



Continuous user identity verification using keystroke analysis

Steven Furnell

Network Research Group
University of Plymouth
United Kingdom

Overview

Introduction

Concepts and previous work

An experimental study

Conclusions



An introduction to keystroke dynamics

Introduction

- Desirable to increase the security beyond secret-knowledge based approaches
- Potentially beneficial to provide:
 - a method that is resilient to accidental or deliberate compromise by the legitimate users
 - transparent / non-intrusive authentication (where possible)
 - reduce or remove the perceived inconvenience for users
 - continuous or periodic authentication of the user
 - enable confidence in the identity to be maintained beyond login

Keystroke Dynamics

- A behavioural biometric, based upon the ability to recognise characteristic typing rhythms
- Applicable to a variety of keyboard/keypad-based devices
- Requires no additional hardware
 - completely implemented in software
- Can operate in parallel with normal activities
 - ↳ non-intrusive additional security
 - ↳ minimises inconvenience for users

Potential measures

- Digraph latency
- Trigraph latency
- Keyword latency
- Keystroke duration (hold time)
- Typing errors
- Force of keystrokes
- Keystrokes or words per minute



Application scenarios



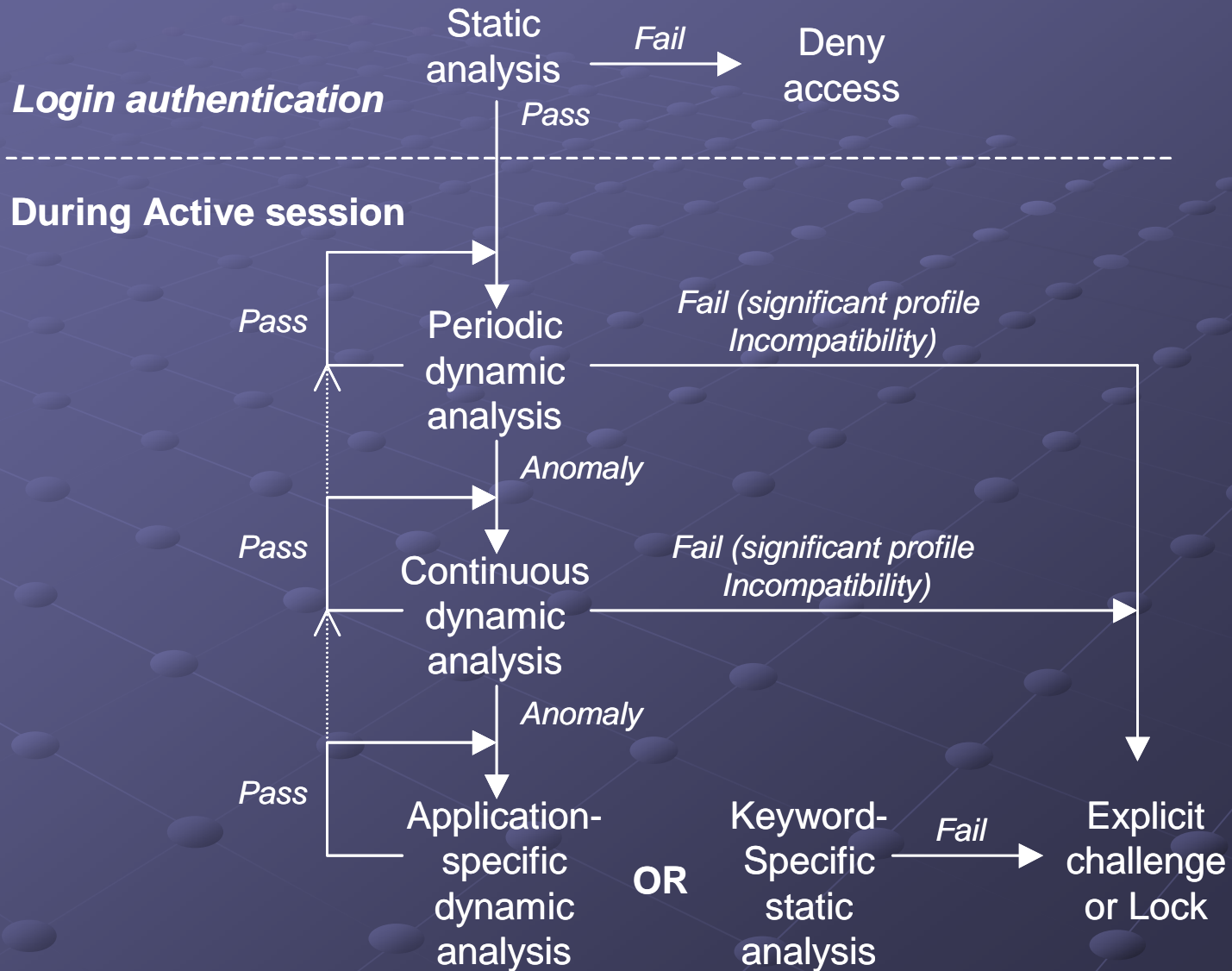
● Static

- At login
- Keyword-specific

● Dynamic

- Periodic
- Continuous
- Application-specific

Application scenarios



Previous work

- Focused upon digraph latencies
- Used statistical / NN approaches
- FAR/FRR rates $\sim < 10\%$
- Data collected in controlled environments

Previous work – examples

Study	Classification Technique		Users	FAR (%)	FRR (%)
Joyce & Gupta 1990	Static	Statistical	33	0.25	16.36
Leggett et al. 1991	Dynamic	Statistical	36	12.8	11.1
Brown & Rogers 1993	Static	Neural Network	25	0	12.0
Napier et al 1995	Dynamic	Statistical	24	3.8% (combined)	
Obaidat & Sadoun 1997	Static	Statistical	15	0.7	1.9
		Neural Network		0	0
Monrose & Rubin 1999	Static	Statistical	63	7.9 (combined)	
Cho et al. 2000	Static	Neural Network	21	0	1
Guven and Sogukpinar 2003	Static	Statistical	12	1	10.7



