Machine and Deep Learning Applications to Mouse Dynamics for Continuous User Authentication

AIMS Research Group

Nyle Siddiqui | Department of Computer Science

Advisor: Dr. Rushit Dave

University of Wisconsin Eau Claire

Introduction

- Static authentication methods, like passwords and PINs, authenticate a user once and only once. Continually authenticating a user even after initial access can drastically increase account security against imposters.
- Unique behaviors, such as mouse movements, are distinct and varied enough between humans to be irreproducible. This makes them a viable biometric to utilize for user authentication.
- Machine and deep learning have exploded in popularity due to their superior ability to process large amounts of data. We train and evaluate three machine learning and three deep learning algorithms on our own novel mouse dynamics dataset.

Methodologies

Our novel dataset contains mouse movement data from 40 users as they played the video game *Minecraft (Fig. 1*) for 20 minutes.

Machine Learning

- A Random Forest (RF), K-Nearest Neighbor (KNN), and a Support Vector Machine (SVM) were the three machine learning models tested.
- Raw mouse movement data is not sufficiently granular for machine learning algorithms to accurately perform on. Thus, raw mouse data is combined into groups of 10 , and additional features and summary statistics are extracted from these groups.

Deep Learning

- A 1-Dimensional Convolutional Neural Network (1D-CNN), Long Short-Term Memory Recurrent Neural Network (LSTM-RNN), and an Artificial Neural Network (ANN) were the three deep learning models tested.
- No preprocessing was required for the deep learning algorithms. Thus, the raw mouse movement data of each group of 10 *(Fig. 3)* was sufficient as input for the model.

Results

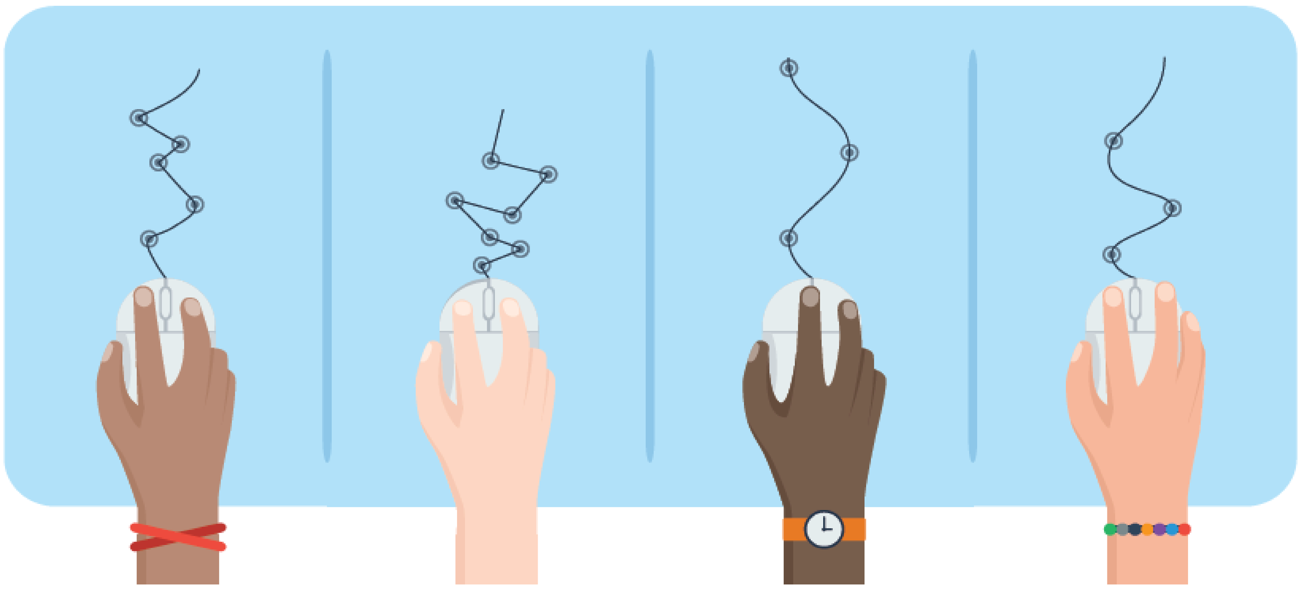
For brevity, results of the top-10 users are reported. Evaluation metrics consisted of Accuracy (ACC), False Positive Rate (FPR), and F1-Scores. Results of the best deep learning and machine learning algorithm are shown in Table 1 & 2 the ROC Curve for the machine learning algorithm is shown in Figure 5.

