OULU 2019
C 707

ACTA
UNIVERSITATIS OULUENSIS

Aku Visuri

WEAR-IT: IMPLICATIONS OF MOBILE & WEARABLE TECHNOLOGIES TO HUMAN ATTENTION AND INTERRUPTIBILITY

UNIVERSITY OF OULU GRADUATE SCHOOL;
UNIVERSITY OF OULU,
FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING;
UBIQUITOUS COMPUTING (UBICOMP)

ACTA UNIVERSITATIS OULUENSIS
C Technica 707

AKU VISURI

WEAR-IT: IMPLICATIONS OF MOBILE & WEARABLE TECHNOLOGIES TO HUMAN ATTENTION AND INTERRUPTIBILITY

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Information Technology and Electrical Engineering of the University of Oulu for public defence in the OP auditorium (L10), Linnanmaa, on 27 May 2019, at 12 noon

UNIVERSITY OF OULU, OULU 2019

Acta Univ. Oul. C 707, 2019

Supervised by
Doctor Denzil Ferreira

Reviewed by
Doctor Martin Pielot
Doctor Fahim Kawsar

Opponent
Professor Kasper Hornbaek

ISBN 978-952-62-2270-7 (Paperback)
ISBN 978-952-62-2271-4 (PDF)

ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)

Cover Design
Raimo Ahonen

JUVENES PRINT
TAMPERE 2019

Visuri, Aku, Wear-IT: Implications of mobile & wearable technologies to human attention and interruptibility.
University of Oulu Graduate School; University of Oulu, Faculty of Information Technology and Electrical Engineering; Ubiquitous Computing (UBICOMP)
Acta Univ. Oul. C 707, 2019
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

This thesis explores different ways of leveraging mobile sensing to understand how end users use and interact with their smart technologies, namely smartphones and smartwatches. These topics are extensively explored in other parallel research; however, numerous gaps still exist within the literature. The use of *mobile sensing* to collect quantified ground-truth information of device use *in-the-wild* is critical to collect unbiased experiences and usage traces.

This thesis covers three main themes: (a) the way our affect influences our smartphone use, and how our smartphone usage can also be analysed from our usage habits; (b) revealing quantified exploration of smartwatch usage traits, and how these relate to smartphone use, and (c) novel ways to mitigate interruptions during smartphone or smartwatch use. The thesis begins by explaining the related work and the overall theme of mobile sensing and how device usage influences attention; it then proceeds to elaborate on the contribution of each included article to the overall scope of the thesis. The thesis then concludes with a summary of how the presented articles tie together in a broader scope.

Considering the vast amount of research in this field by this thesis' author as well as other researchers, this type of work can potentially improve the use of novel wearable technologies in the future. By the end of the thesis, the reader should have a broad understanding of what mobile sensing is, and how it can be applied to comprehensively uncover technology use as well as leveraging mobile sensing to enhance the use of technology.

Keywords: analysis, attention, device usage, interruptibility, mobile sensing, smartphones, smartwatches, ubiquitous computing

Visuri, Aku, Wear-IT: Puettavien ja mobiiliteknologioiden vaikutus valppauteen ja häiriintymiseen.
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Tieto- ja sähkötekniikan tiedekunta; Jokapaikan tietotekniikka (UBICOMP)
Acta Univ. Oul. C 707, 2019
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä

Tässä väitöskirjassa tarkastellaan erilaisia tapoja hyödyntää mobiilikäytön tunnistamista ymmärtääkseen, miten loppukäyttäjät käyttävät ja ovat vuorovaikutuksessa älykkäiden teknologioidensa, esimerkiksi älypuhelimien ja älykellojen kanssa. Näitä aiheita tutkitaan laajasti muissa rinnakkaisissa tutkimuksissa, mutta kirjallisuudessa on vielä lukuisia aukkoja. Matkaviestinnän käytöstä kerätään kvantitatiiviset tiedot, jotka koskevat laitteen käyttöä luonnossa. Tämän tiedon kerääminen on kriittistä jotta voidaan kerätä puolueettomia kokemuksia ja käyttöjälkiä.

Tässä työssä käsitellään kolmea pääteemaa; i) miten älypuhelinkäyttöömme vaikuttaa meidän mielialamme ja miten älypuhelinkäyttöämme voidaan analysoida käyttötapojen perusteella, ii) paljastaa älykellon käyttöominaisuuksien määrälliset tutkimukset ja miten nämä tulokset heijastuvat älypuhelimen käyttöön ja iii) uusia tapoja lieventää katkoksia älypuhelimen tai älykellon käytön aikana. Työ aloittaa selittämällä siihen liittyvää työtä ja mobiilin tunnistamisen yleistä teemaa ja sitä, miten laitteen käyttö vaikuttaa huomiokykyyn, ja jatkuu sitten yksityiskohtaisesti jokaisen mukana tulevan artikkelin osuuden yleiseen käsittelyyn.

Työssä päädytään yhteenvetoon siitä, miten esitetyt artikkelit sitovat yhteen laajemman kokonaisuuden ja ottavat huomioon tämän alan tekijän ja muiden tutkijoiden tämän alan tutkimukset, ja miten tällaista työtä voitaisiin mahdollisesti parantaa edelleen tulevaisuudessa käyttämällä uusia tekniikoita. Työn päätyttyä lukijalla on laaja käsitys siitä, mitä mobiili-tunnistaminen on ja miten sitä voidaan soveltaa sekä teknologian käytön kattavaan paljastamiseen että mobiilidatan tunnistuksen hyödyntämiseen teknologian käytön tehostamiseksi.

Asiasanat: huomiokyky, häiritsevyys, jokapaikan tietotekniikka, laitteiden käyttödata, mobiiliaistiminen, älykellot, älypuhelimet

Wear-IT: Implications of Mobile and Wearable Technologies for Human Attention and Interruptibility

Acknowledgements

Foremost, I would like to express my sincerest gratitude to my advisor Dr. Denzil Ferreira for the continuous support of my Ph.D studies and related research. For his motivation, patience, knowledge, and encouragement during the process of writing this thesis and the related work. His guidance helped me evolve from in all areas, and I could not have imagined having a better suited advisor and mentor for my Ph.D studies.

Besides my advisor, I would like to thank my follow-up group, Adj. Prof. Susanna Pirttikangas, Adj. Prof. Heli Koskimäki, and Ph.D Teemu Leppänen for their encouragement, insightful comments, and tough questions during the process.

I thank my colleagues who I have worked with during the years, either in Oulu or remotely, wherever you may now be: Niels van Berkel, Simon Klakegg, Jorge Goncalves, Zhanna Sarsenbayeva, Reza Rawassizadeh, Yuuki Nishiyama, Chu Luo, Kennedy Asare, Valerii Kan, Elina Kuosmanen. I want to thank you all for stimulating discussions, for the time spent stressing over deadlines, and for all the fun we had outside of your typical work hours. I also want to thank my friends for helping me cope with life outside of work. Lastly, I am grateful for Prof. Vassilis Kostakos for offering me initial opportunity and introducing me to the possibility in working in research.

I would like to thank my family: my mother Pirjo Nuottimäki and sister Milla Visuri, as well as my mother- and father-in-law – albeit not officially – Sinikka and Mauri Raudaskoski. You have offered important life lessons, and just been there whenever we have needed anything.

Last but definitely not least, I would like to thank my significant other and life-long partner, Jaana Raudaskoski. Words cannot describe the support and encouragement you have offered throughout our shared lives.

8/4/2019 Aku Visuri

Abbreviations

API	Application programming interface
ESM	Experience sampling method
FFT	Fast Fourier transform
GDPR	The European Union general data protection regulation
GPS	Global positioning system
HCI	Human-computer interaction
MAC	Media access control address
MDA	Mean decrease accuracy
MDI	Mean decrease impurity
ML	Machine learning
OS	Operating system
QS	Quantified-self
RSSI	Received signal strength indicator
SMS	Short message service

Original publications

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:

I Visuri, A., Sarsenbayeva, Z., Goncalves, J., Karapanos, E., & Jones, S. (2016, September). Impact of mood changes on application selection. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (pp. 535-540). ACM.

II Visuri, A., van Berkel, N., Luo, C., Goncalves, J., Ferreira, D., & Kostakos, V. (2017, July). Challenges of quantified-self: encouraging self-reported data logging during recurrent smartphone usage. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference* (p. 81). BCS Learning & Development Ltd.

III Visuri, A., Sarsenbayeva, Z., van Berkel, N., Goncalves, J., Rawassizadeh, R., Kostakos, V., & Ferreira, D. (2017, May). Quantifying sources and types of smartwatch usage sessions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3569-3581). ACM.

IV Visuri, A., van Berkel, N., Goncalves, J., Rawassizadeh, R., Kostakos, V., & Ferreira, D. (2019) Understanding usage style transformation during long-term smartwatch use. Manuscript.

V Visuri, A., van Berkel, N., Luo, C., Goncalves, J., Ferreira, D., & Kostakos, V. (2017, September). Predicting interruptibility for manual data collection: a cluster-based user model. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (p. 12). ACM.

Contents

1 Introduction and motivation

The concept of ubiquitous computing dictates the proliferation of digital devices throughout our infrastructure, society, environment and within our daily life. This process started with the personal computer during the early 1990s and continued with the rise of the Internet into today's world, where we carry powerful computers with us at all times in the form of mobile devices such as smartphones and wearable devices. Many services such as booking flights and hotels, ordering food, ordering all sorts of products and navigation that would traditionally require a physical presence can now be accessed via the Web with the touch of a button. Many of these services have also reduced the need for human assistance to a minimum or none (e.g. for fully automated services).

While the world has undoubtedly become more convenient for anyone with a computer or a smartphone and a Wi-Fi access point, digital devices are sometimes overused. Constant connectivity creates symptoms such as anxiety and stress for users. For example, most of us at one point or another became agitated when we did not instantly receive a reply in our instant messaging communications, despite the iconic double tick indicating that the receiver has already read our message. The overuse of social media creates unrealistic expectations for our lives and, according to Microsoft Research, the human attention span has been reduced by the digitalization of our brain to become second to that of a goldfish. This thesis explores the effect of mobile devices on human attention and how we can mitigate potential unwanted interruptions during our daily lives.