A long-term experimental study

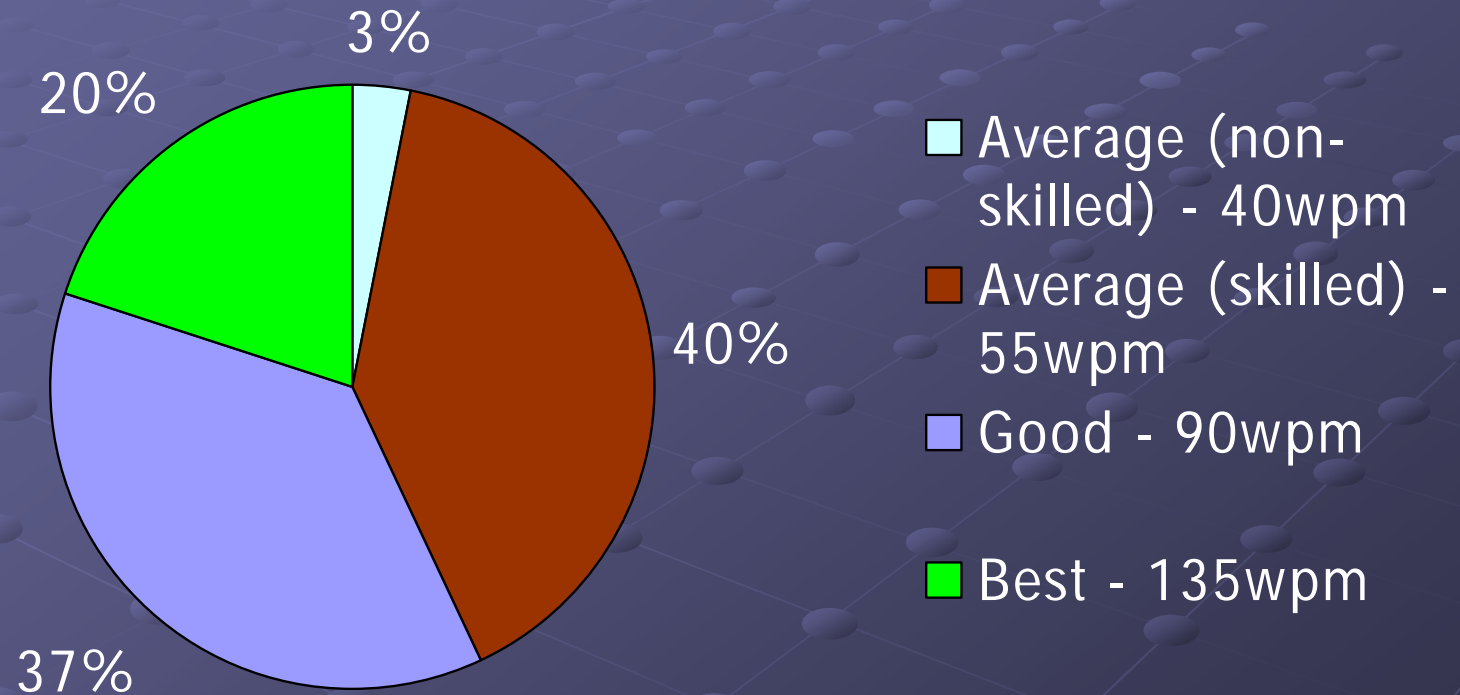
Experiment overview

- Digraph, trigraph and keyword logging
- Applications being used also logged
- 35 subjects
- 3 month logging period
- Nearly 6 million samples logged

