CS259D: Data Mining for Cyber Security

Re-authentication: Practical requirements

- Accuracy
- Quick response
- Difficult to forge

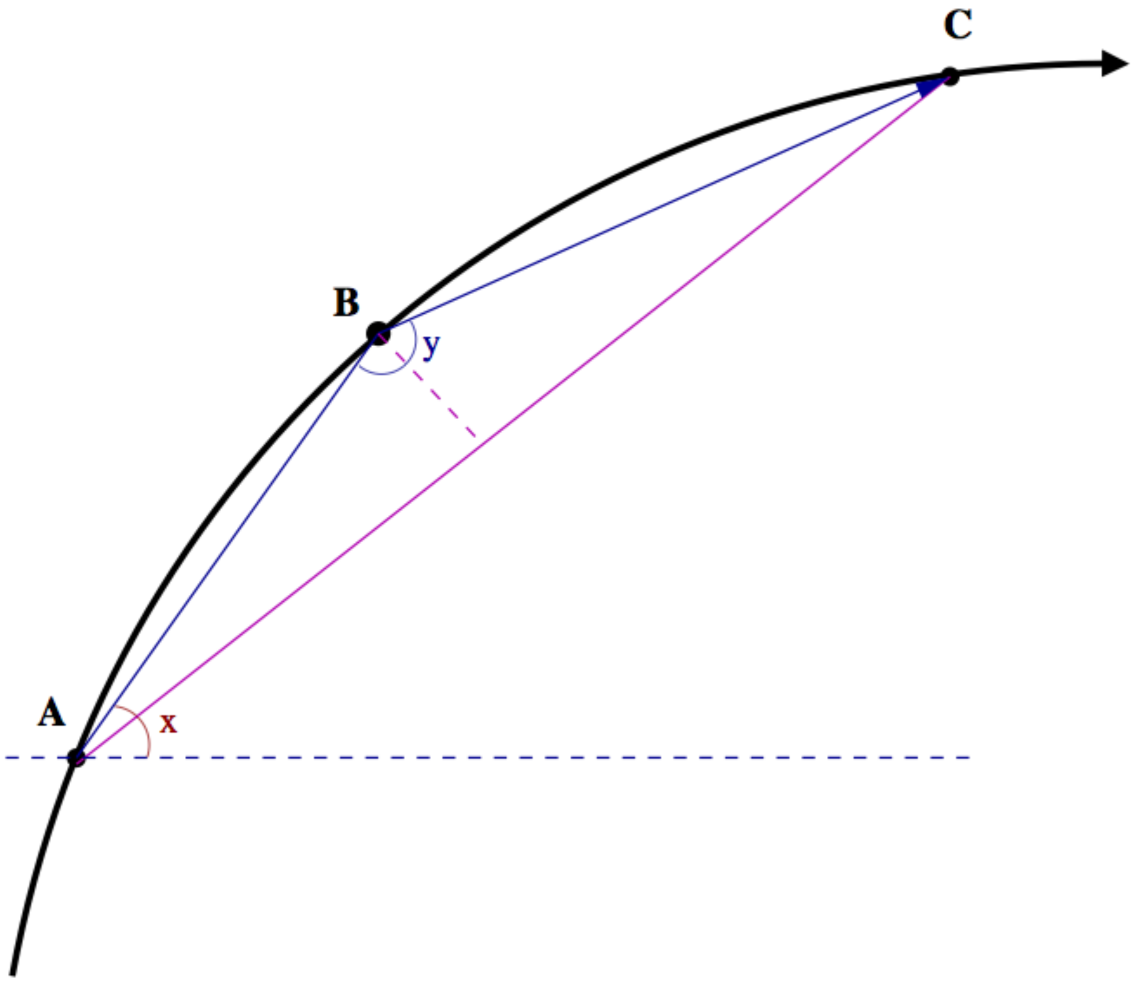
Data

- Controllable set
  - 30 controllable users
- Field set
  - 1000 real field users
- Raw data: <ACTION-TYPE, t, x, y>
  - ACTION-TYPE: mouse-move, mouse click
- Data preprocessing
  - Identify every point-and-click action
    - Continuous mouse movement followed by click

Metrics

- Direction
  - For consecutive points A, B: angle between line AB and horizontal line
- Angle of Curvature
  - For any three consecutive points A, B, C: angle between AB and BC
- Curvature Distance
  - For any three consecutive points A, B, C: ratio between length of AC to length of perpendicular distance from B to AC

