Cognitive Load Inference for Ubiquitous Computing Adaptation

Credit: Shutterstock.com

Dr Veljko Pejović

@veljkoveljko

Faculty of Computer and Information Science, University of Ljubljana

University of Ljubljana
Faculty of Computer and Information Science

University of Novi Sad, June 2019

Background

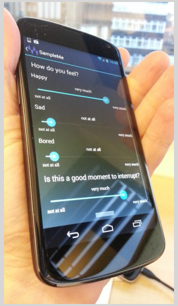
- PhD University of California Santa Barbara, CA, USA
  - Resource-efficient wireless solutions for rural areas
- Postdoc Uni. of Birmingham, UK
  - Mobile sensing for behaviour change interventions
- Assistant Professor at University of Ljubljana
  - Leading mobile computing research

Research Areas

More at: lrss.fri.uni-lj.si/Veljko

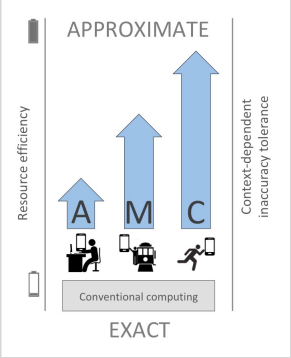
- Mobile sensing and user behaviour modelling

We built mobile apps, conduct user studies, and engineer innovative approaches for notification management



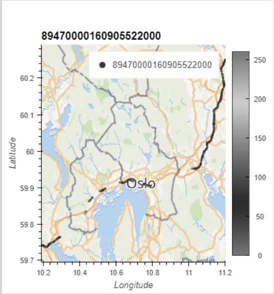
- Resource-efficient computing

We propose approximate mobile computing where the accuracy of computation changes with the context of use



- Mobile network data analysis

We build data mining tools for network traffic analysis



University of Ljubljana
Faculty of Computer and Information Science

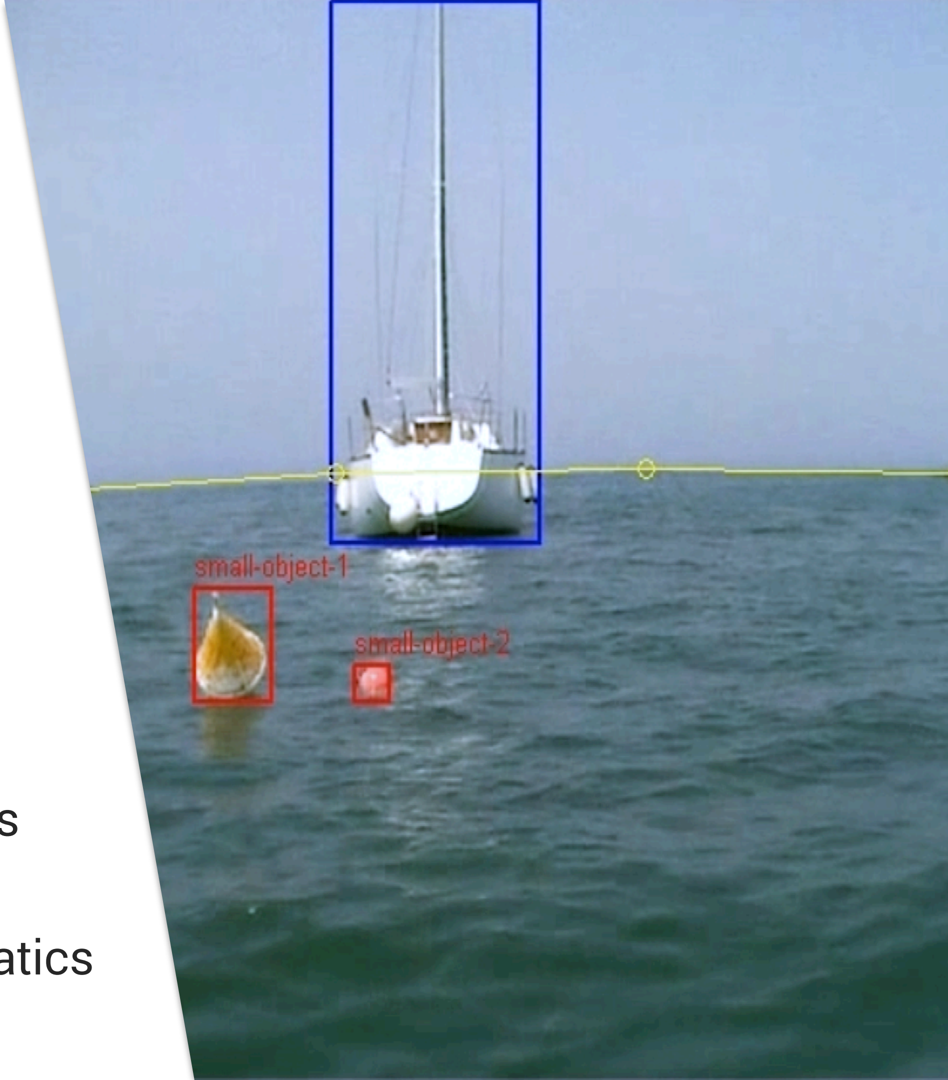
University of Ljubljana

- The oldest, largest, best-ranked university in Slovenia
- Top 500 on Shanghai, Webometrics, Times ranking lists
- Faculty of computer and information science (FRI)
  - 170 employees
  - 1400 students
  - BS/MS/PhD

University *of Ljubljana*
Faculty *of Computer and Information Science*

Research @ FRI

- 19 labs - a range of CS fields
- Traditionally strong in AI
  - Machine learning
  - Data mining
  - Computer vision
  - Bioinformatics
- Highlights:
  - Visual tracking challenge winners
  - Kaggle winners
  - Nature publications in bioinformatics



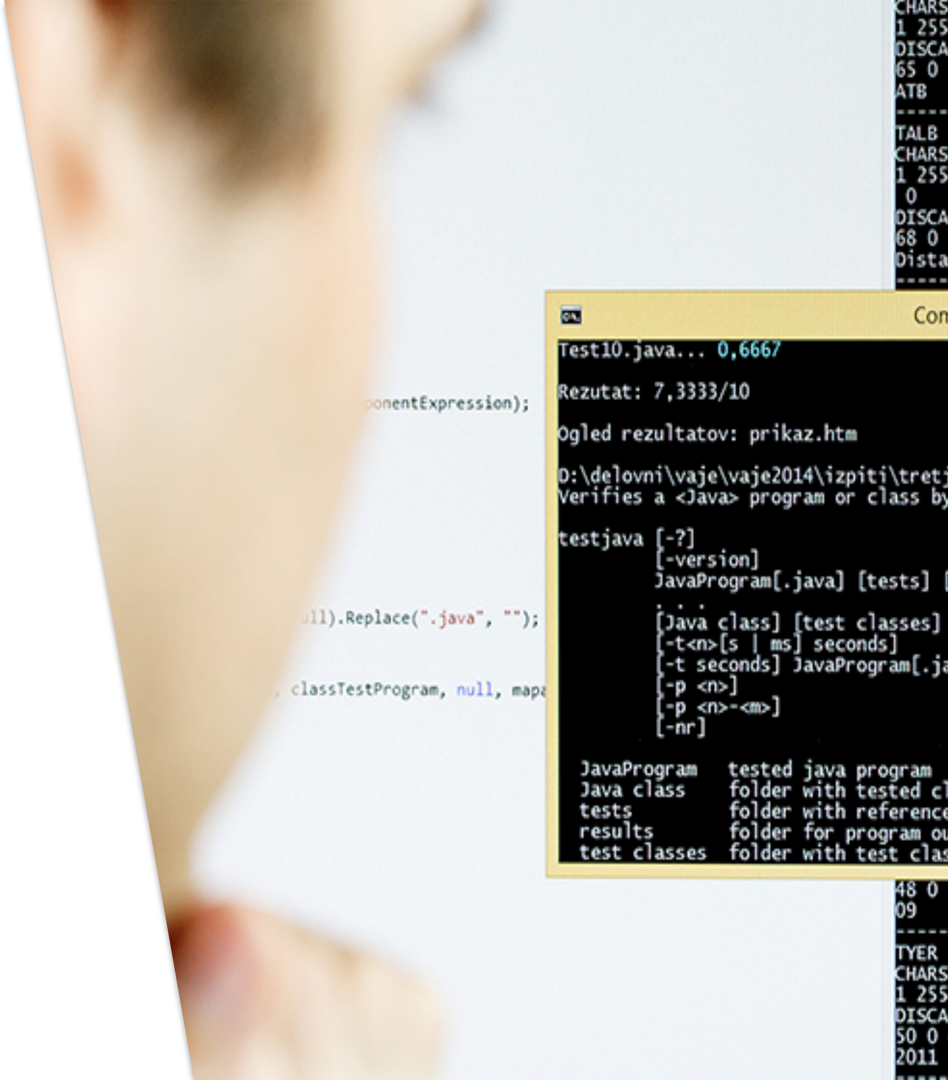
Mobile Systems Research – HCI Perspective

Mobile Notifications

- Increasingly interactive lives
  - 100 notifications/day per user
- For recipients, a means of information awareness
  - Anxious without notifications
- For senders, a way to initiate remote communication

Poor Notification Timing

- Reduced work efficiency



University of Ljubljana
Faculty of Computer and Information Science

Poor Notification Timing

- Reduced work efficiency
- Missed marketing opportunities

University *of Ljubljana*
Faculty *of Computer and Information Science*

Poor Notification Timing

- Reduced work efficiency
- Missed marketing opportunities
- Critical safety consequences

"There is more information available at our fingertips during a walk in the woods than in any computer system, yet people find a walk among trees relaxing and computers frustrating. Machines that fit the human environment instead of forcing humans to enter theirs will make using a computer as refreshing as taking a walk in the woods."