Experiment overview

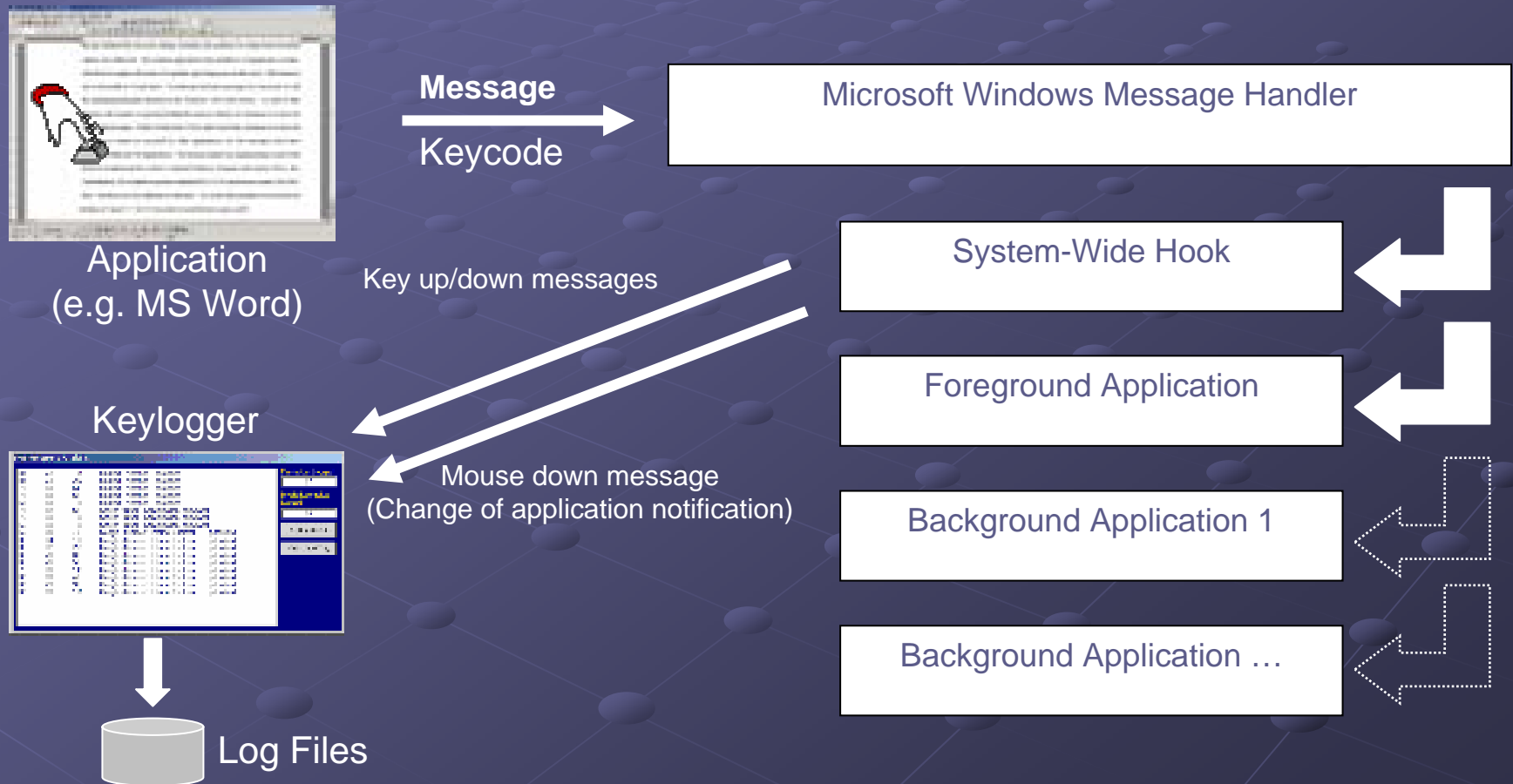
- Keylogger installed on client PC
- Transparently monitored latency of:
 - Digraphs (e.g. T-H)
 - Trigraphs (e.g. T-H-E)
 - Keywords (e.g. T-H-E-R-E)