The use of smartphones and mobile devices spans through the spectrum of our daily lives. This provides an exciting research agenda, as observing human behaviour through these devices can offer valuable insights into all aspects of life. People also observe their lives with the aid of these omnipresent devices, collect information about themselves and transfer mediums such as diaries into a digital, portable format. Social media outlets such as Instagram or Snapchat offer tools to generate (and share) constant flow of lifelogging material. Mobile devices additionally offer the capabilities to effortlessly collect self-reported information in the form of exercise trackers, mood trackers or health-related applications such as allergy symptom trackers or period trackers. When taken to an extreme, this type of information collection transforms into the Quantified-Self movement—constituting a collection of individuals who are eager to collect information about all aspects of their life to better themselves.

The emergence of wearable technologies such as Android Wear and iOS smartwatches, aimed to work in tandem with the wearer's smartphone, and health orientated devices (with fewer features), coined as Activity or Fitness Trackers, was to ease receiving smartphone information without the need to interact with the smartphone physically. There are also some issues that slow down the increase in popularity of wearable devices, such as battery constraints, high prices and reduced attention span caused by an overload of constant interruptions. While smartphone owners like their devices due to their capabilities to receive and transmit information, many still choose to disable specific functions simply because they find them overwhelming.

Many of the discussed aspects of smartphones (e.g. wearable devices and their use) are being researched continuously. As these are broad topics, this thesis approaches these areas in tandem with other works, aiming to locate and fill gaps within the literature and offering new ideas and researching unexplored aspects. This thesis aims to consult the presented topics by exploring three associated themes: (a) understanding smartphone use, (b) early quantitative studies of smartwatch use and adaptation and (c) human interruptibility and attentiveness as the overarching theme alongside mobile sensing. The themes of each original article are listed in Table 1. This thesis takes a data-driven approach to analysing usage traces from mobile devices to uncover both underlying reasons as well as potential solutions to disruptions to our attention.

Table 1. Individual themes within the overarching theme of mobile sensing explored in this thesis.

Theme	Investigated in	Contribution
Smartphone usage behaviour	Article I	How does our mood affect our smartphone application selections?
	Article II	Methods for improving the self-report frequency
Smartwatch usage behaviour	Article III	Quantifying frequency and type of smartwatch usage
	Article IV	How does our exploratory smartwatch usage evolve?
Human interruptibility and attentiveness	Article V	Methods for improving intelligent prediction mechanisms for preventing interruptibility

1.1 Articles

Five articles are included in this thesis, each published in relevant international conferences or peer-reviewed journals in the field of human-computer interaction and/or ubiquitous computing. The articles are presented thematically rather than chronologically, and each entry includes a main research question explored within the scope of the article.

The author of this thesis acts as the principal author in all five works; each article was written and conducted in collaboration with other researchers, many of whom have at some point worked at the University of Oulu or have contributed to the articles through data sets, collaboration, brainstorming and sharing ideas for data analysis. Mr. Niels Van Berkel and Dr. Jorge Goncalves offered substantial support for all five articles, who, at the time of writing this thesis, were based at the University of Melbourne but spent a significant amount of time cooperating with Oulu. All five articles were jointly supervised by Prof. Vassilis Kostakos and Dr. Denzil Ferreira who gave generous feedback by sharing ideas and assisting during the writing process. Lastly, Dr. Reza Rawassizadeh is the author of the data sets concerning smartwatches and offered his expertise for preparing these two publications. The following five articles are included in this thesis:

The first article studied the usage behaviour of smartphones. The article collated insights from two separate experiments on human behaviour, namely mood and its effect on the selection of smartphone applications. Both studies lasted for two weeks (14 days) and consisted of 15 and 21 participants. The objective was to explore how either happiness or activeness (or a combination of both according to the circumplex mood model) influenced application choices. Each participant logged their mood multiple times per day, and the subsequent application launches (or uses) were logged afterwards. The findings showcase several generalisable application choices being influenced by the participant's mood, similar in both studies. The use of application types such as games, media applications (e.g. music or video) and travel applications are associated with higher happiness, while social media are slightly more likely to be used when unhappy—a result highlighted in other studies as well. The article concludes by stating that even these types of simple associations could be used to predict or propose user's mood, based solely on their application choices—something that could be considered helpful in lifelogging tools.

Research question: *Does our current mood influence our application choices on mobile devices?*

The second article was also concerned with the topic of smartphone use. Previous research has studied the effectiveness of several different reminder systems for self-reporting applications, ranging from text messages to on-screen reminders via the unlock screen. The article proposes an on-screen dialogue that converts active smartphone usage time to data contributions for Quantified-Self applications, relying on self-reported information. The article highlights how selected applications could benefit from alert dialogues as reminder mechanisms to increase data precision and frequency and reduce forgetfulness in data logging. The article also explores how users are more likely to contribute data via such dialogues during different types of smartphone usage sessions (e.g. how users are more likely to have excess time for data contribution during more extended sessions). The use of the proposed reminder mechanism should be carefully considered; however, perceiving it as being highly distractive or interruptive and continually demanding user's attention can cause a significant burden during long-term use.

Research question: *Which mobile device usage-related metrics can be leveraged to facilitate more frequent self-reported data logging?*

The third article was an investigation on smartwatch use, using an extensive data set from 308 users over six months. This research, first of its kind, was a quantitative study on how smartwatches (Android Wear) were used in-the-wild. The data set was collected by the Insight 4 Wear application, which offers details to users about their smartwatch use (e.g. battery life). The findings showcased similarities in interactions, preferred application sources and response times with notifications between smartwatches and smartphones. The smartwatch data set was compared against the smartphone data set collected and combined from three previous smartphone studies. The smartwatch usage sessions were more frequent and briefer, which indicated that smartwatches mainly prefer consuming information rather than interacting with their devices.

Research question: *How do sensor-level observations of smartwatch usage reflect similar metrics on smartphones, and what insights do these metrics give about overall smartwatch use?*

The fourth article continued investigating smartwatch use and how it adapted over time. Using the same source for the data set in Article IV, although over an extended period, the article investigated how smartwatch users inherently opted for distinct smartwatch usage styles in terms of session duration and frequency, battery use considerations, notification sources, etc. The individual monthly usage behaviours were then analysed for consistently accepted behaviours—those that are returned to and used for long periods as opposed to those behaviours that are deemed exploratory, indicating shorter patterns and no return to such behaviours at a later time. Finally, we revealed differences in usage traces within the two behaviours, indicating usage behaviour that might lead to prolonged and more beneficial device use.

Research question: *How does smartwatch usage evolve in the long term, and what types of sensor-level usage metrics are ultimately preferred to those that are merely exploratory?*

The last article explored and expanded the theme of interruptibility. The article studied the use of intelligent methods (i.e. machine learning) for applications attempting to provide individual users' intelligent insights or predictions (e.g. would the user be interrupted? Would the user answer a call? Or, which types of news articles would the user prefer?). Traditionally, the so-called general models have been used for predictions, which mean that all users behave alike, and that one prediction model is sufficient to cover all behaviours and preferences. Alternatively, when the differences between users are considered too large, each user will ultimately receive a personalised decision-making model. Due to differences in each user's usage behaviour with mobile devices, the use of general models is often significantly sub-optimal. On the other hand, the training data required to generate personalised models often take time; thus, decision-making is not available from the beginning of device or application use. Article V presents and analyses the efficiency of a method where similarly behaving users are grouped, and a group model is generated. The findings showcase that the group model performs better than either the general or personalised model, considering cases where training duration can be considered limited. Over long periods, the personalised models should succeed in providing the most accurate predictions. The provided methodology could be implemented in any application which collects usage data and provides users with an intelligent decision-making system.

Research question: *Can the inclusion of a grouping algorithm in data separation and pre-processing increase the overall prediction accuracy in low-data scenarios?*

1.2 Thesis outline

This thesis is organised as follows: Chapter 2 reviews the related literature on the topics covered in this thesis. We start with the overarching theme in each of the articles—mobile sensing, the basics, how it is conducted and applied and the concept of context and contextual sampling. We then cover similar works in the field of interruptibility and attention and finally survey smartwatch-related literature. Chapter 3 examines the theme of mobile sensing and explains the collected data types, collection methods and applications for mobile sensing; it also introduces the data collection frameworks and applications used and what data are collected in the thesis' articles. Chapter 4 scrutinises Articles I–IV and their related theme—sensor-based details of smartphone and smartwatch usage habits. The chapter discusses in depth similar works and results and presents a summary of the answered hypotheses and contributions of each article. Chapter 5 continues to discuss the field of interruptibility and user attentiveness (investigated in Article V), which leverages concepts introduced in Chapter 4. Chapter 6 concludes the thesis and offers recommendations for future work arising from this thesis.

2 Related work

2.1 Context and mobile sensing

The concept of context relates to natural language, where context describes 'the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood'. In computer science, context has a similarly broad definition. The now-famous quote by Dey (2001, p. 2) summarises the use of context in computer science:

> "Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

Furthermore, in human behavioural studies within human-computer interaction, the interaction can be replaced by the state of the individual, thus leading to the context (the user state, state of applications and devices, usage information, etc.) and describing and annotating the person and his state (e.g. mood).

Context-awareness is a term that describes a system or an application that leverages the understanding of context and attempts to make decisions and modify its behaviour based on the enveloping context. First thoroughly explored by Schilit and Theimer (1994), some level of context-awareness has since become a standard in our everyday applications, especially on our mobile devices. One could not envision the use of an application such as Uber without access to detailed real-time location information or text prediction and auto-completing texts with no understanding of where or why the user is typing.

The different factors considered within the usage context of a mobile device are collected via a combination of sensors. Thus, the concept is coined as mobile sensing. The details offered by the sensors vary by sensor type and source, ranging from simple sensors such as monitoring the screen state of the device (on or off and whether the screen or interface is locked via a PIN code, for example) to reading the physical speed and acceleration of the device via the accelerometer (often used to count the user's steps taken), monitoring the magnetic fields around the device via the magnetometer, tracking the device's physical orientation or environmental measurements such as ambient temperature and light levels. These types of sensors and readings are generally classified as hardware sensors (i.e. the sensor readings originate from the device hardware). The works published as part of this thesis focus on the Android Operating System (OS); thus the sensor readings are limited to those offered by the Android OS.