Mark Weiser, 1991

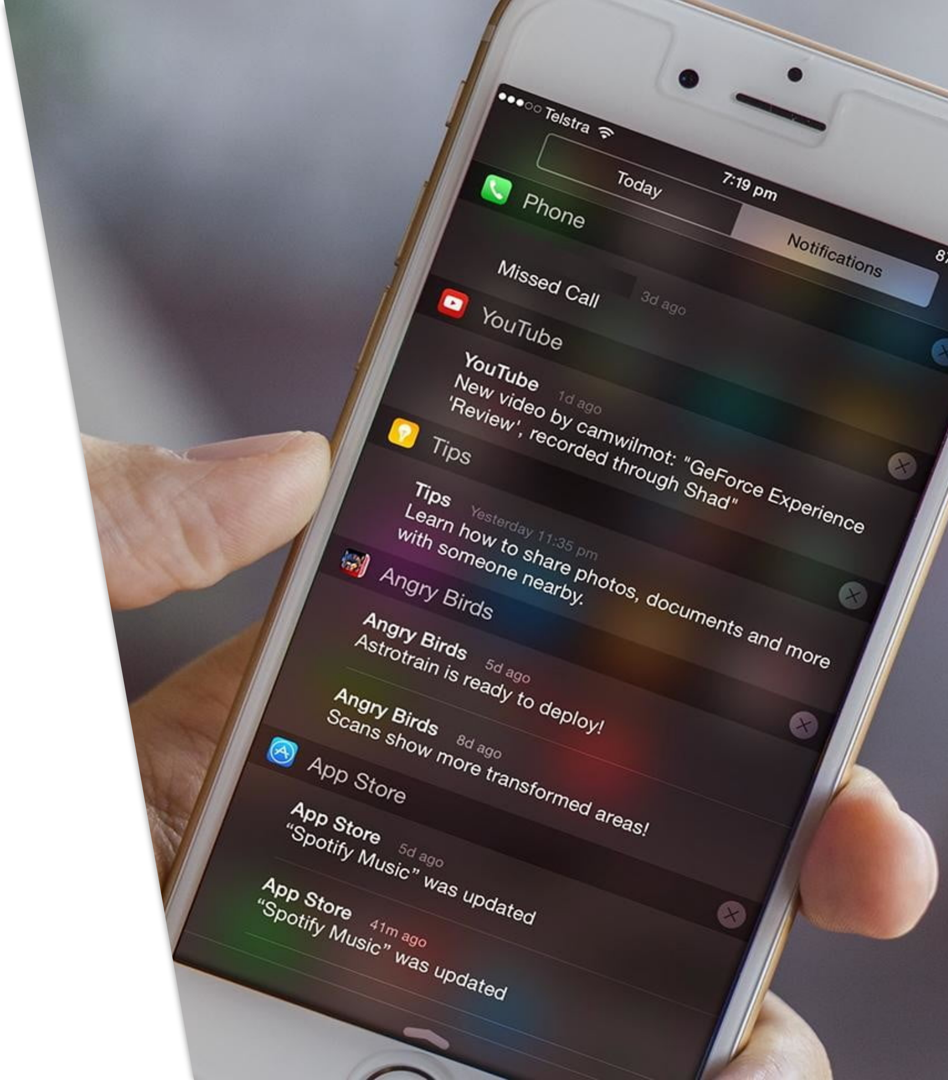
Building a system for intelligent notification scheduling

University of Ljubljana
Faculty of Computer and Information Science

V. Pejovic and M. Musolesi
InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications
UbiComp'14, Seattle, WA, USA

Towards Timely Interaction

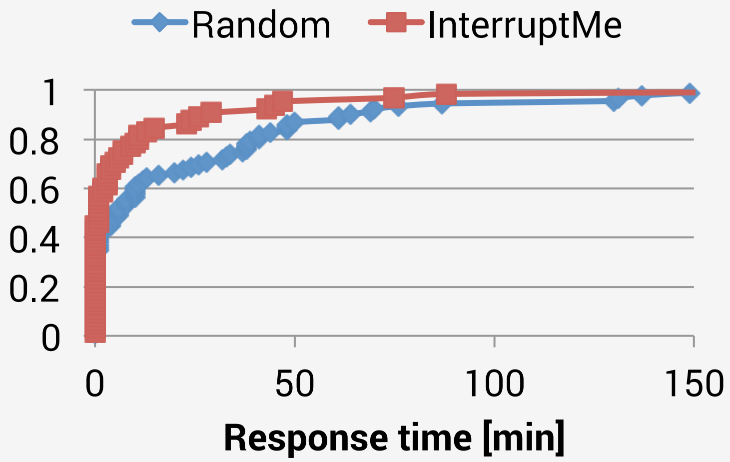
- **Premise:** notification timing is the key!
- **Path:** identify opportune moments to deliver information
- **Hypothesis:** sensed context reveals interruptibility

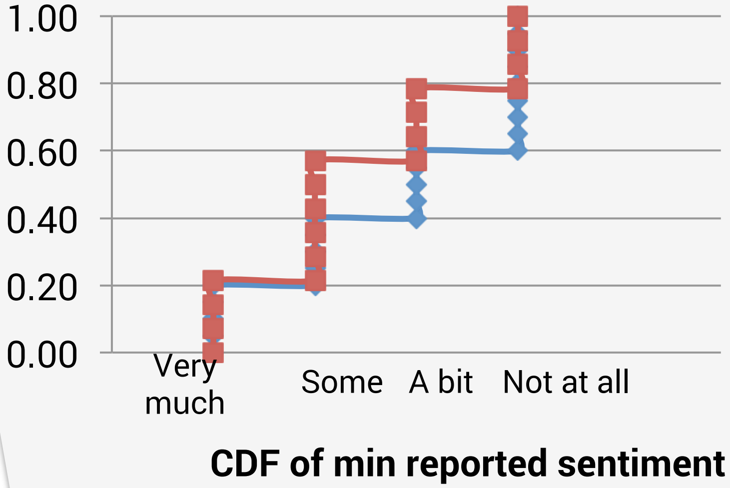


InterruptMe

- Android library for notification management
- Senses
  - accelerometer
  - location
  - time of day
- Machine learning model learns a user’s interruptibility patterns

bitbucket.org/veljkop/intelligenttrigger



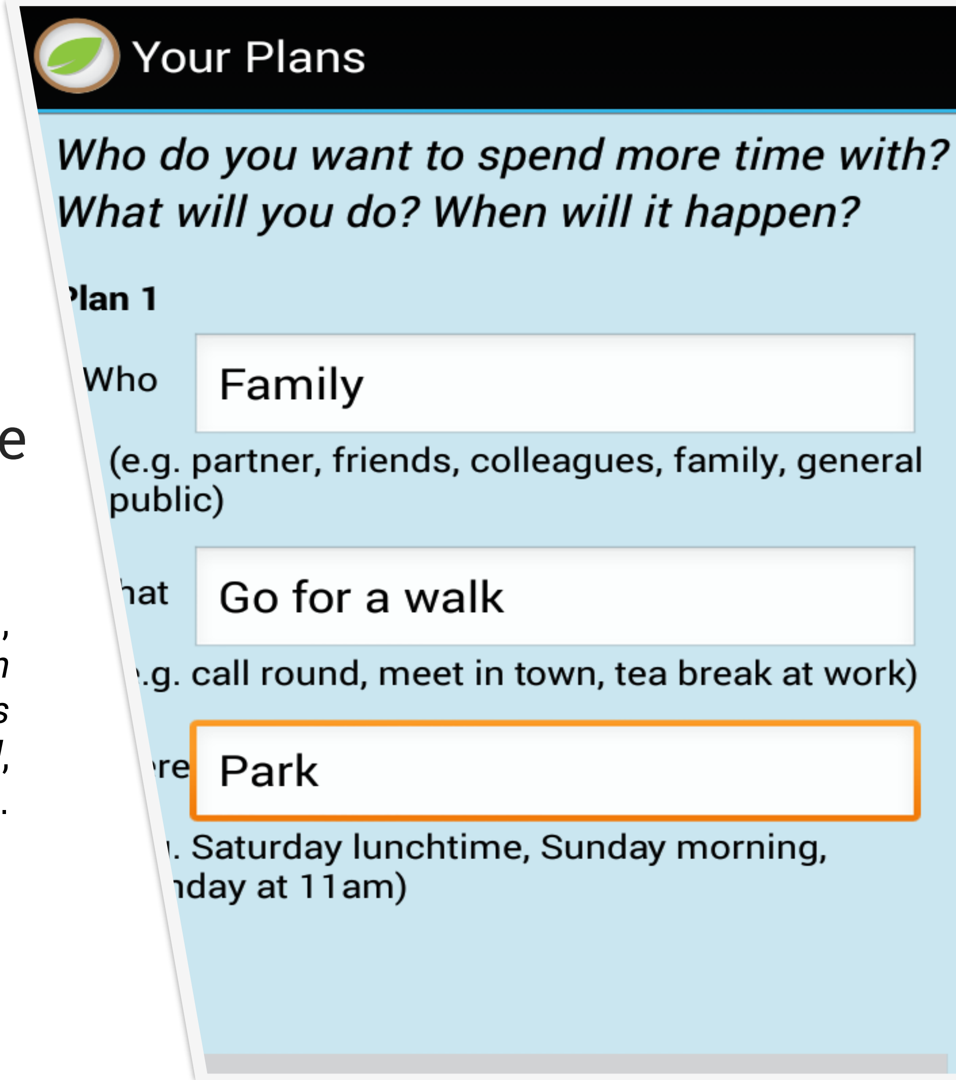


Problem solved?

