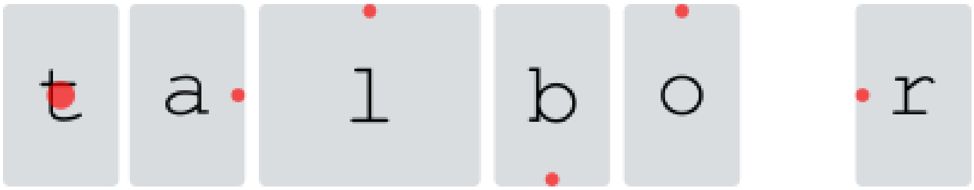
Exploring Intentional Behaviour Modifications for Password Typing on Mobile Touchscreen Devices



Lukas Mecke[1,2†], Daniel Buschek[2†], Mathias Kiermeier[2‡], Sarah Prange[1,3,2†], Florian Alt[3]
[1]*University of Applied Sciences Munich, Munich, Germany, {firstname.lastname}@hm.edu*
[2]*LMU Munich, Munich, Germany, †{firstname.lastname}@ifi.lmu.de, ‡mathias.kiermeier@gmail.com*
[3]*Bundeswehr University Munich, Munich, Germany,{firstname.lastname}@unibw.de*

Motivation

- Premise for behavioural biometrics: behaviour is **hard to intentionally change and imitate**
- But: Successful mimicry attacks on behavioural biometric systems using technical support [1]

[1] Hassan Khan, Urs Hengartner, and Daniel Vogel. Augmented reality-based mimicry attacks on behaviourbased smartphone authentication. In Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services, pages 41–53. ACM, 2018.

Motivation

- Potential assuming behaviour is controllable:
  - **Extending password space** for additional security
  - Actively **protecting biometric traits** by modifying them
  - **Recover from leakage** (problem with immutable traits)

Are people capable to *intentionally modify* their (**keystroke**) behaviour?

Roadmap

- Choose suitable **keystroke features**
- Find **visualisation** to communicate feature modifications
- Study design to foster **exploration** of
  - Participants *ability* to modify their behaviour
  - *Factors* influencing this ability

Keystroke Feature Selection

- 24 Features proposed by Buscheck et al. [1]
- Correlation analysis by Khan et al. [2] → 6 features
- Reduction to 4 features:
  - **(touch) area** (← preasure)
  - **flight time**
  - **hold time**
  - **(touch-to-key) offset** (← x,y)

Feature(s)	Description
key hold interval^t	interval between $\downarrow K_1$ and $\uparrow K_1$
inter-stroke interval^t	interval between $\uparrow K_1$ and $\downarrow K_2$
up-up^t	interval between $\uparrow K_1$ and $\uparrow K_2$
down-down^t	interval between $\downarrow K_1$ and $\downarrow K_2$
down & up pressure^c	touch pressure at $\downarrow K_1$ and $\uparrow K_1$
down & up area^c	touch area at $\downarrow K_1$ and $\uparrow K_1$
down & up axis^c	ellipses axis at $\downarrow K_1$ and $\uparrow K_1$
down x & y^s	x & y coordinate at $\downarrow K_1$
up x & y^s	x & y coordinate at $\uparrow K_1$
offset x & y^s	tap offset x & y from key centre
jump x & y^s	x & y distance between K_1 and K_2
drag x & y^s	x & y drag between $\downarrow K_1$ and $\uparrow K_1$
jump & drag angles^s	jump & drag angles
jump & drag dist.^{s}	jump & drag distances