The concept of mobile sensing can also incorporate software sensors—transforming raw data from hardware to less abstract form by developing software solutions and social sensors—potentially leveraging information about the user from multiple sources outside of the immediate physical device (e.g. calendar events or social media profiles and posts). Real-world applications frequently leverage this type of information over the raw data offered by the sensors. Knowing the device posture by the numbers offers limited information (e.g. angle measurement on the x-, y- and z-axis) as opposed to knowing whether the device screen is facing the user or not, or whether it is held in the typical device use posture (slightly tilted with the screen facing upwards). Similarly, measurements such as step counters rely on software solutions to understand the raw accelerometer information (acceleration with the effect of gravity included) as steps, transforming the data via fast Fourier transformation (FFT) to human-understandable value (i.e. steps taken) (Dirican & Aksoy, 2017). The usage of software sensors over hardware sensors also enables application developers to develop context-aware applications with proper intelligence and triggered events by defining situations (Dey, 2001)—a combination of different variables being in different states. The states can be more accurately defined with the use of software sensors over raw data.

Liu (2013) studied the use of mobile device sensors for collecting finer-grained information and hypothesised that in the future the vision sensor (camera) would be utilised more. Their work also encompasses the use of a smartphone as it is without the need for any external sensors or devices. The state recognition is a common approach in mobile sensing, and recently more efforts have been put into creating energy-efficient frameworks. One of the early efforts in state detection and state transition detection using methods that ensure normal levels of battery drain was performed by Wang et al. (2009). A common criticism for any system aimed for real-world deployment is whether the battery consumption is appropriately optimised.

Further advances in mobile sensing focused on different forms of activity recognition systems. Sun (Sun, Zhang, Li, Guo, Li, 2010) summarised the advances in activity recognition from the recent years, while Choudhury et al. (2008) discussed the use of human activity recognition systems for more general topics, such as fitness monitoring, eldercare support, long-term care and cognitive assistance. McClernon & Choudhury (2013) offered an assistance method focused more on everyday use by proposing the use of smartphone sensors to detect smoking habits to tackle unhealthy habits.

Understanding different forms of behaviour, both human behaviour and device usage-related behaviour, is often explored via mobile sensing. Works by Ferreira et al. (2014) investigating smartphone application micro-usage, how smartphone users revisit their applications by Jones, Ferreira, Hosio, Goncalves, and Kostakos (2015) and the analysis of the breaks between usage sessions and the continuity in task completion by van Berkel et al. (2016) all contribute to our understanding of smartphone use via sensed information. In terms of human behaviour, the perceived mood of a user can be associated with mobile sensing data, either by investigating and attempting to predict and model mood levels such as stress (LiKamWa, Liu, Lane & Zhong, 2013; Sano & Picard, 2013; Servia-Rodríguez et al., 2017) or boredom (Matic, Pielot & Oliver, 2015; Pielot, Dingler, Pedro & Oliver, 2015) or by directly investigating the association between sensed variables and user's emotional states (Mehrotra, Tsapeli, Hendley & Musolesi, 2017). Kushlev, Cardoso, and Pielot (2017) took a step further and investigated the influence of affect (how you are feeling) on smartphone interactions with different types of content.

When investigating the association between specific human (or device use) behaviour, the approach starts by selecting a set of variables (e.g. time spent using the device and in a different form or distinct application choices, among others) and the behaviour we wish to observe. The most straightforward association is the relationship between any two variables, as illustrated in the middle of Figure 1.

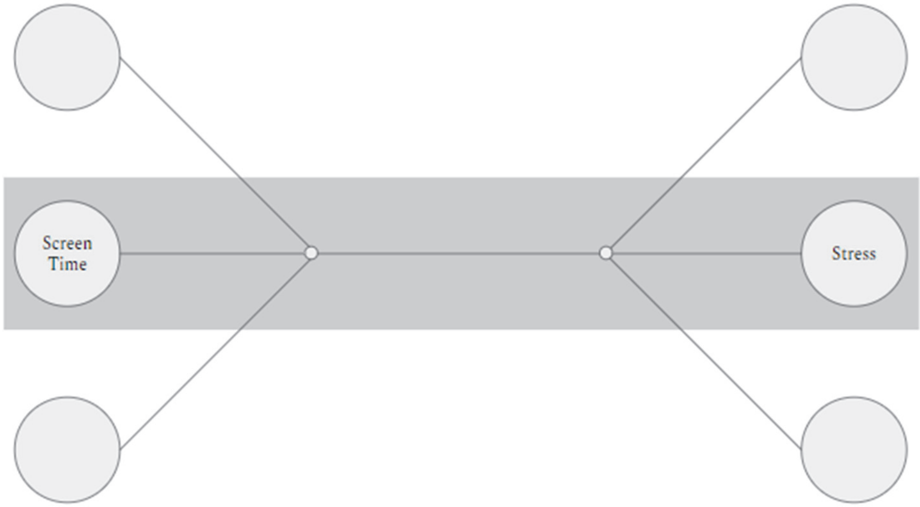

Fig. 1. Conceptual representation of measured states (right) and sensed variables (left) and the multivariate association between them.

The problem in this assessment, especially if attempting to make predictions in the said states (e.g. stress), is that both sides of the association are influenced by other unseen variables (see Figure 1). A mere change in screen time (details related to time spent using the device) can be an indication of stress; however, it can also be an indication of multiple other factors. Thus, a detailed prediction or assumption needs the analysis of multiple variables.

Figure 2 illustrates the multivariate association between different measurable variables and their potential causes (or outcomes). Measuring any behaviour is by no means a straightforward issue because of the sheer number of influencing independent variables, and thus careful consideration should be taken when (a) making decisions on the tracked variables and (b) interpreting the results. The contributions in this thesis aim to mitigate these potential problems by maximising the number of features collected, especially when collecting complex human sensed data (e.g. affect). These factors are considered in the design and experimental implementation of Articles I–V.

All in all, collecting sensor data via smartphones, often carried with us throughout the day and during work hours, leisure time, workouts and even when monitoring our sleep, offers invaluable insights that were previously impossible to collect accurately. Our everyday applications have thus seen the emergence of

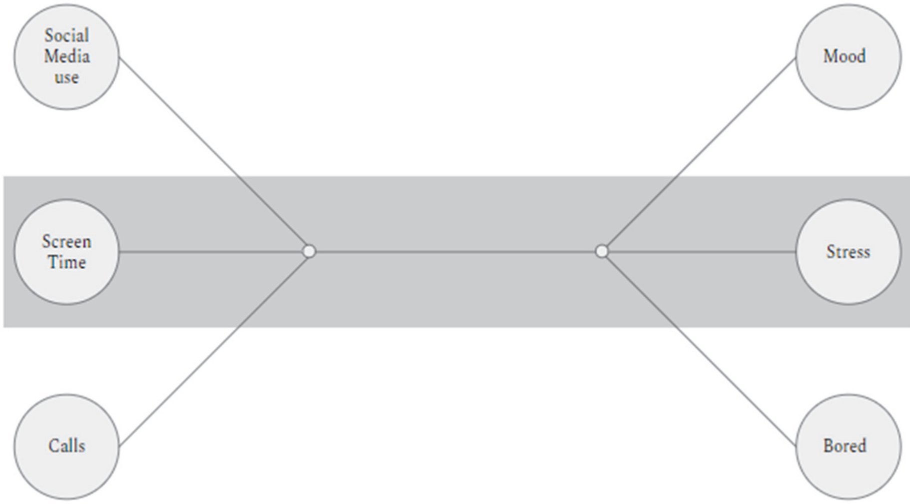

Fig. 2. The combinatory effect on both specific sensed variables (left) and measured states (right), and how factors outside the scope of measurements (grey area) can and usually influence each other.

context-awareness in recent years, as both the technologies and developers have advanced considerably.

2.2 Interruptibility studies

The human understanding of social and situational context offers humans great insights into understanding how available or attentive the other person is. Our assessment of the availability allows behaviour which can be considered natural or polite and thus does not interrupt the other person. Interruptions affect the user in various forms such as information overload (Okoshi, 2015), delayed task completion (Leiva, Böhmer, Gehring & Krüger, 2012) or reduced quality of logged information (Mehrotra, Musolesi, Hendley & Pejovic, 2015). Computer systems are often unaware of this context and thus end up interrupting a user or prompting at inappropriate times. The concept of bounded deferral—deferring an interruption to a later time with a specified upper limit—was first investigated by Horvitz, Apacible, and Subramani (2005). Later, Okoshi (2015) showcased two methods to address interruptions: either rescheduling interruptions to arrive at another time (e.g.

during next device use similar to the bounded deferral) or lowering the number or frequency of such interruptions or bundling messages together (already in use in instant messaging applications, for example). Similarly, Dingler and Pielot (2015) analysed user attentiveness to instant messaging using the bounded deferral approach. According to Iqbal and Bailey (2008), interruption timing can have a positive effect on both time management and task efficiency; thus it is an important research agenda in the modern-day world of increased (or even ubiquitous) connectivity.

Context-aware systems are used in the field of interruptibility to attempt to determine user's availability, readiness and interest in given content (Fogarty et al., 2005) according to their historical responses in different contexts (i.e. training data). The training data often include variables such as interaction choices (e.g. click or dismiss notification) or measurements of user's preference via short surveys or questionnaires, for example. The responses are then associated with the event's context. Simple external sensors, installed within the environment, can successfully predict a person's interruptibility (or availability, although the terminology distinguishes interruptibility from user availability). However, purely modelling this based on mobile sensors is more challenging. Intelligent algorithms (e.g. machine learning) have proved to be a successful method in predicting interruptibility. The combination of mobile phone sensor data (e.g. device posture) (Poppinga, Heuten & Boll, 2014), acceleration sensor (Ho & Intille, 2005), wearable devices with embedded sensors (Kern, Schiele, and Schmidt, 2007) and high-level data such as application usage patterns or location data (Mehrotra, Musolesi, et al., 2015; Pielot, 2015; Poppinga et al., 2014) have been shown to be successful. Many of these works inspired designing the contextual sampling used in Articles I, II and V.

Several different machine learning classifiers have been created over the years, and as the computational capabilities of computers (and hand-held devices) have increased, the complexity, scale and computation speed of the classifiers have also increased. Bayesian classifiers (e.g. Naive Bayes) are considered the simplest and often used as a baseline. More complex algorithms include Support Vector Machines (SVM) (e.g. LibSVM) and tree-based algorithms (e.g. the C4.5 algorithm or Random Forest), which often outperform other classifiers in situations where the training data are not complete. Incomplete training data are described as a training data set without access to all (or majority) of the permutations available for the algorithm. For example, in a four-week deployment with a combination of smartphone sensors, it is not feasible to go through all different permutations of collected variables (inputs) and user responses (outputs). Recently, boosted trees (Chen & Guestrin, 2016) and Random Forests (Mishina, Murata, Yamauchi, Yamashita, and Fujiyoshi, 2015) are being experimented with as more effective predictors, and it seems that they further increase the prediction accuracy and reduce memory use (Pielot et al., 2017).