Participant typist skills



Based upon typist speed
Classification from Card et al (1980)

System-wide keystroke logging



Keystroke Logger

IMS Keystroke Analysis

C1	C2	Duration	Application	Date/Time
U	S	19	Chapter7.doc - Microsoft Word	09/02/2004 15:46:57
S	E	62	Chapter7.doc - Microsoft Word	09/02/2004 15:46:57
E	D	107	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
N	T	29	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
T	H	60	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
H	E	54	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
E	A	132	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
A	R	124	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
R	L	73	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
R	I	133	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59

C1	C2	C3	Duration	Application	Date/Time
U	S	E	169	Chapter7.doc - Microsoft Word	09/02/2004 15:46:57
S	E	D	215	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
T	H	E	152	Chapter7.doc - Microsoft Word	09/02/2004 15:46:58
E	A	R	309	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
A	R	L	252	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
A	R	I	311	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
L	I	E	81	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
I	E	R	163	Chapter7.doc - Microsoft Word	09/02/2004 15:46:59
T	R	I	209	Chapter7.doc - Microsoft Word	09/02/2004 15:47:00
R	I	A	198	Chapter7.doc - Microsoft Word	09/02/2004 15:47:00

Word	Duration	Application	Date/Time
word	742	Chapter7.doc - Microsoft Word	09/02/2004 15:47:18

Keystrokes Logged:

Invalid Keystrokes Logged:

Average Typing Speed (ms):

