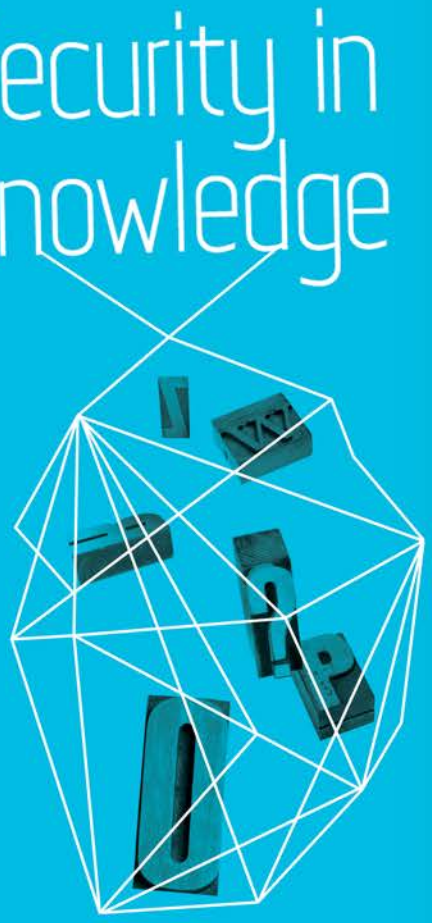
RSA CONFERENCE ASIA PACIFIC 2013

DARPA ACTIVE AUTHENTICATION PROGRAM: BEHAVIORAL BIOMETRICS



Dr Neil Costigan
BehavioSec

Ingo Deutschmann
BehavioSec

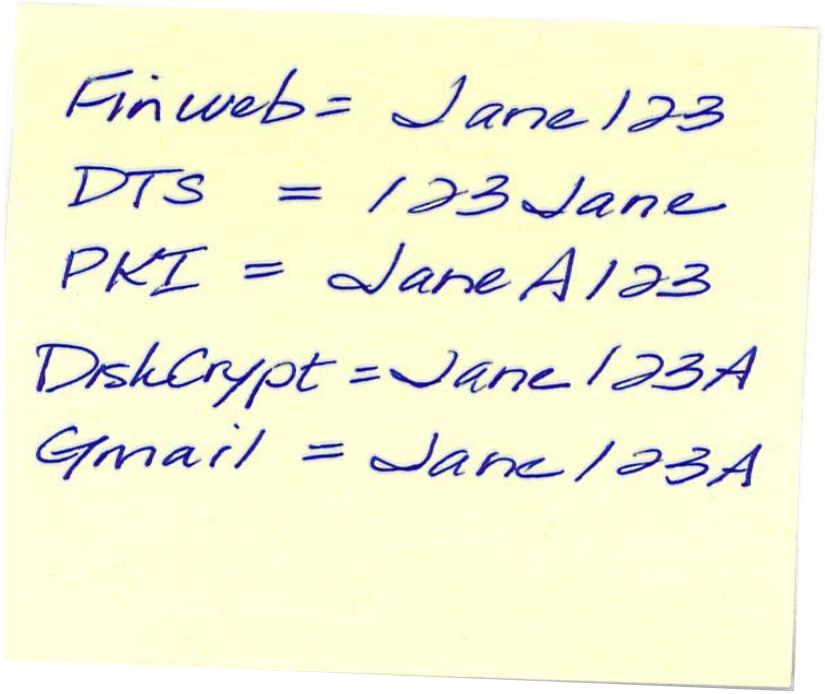
Session ID: SEC-105

Session Classification: Intermediate

Overview

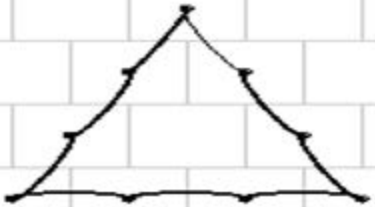
- DARPA’s Active Authentication program
  - Goal
- BehavioSec
  - Who we are
- The participant projects
  - High level overview
- Detail on Behaviosec's project
  - Our trial & results
- Q&A