2.3 Challenges of intelligent methods

One issue with intelligent predictor mechanisms, such as the aforementioned machine learning, is the understandability of the results. Many algorithms function as a black box system, implying that the inner workings and decision-making are effectively masked from the user. Thus, it can be difficult to understand the underlying reasons for the predictions and what ultimately influences the decisions. This can cause difficulties in communicating the results to the user in a trustworthy manner (as it can be very difficult to pinpoint a reason for a single prediction); it can also cause difficulties in communication in the research society (e.g. with colleagues, co-researchers or reviewers lacking sufficient knowledge to understand the limitations of a system or simply misunderstanding the details).

Many approaches leverage a multivariate prediction analysis, similar to the concept in Figure 2, where substantial data are leveraged as a decision-making feature set. Each feature acts as a binary or multivariate state (e.g. true or false, or whether the screen is on, off, unlocked or locked) or as a value (e.g. number or text string). The decisions are then based on the combination of these features, and a simple change in one feature could cause the result to differ vastly. For example, even if all other conditions remain the same, being at home makes a person more likely to answer incoming phone calls as opposed to when he/she is at work. The actual decision-making mechanism can be observed in some cases, such as tree algorithms (e.g. C4.5); however, with multiple features, the mechanism will very quickly bloat in size and become difficult for a human to grasp. The importance and effect of each feature are hard to communicate accurately, simply because the effect of each feature has an intertwined relationship with another feature or with a set of features. The typical means of assessing the impact of individual features are feature importance measurements, measured by either mean decrease impurity (MDI) or mean decrease accuracy (MDA)—how much worse does the prediction model perform when features are removed from the training set or split and gain-based measures—split measures imply how often the feature is used in a single decision process (for tree-based algorithms) and gains measures measuring the increase in performance when the feature is used as a 'cover' feature.

Using approaches such as MDI or MDA that generalise the feature importance into a ranking from 'most important' to 'least important' can easily lead to confusion. While a measurement such as battery charge level can be deemed as the essential feature, the actual relationship between battery level measurement and other features is still very complicated. In one of the many experiments on the association between smartphone sensors and user interruptibility, Poppinga et al. (2014) revealed the time of day and screen coverage of good predictors for opportune moments to present smartphone notifications. Pielot (2014) investigated mobile users' availability to answer incoming calls and showcased a variety of contextual data (e.g. physical activity, screen status, the day of the week) in addition to similar hardware sensors to those in Poppinga et al. (2014). Investigating more high-level variables (software sensors), users tend to have individually distinguishable traits of device usage in terms of session frequencies and durations (van Berkel et al., 2016), application selection (Welke, Andone, Blaszkiewicz & Markowetz, 2016; Zhao et al., 2016) and interactions (Falaki, Mahajan, Kandula, Lymberopoulos, Govindan & Estrin, 2010). These high-level features can reconstruct the user's personal preferences in detail. This past behavioural usage data was used by Pielot et al. (2014) to model user's attentiveness to instant messaging applications.

More detailed investigations on the associations between single or multiple features in end-user decision-making have also been carried out. The upside of using fewer features and more simplified analysis methods is the increase in communication and transparency. Most of Mehrotra's work has been on the area of notification management and understanding the details of user's notification preferences and habits. Highlighted works include (a) PrefMiner (Mehrotra et al., 2016), a non-black box system aimed at understanding and automating user's notification management settings based on a human-understandable intelligent proposal system; (b) MyTraces (Mehrotra, Tsapeli, Hendley & Musolesi, 2017), an investigation on the correlation and causality between mental states and smartphone interactions; and (c) My phone and me (Mehrotra, Pejovic, Vermeulen, Hendley & Musolesi, 2016), which investigates attentiveness and response times to notifications. These works take a more simplified approach in automatic management but do it with increased understandability and transparency to the end user. Other approaches to mitigating the disruptive nature of notifications, for example, have also been explored. Leiva et al. (2012) attempted to prepare the user for disruption before the actual interruption arrives and ways to guide the user back to the on-going task (e.g. by replaying the prior interactions with the interface).

Article V offers a novel approach to mitigating the potential problems that arise when relying on either *general* or *personalised* models. Article V introduces a method for recommender and prediction systems to rely on group-based modelling that reduces the initial downtime of a prediction system or the inaccuracy of relying on generalised approach in situations where individual differences can highly influence accuracy.

2.4 Smartwatch studies

Smartwatches have re-emerged in the past years as a viable and usable product, sitting between wearable products such as activity trackers with somewhat limited capabilities and our smartphones. This re-emergence has caused the research community to investigate this new trend in multiple areas, from a marketing and consumer adoption standpoint to topics closer to those explored within this thesis: *sensor-based information collection on user habits and usage behaviours*. The current iteration of smartwatches allows users to synchronise information to their smartphones, interact with their applications and receive information in the form of instant messages or emails and for navigation or weather purposes, for example. Articles III and IV fill the gap in smartwatch literature by introducing a quantitative analysis of smartwatch use that leverages mobile sensing. The primary motivation for including these works within the scope of this thesis is (a) close the connection between Android smartwatches and smartphones, as the devices are meant to be used in tandem and (b) observing the similarities and differences in the usage of these two device types.

Part of the research on smartwatches and their use is focused on understanding consumer acceptance of smartwatches. While the technology has been available for several years in its latest form, it is yet to attain vast popularity. Wu, Wu, and Chang (2016) surveyed 212 smartwatch users and noted that users are accustomed to the form factor of the smartwatch (i.e. similarity to the wristwatch), although they require result demonstrability. The benefits of owning and using a smartwatch should be tangible to the user, both as an extension to the smartphone and also as a standalone device. In another work, Kracheel, Bronzi, and Kazemi (2014) highlighted the dual nature of the smartwatch as an intelligent, connected device but also as a traditional wristwatch. Kracheel et al. (2014) also noted that the lack of widespread adoption of smartwatches is likely due to the lack of a ‘killer app’, which would sufficiently distinguish the smartwatch from the smartphone or mere activity trackers. Alternatively, technological limitations can also withdraw potential users from obtaining a smartwatch, with battery limitations often cited as the most significant issue (Rawassizadeh, Price & Petre, 2015).

Min et al. (2015) investigated the differences between inexperienced (with less than three months of experience) and seasoned smartwatch users and revealed that if only one feature was to be used, the more experienced users prioritised time-keeping capabilities over notifications, which were favoured by the inexperienced smartwatch owners. Overall, however, 98% of the respondents perceived notifications as the main functionality of a smartwatch. Similar results are highlighted in Maier and Wörndl (2015). The main reason is the ability to covertly access information (Cecchinato, Cox & Bird, 2015) without the social awkwardness and the negative associations of constantly peeking at your smartphone (Palen, Salzman & Youngs, 2000). The increased connectivity and availability are seen as a benefit by the end users, as unavailability is a common problem in mobile communication (Salovaara, Lindqvist, Hasu & Häkkilä, 2011).

The ubiquitous nature of smartwatches also contains other benefits. Wearable devices have emerged as an essential tool in monitoring health issues (Mortazavi, Pourhomayoun, Alsheikh, Alshurafa, Lee & Sarrafzadeh, 2014) due to their ability to monitor activity. Smartwatch can also collect biological, environmental or behavioural information and is likely to be carried by the user throughout the day. The location of the smartwatch also allows more precise activity tracking, such as tracking meals (Dong, Scisco, Wilson, Muth & Hoover, 2014), smoking (Scholl & Laerhoven, 2012), having a cup of coffee or giving talks (Shoaib, Bosch, Scholten, Havinga & Incel, 2015). Research findings suggest that accurate logging of bad habits can help mitigate drinking too much coffee, smoking too much or skipping meals, for example.

The research on smartwatch usage is still emerging. This justifies the inclusion of the works in this thesis, as the data collected from smartwatch users in-the-wild are still scarce, and the owners of smartwatches are still adapting the way they use their devices in everyday life.

3 Mobile sensing

3.1 Mobile sensing frameworks and data sets

The Android OS offers access to several hardware sensors. While the data are not raw sensor data (i.e. values such as current read directly from the on-board chip; the information provided by the OS is often referred to as raw data). The Android OS on-board sensor chips are categorised into three different main categories; motion sensors, position sensors and environmental sensors. The sensors are further divided into hardware sensors (raw data) and software sensors (processed data), although in a sense the hardware sensors are also already processed from the data provided by the chip. Table 2 showcases the different smartphone sensors and the raw data provided by those sensors.

Several other sensors offer information directly about the state of the smartphone. Although not listed in the official documentation, the sensors still provide similar information from on-board chips or other parts of the hardware. For reference, a full overview of these sensors is listed in Table 3. As both of the tables contain considerable information, the sensors used within the articles in this thesis appear in **bold** type.

Table 2. Android hardware sensors. Sensors used in the articles of this thesis are in bold type.

Hardware Sensor	Description	Common Uses
Accelerometer	Measures the acceleration force in m/s^2 that is applied to a device on all three physical axes (x, y and z), including the force of gravity.	Motion detection (shake, tilt, etc.).
Ambient temperature	Measures the ambient room temperature in degrees Celsius (°C).	Monitoring air temperatures.
Gravity	Measures the force of gravity in m/s^2 that is applied to a device on all three physical axes (x, y, z).	Motion detection (shake, tilt, etc.).
Gyroscope	Measures a device's rate of rotation in rad/s around each of the three physical axes (x, y and z).	Rotation detection (spin, turn, etc.).
Light	Measures the ambient light level (illumination) in lx.	Controlling screen brightness.
Linear Acceleration	Measures the acceleration force in m/s^2 that is applied to a device on all three physical axes (x, y and z), excluding the force of gravity.	Monitoring acceleration along a single axis.
Magnetic field	Measures the ambient geomagnetic field for all three physical axes (x, y, z) in μT.	Creating a compass.
Orientation	Measures degrees of rotation that a device makes around all three physical axes (x, y, z).	Determining device position.
Pressure	Measures the ambient air pressure in hPa or mbar.	Monitoring air pressure changes.
Proximity	Measures the proximity of an object in cm relative to the view screen of a device. This sensor is typically used to determine whether a handset is being held up to a person's ear.	Phone position during a call.
Relative humidity	Measures the relative ambient humidity in percent (%).	Monitoring dewpoint, absolute, and relative humidity.
Rotation vector	Measures the orientation of a device by providing the three elements of the device's rotation vector.	Motion detection and rotation detection.