Real-world Trial

- … no significant effects of notification scheduling on the usage of a behavioural change intervention app

L Morrison et al.,
The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial,
PLoS ONE, Vol 12, (2017).



Understanding factors affecting notification acceptance

A. Mehrotra, M. Musolesi, R. Hendley and V. Pejovic

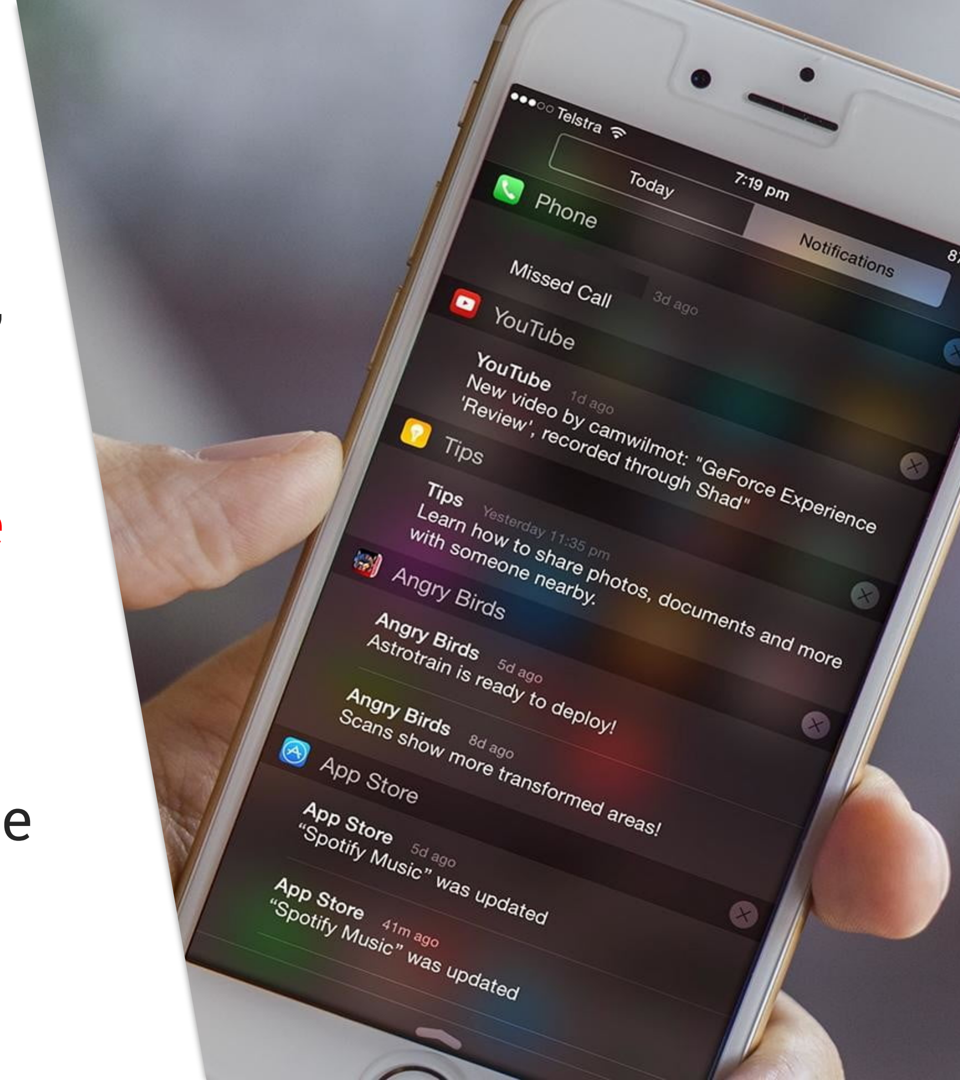
Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications

UbiComp'15, Osaka, Japan, September 2015.

University *of Ljubljana*
Faculty *of Computer and Information Science*

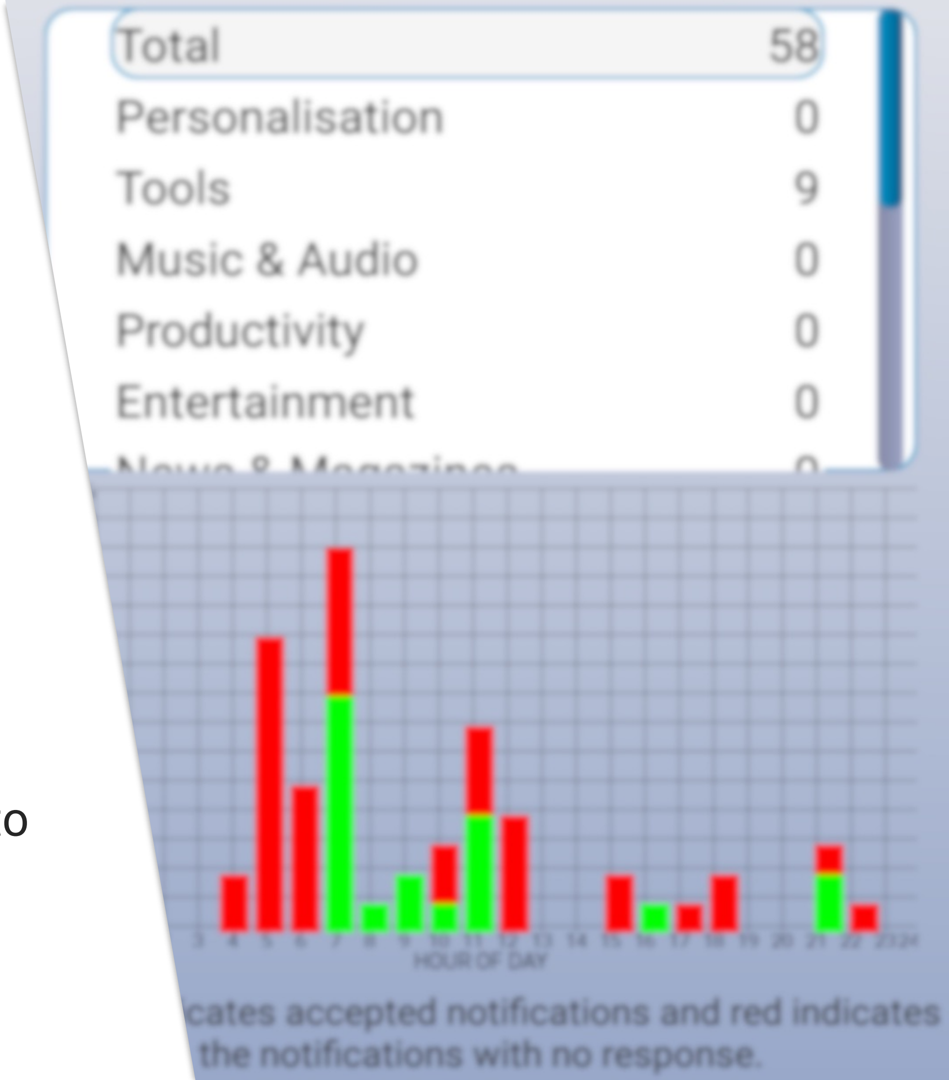
Towards Timely Interaction

- **Premise:** location, movement, and time sensing is not enough
- **Path:** monitor other on-device factors that may impact interruptibility
- **Hypothesis:** application type, content, sender, etc. determine a user’s reaction



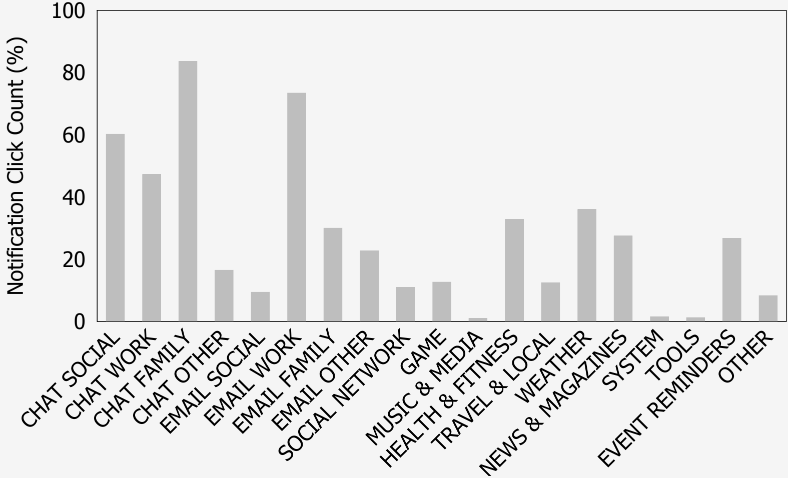
NotifyMe Mobile App

- Senses context
- Records reaction to a notification
  - Notification data
  - Category
  - Sender ID
- Gathers user preferences
  - Where and when would you like to receive notifications with similar content