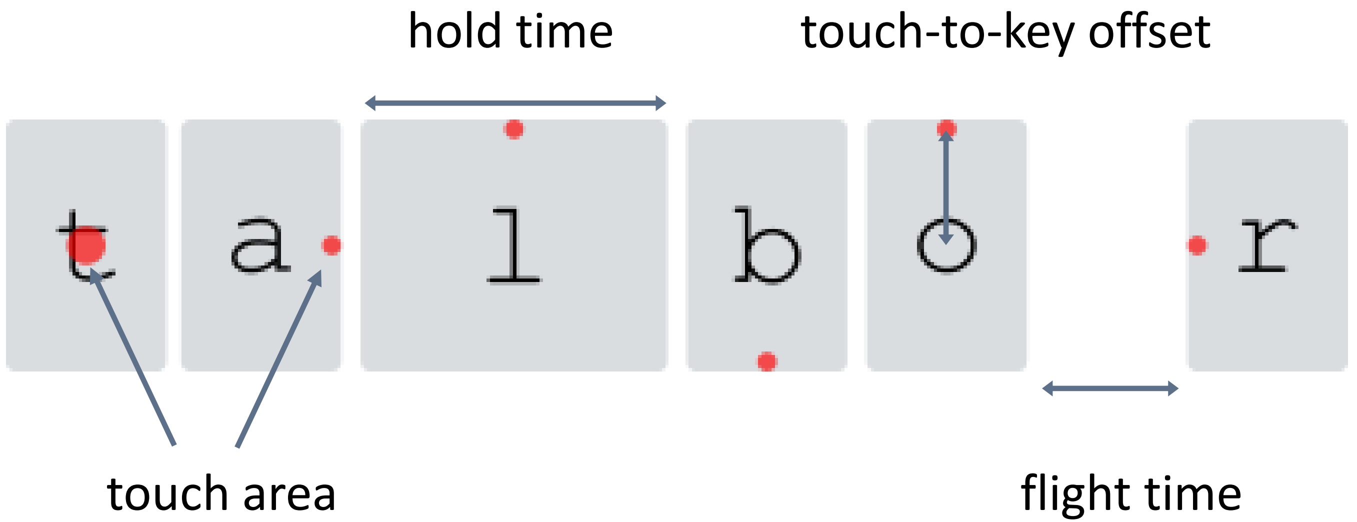
[1,2]

[1] Daniel Buschek, Alexander De Luca, and Florian Alt. Improving accuracy, applicability and usability of keystroke biometrics on mobile touchscreen devices. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, pages 1393–1402, New York, NY, USA, 2015. ACM.
[2] Hassan Khan, Urs Hengartner, and Daniel Vogel. Augmented reality-based mimicry attacks on behaviourbased smartphone authentication. In Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services, pages 41–53. ACM, 2018.

Pre-study

- Goal: **Communicate behaviour modifications**
- Exploration of *mark-up* and *pictorial* designs
- **Online study** (N=114) with two designs:
  Task: *Associate visualisation with given features*
- **Results** for winning design:
  - *Correct attribution* rate > 80% for all features
  - Rated *intuitive* and *readable* (agree)
  - Preferred by 59% of the participants

Proposed text annotations



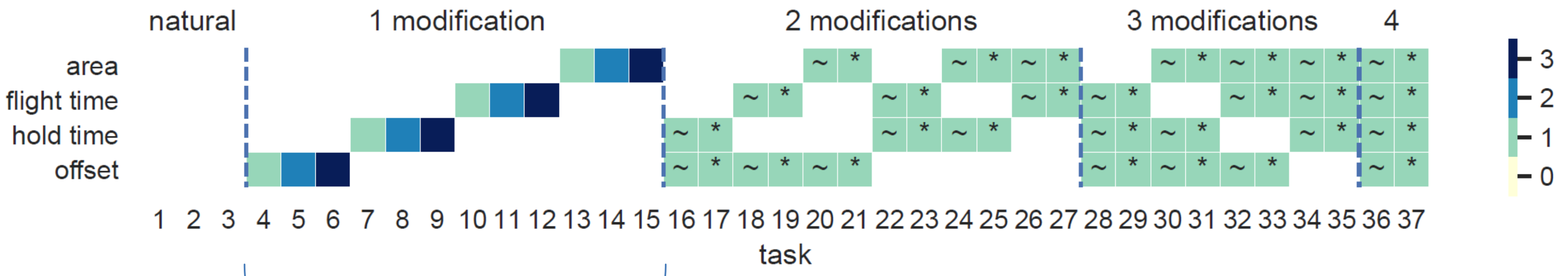
Study design

- Within subject lab study
- 24 participants (14 female, mean age 27)
- 37 Tasks to explore:
  - Different **passwords** *(password, football, princess)*
  - Different feature **modifications** *(offset, flight time, hold time, area)*
  - Different **locations** *(start, middle, end)*
  - Different feature **combinations** *(0-4)*
  - Different **distribution** (distributed or co-located)