Stop Logging

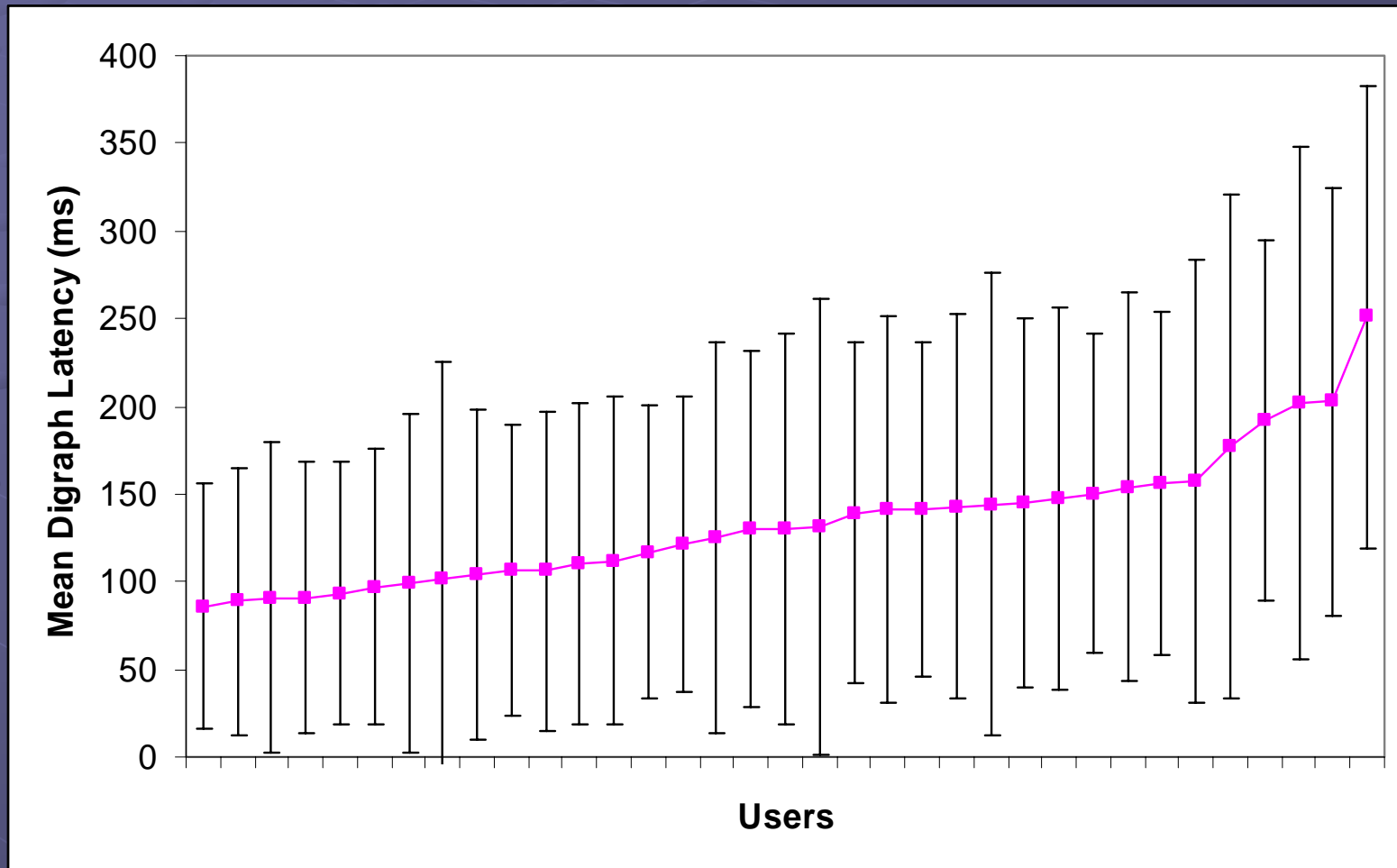
Minimise to Tray

Exit

Results

- Profiles were generated for each participant
 - one each for digraphs, trigraphs and keywords
- Sessions were then re-played to a comparator:
 - legitimate user samples (to measure FRR)
 - other user samples (to measure FAR)

Users' keystroke ranges



Results – The good

- Found to be very effective for some users, for example:
 - comparator suggested that User 11 would be able to enter 227,660 keystrokes before a false rejection
 - almost half of impostors against User 11's profile would be challenged before 50 keystrokes

Results – The bad

- For other users, it was much less effective:
 - comparator suggested that User 23 would only enter 23 keystrokes before a false rejection
 - meanwhile, several impostors would be permitted to enter several thousand keystrokes

Overall results

(average across users)

- System optimised for 0% FRR
 - desirable in a dynamic monitoring scenario
- Resulting average FARs for the three metrics:
 - Digraphs – 4.9%
 - Trigraphs – 9.1%
 - Keywords – 15.2%
- Suggests that accuracy is linked to the number of samples that were used in generating the profiles



Conclusions

Conclusions

- Keystroke dynamics can provide transparent authentication/supervision
- The technique showed promising performance in user trials
- Not as powerful as physiological biometrics, but nonetheless very effective for some users
 - Does not work equally well for all users

Future development

- Need to consider complementary measures for a composite approach:
 - Keystroke dynamics and mouse dynamics
 - Keystroke dynamics and facial recognition
 - Mouse dynamics and voice recognition
- Statistical analysis techniques
- Application-specific profiling
- Even larger scale trials
- Evaluation of impairments (e.g. fatigue, illness)



Dr Steven Furnell
sfurnell@plymouth.ac.uk

Network Research Group
www.network-research-group.org