Users are the weak link…



Passwords will always be a

*Keyboard



6tFcVbNh^TfCvBn

*Keyboar
d

R%t6Y&u8I(o0P-[

*Keyboard

#QWqEwReTrYtUyI

Source: *Visualizing Keyboard Pattern Passwords*, US AF Academy 11 O
2009

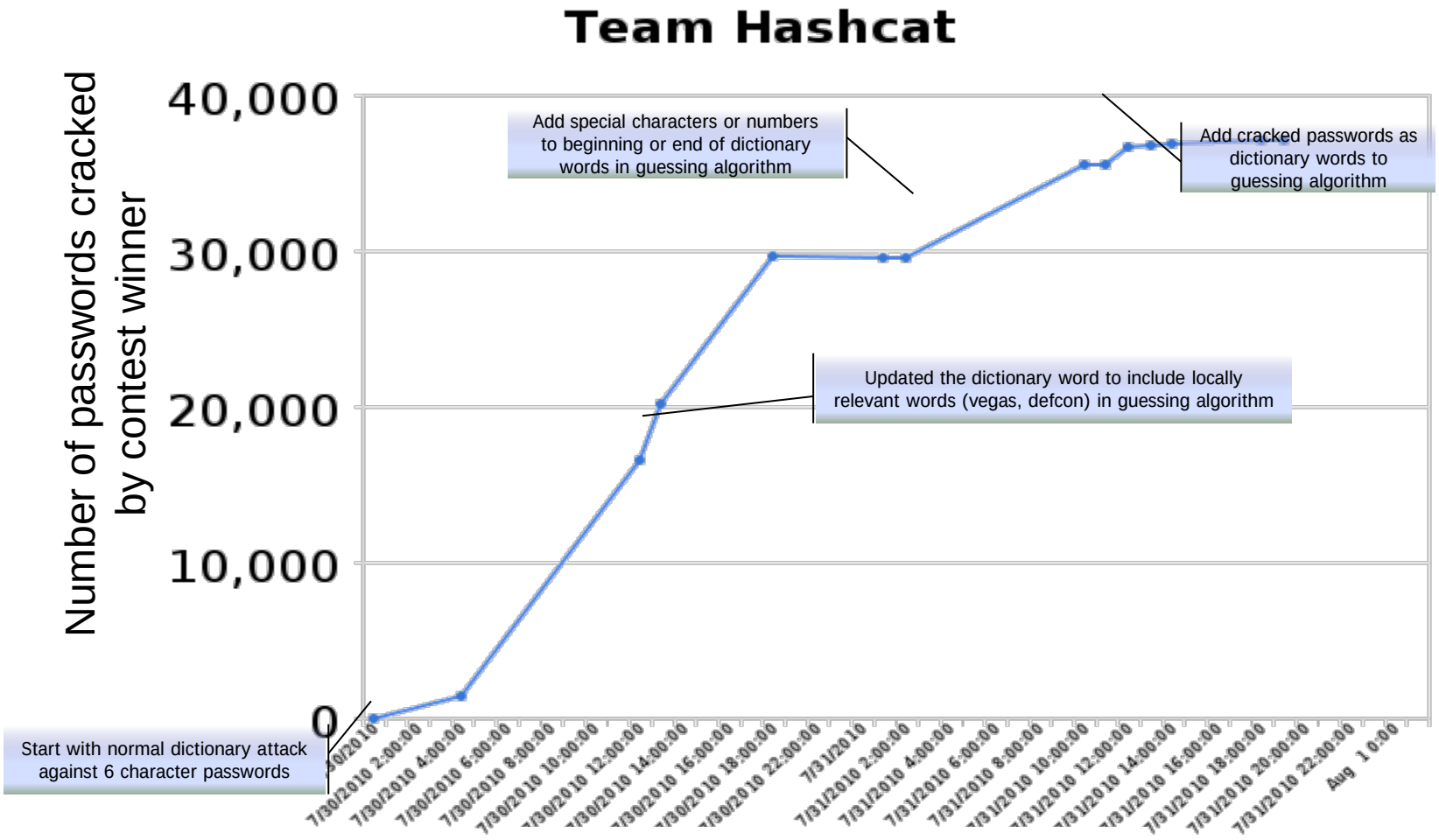
How many passwords do we

DoD IT Asset Type	DARPA Reference System
NIPRnet	Windows DMSS
Laptop Encryption	Guardian Edge
DARPA VPN	Nortel
PDA	Blackberry/iPhone
SIPRnet	Windows DSN
JWICS	Windows DJN
Source Selection	TFIMs, I2O BAA Tool
Contract Management	GSA Advantage, SPS
Contract Invoicing	Wide Area Workflow
Payroll	MyPay
Benefits	Benefeds.com
HR	hr.dla.mil
Training	DAU
Collaboration	Defense Connect Online
Financial System, Local	Momentum
Financial System, Agency	DFAS
Credit Union	PFCU, NCU, etc.

Non-DoD IT Asset Type	Hacked on	Credentials lost
American Honda Motor Co.	27-Dec-10	4.9m
Bank of America	25-May-11	1.2m
Carnegie Mellon University	8-Oct-07	19k
Citigroup	27-Jul-10	30m
Clarkson University	10-Sep-08	245
Countrywide Financial Corp.	2-Aug-08	17m
Fidelity Investments	24-Sep-07	8.7m
Heartland Payment Systems	20-Jan-09	130m
IBM	15-May-07	2k
Johns Hopkins Hospital	22-Oct-10	152k
SAIC	7-May-08	630k
Sony	27-Apr-11	12m
Stanford University	6-Jun-08	82k
TD Ameritrade Holding Corp.	14-Sep-07	6.5m
Texas A&M University	9-Nov-08	13k
TJMax Stores	17-Jan-07	100m
U.S. Depart. of Veteran Affairs	14-May-07	103m
U.S. Marine Corp – PSU research	26-Jul-07	208k
Visa, MasterCard, and American Express	27-Dec-10	4.9m

Source: www.privacyrights.org/data-breach

Patterns will always be hackable



Source: http://contest.korelogic.com/

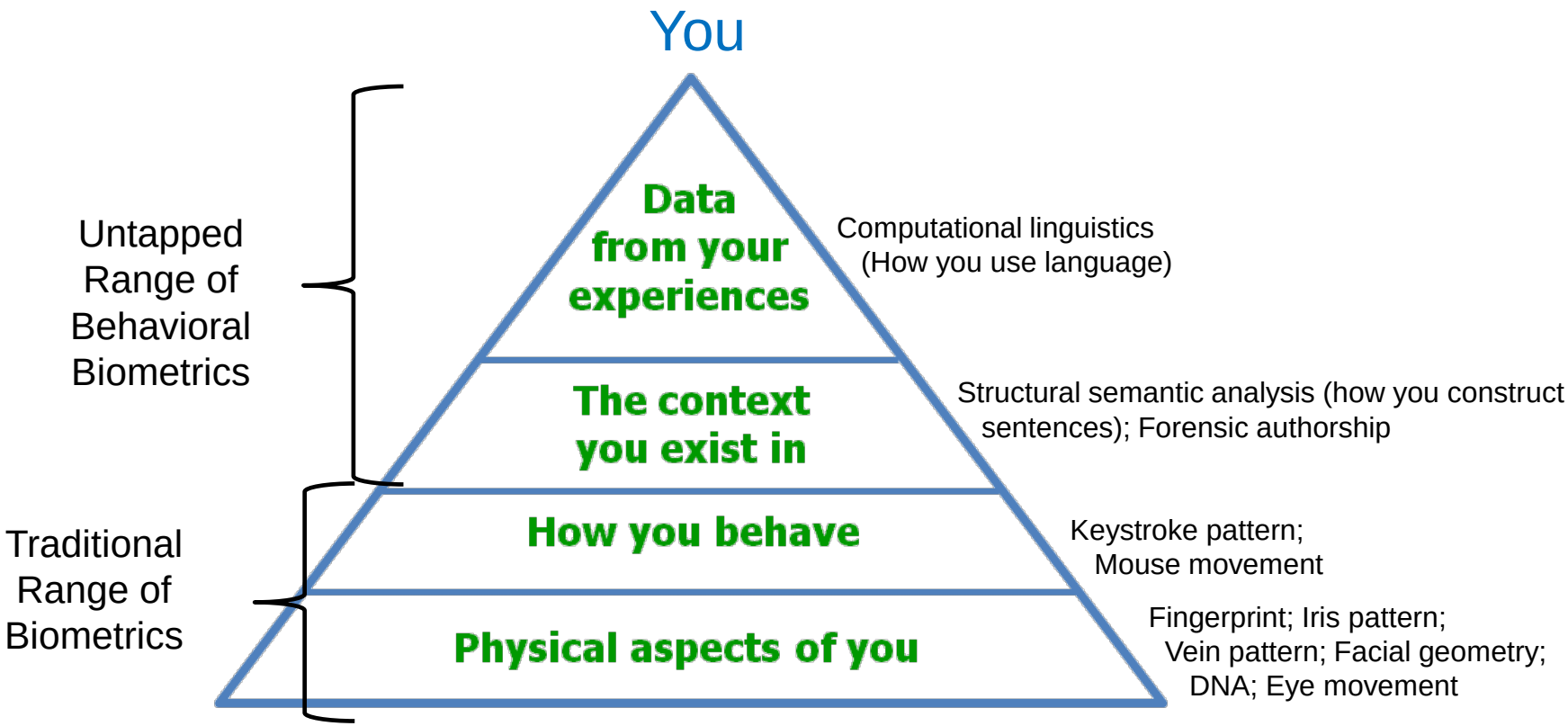
DARPA

- US **D**efense **A**dvanced **R**esearch **P**rojects Agency
- ‘Moon shot’
- Next generation DoD workstation security
- Active Authentication program
  - Transparent. Out of the hands of the end user.
  - Remote, real-time, managed security.
  - Today DoD…. Tomorrow mainstream.