Study design

Unmodified behaviour

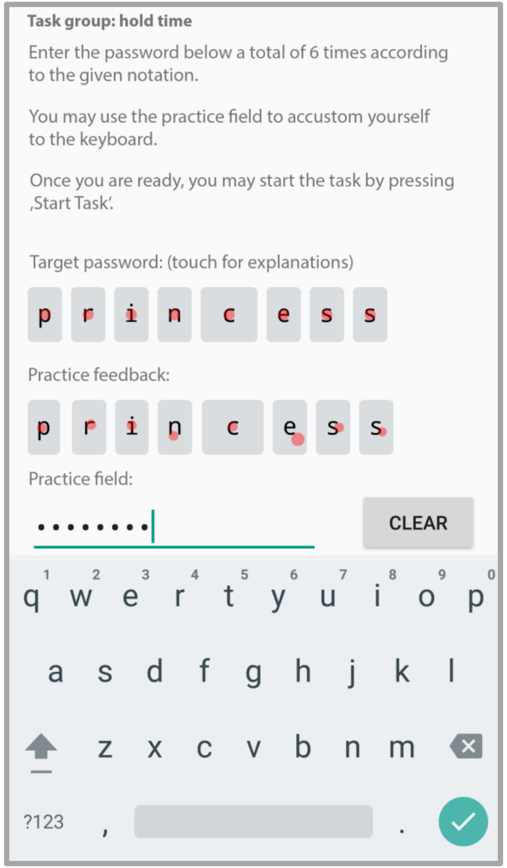
Repeated measures design with *number x distributed x session*



All features **in isolation**

Procedure

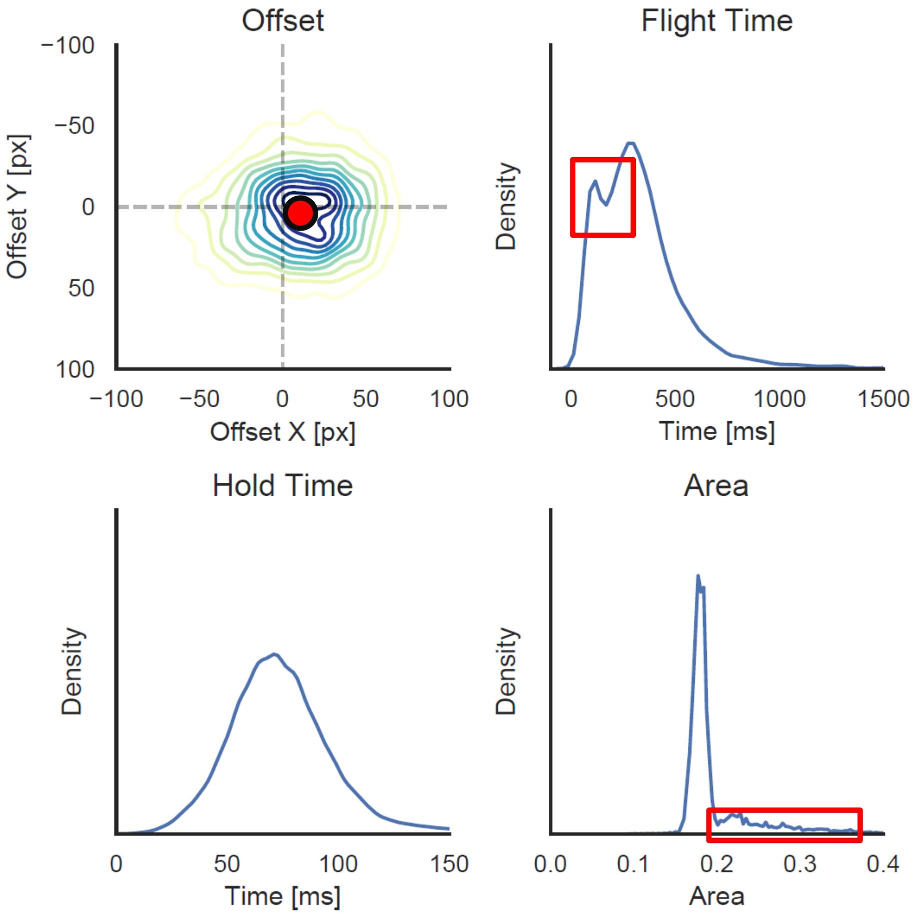
- Two sessions with each
  - Execute tasks (counterbalanced) on our test device with the right thumb (training with feedback, task without)
  - *Experience sampling* after each task
  - Create or reproduce a custom password
- Concluding Interview