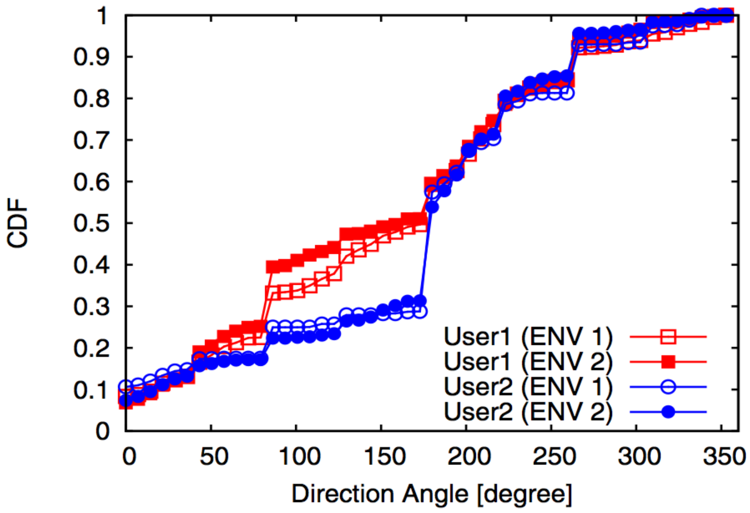
Metrics

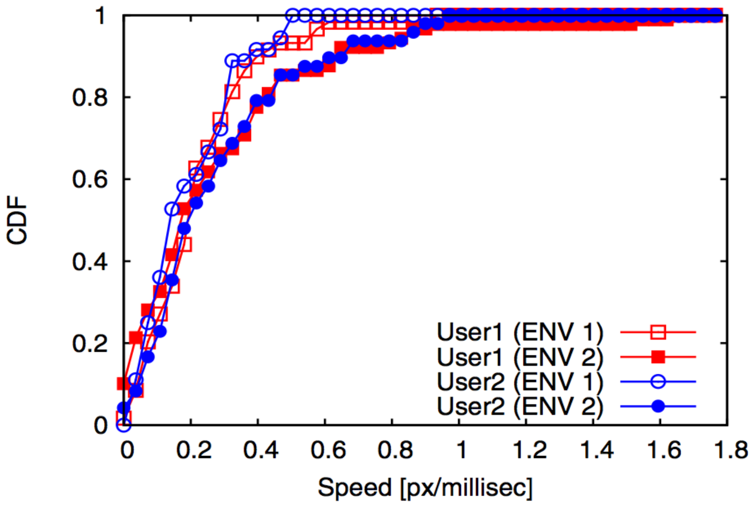


Mouse movement characterization

- Dependence on different platforms
  - OS, screen size & resolution, mouse pointer sensitivity, brand of mouse, desk space available near mousepad
  - Affects measurements such as speed, acceleration
- Uniqueness of angle-based metrics across users

Mouse movement characterization





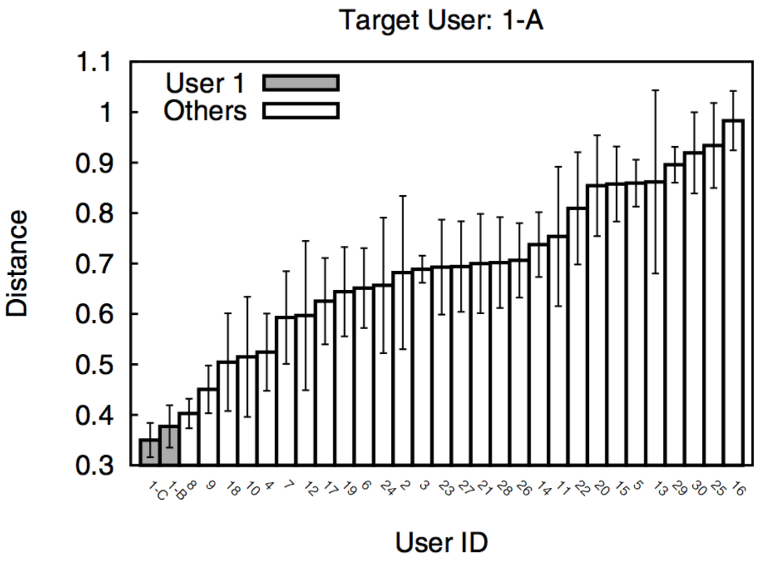
Distance between distributions

- Binned PDFs: $\{p_1, p_2, \ldots, p_n\}, \{q_1, q_2, \ldots, q_n\}$
- Distance:

$$\sum_i | p_i - q_i |$$

Distance between distributions

Setting	Machine Type	Mouse Type
1-A	Dell Precision T3500	Dell MOC5UO Two-Button Scroll-Wheel
1-B	Apple Macbook MB990LL/A	Apple A1152 One-Button Trackball
1-C	Apple Macbook MB990LL/A	Dell MOC5UO Two-Button Scroll-Wheel

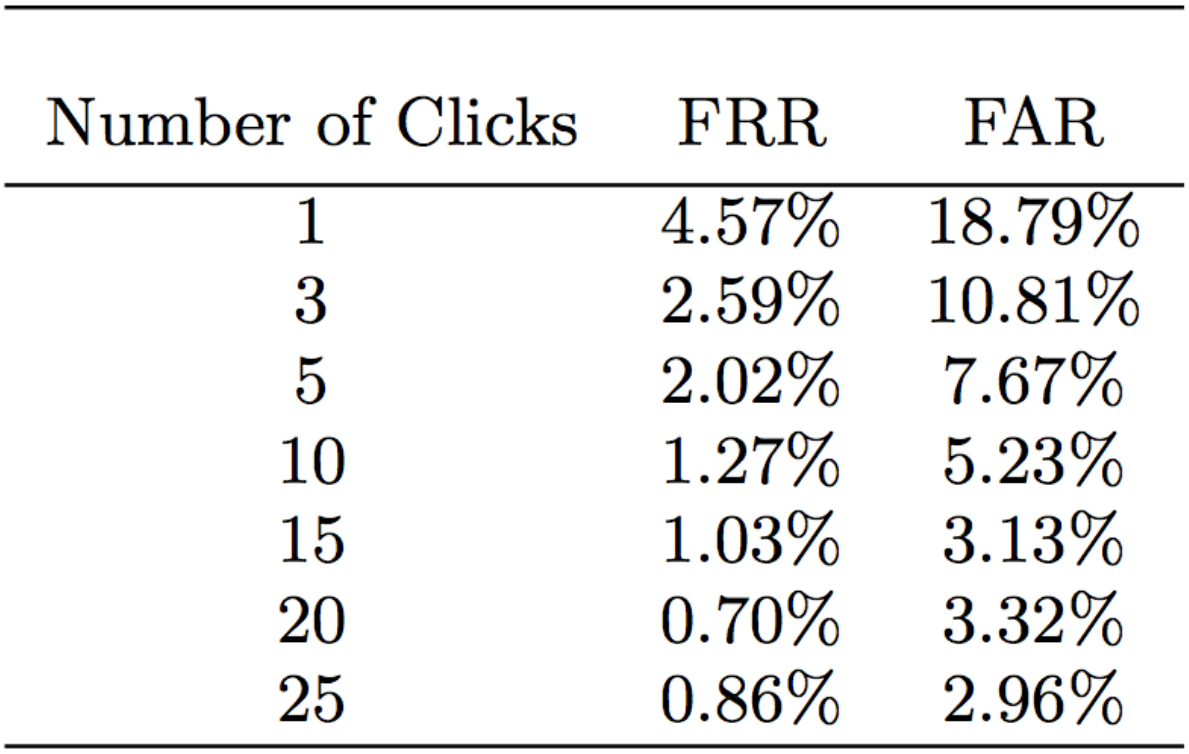


Classifier

- 2-class SVM
- RBF kernel
- Decision:
    - Threshold
    - Majority vote
        - Multiple models using sampled data

