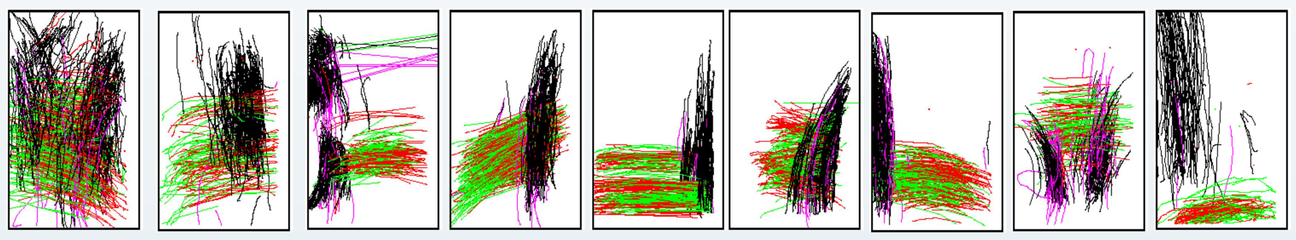
Touchalytics
On the Applicability of Touchscreen Input as a Behavioral Biometric for Continuous Authentication

Mario Frank, Ralf Biedert, Eugene Ma, Ivan Martinovic, Dawn Song
IEEE Transactions on Information Forensics and Security (Vol. 8, No. 1), pages 136-148, Jan 2013

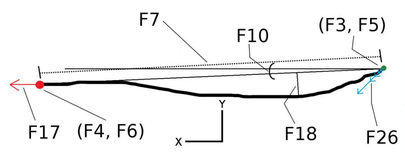


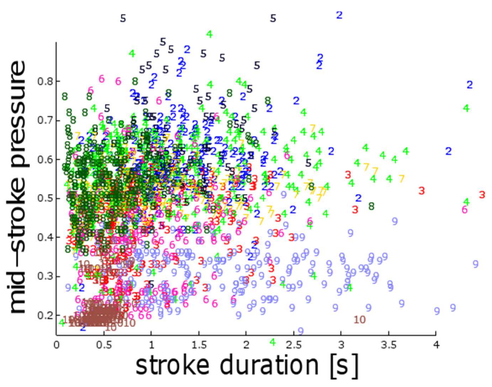
Summary

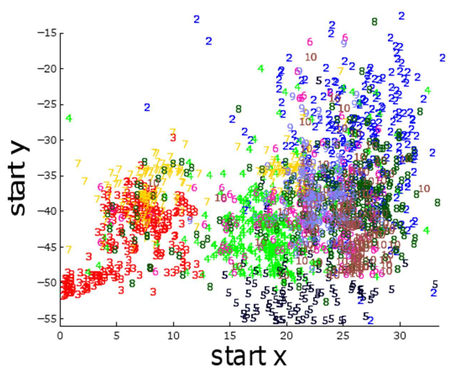
- **Users exhibit unique patterns when interacting with a touch screen**
- **Can this be used for authentication beyond password entry?**
- **Experimental study with 41 users**
- **30 behavioral features from 11 strokes achieve an equal error rate (EER) of 2-3% for authentication across sessions.**

Features

- Extract 30 features from each stroke (see list)

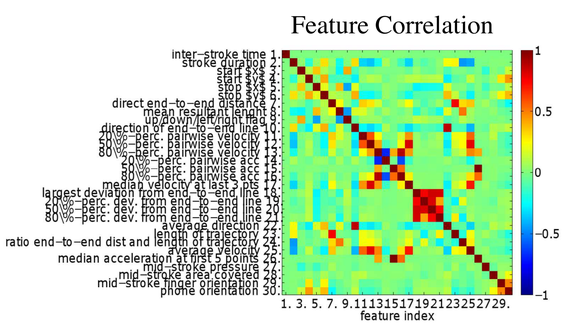






Feature selection

Rel. mutual information	Feature description
20.58%	mid-stroke area covered
19.63%	20%-perc. pairwise velocity
17.28%	mid-stroke pressure
11.06%	direction of end-to-end line
10.32%	stop x
10.15%	start x
9.45%	average direction
9.43%	start y
8.84%	average velocity
8.61%	stop y
8.5%	stroke duration
8.27%	direct end-to-end distance
8.16%	length of trajectory
7.85%	80%-perc. pairwise velocity
7.24%	median velocity at last 3 pts
7.22%	50%-perc. pairwise velocity
7.07%	20%-perc. pairwise acc
6.29%	ratio end-to-end dist and length of trajectory
6.08%	largest deviation from end-to-end line
5.96%	80%-perc. pairwise acc
5.82%	mean resultant lenght
5.42%	median acceleration at first 5 points
5.39%	50%-perc. dev. from end-to-end line
5.3%	inter-stroke time
5.14%	80%-perc. dev. from end-to-end line
5.04%	20%-perc. dev. from end-to-end line
5.04%	50%-perc. pairwise acc
3.44%	phone orientation
3.08%	mid-stroke finger orientation
0.97%	up/down/left/right flag
0%	change of finger orientation



