Use of Hidden Markov Model for keystroke biometric studies

Dibyajyoti Nayak
Mahesh Tom

IISc Bangalore

02-05-2022

Problem Definition

- Classification of a user as genuine/impostor by modeling user typing behaviour with a Partially Observable Hidden Markov Model (POHMM)
- Benchmarking performance of POHMM with HMM on the same dataset
- Implementation of an alternate model for modeling keystroke behaviour of a user that shows better performance over POHMM

Technical details

- The CMU keystroke benchmark dataset that contains keystroke information for 51 users is used for model training.
- A POHMM is formally defined as:
  $P(x_1^{n+1}, z_1^{n+1}) = P(x_1^n, z_1^n) P(x_{n+1}|z_{n+1}, \Omega_{n+1}) P(z_{n+1}|z_n, \Omega_n, \Omega_{n+1})$
  where Ω is event type, x is emission and z is the hidden state.

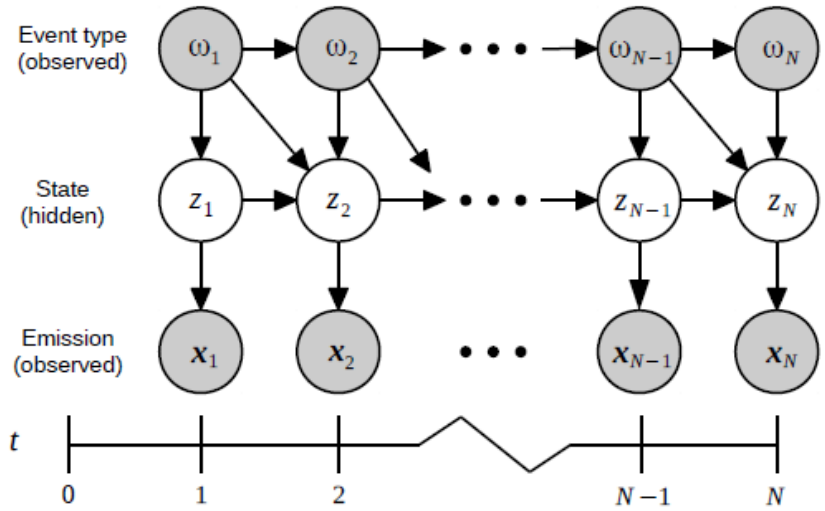

Figure 1: POHMM representation

- In the current study, each event type is taken as a character of the password **".tie5Roanl"**

Results

- An average accuracy of 96.4% is obtained with the POHMM model (convergence in 22 iterations)
- POHMM has a lower False Acceptance Rate (FAR) and False Rejection Rate (FRR) than the HMM model. This is observed in the ROC curve as shown below:

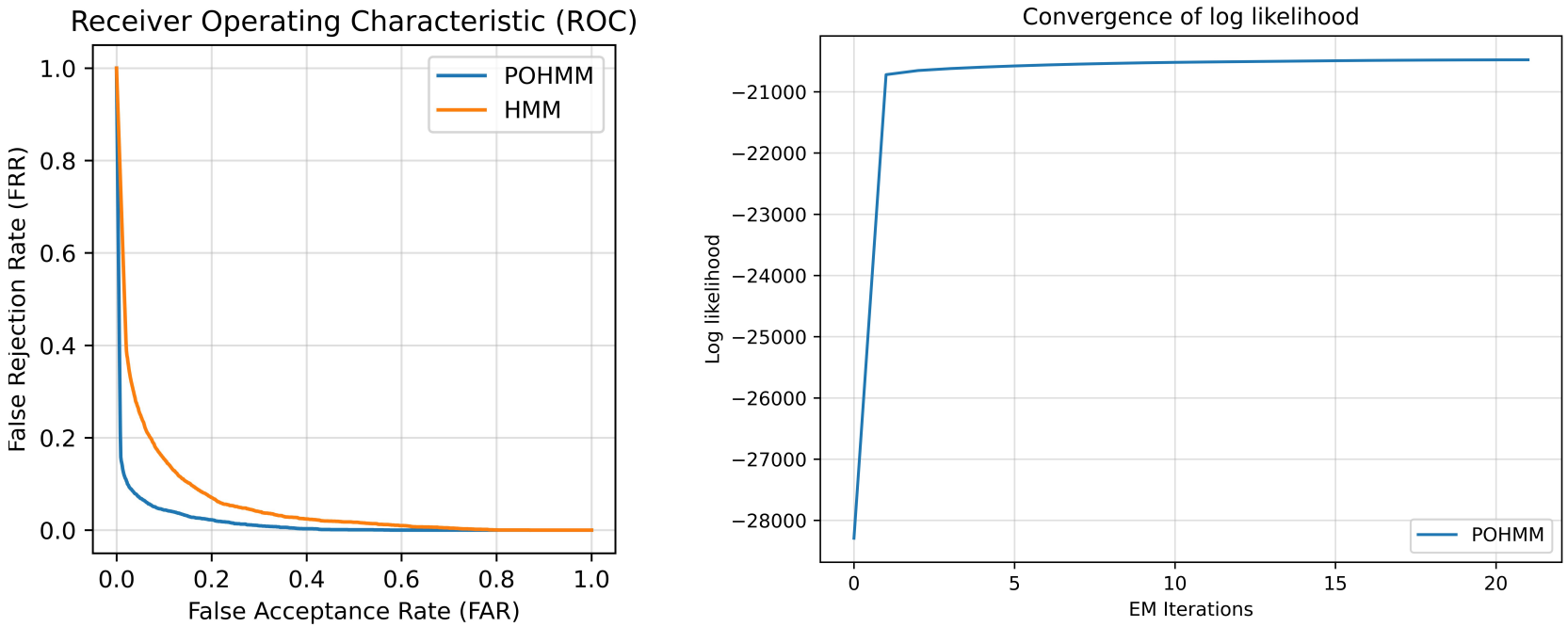

Figure 2: Comparison of ROC curves for HMM and POHMM (left); Convergence of log-likelihood for POHMM (right)

Novel Contributions

- Modeling keystroke behaviour for each user with a Deep Neural Network (DNN) which gives a higher accuracy (99.1%) compared to a generative models like POHMM and model in Reference 4.
- Training the POHMM separately for different users is computationally expensive. In order to mitigate this problem, we parallelised the POHMM training process.

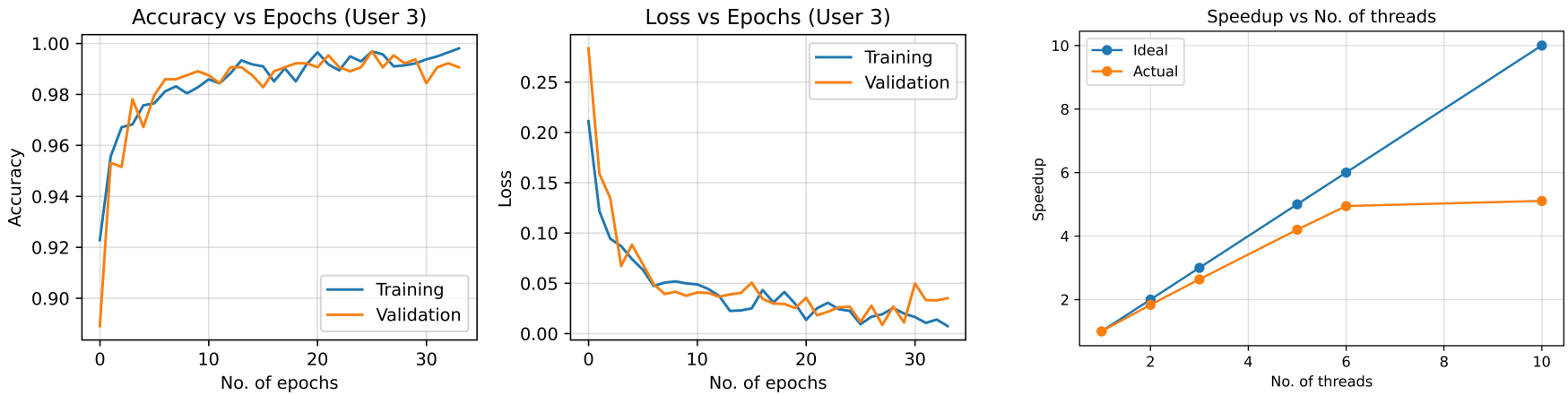

Figure 3: Accuracy and loss curves obtained during training of DNN(left, middle); Strong scaling performance with thread-level parallelism (right)

References

- The Partially Observable Hidden Markov Model and its Application to Keystroke Dynamics - John V. Monaco, Charles C. Tappert
- Keystroke Biometric studies with Hidden Markov Model and Its Extension on Short Fixed-Text Input - Md Liakat Ali,John V. Monaco,Charles C. Tappert
- Hidden Markov Models in Keystroke Dynamics - Md Liakat Ali, John V. Monaco, and Charles C. Tappert
- Comparing anomaly-detection algorithms for keystroke dynamics - K. S. Killourhy, and R. A. Maxion