(Translated from German)

Results

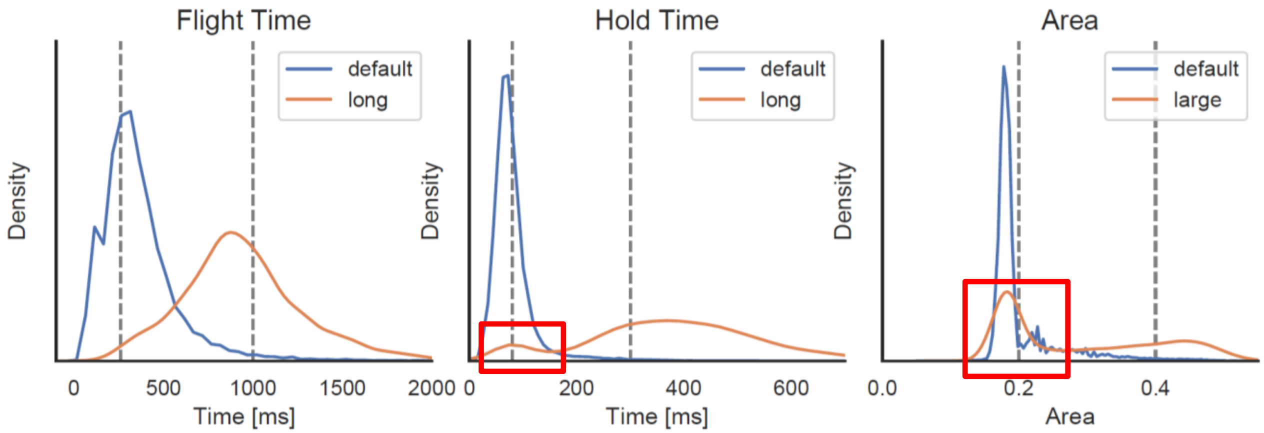
- Natural behaviour
  - Offset towards *bottom right* [1]
  - Secondary peak in flight time for *double letters*
  - Correlation of touch area and key x-position (*thumb stretching*)

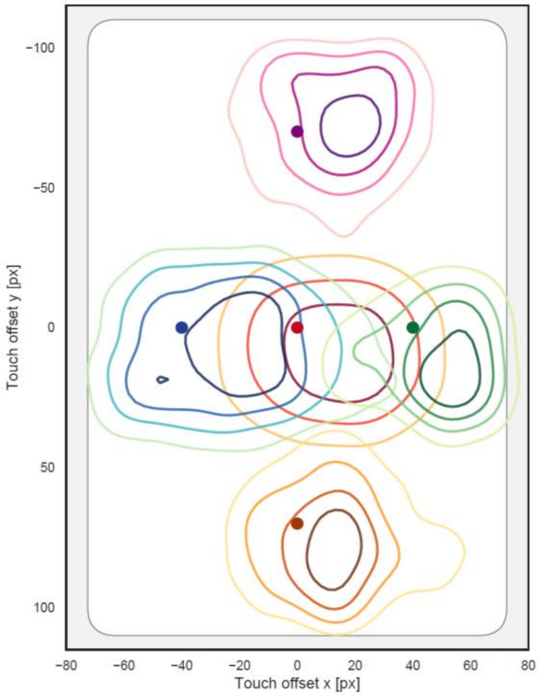


[1] Daniel Buschek and Florian Alt. TouchML: A machine learning toolkit for modelling spatial touch targeting behaviour. In Proceedings of the 20th International Conference on Intelligent User Interfaces, IUI ’15, New York, NY, USA, 2015. ACM.