Table 3. Android software (state) sensors. Sensors used in the articles of this thesis are in bold type.

State Sensor	Description
Screen	Keeps track of the screen state of the smartphone, using four different states: ON, OFF, locked, and unlocked.
Applications	Whenever the foreground application changes, the application’s package name (unique identifier created by the app developer (e.g. com.facebook.katana for the Facebook main application) is logged.
Bluetooth	List of nearby Bluetooth devices, their signal strength (RSSI) and MAC address (unique hardware identifier for each device).
Network	The state of different network devices; mobile network ON or OFF, Wi-Fi activated or deactivated, airplane mode (ON/OFF), Bluetooth (ON/OFF), GPS activated or deactivated, Internet access available or unavailable and the amount of network traffic in both bytes and packets.
Battery	Battery level (%) and the charging state, voltage, ampere and battery temperature.
Calls	Incoming and outgoing calls, their result (e.g. answered or unanswered) and either source (if incoming) or target (if outgoing). The duration of a call can also be calculated from time between start and end of a call event.
SMS	Sent and received SMS messages and the corresponding sender or target.
Location	The latitude and longitude of the smartphones location, the current bearing (direction), movement speed, altitude and the location provider (GPS or via network) and estimated location accuracy.

The information provided by the hardware chips, monitoring both the environment and the user’s actions, offers limited insights into its raw form. Further processing is required for it to become a meaningful measurement of anything besides very simplified information.

3.1.1 Sensor data processing

In this example, we start with the raw data from the screen sensor that broadcasts the state changes (the screen turns on or off and becomes locked or unlocked) in the smartphone screen, and when (timestamp) each change occurred. The change from a switched off screen to a switched on one can happen automatically (e.g. due to an arriving message or a call); however, unlocking the device requires user interaction. From this raw data, we can explore measurements such as how often the user interacted with their smartphone over a day and at what times of the day.

For furthering this process, the information can be analysed by combining multiple samples and measuring time differences (e.g. to assess the length of an individual usage session): the time between the screen becoming unlocked—requiring specific user interaction to do so—and the time when the screen turns (or is turned) off again. Inactivity can also cause the device to become locked at a certain (usually on the minute) interval, which can be identified, and thus usage sessions consisting of some level of inactivity can also be identified. The opposite of a usage session can also be identified; this was coined as *usage gaps* by van Berkel et al. (2016) in their thorough study of the existence and influence on smartphone use from the information between times when devices become unlocked.

Furthermore, different sensor data can also be combined (as discussed previously) to either gain a more detailed understanding of an event, or observe the influence of one event factor on another. In Article III, the association between incoming notifications and usage sessions is analysed by measuring the longitudinal nature of both variables. The analysis evaluates whether the usage session originated from an arriving notification (and vice versa) or whether an incoming notification was observed by the user (indicated by the following usage session) swiftly. This approach was enabled by both variables sharing a timecode which were then cross-referenced and matched accordingly within the specified time window (60 seconds). Figure 3 is from Article III and showcases how notifications are categorised based on how swiftly they were observed. In addition, each notification and the following usage session were used to categorise a usage session into user or notification-initiated usage session. Lastly, the usage sessions were also categorised into peek or interaction sessions, based on whether the session contained any user interaction (interaction session) or not (peek session).

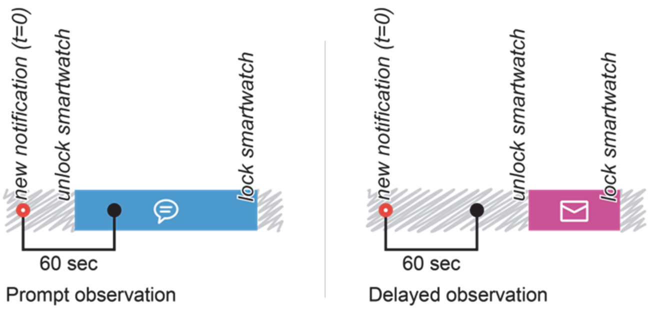

Fig. 3. Our methodology from Articles III & IV for classifying notifications based on their observation delays (Reprinted by permission from Article III ©2017 Authors).

Other commonly used sensors for determining device usage styles involve the applications sensor; it indicates which applications have been actively used (appears as the *foreground* activity on the device) and when. For example, the battery indicates different types of periods of device's use based on battery depletion rate (or specific time or *context* for charging the battery). Article I explores user's mood that affects the application choices, and specifically in the order of choices being influenced by the current mood and not vice versa. In Article I, the mood was self-reported by the user after unlocking his device, and the timestamps were then cross-referenced with the applications that launch within the following usage session, essentially combining three different sets of data (applications, screen, and mood) based on a similar timeframe in which the events occur.

The battery sensor contains simple information of the battery level and charging state, logging a database entry every time the battery level or the charging state changes. This limited and simple information can be processed further into measurements for battery drain rate over time, charging rate over time, and overall battery use, indicating heavier and more frequent device usage. In their work, Min et al. (2015) explored the current practices for smartwatch battery use and management. A more human-centric investigation combining the sensor with human interaction is also feasible, as demonstrated by Hosio et al. (2016) in their insightful work on understanding the (monetary) value of a smartphone battery charge level.

The approaches to data analysis above from just the screen sensor are used throughout the works that look into device usage and are a person's typical go-to methods for inferring usage sessions and general device use. More in-depth analysis has also been performed, which is often more common in the form of a post hoc analysis. The approach used by Meyer, Wasmann, Heuten, El Ali, and Boll (2017) inspired one of the analysed features in Article IV. Their work analysed long-term usage patterns in activity trackers, and how users adapt their use over long periods, ranging from intense and frequent usage to periods of inactivity, but with the users periodically returning to their devices after inactive periods.

Meyer et al.'s (2017) work analysed the data from an activity tracker, a carried or worn device that uses the accelerometer sensor to track *steps*. A typical step goal for a healthy active lifestyle is 10,000 steps per day. Each day that the device logs more than 500 steps is considered a *use day*, indicating that the device was used during that day. The information is then processed further by measuring the density of use within a period or in total (ranging from 0% to 100%), uninterrupted series of *use* or *non-use days* ('streaks' and 'breaks') and use days per week (from 0 to 7). The authors then identified and labelled different use patterns (e.g. 'power use', 'slow start', etc.) based on the measurements.

3.2 Data collection: frameworks and conducted studies

Collecting sensor data from mobile devices can pursue two goals. One approach is enhancing the user experience of application use by adapting the application directly through sensed information. Alternatively, the possible adaptations can be first investigated remotely by *collecting* the sensed information and performing the analysis before making any direct adaptations to the application. Broadly speaking, the approaches can be considered as more business orientated (first approach) or research orientated (second approach). However, a post hoc analysis is required to conduct the first approach effectively or investigate the adaptations' effectiveness.

With the increased connectivity of smartphones, the research community has benefited greatly from the ease of directly collecting data from end users' devices through cloud services and servers. Data collection methods for mobile devices usually leverage this connectivity to share data from the end devices with the centralised servers. Although each platform, framework, application and study conduct the data collection distinctively, the basic principle applies to most. Most of the applications we use daily communicate with a server to fetch, share or store information; so end users are less likely to see such communications as unwanted or dangerous when they are used to collect information, although the emergence of personal data protection laws, such as the GDPR (the European Union General Data Protection Regulation), has caused the general population to exercise more caution in their data-sharing habits.

In the mobile environment, data are collected to capture the *context* in the domain of mobile computing. Drawing on Dey (2001) and Abowd et al. (1999), Ferreira (2013) extends the categories of context for mobile instrumentation, offering the following categorisations:

- *Who*: the unique identity of the entities (e.g. sensor or application).
- *What*: the characteristics of the entities that can be labelled, measured or inferred (e.g. a label for a geo-location coordinate, currently engaged physical activity for a person, etc.).
- *When*: the instance of time in which the event is occurring.
- *Where*: the location (e.g. place, application, sensor) of the event.
- *Why*: the intelligibility of the system or application, the user intent and accountability of the system, application and user.

Ferreira (2013) adds that the category of *why* is the most difficult to capture, as it implies personal and informed choice based on the information available to humans but not necessarily to machines.

Two different data collection frameworks were used in the thesis' articles. The smartphone-based data were collected using the AWARE framework (Ferreira, Kostakos & Dey, 2015). AWARE allows application developers to include sensor data within their applications and create studies for their users which automatically offload the data from their users to a centralised server. Similarly, researchers can set up studies to collect specific hardware or software sensor data from their study participants. AWARE has been used for numerous research projects, studies and publications throughout the world since its emergence in 2013 and can be considered the most popular mobile sensing framework currently in use. Articles I–II and V leverage the AWARE framework and collect sensor data during device use; each distinct smartphone then uploads the data to the centralised server for further access and analysis.

Article I is based on two distinct studies. The first study was conducted by Sarsenbayeva et al. (2017). It investigated the feasibility of using a Vision API to infer user's happiness from self-portraits. The used application is named MoodTracker, and the application details are described in the cited article. The self-reported happiness value was collected when the user unlocked his device once during three different periods of the day: once in the morning (8 a.m.–12 p.m.), once in the afternoon (12 p.m.–18 p.m.), and again in the evening (18 p.m.–24 p.m.). In case the participant did not unlock his/her device at a designated time slot, we would discard the data logging request. The study lasted for two weeks and consisted of 15 participants (aged 21–30, M = 26.47, SD = 2.13). The second study used in Article I was unpublished at the time of writing of the article. It consisted of 21 participants who for two weeks reported their happiness and activity level when prompted after unlocking their devices. Both studies passively collected application usage data continuously during the two-week study period.

Articles II and V report data collected initially in the author's master's thesis work (Visuri, 2016)—the LifeTracker application. The conducted study was based on a self-monitoring application that enabled users to track their health, exercise or mood daily. Forty-eight participants participated in a four-week long study, during which they were instructed to use the application daily. The participants also interacted with the application prompts by accepting or dismissing each prompt—behaviour that is investigated in detail in Article V. The application was built on top of the AWARE framework, enabling the base data collection from the participants' smartphones to function in the background while the application itself offered expansive interfaces and self-reporting options to the users. Figure 4 showcases the self-reporting interfaces of the LifeTracker application.