Notification Reaction Analysis

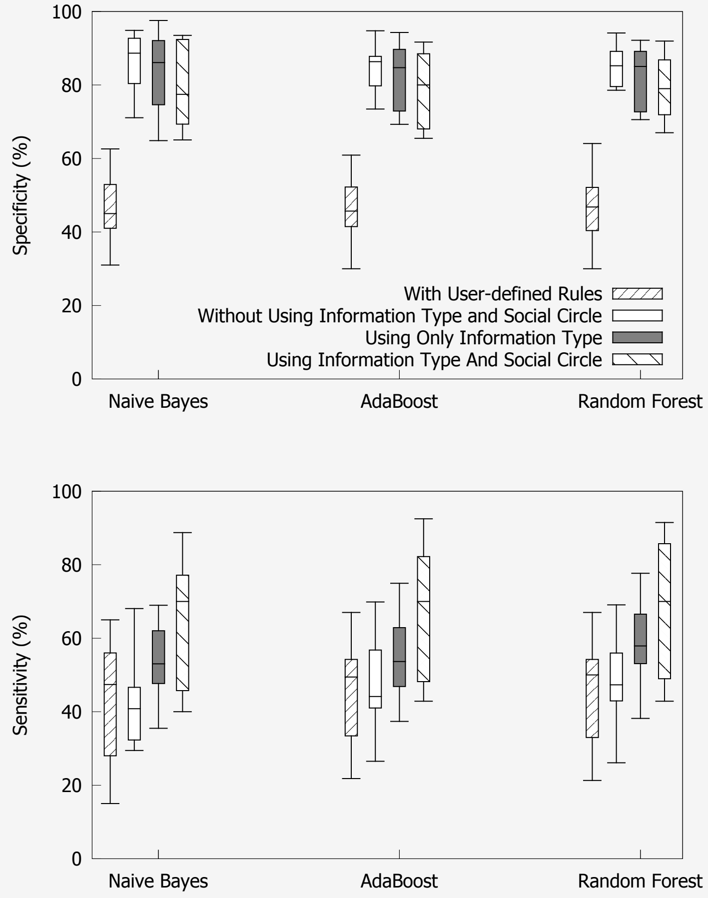
- Notification click count differs between application types (i.e. content type) and sender-receiver relations



University of Ljubljana
Faculty of Computer and Information Science

Notification Reaction Prediction

- By using information type and social circle we were able to predict the acceptance of a notification within 10 minutes from its arrival time with an average sensitivity of 70% and a specificity of 80%
- Better than user-defined rules



University *of Ljubljana*
Faculty *of Computer and Information Science*

User reaction does not imply user satisfaction

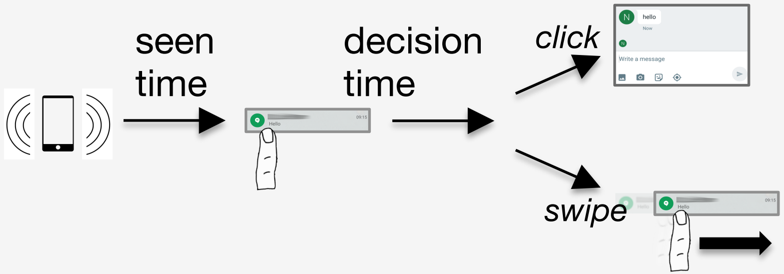
A. Mehrotra, V. Pejovic, J. Vermeulen, R. Hendley and M. Musolesi
My Phone and Me: Understanding User's Receptivity to Mobile Notifications
ACM CHI'16, San Jose, CA, USA, May 2016.

Towards Timely Interaction

- **Premise:** we identified a number of factors that impact reactions, but reactions are diverse
- **Path:** monitor users' actions and the surrounding factors
- **Hypothesis:** sensed context reveals reaction and disruption

My Phone and Me App

- Automated logging:
  - Notification time of arrival, seen, removal
  - Notification response
  - Notification details (title, app)
  - Alert type
  - Context (activity, location, etc.)
- Experience sampling:
  - Sender-receiver relationship, personality, task engagement



University of Ljubljana
Faculty of Computer and Information Science

Disruption Analysis

- Task complexity and interruptibility:
  - More disruptive if it arrives when the user is in the middle of or finishing a task
  - Perceived disruption increases with the complexity of an ongoing task
  - Faster to react if engaged in a complex task

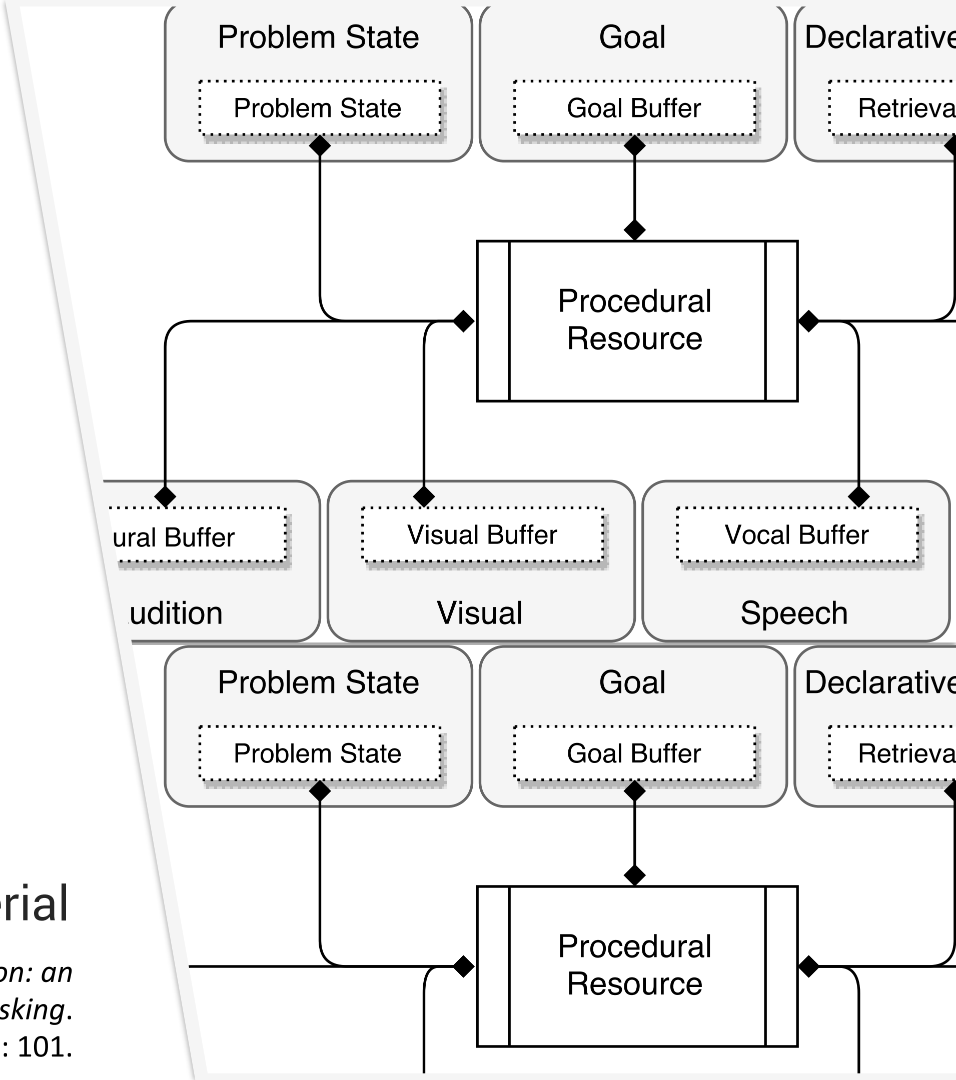
Also confirmed:

Pejovic et al., "Investigating The Role of Task Engagement in Mobile Interruptibility", Smarttention workshop with Ubicomp'15

How does a thought get disrupted?

Theory of Multitasking

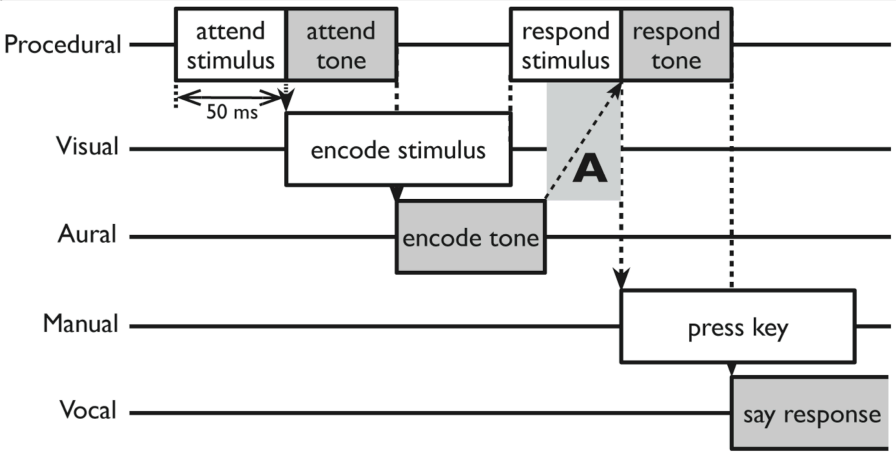
- Resources:
  - Perceptual and motor
  - Cognitive
    - Procedural memory
    - Declarative memory
- Mechanisms:
  - Resource use is exclusive – one task at a time per resource
  - Multiple problem threads run in parallel, but processing is still serial



Salvucci and Taatgen. *Threaded cognition: an integrated theory of concurrent multitasking*. Psychological review 115.1 (2008): 101.

