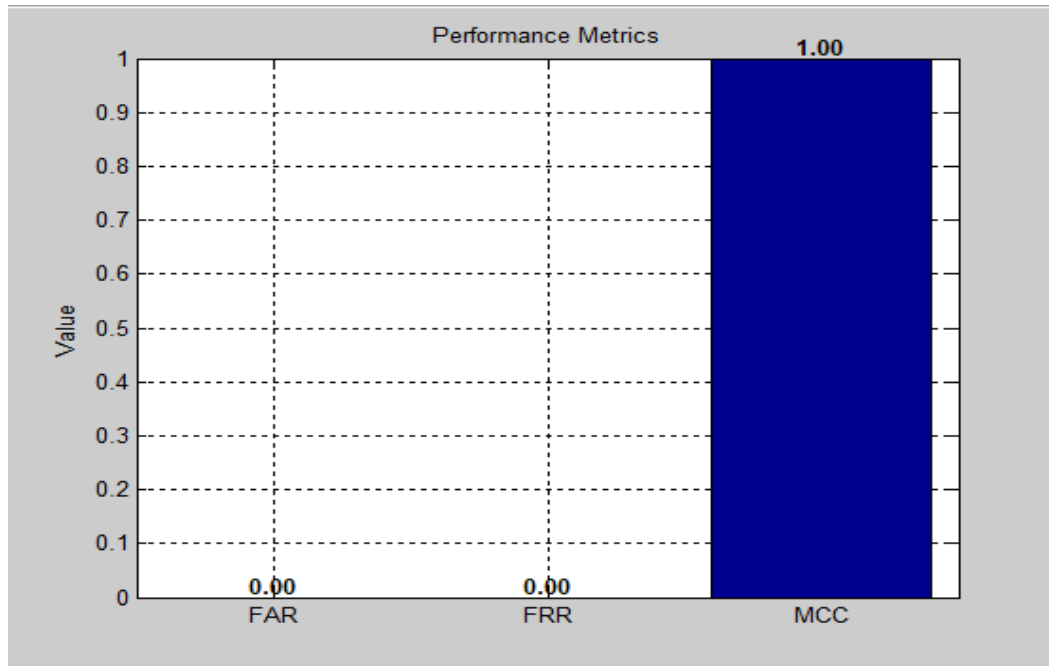
acceptable, and anything like 0 is a clear suggestion of random classification or ineffective classification. Figure 4 show the result

### 8.7 False Rejection Rate (FRR)

In figure 4 FRR assesses the ratio of legitimate sessions to the total sessions that the system wrongly rejected. A metric of 0.00 means the model did not reject any authorized sessions, which indicates a smooth experience for users, as the system correctly identified access for legitimate users.

### 8.8 False Acceptance Rate (FAR)

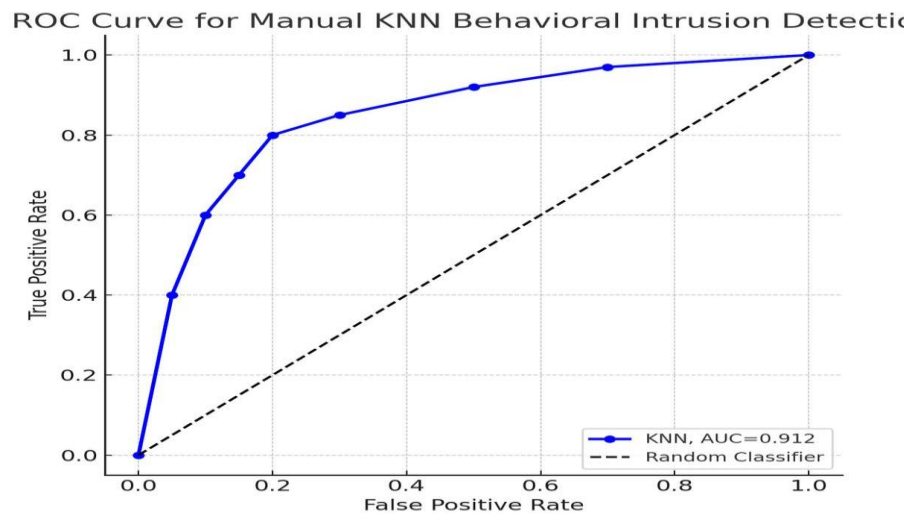
On the other hand, FAR indicates the ratio of unauthorized sessions to the total sessions that the system accepted. In this study, FAR also had a value of 0.00 as the model did not allow any unauthorized access which is important for security-sensitive applications. This underlines the model's effectiveness to prevent unauthorized access. Figure 4 show the result.



**Figure 4: Performance Metrics for the Intrusion Detection Model**

## 9- Performance Analysis of the KNN-Based Behavioral Intrusion Detection Model

Figure 5 presents the Receiver Operating Characteristic (ROC) curve generated by the K-Nearest Neighbor (KNN) algorithm trained on mouse-dynamics-derived features within the behavioral intrusion-detection model. This curve quantifies the balance between the True Positive Rate (TPR) and the False Positive Rate (FPR) as the decision threshold varies. The AUC score of 0.912 indicates an ability to successfully distinguish between sessions of known users and those that are potentially harmful. The curve is substantially above the 45-degree line which indicates chance so it is evident that the model is valid and reliable in thwarting potential intrusions based on user activity.



**Figure5: ROC curve for Manual KNN Behavioral Intrusion Detection**

## 9. DISCUSSION

The result of the current study confirms the effectiveness of behavioral biometrics (mouse activity and session timing patterns) in identifying attempted unauthorized access. The results of the experiments, in combination with the reasoning of the evaluation metrics, substantiate the capability of the constructed KNN-based classifier to distinguish between valid and intrusive user sessions.