- A tool for all enterprise security desktops & professionals
- 2013 has mobile focus.

The Active Authentication Program

A continuous authentication solution that takes the data available on a DoD computer system and makes an informed decision on the identity of the user of the computer



Non-cooperative behavioral biometrics allow the validation of identity simply by the user acting normally, not requiring interruption of the user

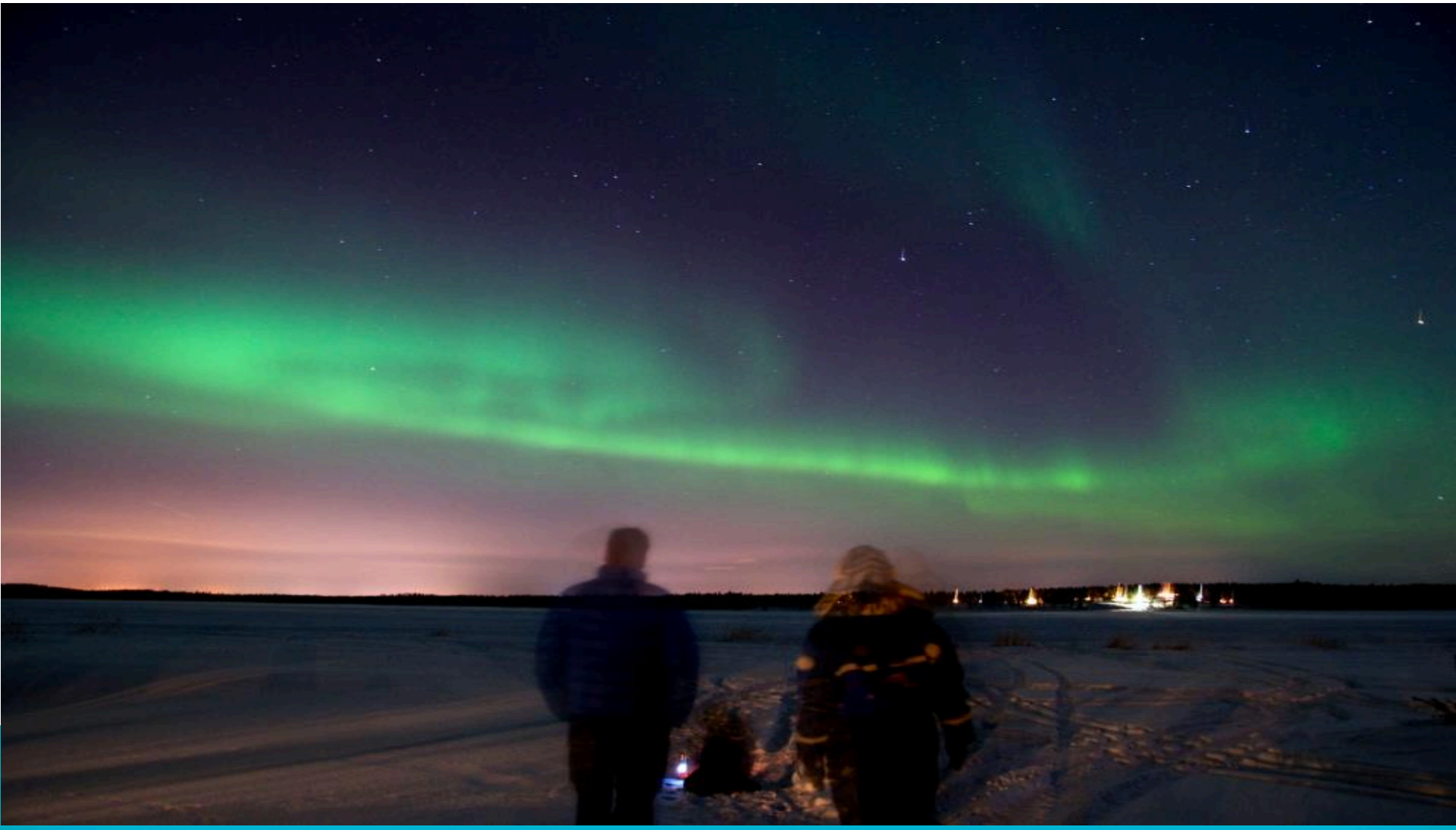
BehavioSec

- Swedish IT-Startup. University spinout
  - Luleå Technical University
- Offices in
  - Luleå (R&D) & Stockholm, Sweden.
  - Germany & US.