Results

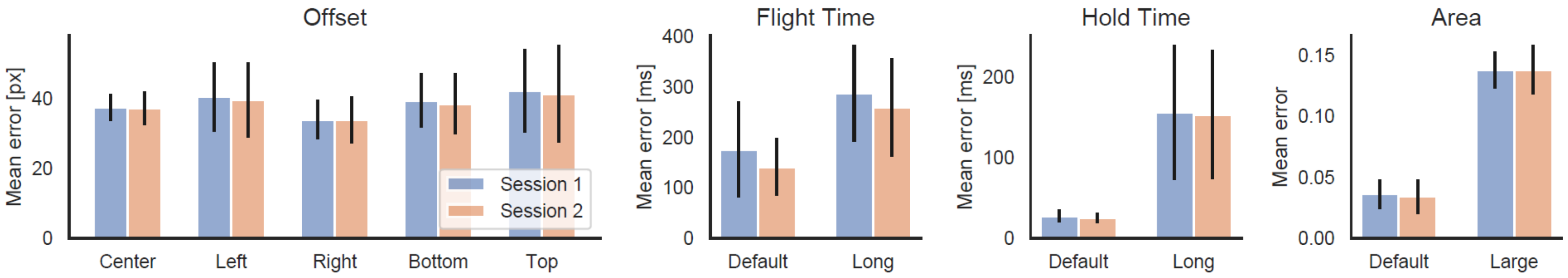
- Modified behaviour
  - Successful modification for all features
  - Secondary peaks indicating user errors





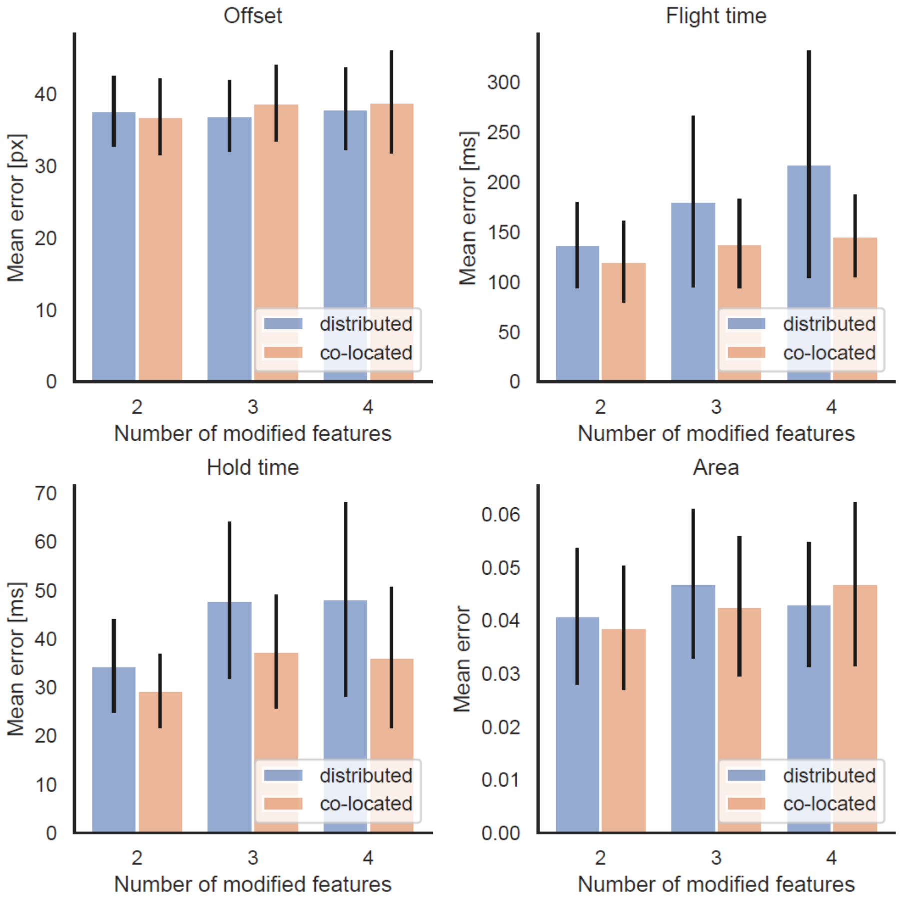
Results

- Errors by *target* and *session*
  - Error for offset right significantly smaller than the others
  - Significant session effect for flight time
  - Generally default error was significantly smaller than modified



Results

- Errors by *number* and *distribution* of modifications
  - *Offset remained stable*
  - *Co-located* features resulted in significantly lower error
  - Increased *number* of modifications significantly increased error

