Multifactor User Authentication with In-Air-Handwriting and Hand Geometry

Duo Lu, Dijiang Huang, Yuli Deng, Adel Alshamrani

School of Computing, Informatics, and Decision Systems, Arizona State University

{duolu, dijiang.huang, ydeng19, aalsham4}@asu.edu

This research is supported by NSF (award #1528099)

Objective: Verify an **account** that you claim to possess, not the **identity** that you claim to be.

like a password:

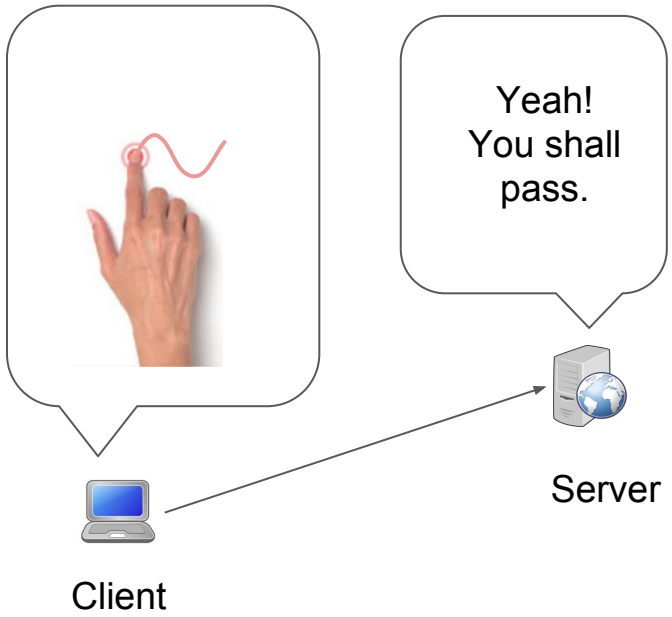
- Created by the user, changeable and revocable;
- The account and the personal identity are not linked.

like a biometric:

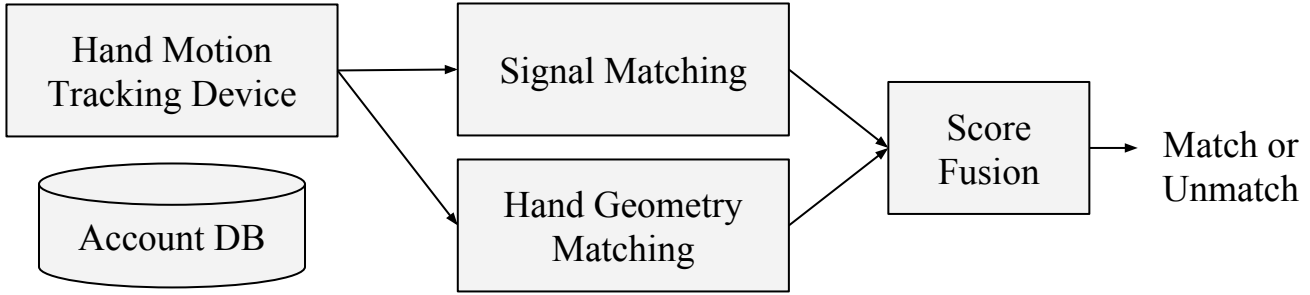
- Can not be given to someone else;
- Can not be easily spoofed.

Technical Challenges:

- Hand movement tracking is difficult;
- Minor variations of writing the same content;
- Lack of understanding of the features.



System Model and Dataset



200 passcodes created and written by 100 users, 5 + 5 repetitions to simulate sign-up and sign-in.
7 impostors write the same content (all 200 passcodes) as legitimate users write, 5 repetitions each.
44 passcodes by 22 users are tracked for 4 weeks, on average twice a week, 5 repetitions each time.

true-user: the same account; false-user: different accounts; spoof: from the impostors

Threshold-Then-Vote (TTV)

Request Signal

Template

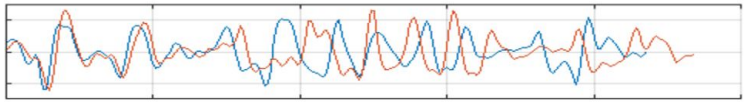
Variance

$$T_{ij} = mean(S_{ij}^1, S_{ij}^2, ..., S_{ij}^k),$$
$$C_{ij} = var(S_{ij}^1, S_{ij}^2, ..., S_{ij}^k).$$