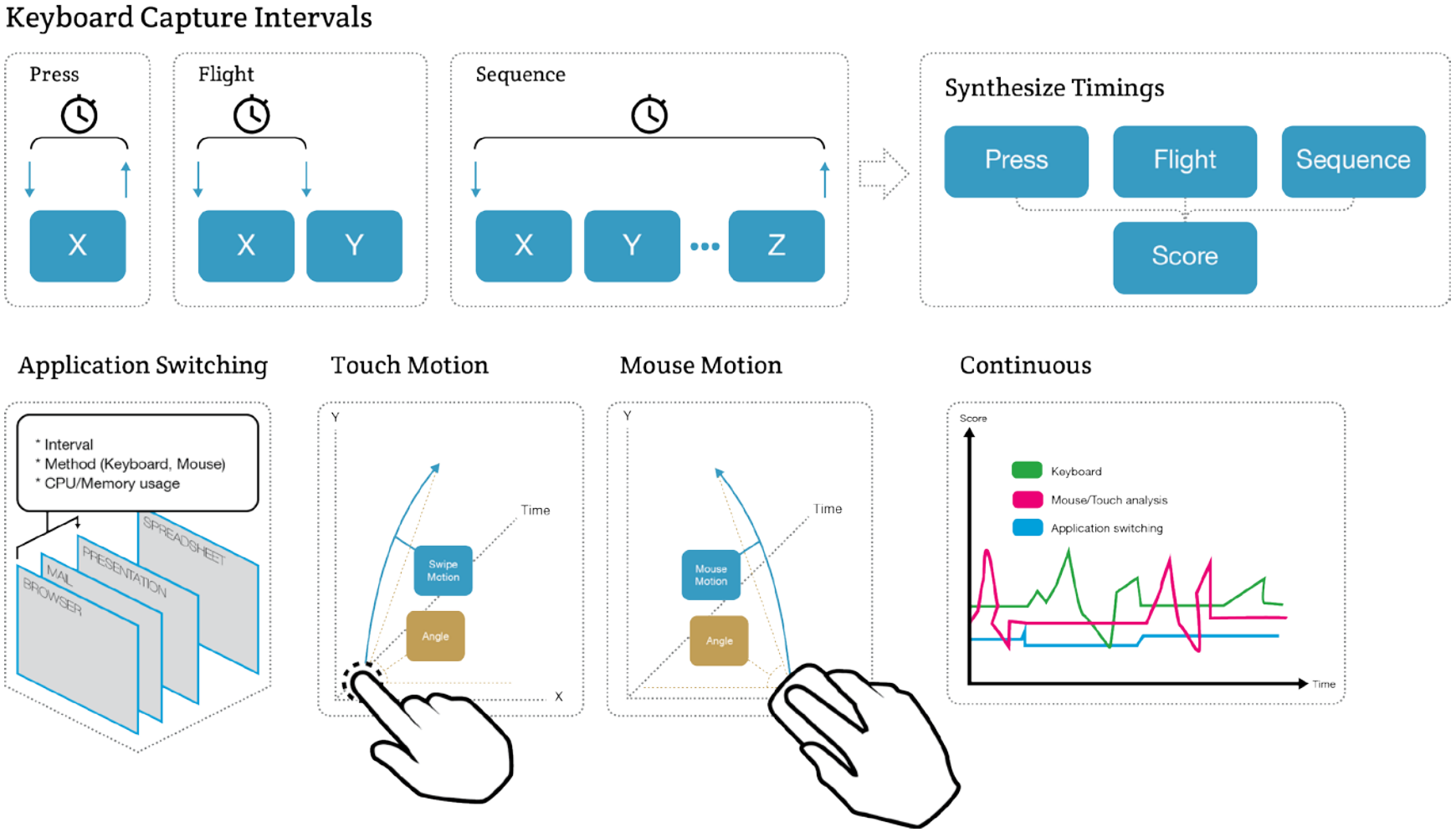
- Web, Mobile & Enterprise products in high value paying customers TODAY

Luleå

Luleå , Sweden



BehavioSec Modalities



Other Performers inside the DARPA AA program

‘New’ biometric modalities (1)

- Neuro-cognitive patterns
  - Naval Post Graduate School
  - Developing digital “cognitive fingerprints” from various biometric sources; potentially developing a framework for identification of other behavioral biometrics.
- User Search Patterns
  - Allure Security Technology, Inc
  - Using the user’s patterns for searching for information on the computer, verified by high volumes of decoy document touches placed in the file system.

‘New’ biometric modalities (2)

- User Behavior Patterns as seen from the Operating System
  - Coveros
  - Using traditional computer based IDS algorithms on user behavior (as seen in OS interactions) to determine when someone other than the authorized user is accessing the system.
- Stylometry
  - Drexel University
  - Using traditional stylometric methods to validate a user based on what they are typing. Also researching how to detect adversaries who attempt to impersonate users through mimicking typing methods.

‘New’ biometric modalities (3)

- Stylometry focused on Cognitive Processing Time
  - Iowa State University
  - Using stylometric methods to validate the user based on natural pauses in the way they type.
- Stylometry focused on Cognitive Rhythms
  - NYIT
  - Using text productivity, pause, and revision behaviors to validate users based on how they type (includes content/language).
- Covert Games
  - Southwest Research Institute
  - Determine the user’s pattern of behavior by introducing patterned system aberrations that the user intuitively learns.

‘New’ biometric modalities (4)

- Screen Interface
  - University of Maryland
  - Using spatio-temporal screen fingerprints to identify the user for authentication.
- Behavioral Web Analytics
  - Naval Research Labs (NRL funded)
  - Identification of the user from Web browsing activities to include semantic (what kind of webpages are visited) and syntactic session features

Details on three…

- User search behavior characteristics
- Stylometry (how people use language when they write) augmented by author classification and verification
- Stylometry, focused on how thought processing impacts keystroke dynamics. Users changes in typing rhythms induced by cognitive factors, especially when it is manifested as natural pauses in typing.

BehavioSec participation 1/2

- Verify/validate our existing software with empirical data on DoD specified workstations with a significantly scaled data set of test users working for a sufficient amount of time.
- Enhance the field of continuous authentication by adding to the understanding of metrics suitable for measurement of continuous biometrics.

BehavioSec participation 2/2

- Validate & extend prior work by BehavioSec & academic researchers on the idea of continuous trust that promised to enhance the accuracy and security level of an active authentication system.
- Introducing and test a new test metric of 'Application usage' as a suitable measure for authentication.
- Propose a common open data format for interoperability.

New metrics for continuous behavioural biometrics

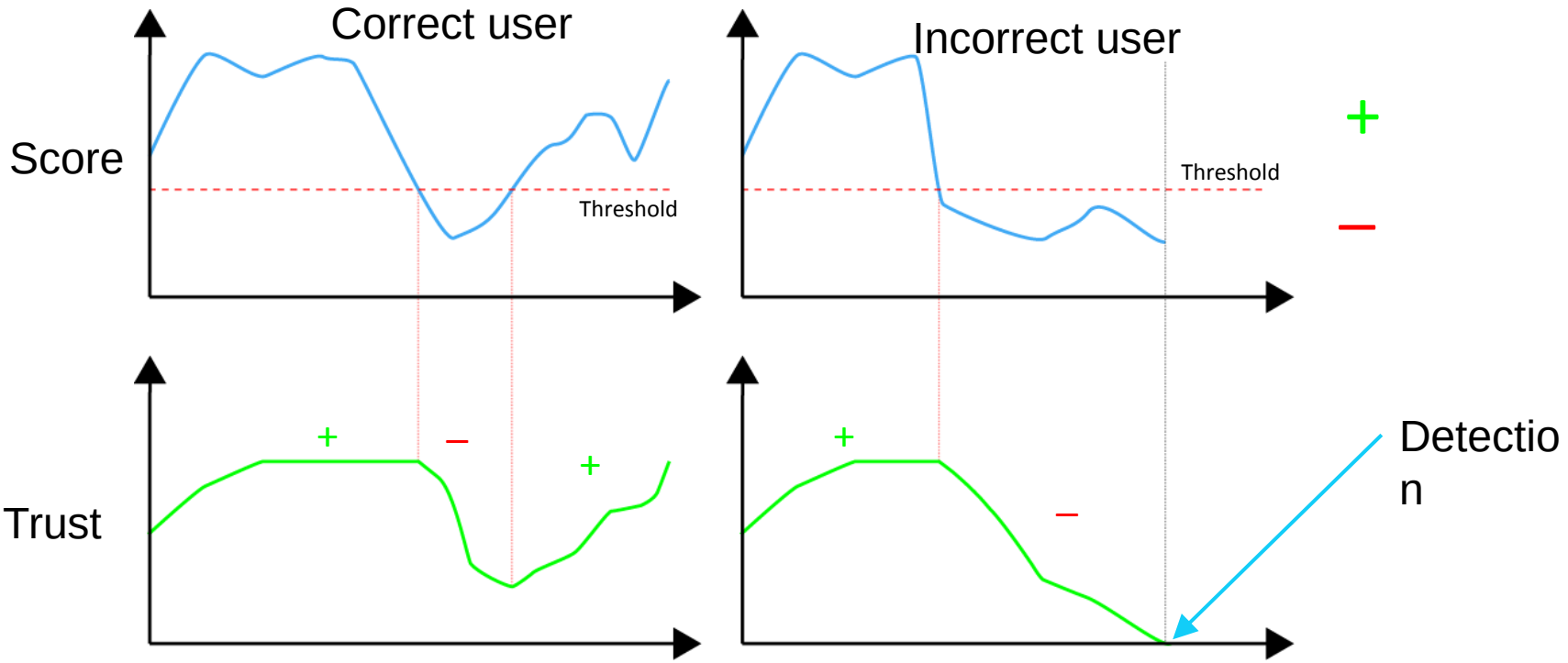
- Extend current biometrical measurement definitions to better fit the unique characteristics of continuous behaviometrics.
- One of the main differences between “One-time” authentication and continuous authentication is the extra time dimension
- Moving window (time span that is analyzed)
- Time / number of events it takes to do a detection