The smartwatch data used in Article III and Article IV were collected via the Insight 4 Wear application, originally introduced in (Rawassizadeh, Tomitsch, et al., 2015). The application was designed for real-life use, offering the user information about his battery health and lifetime as well as other information related to his/her device use. The data were then shared with the research community. It is one of the largest wearable data sets available, especially considering its longitudinal characteristic—some users have tracked their data since its introduction in 2015. The full data set consists of a fixed set of physical activities, heart rate (if available on the wearable device), screen events, battery status, notifications, Bluetooth connections and ambient light. The articles in this thesis leverage battery, screen and notification data sets to understand the device usage characteristics from smartwatch users.

Fig. 4. Self-report interfaces (within an application and from an alert dialogue) from the LifeTracker application.

While the AWARE framework is modular, based on the needs of an application developer or a researcher, monitoring and collecting data from different sensors can be turned on and off; Insight 4 Wear is a standalone application with a fixed set of sensors. The benefits of being standalone include better optimisation and energy efficiency in terms of battery drain and lifetime, which are common issues with modern smartwatches. The application is also reasonably comprehensive to collect high-quality and large data sets from wearable devices. The full data set in Articles III–IV is 10 gigabytes in size and contains information from 681 unique devices.

3.3 Mobile sensing applied to included articles

As previously stated, all articles within this thesis involve mobile sensing. In all cases, the methodology is passive sensing (collecting raw data from hardware sensors), interaction or self-reported variables, combined with a post hoc analysis of the collected usage data. The following table and chapter summarise how mobile sensing is applied in each article for data collection and how the collected data are then used for analysis.

Article I contains simple raw sensor values, and the experiment collects self-reported mood information. The mood information is collected via an interface presented during a device usage session that inquires the user's current mood using the circumplex mood model (Russell, 1980)—using two axes of valence (pleasure) and activation (activity). Sensors then passively collect application usage and screen data. The analysis itself looks for correlations between specific application categories ('Internet and Social', 'Games', 'Maps and Travel', etc.) and the reported affect, which is also categorised into eight separate categories according to the combination of valence and activation ('Stressed', 'Happy', 'Bored', etc.).

Article II continues the theme of manual data collection in the smartphone environment and investigates ways to encourage the number of daily data contributions with a more qualitative approach—using survey responses in addition to mobile sensed data to explore the use of potentially interruptive interfaces in collecting data contributions. It uses the categorisation of screen data (usage sessions) to new and continuing sessions (see van Berkel et al., 2016) and explores the timing of an interruption (interface) in the data collection efficiency. Contrary to the four-tier interaction classification from Article I, only the data contribution is considered here. The articles conclude with a statistical analysis of what types of interface timings and during which types of usage sessions there is a higher likelihood of data contributions.

Table 4. Sensors used in presented works and a summary of the analysis methods. Non-native and non-hardware sensed data are included in italics.

	Data	Analysis
Article I Smartphone Usage Behaviour	Screen Applications Self-reported Affect	Correlation between user reported affect using the circumplex mood model (Russell, 1980) and subsequent application usage.
Article II Smartphone Usage Behaviour	Screen Alert Dialogue Interactions Manual Data Input Frequency Survey Data	User's willingness for manual data contributions via alert dialogues in different forms of usage session types and for different input modalities.
Article III Smartwatch Usage Behaviour	Screen (Smartwatch and Smartphone) Notifications	General investigation of differences between usage sessions on smartwatches and smartphones in session frequency, type and duration.
Article IV Smartwatch Usage Behaviour	Screen Battery Notifications	A combination of usage characteristics derived from sensor data is analysed longitudinally for consistent long-term changes in the participants' device usage styles.
Article V Human Interruptibility and Attentiveness	Screen Battery Calls Network Proximity Physical Activity Alert Dialogue Interactions Manual Data Input Frequency Survey Data	User's receptiveness to interruptions is modelled with machine learning using a vector feature consisting of sensor data (see text for details) which is labelled according to the logged interactions and manual data inputs.

Article III continues to discuss the smartwatch device space and uses a collection of screen and notification data to explore different types of smartwatch usage sessions, which are then compared to a smartphone screen data set collected from earlier published works from the same authors. Much of the data in Article III are focused on usage sessions, and how the usage sessions differ in the used smartwatch and smartphone data sets. The notification data set timestamped is then cross-referenced with the usage sessions, according to the previous illustration in Figure 3. The usage sessions are categorised into *peek* and *interaction* sessions, according to a five-second threshold for the screen to turn back off unless there occurs a specific user interaction (touch, voice command, etc.) during use. The notification data set is further statistically analysed according to source application and its category. Figure 5 is an excerpt from the article and showcases the major differences in smartwatch and smartphone use according to the frequency and duration of usage sessions; showcasing smartphone sessions is usually significantly longer and less frequent.

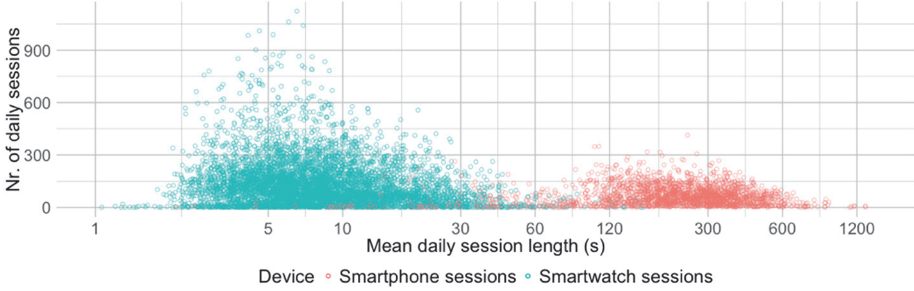

Fig. 5. Differences between usage sessions on smartwatches and smartphones from Article III (Reprinted by permission from Article III © 2017 Authors).

Article IV uses the same smartwatch data set, albeit with extended duration, and incorporates the battery and charging tendencies on top of the screen (usage) and notification-related data. The goal of Article IV is to investigate how smartwatch users adapt their use over time. The paper takes a more complex approach to investigate usage traits, create summaries of user's daily activities and analyse how the activities' focus shifts over time. The paper incorporates further theorems from other researchers, namely the concept of *short, isolated* and *reward-based* device use (Oulasvirta, Rattenbury, Ma & Raita, 2012), and whether battery levels have reached what Hosio et al. (2016) considered low or *critical* levels. The analysis considers daily summaries in terms of (a) ratios (e.g. how many in % of the received notifications were quickly addressed or how many of the sessions were swift glances (*peek* sessions), (b) total number of events (e.g. number of usage sessions) and (c) means and medians (e.g. average duration of a usage session during the day or median time between sessions). The daily summaries are aggregated to 30-day summaries, which average each feature (coined as *usage session characteristics)* within the 30-day time window. Furthermore, Meyer et al.'s (2017) work inspires investigation of overall use (or lack thereof) during these 30-day windows, and the *streak*, *break*, and *phase* measurements are calculated for the 30-day periods. Overall, Article IV presents detailed and complex measurements, going beyond just reading raw sensor values.

As noted in Article V, probably the most complex work in terms of mobile sensing, a machine learning algorithm is applied to the sensed data, which attempts to predict the user's attentiveness to presented interruptions, in this case whether the user is willing to contribute self-reported data and what their response type (*accept* or *reject*) is to an incoming interruption. The selected machine learning algorithm, Random Forest (Breiman, 2015), uses a vector of samples for training and testing its accuracy. The vector is generated from the raw sensor data and used as a snapshot of a moment in time—the moment when the user performs a manual interaction indicative of his interaction choices. The interaction is used as a *label* for the data vector, which indicates both the result in the training data and the predicted interaction choice for the testing data, which otherwise is missing the label. In Article V, the interaction choice is defined as a combination of two binary options: whether the user contributed data via the presented interface or not, and whether the user accepted or rejected the interface. The resulting combination offers four distinct interaction choices (labelled as A, B, C and D) that function as the label for the data vector.

The data vector uses a multitude of sensors to collect the user's *context*. Table 4 describes each presented work and the sensors used. Article V uses a combination of six hardware sensors, their raw data, processed sensor data and interactions to create a simplified idea of the user's current situation. For those interested in details or looking to replicate similar methods, the used sensors in no particular order of importance are as follows:

1. Whether data are contributed in the current usage session.
2. Delay (in seconds) of the input interface from the beginning of the usage session.
3. The number of usage sessions within the last five minutes.
4. Duration of the current usage session.
5. The type of the current usage sessions - whether it is a new or continuing usage session (van Berkel et al., 2016).
6. Time since the last interaction with the data collection application.

7. Time since the last data contribution.
8. Availability of the Wi-Fi connection.
9. Availability of Internet connectivity.
10. Cellular network connection type.
11. Current battery level (in %).
12. The current battery charging state.
13. Whether the smartphone screen is covered or not.
14. The current physical activity of the user.
15. The current hour of the day.
16. The current day of the week.

Points 8–13 are raw sensor data, 14 is a software sensor provided by Google (the Google Activity Recognition API) and 15–16 are from the data timestamps. Points 1, 6 and 7 are from the application usage logs. The usage session points (2–5) use screen events, namely screen becoming unlocked and then subsequently turned off, identifying periods of *device usage* and then further process the information either by comparing data point timestamps (4), comparing with previously processed usage sessions (3 and 5) or comparing with application usage logs (2).

4 Understanding technology and its effects on the user

Understanding the dual nature of both technology use of individuals and how using technology influences individuals are a core aspect of research in HCI. In addition, it is commonly investigated in which ways our technology use is influenced by both external and internal context. A plethora of different methods can be used to assess the effects, including qualitative methods such as surveys and self-reports, interviews and collecting user experiences as well as collecting quantified data using mobile sensing. Often several methods can be used in combination, and the different methods verify results from each other. Two of the articles (Articles I and II) in this thesis probe the topic of technology use (i.e. smartphone use) from the aspect of *what influences our technology use* and how this information could be leveraged. The following two articles (Articles III and IV) investigate technology use and adaption from the perspective of wearable devices (i.e. Android smartwatches).

4.1 Article 1: The influence of mood in application selection

Article I introduces two experiments where participants ($N = 15$ for study A and $N = 21$ for study B) reported their mood levels throughout the day over two weeks. In Study A, the participants reported their *happiness*, and in Study B, the participants reported their *pleasure* and *activeness* values according to the circumplex mood model. The pleasure axis correlates to the happiness value, and using the mood model, the angular sum can be used to derive more detailed mood *categories* (see Figure 6). The mood levels were normalised to use a three-tier ('Low', 'Neutral', 'High') scale. The application use was collected via the on-board sensors, and the data were periodically uploaded to a secure server. The collected data set contains a total of 35K instances of application changes (M = 972 per user) and 9.2K (M = 254 per user) mood reports. The circumplex mood model entries are then mapped to eight distinct mood categories (Figure 6).