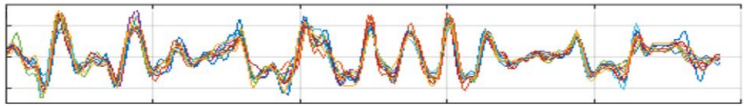
input : R, T, C
parameter: p, q, th_1, th_2
output : dist

signal level matching distance

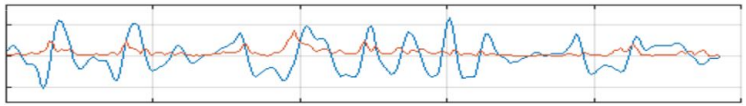
R (unaligned)



S (aligned)



T (blue line)
C (brown line)



preparation

$R' \leftarrow$ preprocess(R).
$S \leftarrow$ align(T, R').

distance

$D \leftarrow$ element-wise-abs(T - S).

scaling [*]

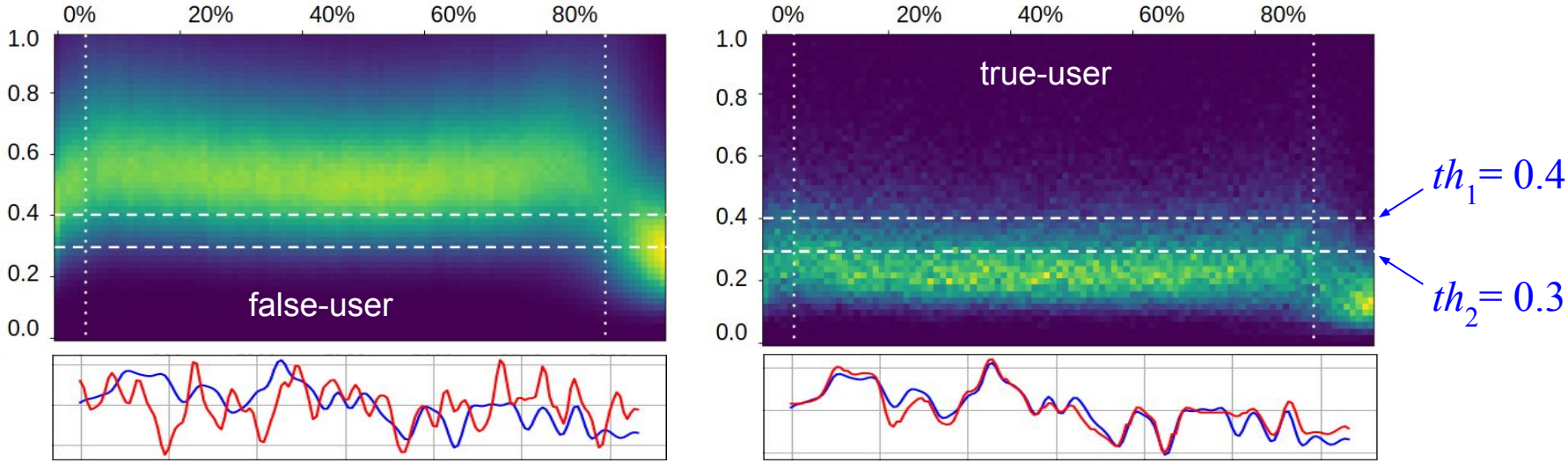
$P \leftarrow$ element-wise-inverse($\mathbf{1} + p \times T$)
$Q \leftarrow$ element-wise-inverse($\mathbf{1} + q \times C$)
$D \leftarrow$ element-wise-multiply(D, P).
$D \leftarrow$ element-wise-multiply(D, Q).

TTV

for i = *1 to* l **do**
  for j = *1 to* d **do**
    if $D_{ij} < th_1$ **then**
      $D_{ij} \leftarrow 0$
    else if $D_{ij} > th_2$ **then**
      $D_{ij} \leftarrow 1$
    else
      $D_{ij} \leftarrow 0.5$
    end
  end
end
dist $\leftarrow \frac{1}{l \times d} \sum_i^l \sum_j^d D_{ij}$