Continuous trust

- Incorporate the ‘time factor’ to increase accuracy of the overall authentication system?
- Previous research
  - Professor Patrick Bours
    - Norwegian Information Security Lab (NISLab)
    - Gjovik University College (HiG)
    - A new metric that keeps a running penalty/reward system of the ‘trust’ of a user in a continuous environment.
    - Bours PAH, Continuous keystroke dynamics: A different perspective towards biometric evaluation, Information Security Technical Report (2012), doi:10.1016/j.istr.2012.02.001
      preprint ? http://www.tapironline.no/last-ned/208
    - Extend that research with empirical data

Trust

- Value between 0 - 100
- Starts at 50
- Altered by confidence (see image)



The BehavioSec model

4.1 BehavioSec Model

The trust in the BehavioSec model is represented as value between 0 and 100, where higher value means higher trust. The initial trust value is configurable and sets how aggressive the system should be from start. Trust is then updated using Equation (4.1), with input from the different continuous tests(keyboard, mouse, application usage etc). A detection is made when the trust decreases below the trust threshold.

$$C := \begin{cases} 50, & \text{Start value} \\ Max\left(\left(C - \dfrac{T-P}{100 * \dfrac{T}{Z}}\right), 0\right) & P < T \\ Min\left(\left(C + \dfrac{P-T}{100 * \dfrac{1-T}{Z}}\right), 0\right) & P \geq T \end{cases} \tag{4.1}$$

C = Trust of the user
P = Probability from one test
T = Threshold that decides if trust should increase or decrease
Z = Constant value, the maximum value that the trust is increased or decreased

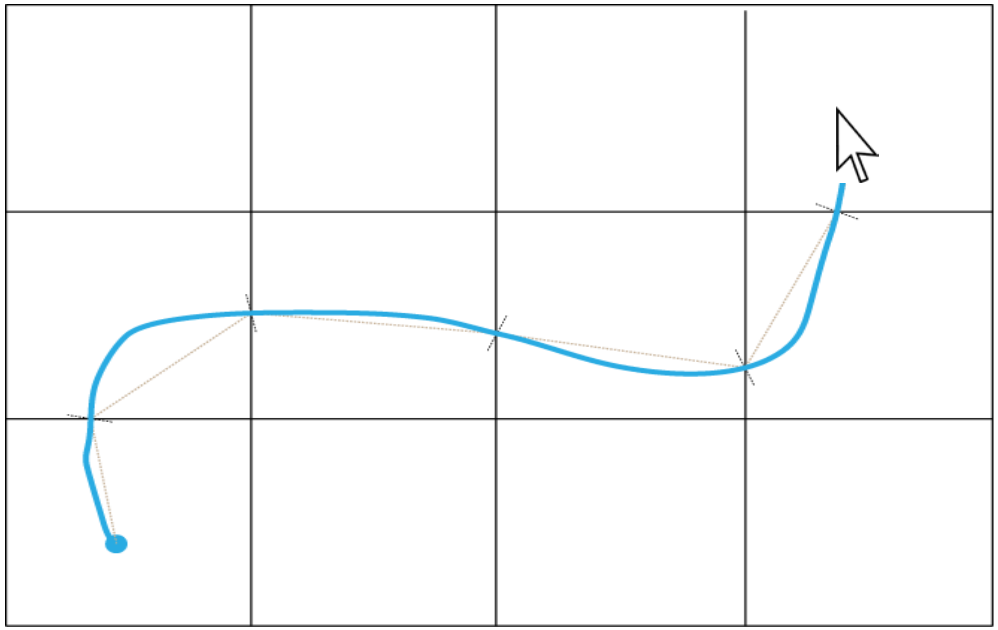
Data Collection

- Open data format
- Implement a data collector
- Install and collect data from 99 users on a DoD-like environment, 20 hours a week for a total of 10 weeks

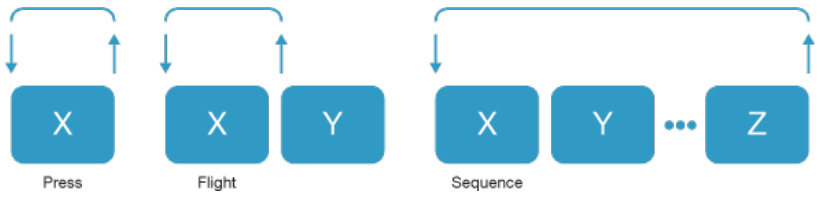
Open data format

- **Keyboard data** *K <action> <key id> <timestamp>*
  K 256 194 123235367
- **Mouse data** *M <action> <x coordinate> <y coordinate> <timestamp>* M 512 645 234 23523622
- **Monitor resolution data** *R <X> <Y>* R 1920 1080
- **Timestamp synchronization data** *T <unix time> <timestamp>* T 1345715086 23523622
- **Program switch data** *P <hash representing new program>* P 226234523462346234
- **Process usage data** *<process usage identifier> <hash process> <usage in parts per thousand> <memory usage> <timestamp>* U 24523623572357 404 205 2351235123
- **Version data** *V <version number>*V 1

Interaction Examples



Example of mouse interactions each crossing of the grid lines triggers a interaction



Example of keyboard interactions each key pressed is an interaction same for flight between keys and keys sequences.