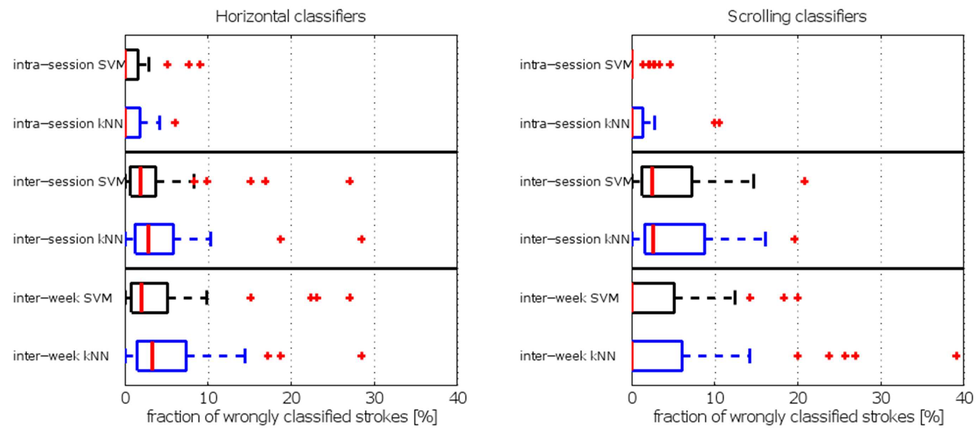
Data collection

- 41 users interacting with our App
- Read three **Wikipedia** articles
- Answer **questionnaire** (interrupt session) between each article
- Play two rounds of "**find the difference**"

download data from:
http://www.mariofrank.net/touchalytics/index.html

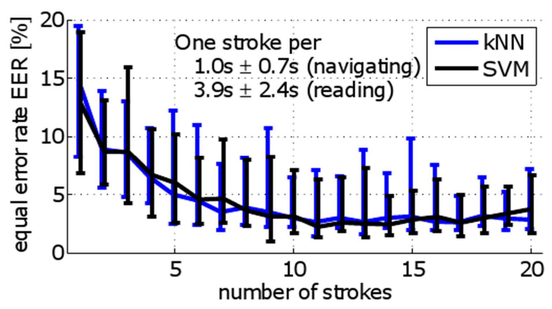
Accuracy

- **Classifiers**: kNN and a one-versus-all rbf-SVM
- 3 scenarios: Intra-session, inter-session (3 min break), inter-week (1 week break)
- Separate classifiers per stroke category: horizontal, vertical (could be extended by clicks)



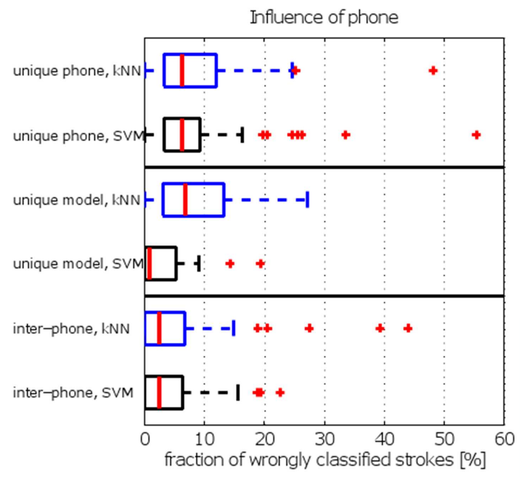
Influence of refresh rate

- Take **majority vote** across x strokes
- No substantial improvement after **11 strokes**

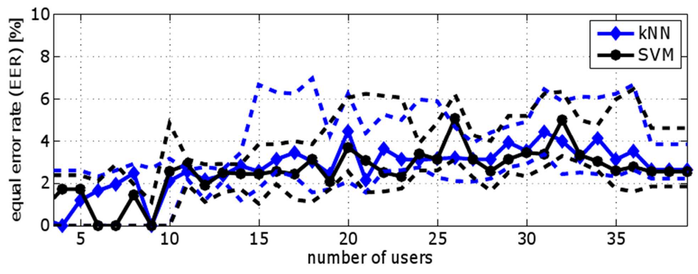


Experimental pitfalls

- Different phone/OS -> different behaviour?
- Different experimenter -> different behaviour?
- Too few degrees of freedom for individual users?
- **Challenge**: keep inter-experiment variability low, keep intra-experiment variability large



- Influence of **sample size**



Deployment

- Accuracy insufficient for standalone authentication across weeks.
- Standalone **theft-detection** (integrate over hours/days) in addition to PIN
- **Combine** with other modalities (gait, content behaviour, GPS, …)
- Activate only in **low-risk scenarios** (detect low risk!), otherwise fall back to PIN