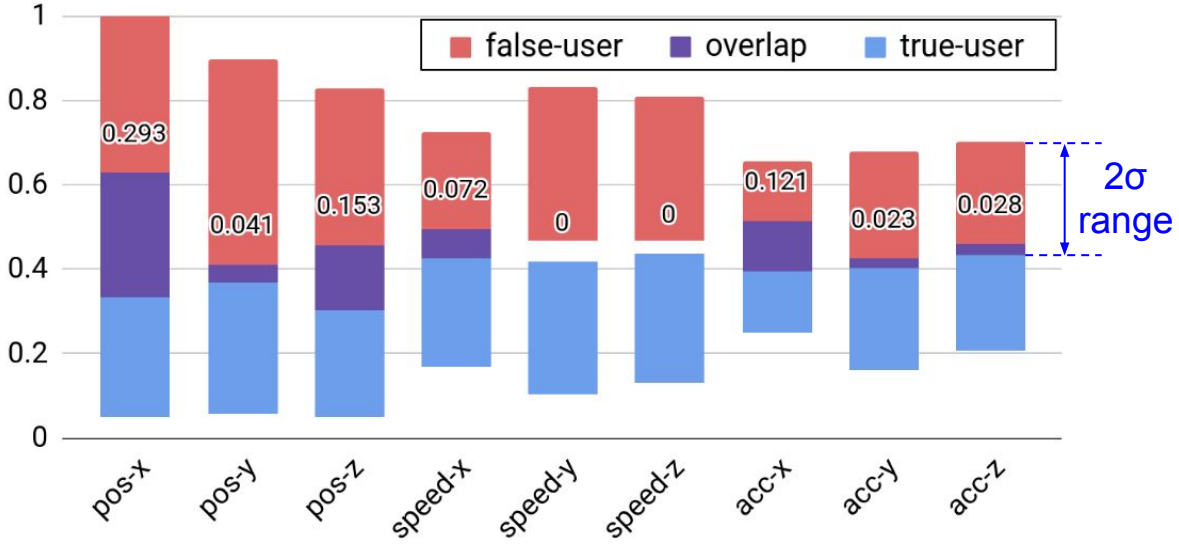
[*] Duo Lu, Kai Xu, and Dijiang Huang, "A data driven in-air-handwriting biometric authentication system". (*IJCB 2017*)

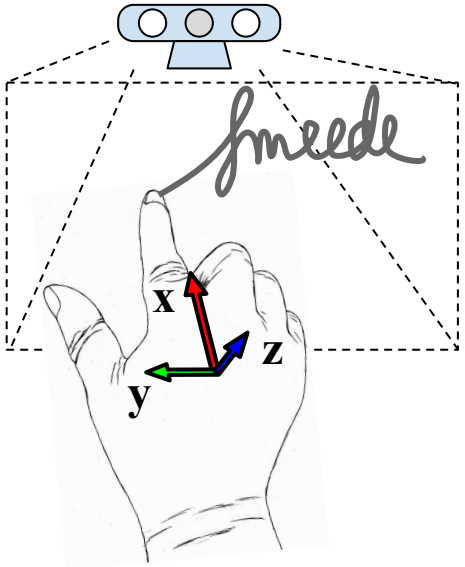
Distance Analysis - rows of *D* - $p(\frac{1}{d}\sum_j^d D_{ij}|c)$



- Anomalies at the beginning 5% and ending 10% (alignment forces the signals mapped together).
- If the signals are generated by the same user writing the same content, distance is small.
- Overlap determines the two thresholds in TTV.

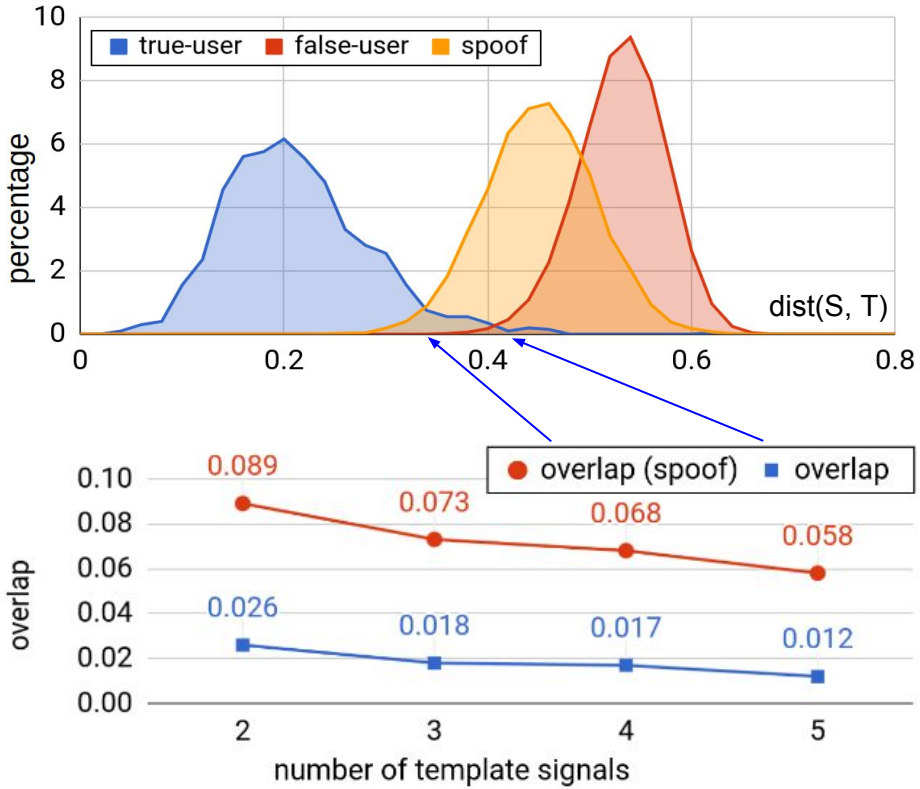
Distance Analysis - cols of *D* - $p(\frac{1}{l}\sum_i^l D_{ij}|c)$