Theory of Multitasking

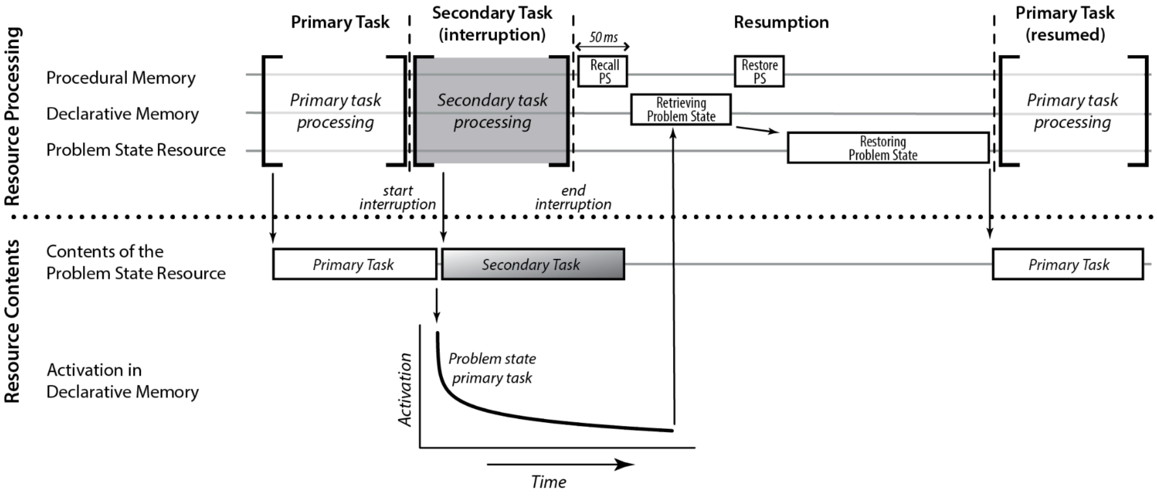
- Interference when two or more threads ask for the same resource at a time



Borst et al.
The problem state: a cognitive bottleneck in multitasking.
Journal of Experimental Psychology: Learning, memory, and cognition 36.2 (2010): 363.

Theory of Multitasking

- Complex tasks require problem state saving/retrieving



Borst et al.
What Makes Interruptions Disruptive?: A Process-Model Account of the Effects of the Problem State Bottleneck on Task Interruption and Resumption.
CHI'15, 2015.

Implications on Mobile Attention Management

- Interruptions are more disruptive if they require problem state switching

University of Ljubljana
Faculty of Computer and Information Science

Implications on Mobile Attention Management

- Make them less disruptive by interrupting:
  - At moments when a task is not fully active (e.g. just starting, or just finished)
  - At moments when a task does not require a problem state
  - At moments when a user is working on a task that is well practiced, a routine


"Mr. Osborne, may I be excused? My brain is full."

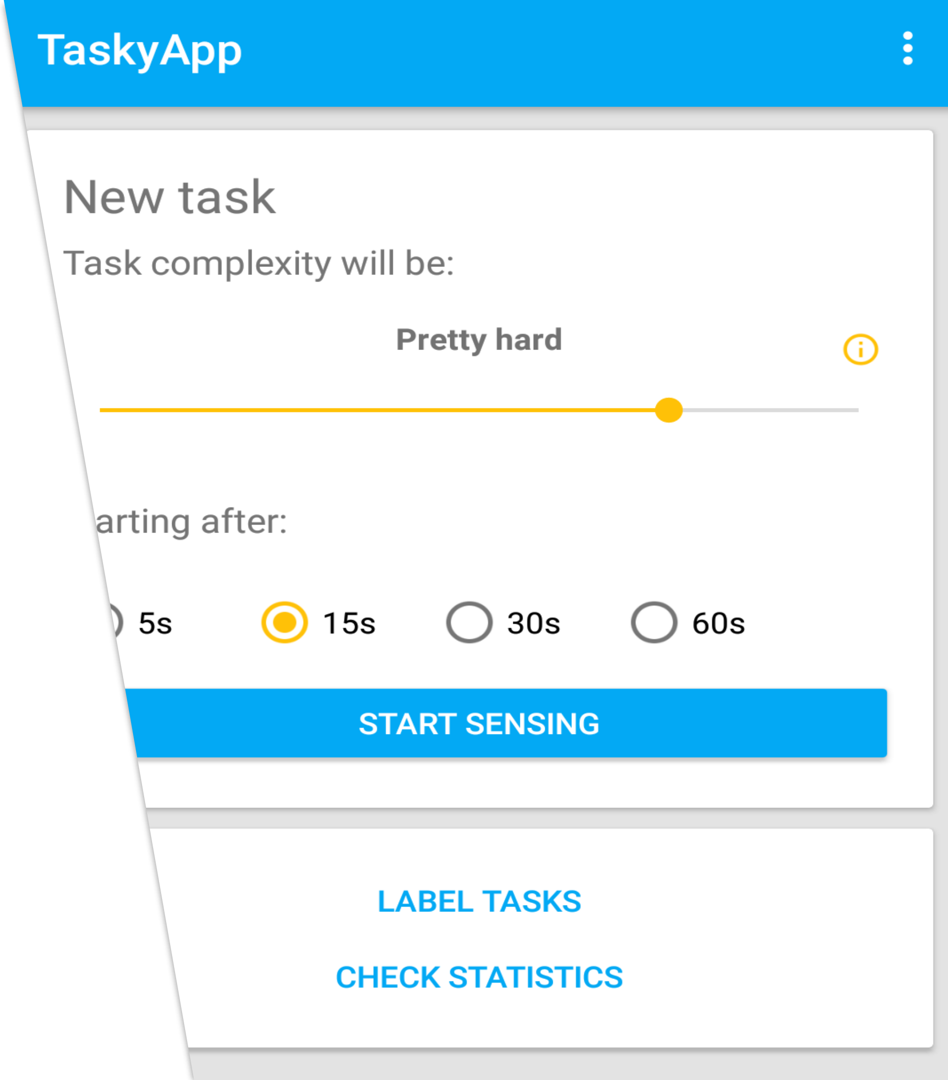
Can we automatically infer task engagement with smartphones?

G. Urh and V. Pejovic,
TaskyApp: Inferring Task Engagement via Smartphone Sensing
Ubittention workshop with ACM UbiComp'16, Heidelberg, Germany.

University *of Ljubljana*
Faculty *of Computer and Information Science*

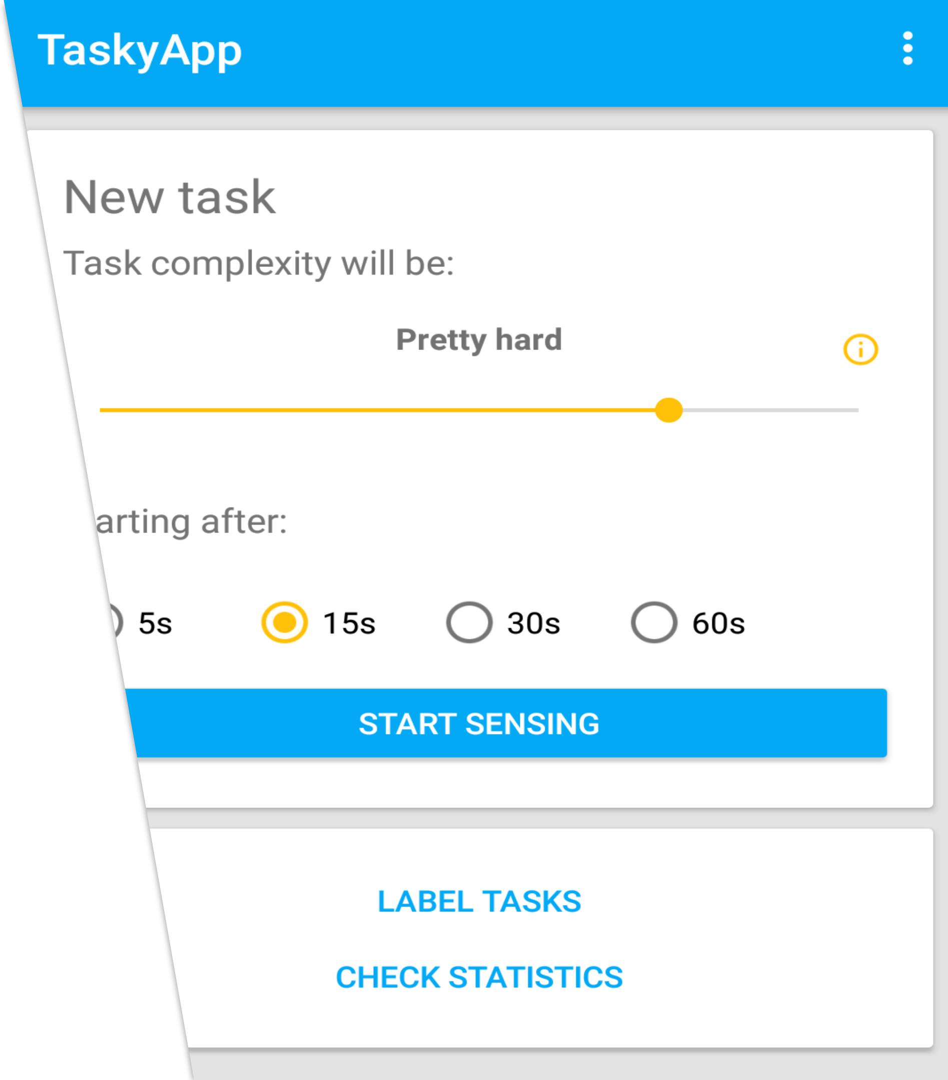
TaskyApp

- Background sensing of device movement, ambient sound, collocation with other devices
- Data labelling via experience sampling and retroactive assisted labelling



TaskyApp

- Recruited eight office workers for five weeks
  - 232 labelled instances (3035 unlabelled)
  - Most data between 8am and 6pm



Data Analysis

- Linear regression (N=232) fit with sensed features as independent variables and task difficulty (1-5) as a dependent variable
  - Movement data gives the most informative features
  - The regression explains only a small part of the data (R^2=0.19)

Variable	Coefficient	t (sig.)
Acc. Y-axis mean	-.038	-1.84 (.068)
Acc. Z-axis mean	.026	1.43 (.153)
Gyro. mean intensity crossing rate	0.003	4.06 (.000)
Gyro. intensity variance	0.200	1.24 (.217)
Hour of day	.067	3.49 (.001)
Majority	0.5	0.5

Data Analysis

- Classify a task engagement moment as either "easy" or "difficult" depending on the sensed features
  - We experimented with different classifiers but Naïve Bayes seems to work best (probably due to the low amount of data)
    - 62.5% accuracy (52.8% baseline)
    - "Favourable" errors

EASY'	DIFFICULT'	
45 (19.4%)	62 (26.7%)	EASY
25 (10.8%)	100 (43.1%)	DIFFICULT

Can we automatically infer task engagement with wearables?