Research has shown an association between different types of smartphone use (e.g. application choices), human affect and mental states (e.g. depression in young populations who habitually use social media) (Lin et al., 2016). Similarly, Niforatos and Karapanos (2014) showed that a specific type of application use leads to a heightened mental state among its users.

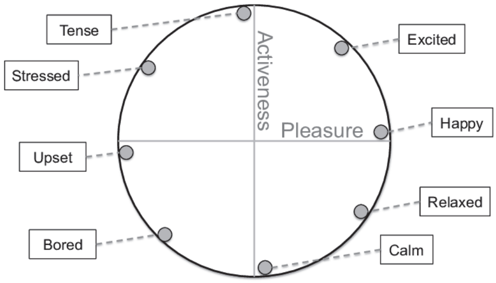

Fig. 6. The circumplex mood model, measuring individual's affect with two axes: activeness and pleasure (Reprinted by permission from Article I © 2016 Authors).

Using Pearson's Chi-squared test, the article initially reveals different moods experienced at different periods of the day. The most likely experienced moods begin with 'Stressed' or 'Upset' during the morning (6 a.m.–12 p.m.), leading to 'Tense' or 'Calm' during the afternoon (12 p.m.– 18 p.m.) and finally to 'Relaxed' during the evening. The daily cycle and how individuals' days are structured influence how they experience day-to-day life. The statistical test verifies this effect ($\chi2(21) = 168.2$, $p < .05$) for different periods of the day. Hardly a novel result, at least in psychology, the work then proceeds to measure how (or whether) these differences influence our technology use.

In both Studies A and B, the happiness level and application choices were significantly related ($\chi2\ (14) = 38.61$, $p < 0.05$), and in Study B, both the activeness level ($\chi2\ (14) = 84.59$, $p < 0.05$) and happiness level ($\chi2\ (14) = 32.90$, $p < 0.05$) were significantly related to the application choices. Furthermore, Chi-squared test reveals a semi-weak (Cramer's V = .23) relation between the experienced mood and application choices ($\chi2(49) = 328.29$, $p < .05$). The experienced *pleasure* level indicates significant differences in application categories of 'Games', 'Internet and Social', 'Maps and Travel' and 'Media', in particular. This indicates that mood plays a critical role in our smartphone use and application selection.

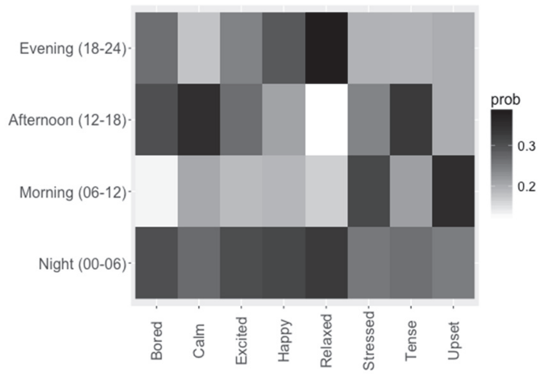

Fig. 7. The variance of an individual's mood during the day. (Reprinted by permission from Article I © 2016 Authors).

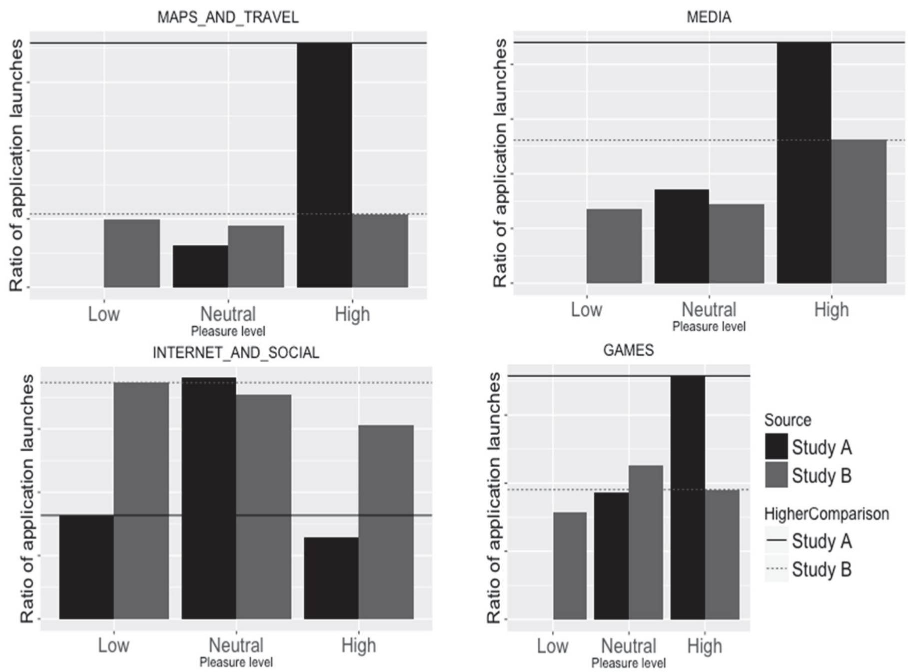

Fig. 8. Use of different application types according to the user's pleasure level. (Reprinted by permission from Article I © 2016 Authors).

The results of Article I indicate that regarding the presented research question, our mood influences our mobile application choices. Similar effects have been discovered regarding social media applications, and how negative emotions are often associated with increased social media use (and vice versa, although this spectrum was not explored in our research). However, many of the relationships between moods and application choices can also be explained through third or fourth variables. For example, the use of 'Maps and Travel' applications is often associated with high pleasure levels; however, this could ultimately be an outcome of having a holiday in a new location, for example.

4.1.1 Limitations

In Article I, we depicted a very broad idea of mood, and how it influenced our technology usage choices. The user-reported grading uses a scale (1 to 10 Likert scale) that might include a lot of variability and ambiguity, especially when trying to generalise over the studies—one person's medium pleasure level might be another's low, and some participants might experience more daily variance in their mood than others. Numerous works have since (and before) analysed mood using more rigorous methods, either with more detailed reporting or focusing solely on a single axis, such as anxiety, boredom or stress. There is also the ongoing ambiguity of self-reported values and the need for either intelligent predictive systems or hardware solutions that can be used to track and model personal affect. The use of such instruments, however, has been limited thus far. As such instruments make further progress and advance the field of affective computing, and its applications in both research and consumer use would greatly benefit.

4.2 Article II: Leveraging traces of technology use for personal improvement

The majority of our smartphone use is unprompted and habitual (Lee, Chang, Lin & Cheng, 2014); much of the use is not deemed as *meaningful* (Lukoff, Yu, Kientz & Hiniker., 2018), and the use is rarely planned in terms of continuing existing goals (van Berkel et al., 2016). Quantified-Self applications, aiming to collect self-senses or self-reported data of the user, generally suffer from lack of long-term user engagement or motivation for use (Brown, McGregor & McMillan, 2014; Gouveia, Karapanos & Hassenzahl., 2015; van Berkel et al., 2015). The goal of Article II was to offer the user a quick and efficient method for data contribution in tandem with these short bursts of micro-usage (Ferreira et al., 2014) to ensure a longer commitment to using such applications.

Article II of this thesis deals with the topic of investigating technology use and takes a more active stance to influence the end-user directly via observing usage habits. Using the same data set as in Article I, although extended through including participants outside of merely logging their mood, the work explores self-reported data contributions from the full participant pool of 48 participants. Smartphone usage characteristics (e.g. distinctly identified session types) are shown to influence the participants' choices in interacting with applications. The data set contains a total of 19.3K traces of application interactions via the presented alert dialogues (see Figure 4) and 395K screen events, which are further aggregated into 49.3K smartphone usage sessions (events between the screen becoming unlocked and locked).

The usage sessions are divided into sessions aiming to either *continue* a previously left task or ones where the user deliberately begins a *new* task, according to a 45-second interval introduced in (van Berkel et al., 2016). Furthermore, each presented alert dialogue includes the resulting outcome for the prompt—whether the user perceived the context of the dialogue as *accepted* or *rejected*, and whether the dialogue was presented to the user immediately at the *beginning* of the usage session ('instant') or whether the dialogue was 'delayed' (five minutes into a usage session). These tracker variables, including the session duration, were analysed for cases where user data contributions were more likely. These data contributions included information of individuals' self-perceived mood, physical symptoms or activity-related information, set at the beginning of the experiment for each participant.

The idea of active continuous self-monitoring, known as the Quantified-Self movement, relies on continuous data collection, and any individual who intends to achieve personal improvement via technology benefits from both continued data collection and increasing the quantity of available data points. The purpose of Article II was to investigate the use of the said alert dialogues in such fashion and investigate how the use of alert dialogues (and other reminder mechanisms such as SMS messages or notifications) could be improved through a more detailed understanding of technology use.

Table 5. The ratio of 'accepted' dialogues according to the delivery timing (upper), and data contribution frequency according to the usage session type (lower) (Reprinted by permission from Article II © 2017 Authors).

% of accepted dialogues (N = 19.3K dialogues generated)	
Overall	80.0%
‘Delayed’ dialogues	89.4%
‘Instantly’ delivered dialogues	79.6%

Table 6. Data contribution frequency according to the usage session type (lower) (Reprinted by permission from Article II © 2017 Authors)

Session (N = 49.3K)	New task during usage	Continuing on a previous task
Overall (N = 49.3K)	71.0%	29.0%
Includes data contributions (N = 5.6K)	12.5%	9.5%

The collected data indicate significant increases in data collection frequency when dialogues were presented during a session (‘delayed’) rather than at the beginning of a usage session (‘instant’). Users tend to be willing (and able) to contribute data when they are starting a new task rather than returning to a previously left off one. These results have implications for the design of reminder mechanisms to increase the frequency of user responses. Identifying moments in time and during device use, where the user essentially has *more free time* (or *idle time)*, will benefit in terms of less unwanted interruptions to the user. More results and discussion regarding the area of interruptions are presented in Chapter 6.

Article II also explored the use of a specific novel medium, the alert dialogue and how it influenced the user experience by surveying the participants’ experiences. Over half (53.3%) of the participants preferred the alert dialogue as an input method for self-reported data snippets using an open-ended questionnaire, and four recurring themes became evident in their reasoning: (1) **improved recall**, (2) **situated data logging**, (3) **low effort** required and (4) **even distribution** of data entries. These findings fit well into the predetermined requirements for *leveraging breakpoints and idle time during smart use* presented in the article: (a) brief interactions, (b) appropriate timing of a prompt, (c) easy dismissal and (d) quick generation.