- Reference frame: x - average pointing direction, y - perpendicular to x in horizontal plane.
- Human users are better at maintaining speed and force (i.e., acceleration) than position.
- Motion in the x direction varies more significantly (because users often write in a virtual plane).

Distance Analysis - signal level - $p(dist(S,T)|c)$

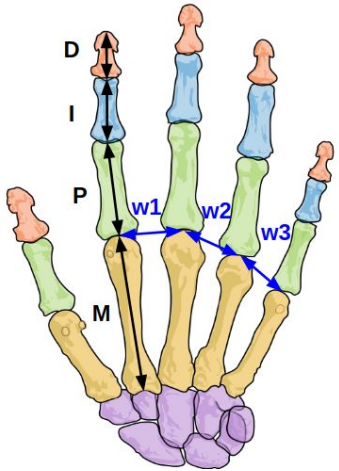


- Distance histogram is an approximation of the class-conditional probability.
- Overlaps between different classes denote the discriminating capability.
- The discriminating capability comes from both different content and different writing convention.
- More signals at registration, better performance. **Only two are also OK.**

Hand Geometry

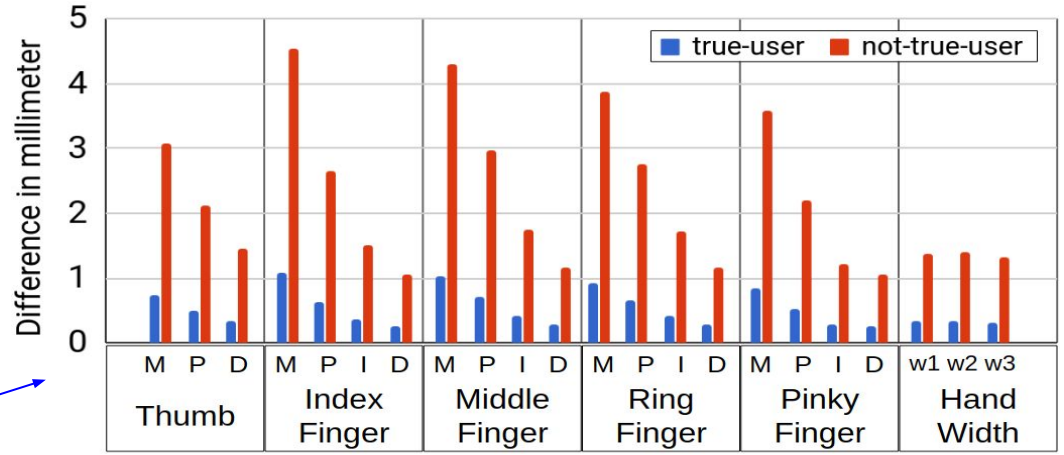
22 Hand Geometry Features:

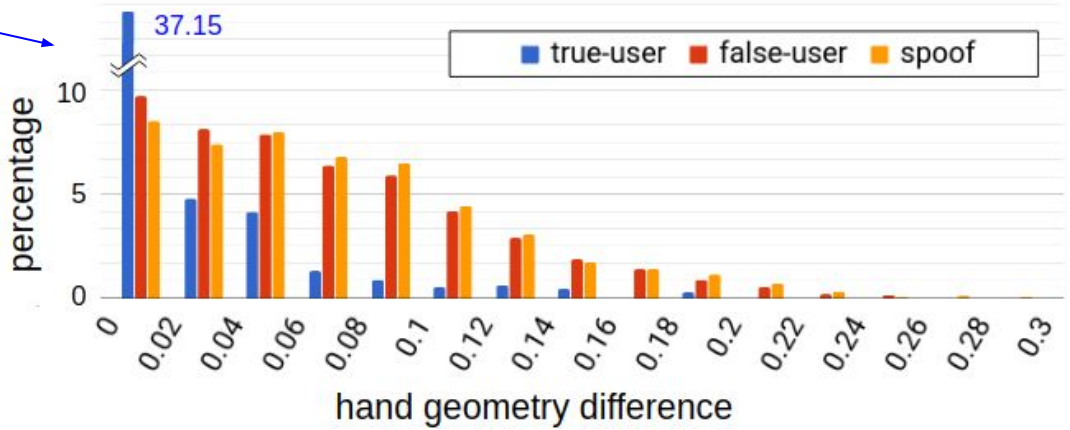
$$\boldsymbol{h} = (M_T, P_T, D_T, M_{IF}, P_{IF}, ..., w_1, w_2, w_3)$$



absolute difference

relative difference





Template: $\mathbf{t} = mean(\mathbf{h}^1, \mathbf{h}^2, ..., \mathbf{h}^k)$

Distance: $dist(\mathbf{h}, \mathbf{t}) = \frac{1}{22} \sum_i \frac{\mathbf{h}_i - \mathbf{t}_i}{\mathbf{t}_i}$

Score Fusion

$$score = dist(S,T) + w_1 dist(\mathbf{h},\mathbf{t}) + w_2 |l_T - l_R| / l_T$$

- if score < decision_threshold

  accept.

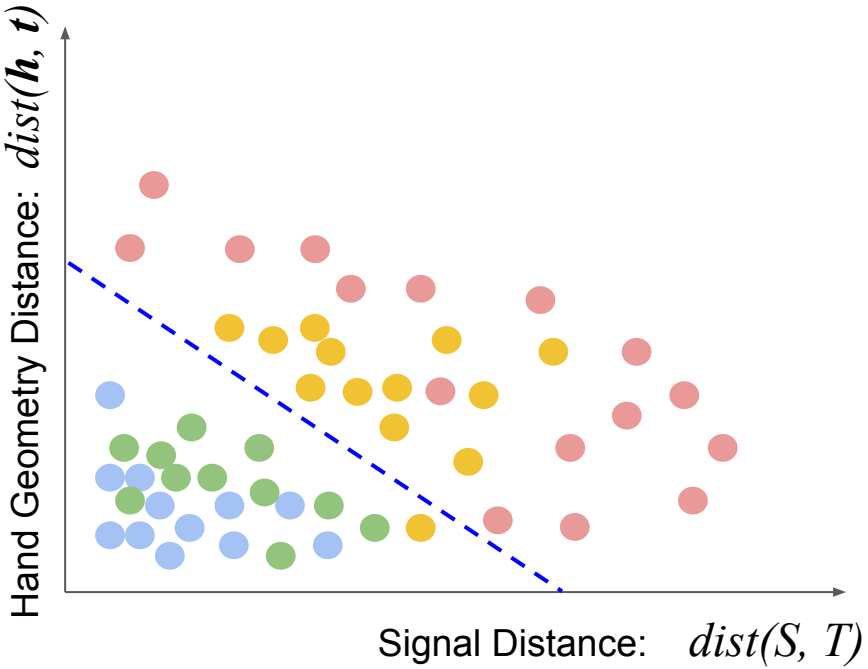
  else

  reject.

- decision_threshold can be adjusted to trade off accuracy and convenience.

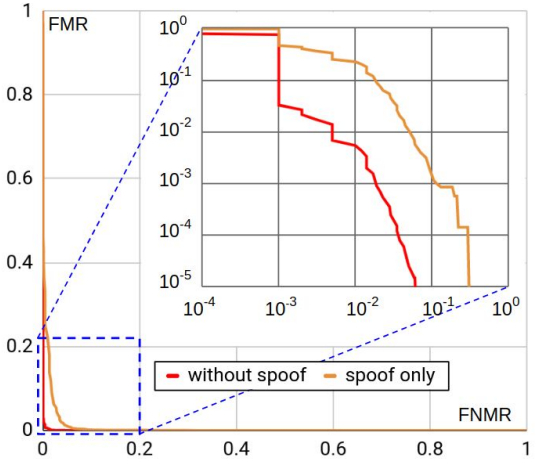
- w_1 and w_2 are parameters.