University *of Ljubljana*
Faculty *of Computer and Information Science*

M. Gjoreski, M. Luštrek and V. Pejović,
My Watch Says I'm Busy: Inferring Cognitive Load with Low-Cost Wearables
Ubittention workshop with ACM UbiComp'18, Singapore.

Physiological Signals for Cognitive Load Inference

- **Premise:** heart rate (variability), electrodermal activity, pupil dilation, EEG changes correlate with CL changes
- **Path:** low-cost wearable sensing devices can capture signals ~ cognitive load
- **Hypothesis:** ML on these data to infer cognitive load

Collected Data

- Preliminary data:
  - Demographics
  - Cognitive capacities (N-back test)
  - Personality (Hexaco) test

Collected Data

- Primary (PC-based) task
  - Adapted from Haapalainen et al.
  - Six task types, each with three difficulty levels
  - NASA TLX after each task
- Physiological measurements
  - Heart rate intervals (R-R), galvanic skin response (GSR) and skin temperature (ST)
- Secondary task

Experiment

Part 1	Demographic Questionnaire	2-back task	3 minutes Rest	3-back task	3 minutes Rest	Personality Questionnaire

Part 2	P-task n Intensity x S-task	Task load Quest. + Rest	P-task n Intensity x S-task	Task load Quest. + Rest	P-task n Intensity x S-task	Task load Quest. + Rest	3 minutes Rest	6 cycles

- Demographics:
  - 25 users (21 completed successfully)
  - 20-58 years old
  - 5 female

Data Overview

- Extracted 81 physiological, demographic, cognitive capacity, and personality features
- Predicting three CL measures:
  - TLX (subjective)
  - Opacity (sec. task performace)
  - Task label (objective)

P-Task	(μ±δ)TLX	(μ±δ)Opacity	r(TLX-DTD)	r(TLX-Opacity)	r(DTD-Opacity)
HP	13.8 ± 4.7	0.1 ± 0.04	0.34	-0.01	0.13
FA	17.9 ± 7.8	0.1 ± 0.03	0.16	-0.08	0.07
GC	17.4 ± 6.1	0.1 ± 0.06	0.48	-0.06	-0.05
NC	17.7 ± 7.7	0.08 ± 0.03	0.34	-0.14	-0.01
SX	17.1 ± 7.7	0.12 ± 0.1	0.40	-0.21	-0.33
PT	17.4 ± 9.0	0.14 ± 0.16	0.43	-0.08	-0.27
Overall	16.9 ± 7.4	0.1 ± 0.08	0.34	-0.09	-0.13

Secondary task shows very weak correlation with TLX or DTD

University of Ljubljana
Faculty of Computer and Information Science

Cognitive Load Prediction

- Cast into classification task
- Classifiers: Naïve Bayesian, Random Forest, Gradient Boosting, AdaBoost, SVM, KNN, Trees
- Modestly better than the baseline
- Confuses neighbouring difficulties

Target	µ Majority	Best model	Best model µ Accuracy	Accuracy increase relative to Majority						
				HP	FA	GC	NC	SX	PT	µ
TLX	40%	RF	47%	6%	-5%	5%	6%	21%	10%	7%
DTD	33%	NB	51%	27%	11%	10%	22%	14%	24%	18%
Opacity	36%	GB	46%	16%	5%	13%	6%	3%	20%	10%

	Easy	Medium	Difficult
Easy	158	101	65
Medium	98	163	63
Difficult	69	91	164
Precision	49%	46%	56%
Recall	49%	50%	51%
F1	49%	48%	53%
Accuracy		51%	

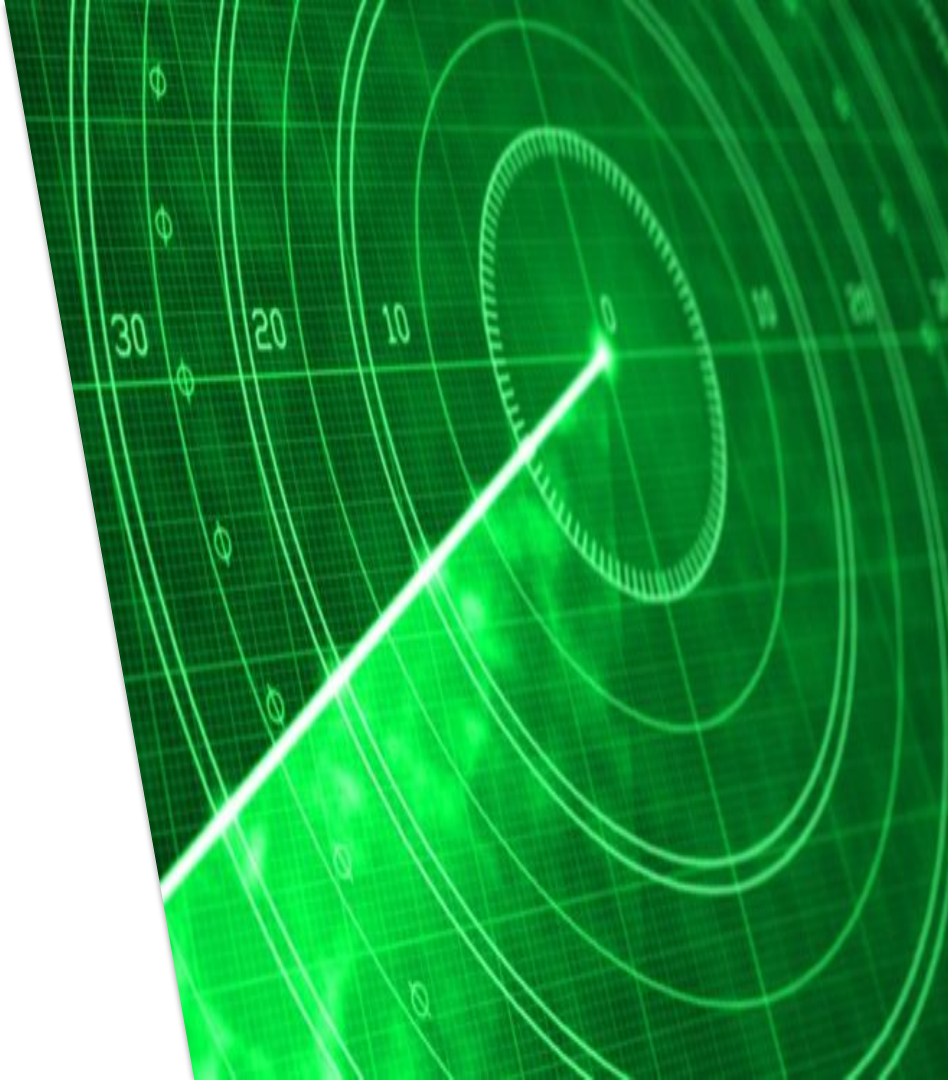
Fully unobtrusive task engagement inference

T. Matkovič and V. Pejović,
Wi-Mind: Wireless Mental Effort Inference
Ubittention workshop with ACM UbiComp'18, Singapore.

University of Ljubljana
Faculty of Computer and Information Science

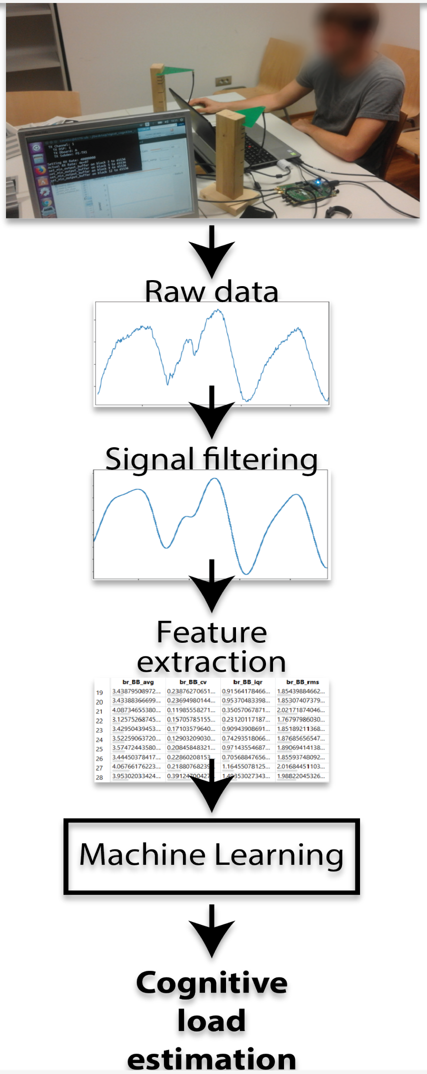
Wireless Cognitive Load Inference

- **Premise:** radar can detect breathing and heart beat related body movement
- **Path:** custom radar
- **Hypothesis:** filtered radar signals as a basis for ML models of CL