How the metric is calculated

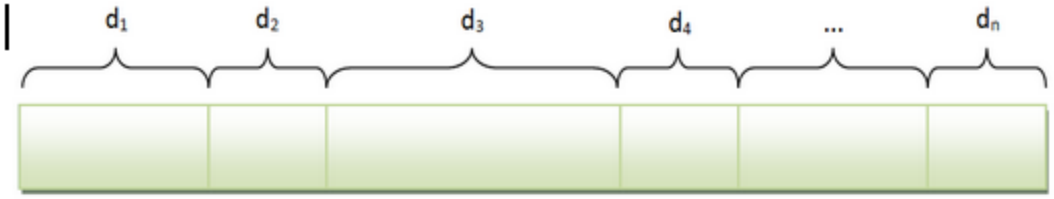


figure: Log file of a user.

n = number of detections.

di = number of interactions between a specific detection.

$I = \sum_{i=0}^{n} di$ Sum of all interactions.

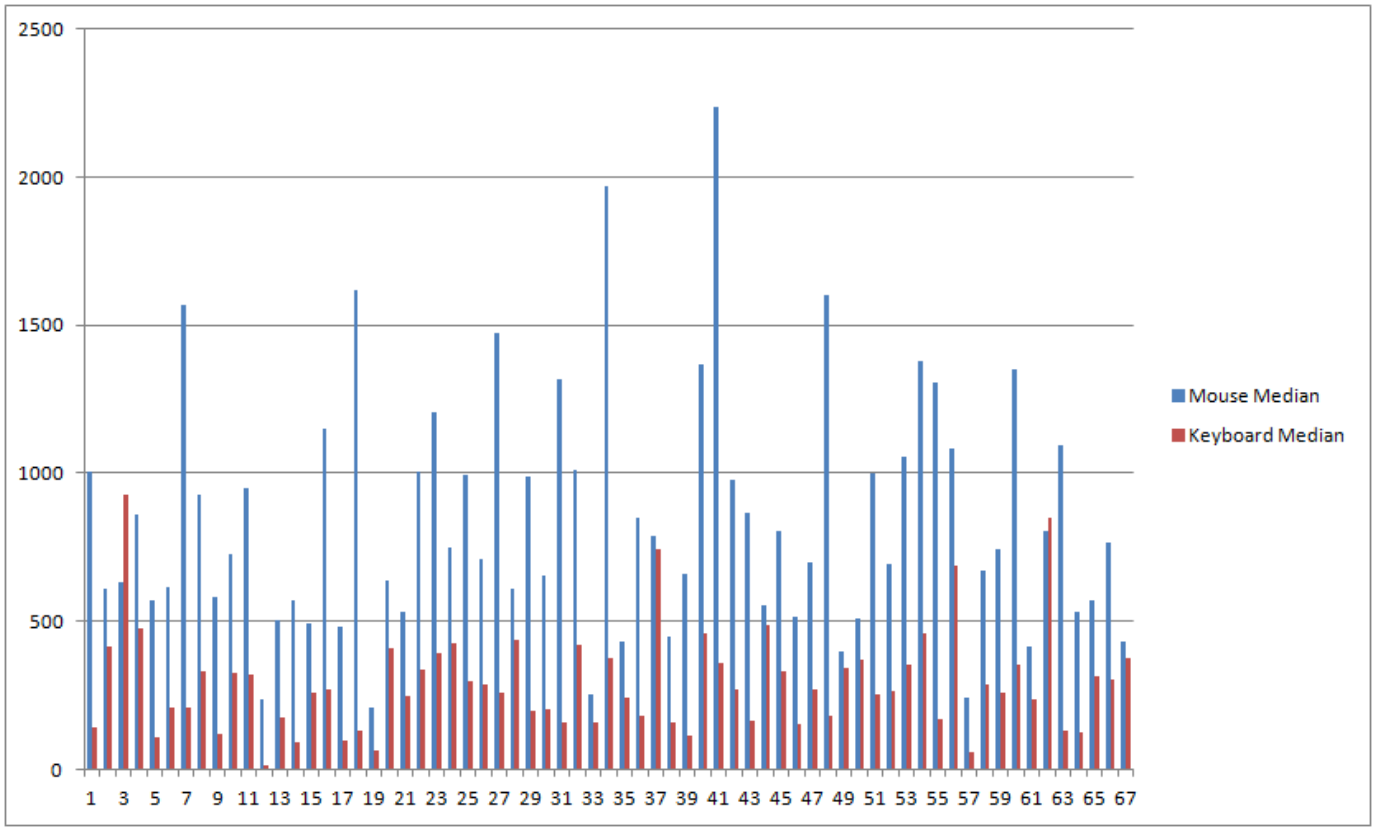
$D = \frac{n}{I}$ The average nr of detections per interaction.

The average nr of detections per interaction D is one metric for the system. D ranges from 0 (never detects anything) and 1 (detects on every interaction). The wrong user should have a high value of D while the correct user would ideally have a low value of D .

Test group

- Sample base: 99 users for 10 weeks
- 67 fulfilled the 20 hours a week for 10 weeks requirement
- 22.2GB data in the latest dump.
- Active time is the amount of time the user have been constantly active. After 10 seconds of inactivity the user is no longer considered to be active.
- Collected data contain 2302.66 hours active time so far which is corresponds to 0.26 years worth of data. 2.8M interactions.
- Simulated attackers (cross comparisons) results in 92577.32 hours active active time which corresponds to 10.57 years. 3906 comparisons totaling 120M interactions.
- We Sampled keystrokes, mouse movements, and OS events (applications used, system footprint etc)

Mouse and keyboard in median on a typical work day



Results for Mouse/Keyboard

- Profile is built dynamically during the analysis and this is what the correct user is matched against. The simulated attacker is then attacking the fully trained profile.
  - The first 5000 interactions is hardcoded to be the training phase and is not included in the score for the correct user. The actual training time frame is to be evaluated in the next stages.
- An interaction is an event such as a mouse move or a key press.
  - If the time between two interactions exceeds 10 seconds it is not counted as active time. We think that the current inactive time used in following results is too low for real world.

Analysis

- Analysis software leveraged existing tools & extended. Updated for new formats and tests
- Tweaking activity
  - Random select a number of users to use as sample base throughout the tweaking activity.
  - Run the software and analyze the results
    - Identify the weak tests and tune the variables for the individual tests to make them stronger.
    - Implement filters if needed.
    - Iterate until goal is met.
  - Test against full sample base.

Hard Results

	Typical day	Correct	Incorrect
Mouse	743	810	86
Keyboard	267	268	6
App usage	7690	10979	88

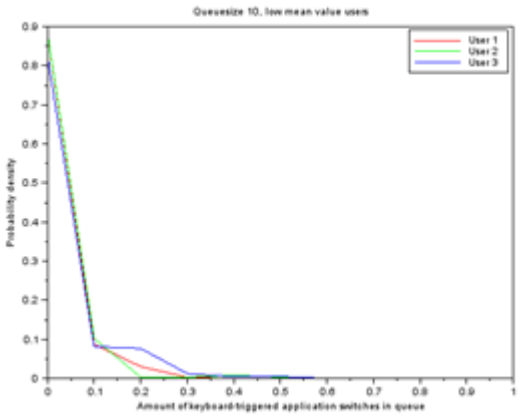
Interactions!