Future Work

Limitations are important to address, especially in a relatively niche area like mouse dynamics, to encourage continued research and developments. Future work could include evaluating larger deep learning models, creating larger group sizes during the preprocessing phase, and/or improved hyperparameter tuning.

Acknowledgements

Special thanks to the High-Performance Computing Cluster at UWEC, specifically for the generous access to the BOSE supercomputing cluster for training the models in this research.



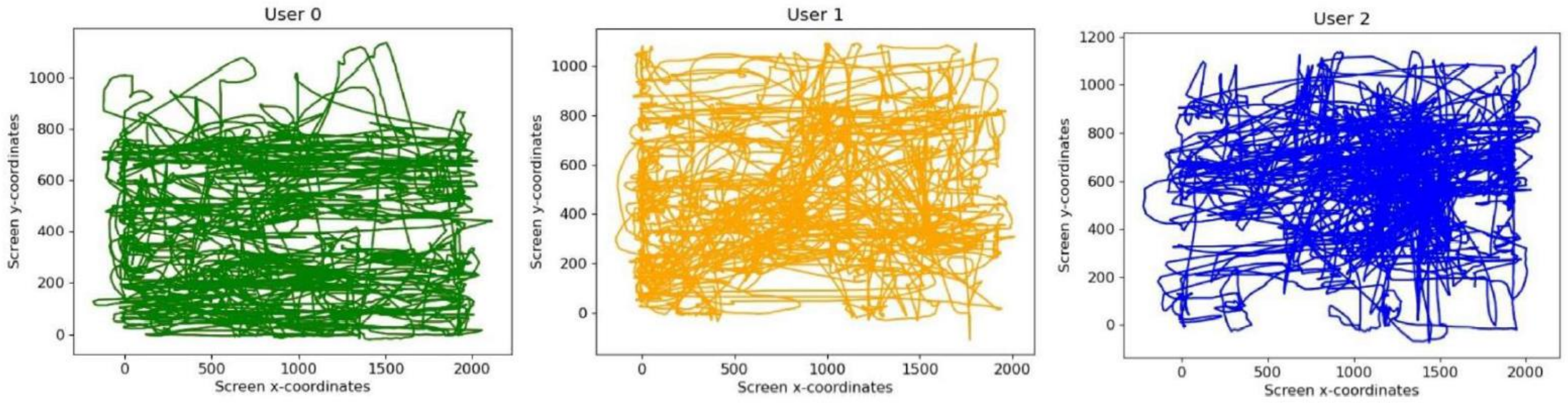


Fig 1. Raw mouse movement maps from three users' gaming sessions

Table 1. Results from the Random Forest.

User	ACC	FPR	FNR	F1 Score
8	0.6827	0.2582	0.3764	0.7016
12	0.6740	0.3076	0.3442	0.6739
35	0.6580	0.2732	0.4108	0.6564
11	0.6575	0.3692	0.3158	0.6673
6	0.6548	0.3171	0.3734	0.6745
34	0.6507	0.2820	0.4165	0.6501
32	0.6408	0.3923	0.3259	0.6401
30	0.6355	0.3761	0.3569	0.6575
1	0.6305	0.2454	0.4937	0.6247
18	0.6212	0.4022	0.3553	0.6210
Average	0.6506	0.3223	0.3768	0.6567
Standard Deviation	0.01912	0.05838	0.05232	0.02440

Table 2. Results from the 1D-CNN.

User	ACC	FPR	F1 Score
13	0.8637	0.1563	0.9197
24	0.8634	0.1602	0.9202
2	0.8587	0.1839	0.8551
39	0.8581	0.1588	0.9181
15	0.8577	0.1607	0.9225
30	0.8575	0.1194	0.9191
19	0.8567	0.1513	0.9171
14	0.8525	0.1652	0.8934
0	0.8524	0.1488	0.9182
9	0.8521	0.1419	0.9151
Average	0.8573	0.1546	0.9099
Standard Deviation	0.004156	0.01668	0.02091

We thank the Office of Research and Sponsored Programs for supporting this research, and Learning & Technology Services for printing this poster.