Wi-Mind

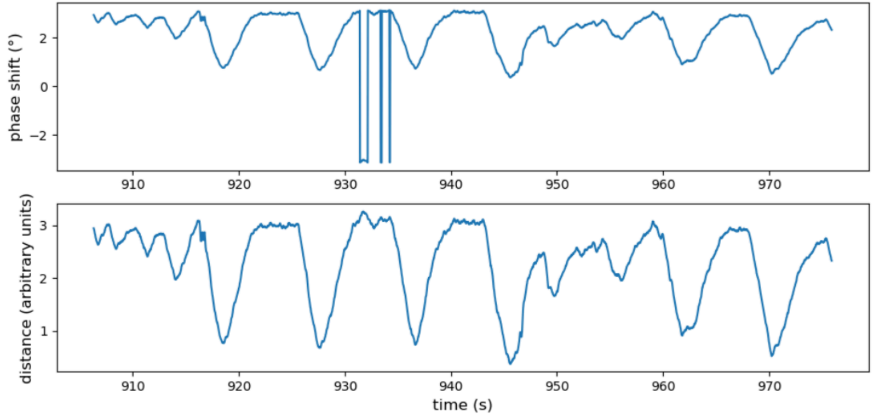
- Software-Defined Radio (SDR) implementation of FMCW radar followed by phase analysis
- Monitor movement as a user is solving tasks of different difficulty
- Extract heart beat and breathing-related features
- Build ML models

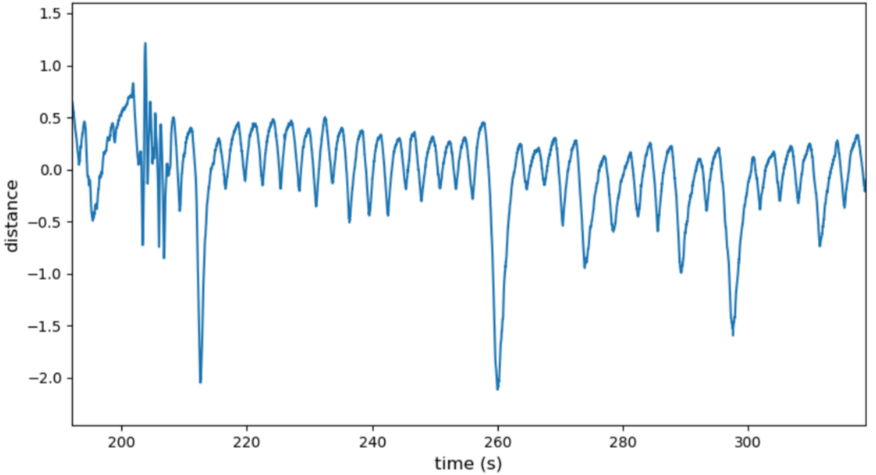


University *of Ljubljana*
Faculty *of Computer and Information Science*

From EM Waves to Physiological Signals

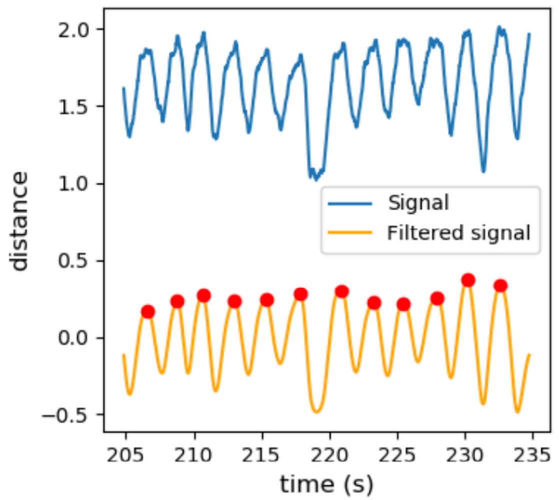
- Preprocessing:
  - Unwrapping phase
  - Filtering HF and LF noise

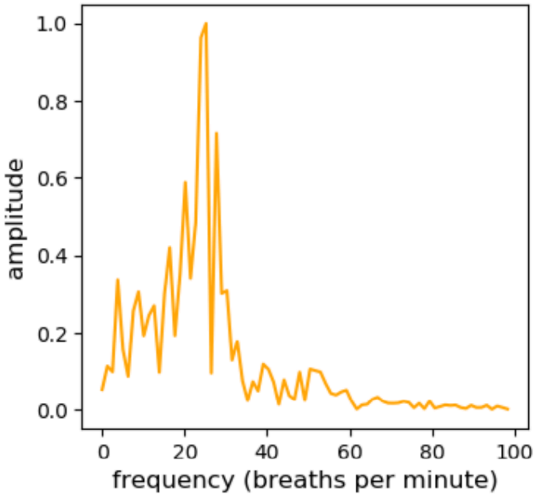




From EM Waves to Physiological Signals

- Preprocessing:
  - Unwrapping phase
  - Filtering HF and LF noise
- Extracting breathing signal
  - Breathing rate (via FFT) features: mean rate, power in different bands, etc.
  - Inter-breath features (peak detection): avg. interval, variation, I:E, etc.
- Metafeature
  - Is the signal “clean”?

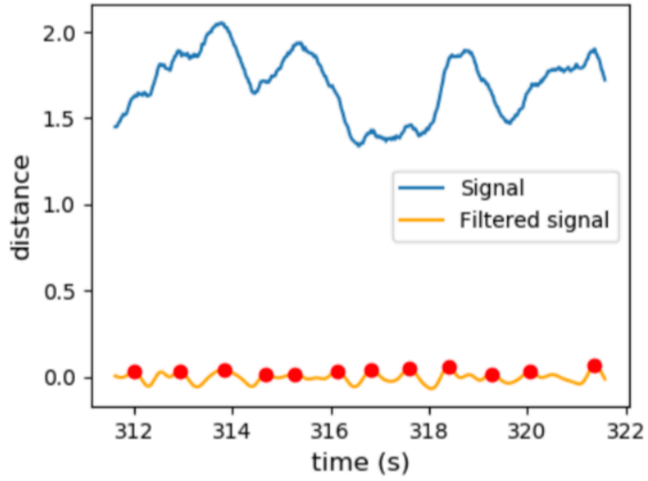


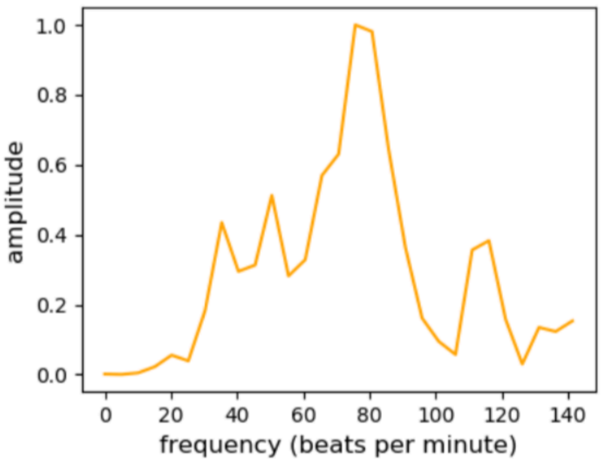


University of Ljubljana
Faculty of Computer and Information Science

From EM Waves to Physiological Signals

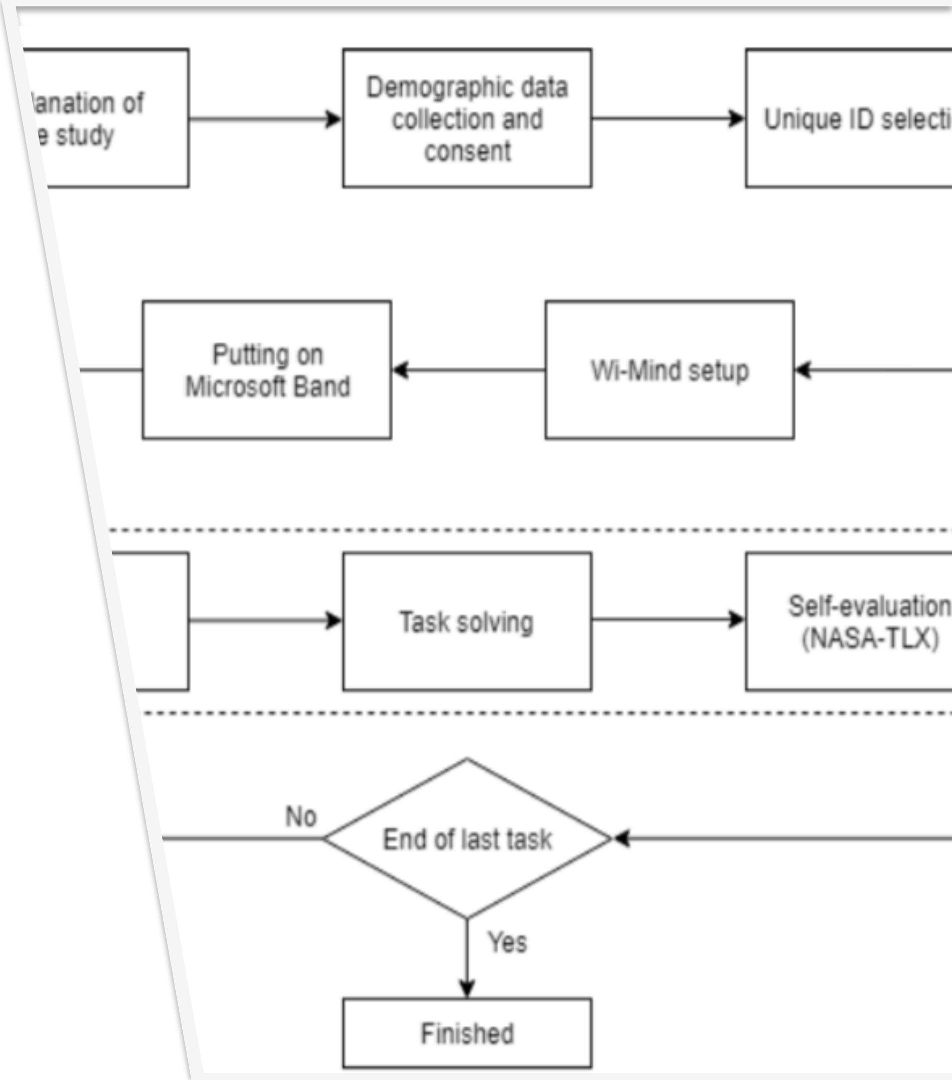
- Preprocessing:
  - Unwrapping phase
  - Filtering HF and LF noise
- Extracting heart beat signals:
  - Heart rate (FFT)
  - Heart rate variability HRV (peak detection + filtering) features: RR intervals, LF and HF HRV