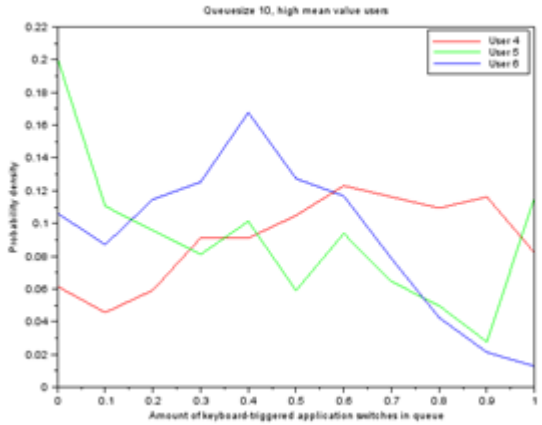
’Soft’ results

- What does it really mean?
- While the correct user can work through a regular workday without being falsely rejected the incorrect user would be detected within 10 seconds using keyboard (6 interactions, roughly 3 keys) or just less than 3.5 minutes using mouse (86 interactions).

Application Usage

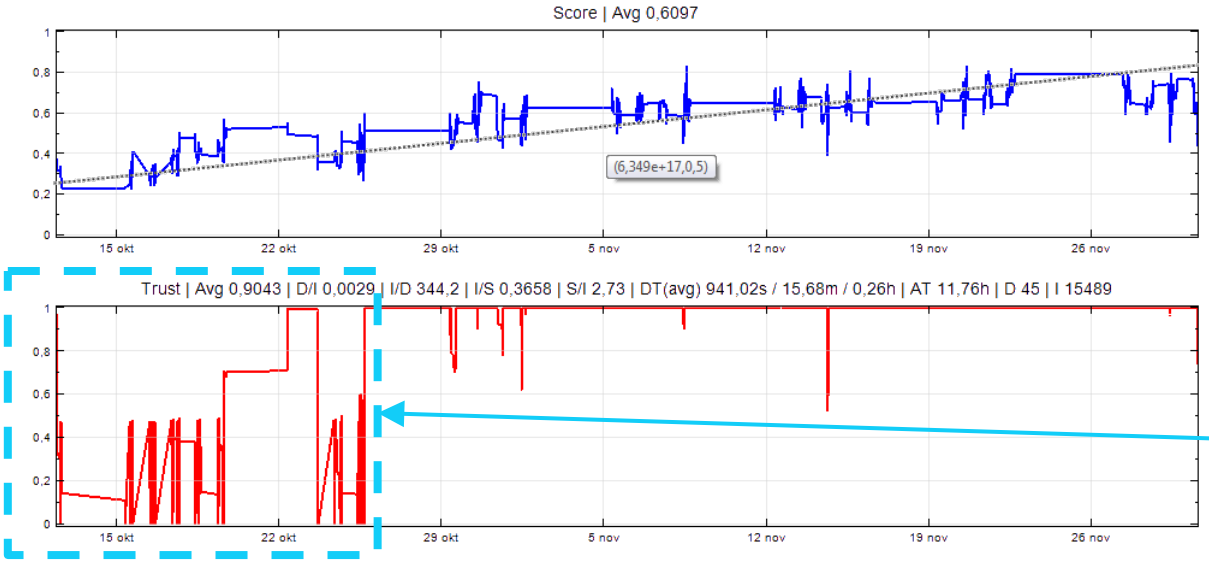
- Events that happened just before application launch
  - ie.. how does someone start an application.
- Can categorize people into groups
  - Three groups: mouse / mouse & keyboard / keyboard
  - The ratio between the groups are 50% / 30% / 20% and users are consistent