Results: discrimination in same environment

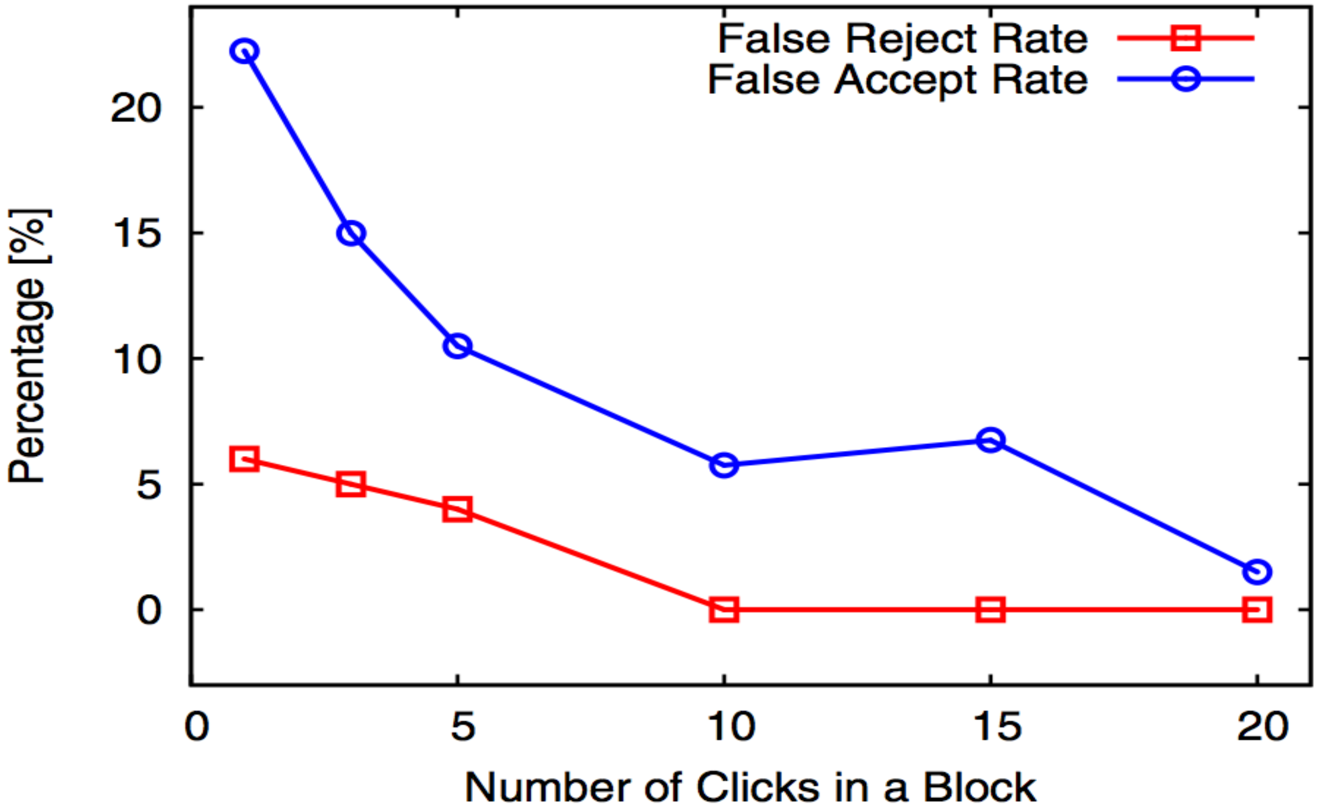
- 500 training blocks, threshold 0.5, 3/5 majority



Number of Clicks	FRR	FAR
1	4.57%	18.79%
3	2.59%	10.81%
5	2.02%	7.67%
10	1.27%	5.23%
15	1.03%	3.13%
20	0.70%	3.32%
25	0.86%	2.96%

Results: discrimination in different environments

- Train model on data from a work desktop
- Test on data from a home laptop



Results: Partial movements

- Continuous mouse movements without ending in a click
- Compare to point-and-clicks
  - More noisy
  - Much more frequent
    - 0.53 mouse clicks per minute
    - 6.58 partial movements per minute

	Equal Error Rate	Verification time
Point-and-click	1.3%	38 minutes
Partial movement	**1.9%**	**3 minutes**

Research problem

- Angel-based metrics + frequent patterns

Administrativia

- No class on Thursday
  - Work on homework instead ☺

Entry-point authentication on mobile devices

- Usability
  - Inconvenient for quick activities
- Security
  - Short passwords
  - Increased screen lock time-outs
  - Disable unlock
  - Higher risk of theft

Trigger actions

- Sliding horizontally over the screen
  - Browse through images
  - Navigate to next page of icons
- Sliding vertically over the screen
  - Reading email, documents, webpages
  - Browsing menus