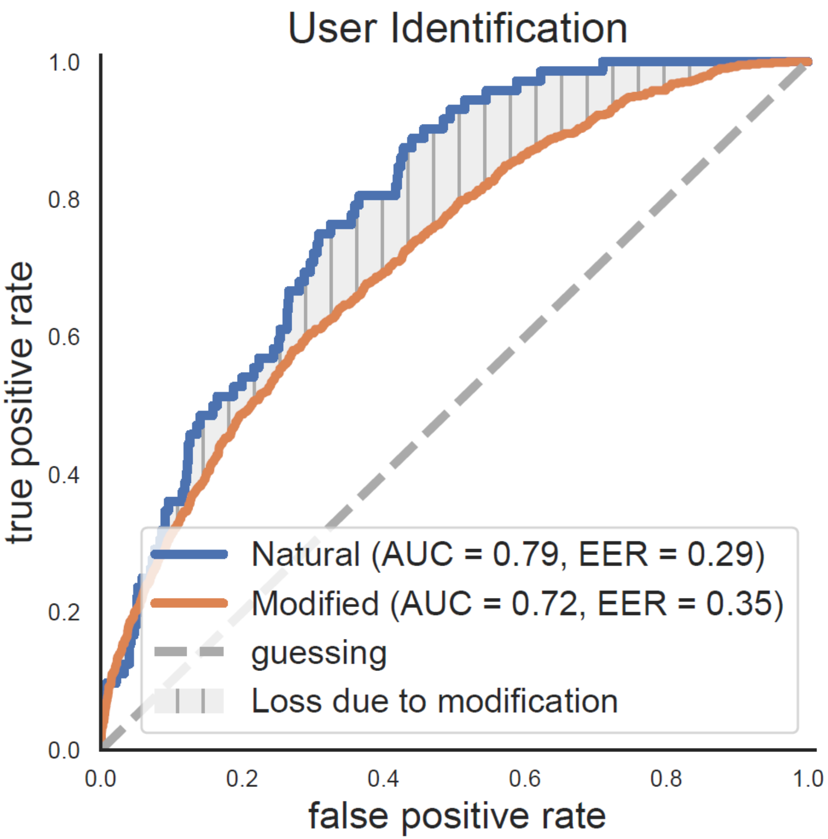


Results

- Meta data and subjective ratings
  - Increased *task completion time* for more modifications and for distributed modifications
  - Decreased *typing speed* for more modifications and for distributed modifications
  - More *incorrect password entries* for distributed modifications
  - Co-located modifications were perceived subjectively easier
    (Likert ratings: better able to adjust, higher success, less difficult)

Results

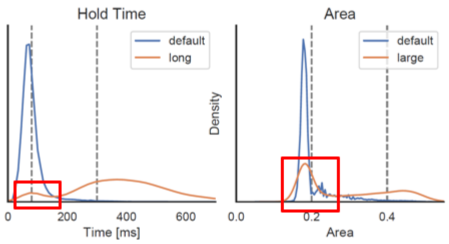
- *Impact on individuality* (Gaussian mixture model for user identification)
  - *Biometric value is decreased* by following modifying towards the same target
  - Some individuality remains



User Feedback

- Hard to control:
  - Offset modifications (hitting the wrong key)
  - Distinguish large area and long hold time
- **Creation Strategies**:
  - Emphasis

    "When I created the password I first typed it and observed what I automatically did. For example I typed a 'g' rather to the left, entered a 'b' rather [long]; That's what I adjusted [the password] to."
  - Salient positions (password)



Extending password space

- Detecting Modifications technically feasible:
  - Random Forest Classification (100 trees) with default parameters
  - Leave-one-out validation across sessions
  - Results: **accuracy > 94% for all features**
- (Upper bound) entropy, assuming random passwords with random modifications (|∑|=72)

password length	8	7	6	5
no modifications	**49.36**	43.19	37.02	30.85
1 modification	**55.14**	48.77	42.38	35.94
2 modifications	59.84	**53.27**	46.63	39.90
3 modifications	63.90	57.10	**50.20**	43.16

Extending password space

- But:
  - Effect of different keyboard layouts and hand postures
  - Potential common patterns reducing entropy
  - Practically: Requires capturing hardware on all devices

→ Questions for future work

Take away

- **Participants are able to intentionally control typing behaviour**
- Using modifications to *extend password* space is possible
- Modifying *less* and *co-located* features is easier
- New *perspective* on typing behaviour (implicit → explicit)

Exploring Intentional Behaviour Modifications for Password Typing on Mobile Touchscreen Devices

Contact: Lukas Mecke
lukas.mecke@ifi.lmu.de