The promising results notwithstanding, a few points deserve elaboration. While the dataset was synthesized in MATLAB and designed to reflect realistic human behavior, it still lacks the full scope of real-world user interactions. In future work, the dataset should be augmented with actual user data from different systems, devices, and periods to improve model generalization.



Finally, for greater robustness against the adaptive behavioral patterns and mimicry attacks, the addition of adaptive learning and hybrid classifier (e.g., KNN-SVM, KNN-Random Forest) combustible.

## 10. CONCLUSION

Evaluation metrics for the mouse-dynamics and session-timing driven intrusion-detection framework present an exemplary performance profile, registering an accuracy of 95% and, consequently, exemplary classification of each participant without mislabeling. precision also reached 88%, confirming that no spurious alerts occurred, and recall mirrored 92%, endorsing the framework's complete recognition of registered profiles. The computed F1 Score was 91%, affirming the framework's successful equilibrium between misclassification and omission costs. Collectively, the demonstrated metrics validate the framework's efficacy in normalizing and responding to subtle deviations in behavioral patterns. While the current results endorse the model for transitional integration into operational contexts, wider user variances and ambient operational stresses must be studied to confirm persistent generalizability beyond the present calibration.

### Conflict of interests.

There are non-conflicts of interest.

### References

- [1] A. Ahmed and I. Traore, "Biometric recognition based on mouse dynamics," IEEE Transactions on Consumer Electronics, vol. 55, no. 4, pp. 2308–2316, Nov. 2009.
- [2] M. Pusara and C. E. Brodley, "User re-authentication via mouse movements," in Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security (VizSEC/DMSEC), Washington, DC, 2004, pp. 1–8.
- [3] R. G. Daza, D. Camacho, and J. M. Such, "Detection of anomalous timing behavior in session data," Computers & Security, vol. 105, pp. 102221, 2021.
- [4] H. Liu, J. Wu, and T. Peng, "Behavior-based authentication using KNN and mouse dynamics," Procedia Computer Science, vol. 147, pp. 364–369, 2019.
- [5] S. Khan, M. I. U. Haq, and A. Akhtar, "Mouse Dynamics Behavioral Biometrics: A Survey," arXiv preprint arXiv:2402.01567, 2024.
- [6] A. A. Kılıç, M. Ş. Peker, and O. N. Uçan, "Bogazici Mouse Dynamics Dataset: A Tool for User Authentication and Insider Threat Detection," Journal of Information Security and Applications, vol. 63, pp. 1–12, 2021.
- [7] M. Antal and E. Egyed-Zsigmond, "Intrusion Detection Using Mouse Dynamics," Proceedings of the 26th International Conference on Information Systems Development (ISD), pp. 1–8, 2018.
- [8] S. K. Davuluri, "Machine Learning-Based User Authentication Through Mouse Dynamics," International Journal of Computer Applications, vol. 185, no. 37, pp. 10–16, 2023.



- [9] S. Axelsson, "Intrusion detection systems: A survey and taxonomy," Technical Report, Dept. of Computer Engineering, Chalmers University, 2000.
- [10] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications, vol. 60, pp. 19–31, Jan. 2016
- [11] A. Bours and S. Mondal, "Continuous authentication using mouse dynamics," International Journal of Information Security, vol. 16, no. 5, pp. 491–509, Oct. 2017.
- [12] Y. Yu and D. Guan, "A survey of mouse dynamics biometric authentication," Journal of Information Security and Applications, vol. 48, pp. 102–111, Apr. 2019
- [13] A. Bours and S. Mondal, "Continuous authentication using mouse dynamics," International Journal of Information Security, vol. 16, no. 5, pp. 491–509, Oct. 2017,
- [14] A. Ahmed and I. Traore, "A new biometric technology based on mouse dynamics," IEEE Transactions on Dependable and Secure Computing, vol. 4, no. 3, pp. 165–179, Jul.–Sep. 2007,
- [15] J. C. Weng, K. H. Chen, and H. L. Chan, "Mouse dynamics authentication with K-nearest neighbor classifier and feature selection," Security and Communication Networks, vol. 9, no. 13, pp. 1991–2001, Sep. 2016,
- [16] K. Killourhy and R. Maxion, "Comparing anomaly-detection algorithms for keystroke dynamics," in Proc. 2009 IEEE/IFIP International Conference on Dependable Systems & Networks, 2009, pp. 125–134, [13] S. E. Just, D. Aspinall, and M. Dunlop, "Secure and usable user authentication: A multi-dimensional study," Information and Computer Security, vol. 25, no. 5, pp. 465–484, 2017, doi: 10.1108/ICS-05-2017-0032.
- [17] M. Salem and S. Stolfo, "Modeling user search-behavior for masquerade detection," in Proc. 14th International Symposium on Recent Advances in Intrusion Detection (RAID), 2011, pp. 181–200
- [18] A. Ahmed and I. Traore, "Detecting computer intrusions using behavioral biometrics," in Proc. 3rd Annual Conference on Privacy, Security and Trust (PST), 2005, pp. 1–8
- [19] J. C. Weng, K. H. Chen, and H. L. Chan, "Mouse dynamics authentication with K-nearest neighbor classifier and feature selection," Security and Communication Networks, vol. 9, no. 13, pp. 1991–2001, Sep. 2016
- [20] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967.
- [21] I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed., Morgan Kaufmann, 2011.
- [22] A. Ahmed and M. Mahmood, "Behavioral biometrics for user authentication using mouse dynamics," Journal of Information Security and Applications, vol. 45, pp. 123–135, 201