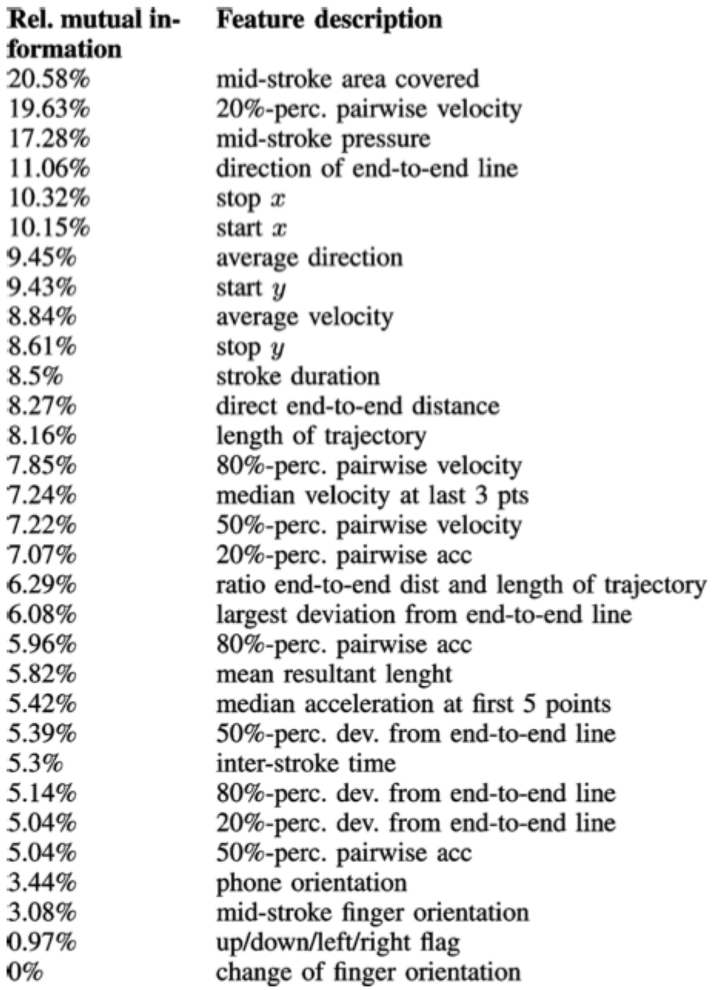
Data acquisition

- Android phones
- Tasks: read documents, compare images
- Raw features:
  - Event code (e.g., finger up, finger down, finger move, multi-touch)
  - Event time
  - Device orientation
  - x, y coordinates of finger
  - Finger pressure
  - Area on the screen covered by the finger
  - Finger orientation with respect to screen orientation

Stroke

- Sequence of touch data starting with touching the screen, ending with lifting the finger
- Sequence of vectors:
  - $s_n = (x_n, y_n, t_n, p_n, A_n, o_n^f, o_n^{ph}); (1 \leq n \leq N)$