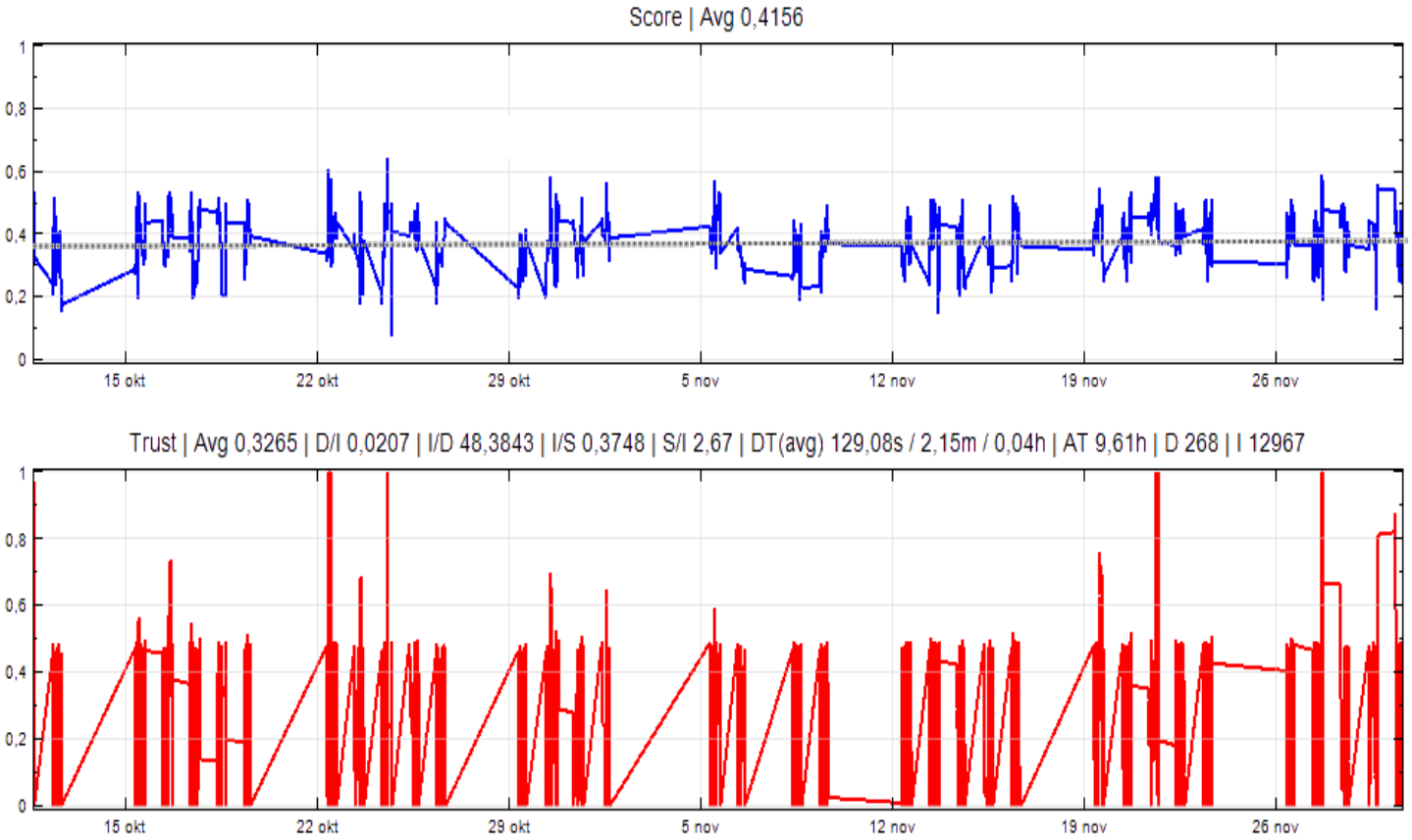
Real world 'trust‘ correct user

Correct user



Training phase

Real world 'trust‘ fraudster

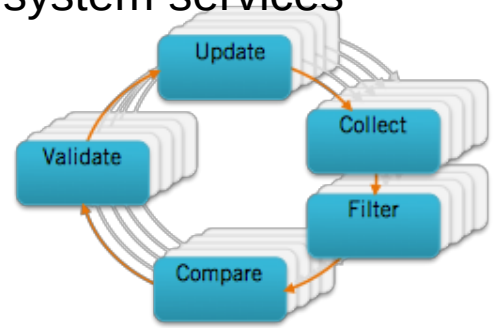


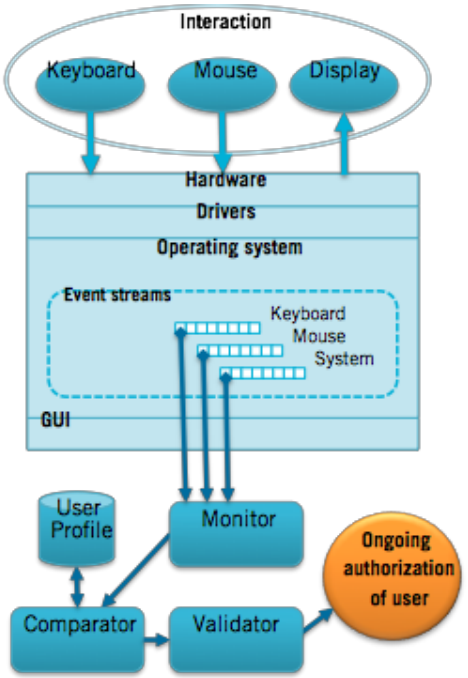
Observations from year one

- Definitely interesting biometric modalities
- Adding modalities together is hard
- ‘Attack’ users are hard
- ‘Time’ is hard
- Trial users are hard

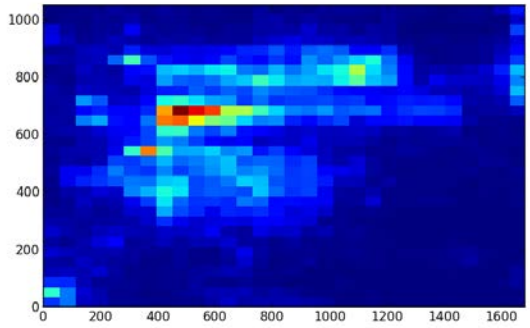
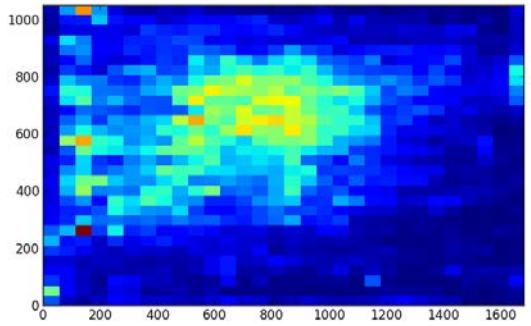
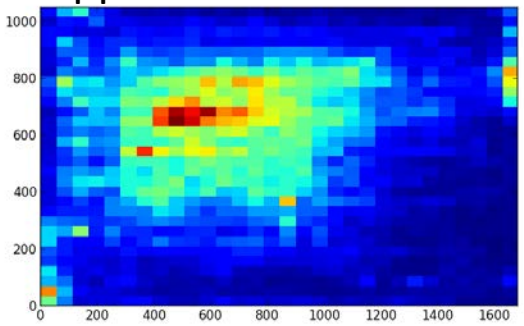
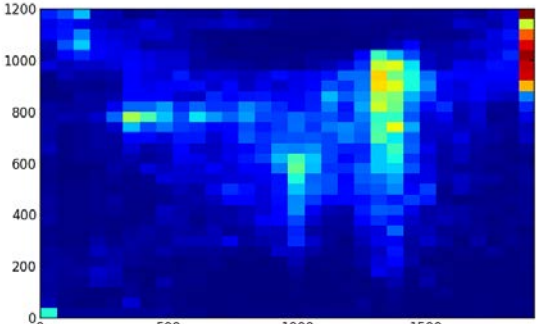
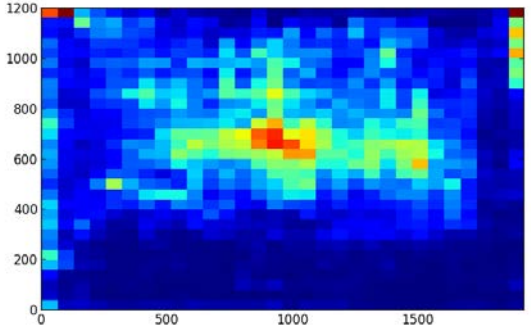
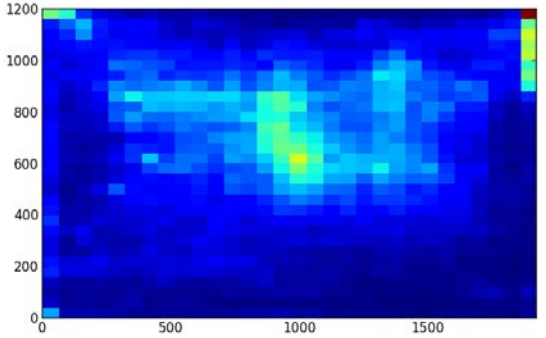
Where we going with all this…

- Anti-virus like system service
- Installed in administrator security space/storage
- Has 'policy'.
- Can be 'polled' by remote administrator.
- Reports via standard system services
  - Performance monitor
  - Event log





backup

	Application A	Application B	All applications
User 1			
User 2			
User 3