WiMind Experiments

- Primary (PC-based) task
  - Adapted from Haapalainen et al.
  - NASA TLX after each task

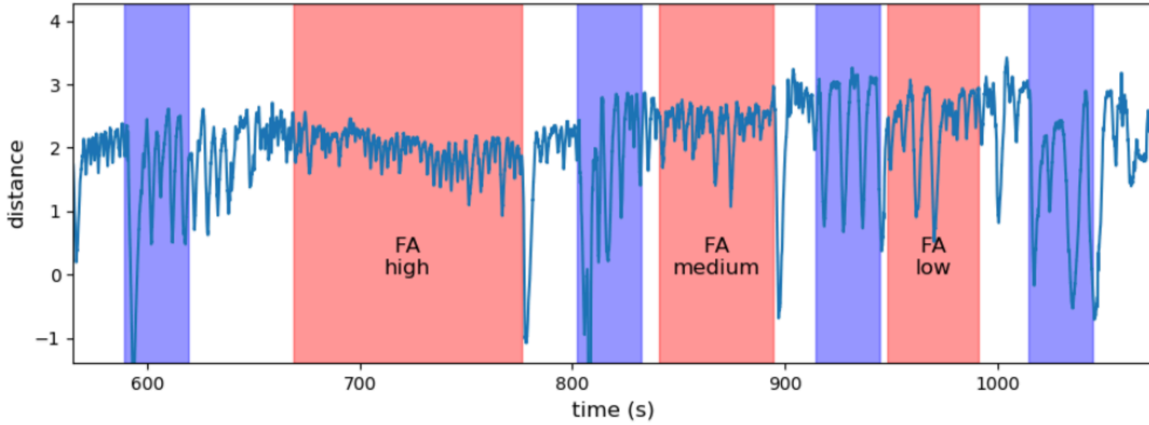


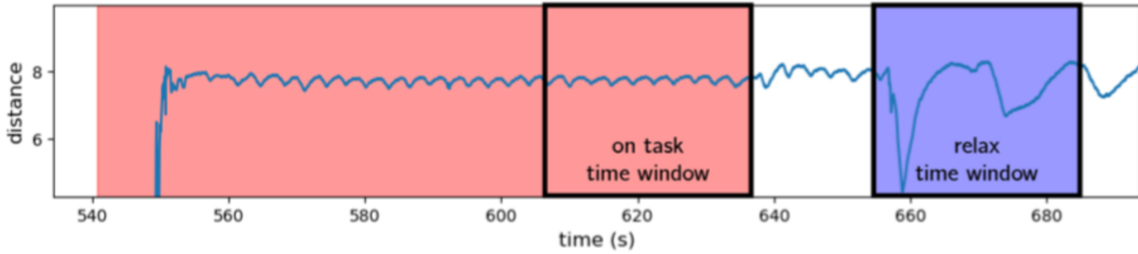
WiMind Experiments

- Primary (PC-based) task
  - Adapted from Haapalainen et al.
  - NASA TLX after each task
- WiMind wireless measurements
- MS Band + Android app
- Demographics
  - 23 users
  - 20-38 years old
  - 6 female, 17 male

Results

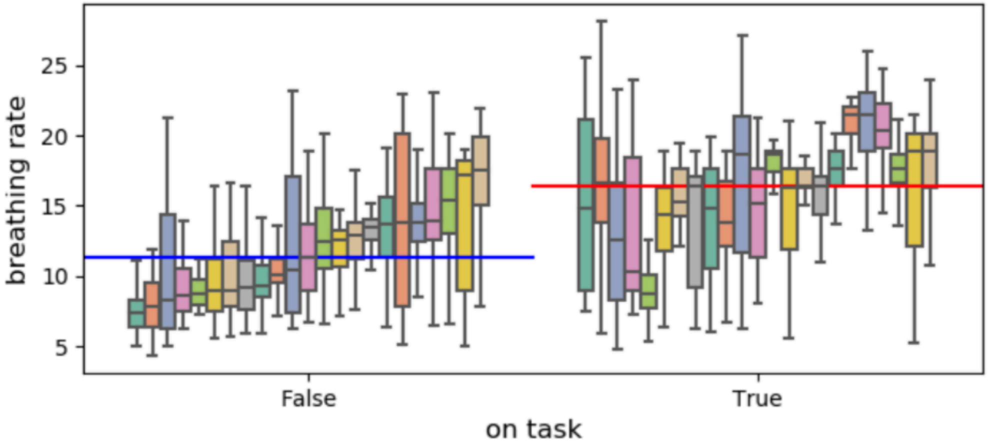
- Labelling signals via time windows:
  - Last 30 seconds of task engagement (label “busy”)
  - 30 seconds of explicit relaxation (label “relax”)



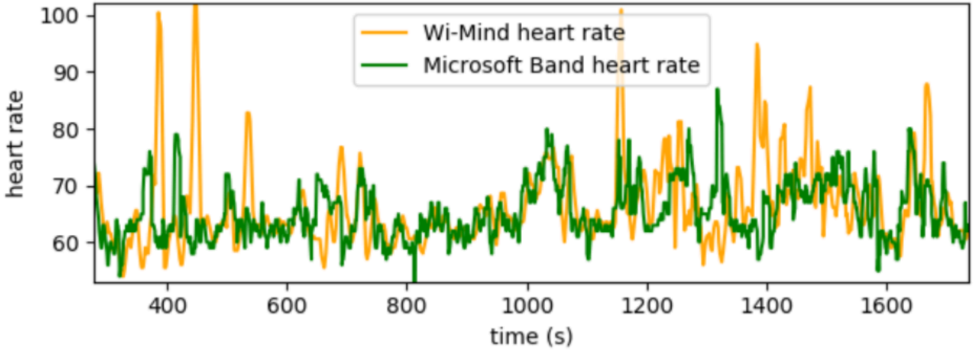


But first…

- Breathing rate validation



- Heart rate validation



University of Ljubljana
Faculty of Computer and Information Science

Inferring Task Engagement (Binary)

- Normalised breathing rate
- Different ML models from the “standard” toolbox
- Leave-one-person-out validation

Method	**AUC**	**Accuracy**
k-NN	0.752	0.704
SVM	0.670	0.580
Random forest	0.806	**0.746**
Naïve Bayes	0.780	0.723
Majority	0.5	0.5

Inferring Task Engagement (Binary)

- Normalised breathing rate
- Different ML models from the "standard" toolbox
- Leave-one-person-out validation
- Personalised models improve performance for some users but overall no improvement

Method	Accuracy
k-NN	0.604
SVM	0.721
Random forest	0.721
Naïve Bayes	**0.734**
Majority	0.5

Inferring Task Engagement (E/M/H)

- Unable to distinguish among different complexity levels
- Results are better if we consider only Easy and Hard tasks
- Linear regression for TLX gives similarly poor results

Method	Accuracy
k-NN	0.343
SVM	0.328
Random forest	**0.369**
Naïve Bayes	0.337
Majority	0.34

Neural Network Approach

- Long Short-Term Memory (LSTM) neural network
- Raw wireless phase signal
- Accuracy results:
  - Binary (busy/relaxed): **0.752** (vs 0.5 majority; 0.746 random forest)
  - No improvement with tertiary (E/M/H) or task-specific models

Towards (very accurate) unobtrusive cognitive load inference

University of Ljubljana
Faculty of Computer and Information Science

Summary

- (Relatively) successfully detect whether a person is engaged in a task or not even with WiMind
- Detecting the level of engagement is challenging even with direct sensing with off-the-shelf wearables
- Secondary task (the way we designed it) is not a reliable proxy for task complexity or TLX

Expanding Our Approach

- The role of personality traits
- Heterogeneous data sources:
  - Phone: accelerometer, calendar info, screen on/off
  - Wristband: HR(V), GSR, accelerometer, barometer, UV
  - Wireless: breathing, HR(V)
- Task types that elicit the strongest physiological response

Research Directions

- Which type of cognitive load can/should we detect:
  - Intrinsic
  - Extraneous
  - Germane
- Should we infer objective or subjective task difficulty?

Collaboration Opportunities

- Real-world applications
- Research projects
  - Horizon Europe, etc.
- Erasmus+ exchanges for
  - PhD students (up to three months)
  - Staff
- Join us as a doctoral student!

University *of Ljubljana*
Faculty *of Computer and Information Science*

Submit to our conference!

- Human-Computer Interaction in Information Society HCI-IS 2019 conference
  - http://hci.si/hci-is-2019/

Collaborators

- WiMind:
  - Tilen Matkovic, Uni. of Ljubljana
- Wearables:
  - Martin Gjoreski, Mitja Lustrek, Institut Jozef Stefan, Ljubljana
- TaskyApp:
  - Gasper Urh, Uni. of Ljubljana
- Mobile Interruptibility:
  - Mirco Musolesi, Abhinav Mehrotra, University College London
  - Christoph Anderson, University of Kassel

Thank You!

Credit: Shutterstock.com

Find out more:
"A Survey of Attention Management Systems in Ubiquitous Computing Environments" by Anderson et al., ACM IMWUT (Ubicomp) 2018

Dr Veljko Pejović
Faculty of Computer and Information Science
University of Ljubljana, Slovenia
Veljko.Pejovic@fri.uni-lj.si
@veljkoveljko

University of Ljubljana
Faculty of Computer and Information Science