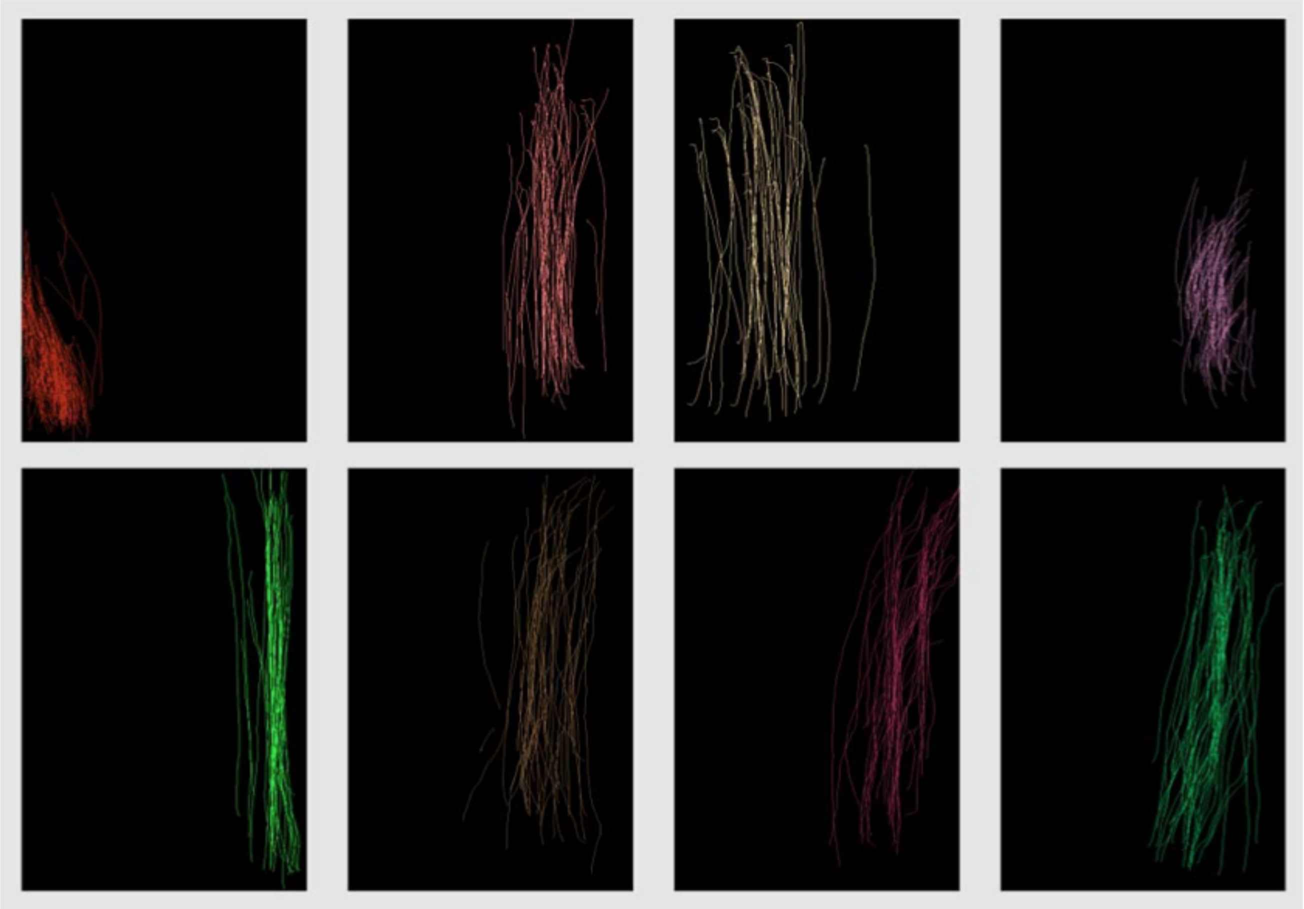
30 Features



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

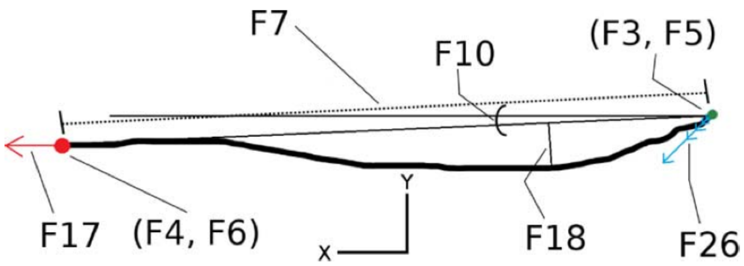
Example feature

- Coordinates of the two endpoints



Example feature

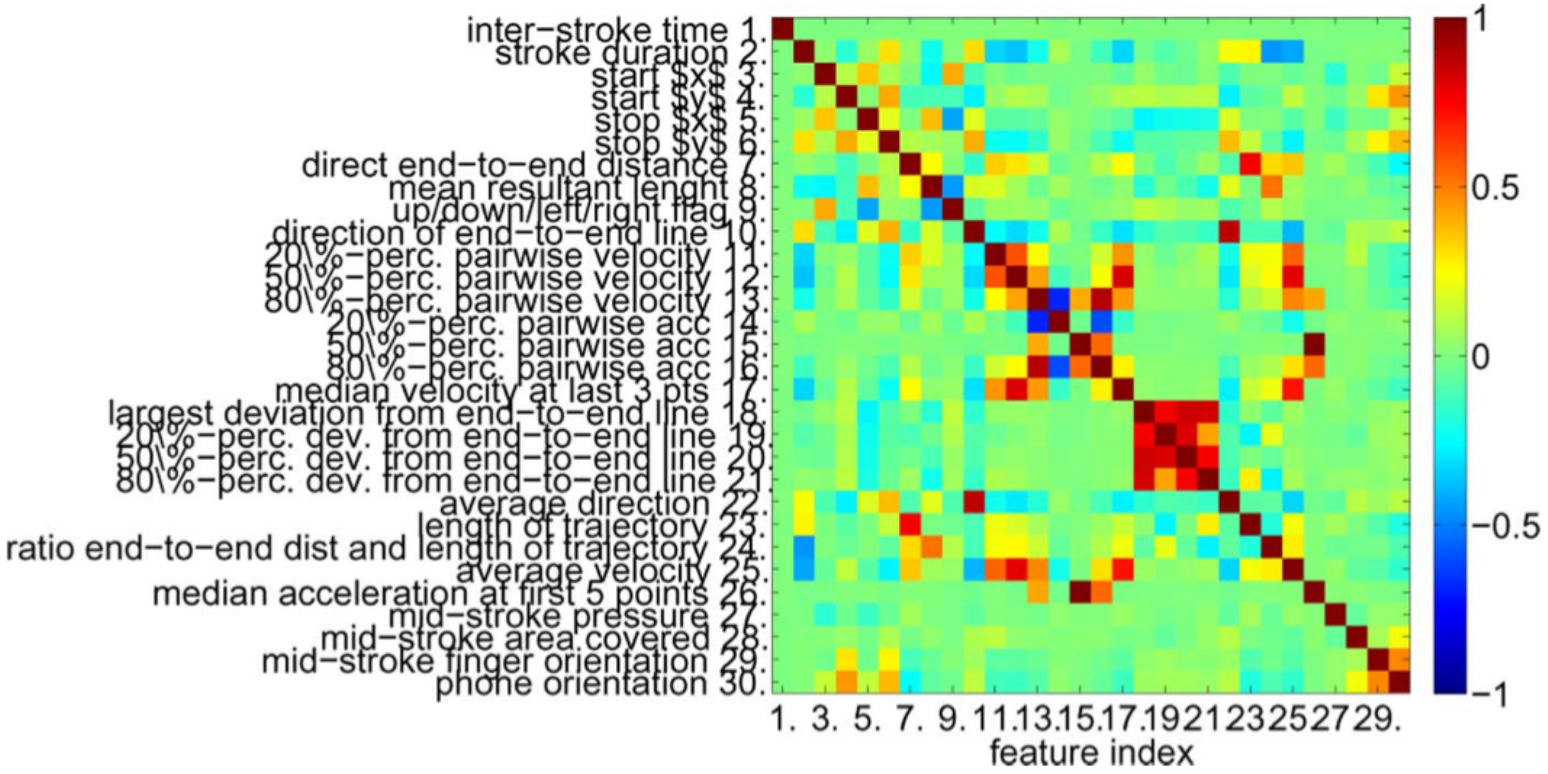
- Median velocity of the last five points
  - “ballistic” scrolling
- Mean resultant length
  - 1 for straight stroke, 0 for random angles
- Length of the trajectory
- Direct distance between endpoints
- Largest perpendicular distance between end-to-end line & trajectory
- Stroke duration
- Inter-stroke time



Feature selection: informativeness

- $I_F = I(F;U)/H(U)$
- Most informative features
  - Area covered by fingertip
  - 20% percentile of stroke velocity
  - Fingertip pressure
  - Direction of the stroke
  - Locations of endpoints
- x coordinate more informative than y coordinate

Feature selection: correlation



Classification

- k-NN
  - Using a k-d tree
  - Euclidian distance
  - k between 1-7
    - Cross-validation
- SVM
  - RBF kernel
    - 5-fold Cross-validation
- Combine scores of multiple strokes
  - Threshold the combined score

Results

- EER = 13% based on single stroke
- EER = 1-2% for 11-12 strokes
- Reading text: Median one stroke per 3.9 sec
- Image comparison: one stroke per 1.0 sec
- Verification time with 11 strokes: 11-43 sec

Results

- inter-week authentication
  - EER = 0-4%
- Inter-session authentication
  - EER = 2-3%
- Short-term authentication
  - EER = 0%

References

- “An Efficient User Verification System via Mouse Movements”, 2011
- “Touchalytics: On the Applicability of Touchscreen Input as a Behavioral Biometric for Continuous Authentication”, 2013