  (w_1 = 0.4, w_2 = 0.05)



Empirical Results

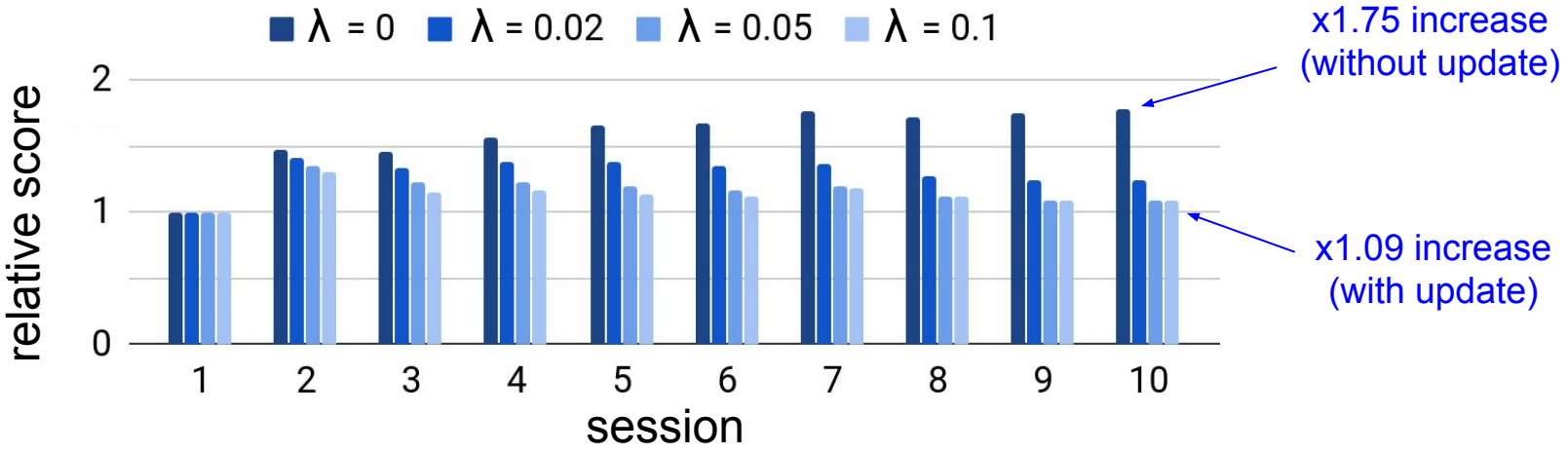
	Algorithm	EER	EER (spoof)	FMR 10K	FMR 100K	Zero -FMR
5 signals at registration	TTV	0.8%	3.7%	3.9%	6.9%	9.1%
	FUSION	**0.6%**	**3.4%**	**3.6%**	**6.2%**	**6.3%**
	DTW	1.3%	4.8%	7.8%	20.0%	22.0%
2 signals at registration	FUSION†	1.3%	5.0%	8.4%	13.0%	18.0%
	DTW†	2.0%	6.5%	11.0%	15.0%	17.0%



Reasons for performance improvement over DTW:

- Preprocess provides robustness against poor signal quality and minor variation in writing behavior.
- Threshold-Then-Vote (TTV) prevents locally mismatched signal segments for legitimate users.
- Score fusion further prevents some false matches with additional hand geometry information.

Long Term Performance



- Template update: $\mathbf{t} \leftarrow (1-\lambda)\mathbf{t} + \lambda\mathbf{h}, \qquad T_{ij} \leftarrow (1-\lambda)T_{ij} + \lambda S_{ij}$
- Update of C (the variance) should be regularized to prevent it growing very large.
- A few signals at registration may not be enough for the inherent complexity of the writing behavior.

Conclusions

- Multifactor authentication with in-air-handwriting and hand geometry has good potentials.

Limitations

- Constraints on user behavior, e.g., user must write within the field of view of the camera.
- Parameter tweaking and template protection on the server.

Future Work

- An in-depth study on the influence of passcode content, length, and strength.
- Larger dataset, more users, longer tracking time (several weeks).
- Advanced score fusion mechanism (beyond weighted sum).

Thank you!

Q & A