The alert dialogue itself is not a delivery medium whose effect can be directly compared with other mediums (e.g. notifications or passive methods). The results still indicate that, within our explored research question regarding leveraging, different types of usage sessions and device usage-related metrics for more frequent data logging considerations taken during device use can definitely have a positive influence on user response frequency (i.e. *Which mobile device usage-related metrics can be leveraged to facilitate more frequent self-reported data logging?*). A potential side effect of this can also be an increase in data quality.

4.2.1 Limitations

While there were clear takeaways in the results of Article II, not all the feedback from the participants was overwhelmingly positive. The frequent prompts were occasionally deemed interruptive, and the quality and reliability of self-reported information could diminish if the user felt burdened or was prompted too often. While the alert dialogue methodology was manageable for the experiment, it might not be suitable for long-term deployment. Additionally, if multiple applications would simultaneously leverage such methods, the burden to the user would rapidly become unbearable. However, the findings in Article II can potentially be applied to any delivery method or consideration of when users are more likely to be willing to contribute to self-reported data.

4.3 Article III: Quantifying smartwatch usage sessions

The research in wearable technologies from both perspectives of usage and design is a re-emerging field, which is mainly industry driven due to the re-emergence of smartwatches and other wearable technologies (e.g. activity trackers). The use of mobile sensing was not available during the era when the smartphone was an emerging product; thus it is challenging to reflect on the research of smartphone use during its early days. Our work sought to investigate and quantify smartwatch use and compare the uncovered details of usage traces to similar characteristics in smartphone use. The smartwatch is often described as an extension to the smartphone; however, in reality, the two devices often function in tandem.

Cited research often reveals that the most desired feature of the smartwatch is the capability to address the notifications without having to interact with the smartphone (Min et al., 2015). Other benefits include inconspicuous use (Palen et al., 2000), as overuse of smartphones is often seen as a negative from the societal perspective, and increased availability (compared with merely using one's smartphone to address notifications, calls, messages, etc.) (Böhmer, Lander, Gehring, Brumpy & Krüger, 2014; Salovaara et al., 2011). While both Yan, Chu, Ganesan, Kansal, and Liu (2012) and Ferreira et al. (2014) have revealed that even smartphone use is habitually short-lived and happens in short bursts, it is hypothesised that smartwatch usage takes this behaviour even further.

In Article III, we analysed a data set collected by the Insight 4 Wear application (Rawassizadeh et al., 2015) (described in Chapter 3). The data set included screen events ($N = 800{,}119$), indicating periods of device use, and information on arriving notifications ($N = 2{,}801{,}082$), namely the source application of the notification and when it arrived. After pre-processing the data into usage sessions, following a similar methodology used by Gouveia (Gouveia, Pereira, Karapanos, Munson, and Hassenzahl, 2016) ($N = 798{,}423$) and removing any outliers (e.g. sessions with abnormally long durations, $N = 1696$ or 0.2% of the entries), the data set was categorised by *session type* and *interaction type* (Figure 9) and by whether each arriving notification was observed within the following 60 seconds (Figure 10). This process was described earlier in this thesis in Chapter 3. Figure 9 and Figure 10 showcase the results of the pre-processing in terms of distributions of user- and notification-initiated use, *peek* and *interaction* sessions and whether each notification was *promptly observed* or not.

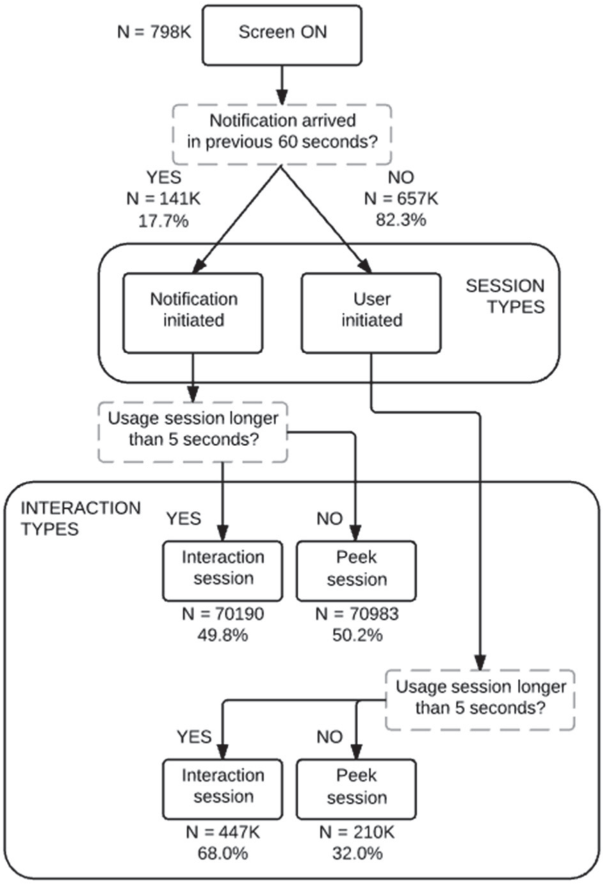

Fig. 9. Categorisation of usage sessions into notification- or used-initiated and peek and interaction sessions (Reprinted by permission from Article III © 2017 Authors).

4.3.1 Key findings

The results presented in Article III are comprehensive. In this thesis, the key findings and underlying reasoning for these outcomes are described. The article is first of its kind to elaborate smartwatch use and uncovers the following key points:

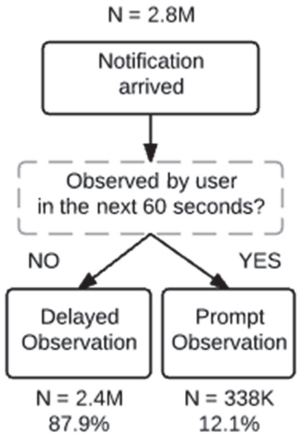

Fig. 10. Distribution of notifications either delayed or promptly observed (Reprinted by permission from Paper III © 2017 Authors).

User-initiated use: While notifications are said to play a significant role in the smartwatches usability and benefits (Cecchinato et al., 2015, Cecchinato, Cox & Bird, 2017), much of the actual use of smartwatches is initiated by the user. The smartwatch screen can be turned on by touch, by moving one's wrist or by incoming notifications as long as the smartwatch is already in a position relative to the user where he/she can observe the incoming notifications. Thus, a set of user-initiated use can merely be attributed to the mistakes of the user or, in some cases, outlier events within the data set. Regardless, the naturalistic use of a *watch*—keeping and observing time—is likely the most significant reason for the majority if use originates from the user. Furthermore, when the user initiates the use, the majority (68.0%) of the use also involves interactions with the device.

Delayed notifications: Of all the smartwatch notifications, 87.9% are not attended to within the first minute of arrival. According to Obuchi et al. (2016), smartphone notifications are addressed within 10 minutes (603.3 seconds) on average. When addressing non-adherence to incoming notifications, Park, Lim, Kim, Lee, and Lee (2017) specify an average deferral time of 51 seconds for delivering notification successfully. While the ratio of 12.1% notifications observed within the first 60 seconds seems low, it is still a significant increase compared with the behaviour elicited by smartphone users. The median duration for observing a notification in our smartwatch data set was 20.0 seconds, and, although Sahami Shirazi et al. (2014) did not report their median, they did specify 50% of the notifications were interacted with within 30 seconds—not a significant difference compared with Park et al.'s (2017) work.

Smartphone use is ubiquitous: Smartwatch users still rely on their smartphones during out-of-office hours, mainly focusing on their smartwatch use during the day (between 8 a.m. and 8 p.m.). This is likely part of the day when their time is limited, and most of the consumed content and information are delivered initially through the smartwatch. Similarly, later during the day (and even during the night), the smartphone offers more useful functionalities (e.g. video and music playback, communication applications) than the smartwatch. A plethora of research has been focusing on designing novel methods to interact with the smartwatch (Gong, Yang & Irani, 2016; Xu & Lyons, 2015); however, without the industry adapting to these designs, in reality, the smartwatch users still neglect to rely on their smartwatches for *producing* content or communication. Furthermore, the session duration and density between smartwatches and smartphones are separated, as indicated in Figure 5.

The research question for Article III has mostly been answered, as the main differences uncovered in the article between smartwatch and smartphone use are already presented. However, there could still be additional significant differences if we would analyse other usage trace types. Application choices, reactions to incoming notifications, response speed and frequency to messages among others could all showcase significant differences between the two device types. Analysing these traces would require parallel data sets for reasonable accuracy, which we did not have at the time of the original publication. The central insight and its outcome in analysing smartwatch usage is that the use is more frequent but less engaging. The idea that the smartwatch functions in tandem with the smartphone is partly correct; however, the watch has minimal capabilities in terms of engagement and interaction. Alternatively, the users merely opt to *not* engage with all the provided functionalities, perhaps due to the difficulty of use.

4.3.2 Limitations

Data analysis based solely on quantified information (i.e. data collected without further knowledge of the evidence surrounding the collected data) always rouses a certain level of interpretation. In contrast to the data collected in Articles I-II and V, the smartwatch-related articles rely on data collected by unknown end users. The scale of the data set is also larger, leading to further interpretations on behalf of the collected data. Odd and unique data traces, if left unidentified, can lead to misinterpretations within the analysis. To prevent such misinformation, in both Articles III and IV, we attempted to identify any outliers and filtered the data prior to analysis. In Article III, in particular, the number of usage traces were discarded in the pre-processing to reduce the number of such outliers; in other words, a high number of notifications from certain applications that used notifications in an unorthodox manner, and which partly deviated from the design guidelines were discarded.

4.4 Article IV: Usage transition over time in early smartwatch adopters

The work in Article III offers a general view of smartwatch use. The results piqued interests in the authors to investigate smartwatch use in more detail. Considering the smartwatch is a new emerging product, it is often claimed in relevant literature that the adoption of smartwatches to everyday use is dwindling and that end users are unsure of the visible and tangible benefits gained from the technology (Wu et al., 2016).

During the late 1990s and early 2000s, the smartphone, as we know it today, was still a novel product. The majority of the devices were *mobile phones* (i.e. the classic Nokia or Motorola devices) that over 90% of people in Finland owned during the early 2000s (Puro, 2002). *Smartphones*, as we know them today, existed mainly as personal digital assistants (PDAs), and only a fraction of the population owned such a device (6–7% during 2005–2006 in US campuses) (Chase & Herrod, 2005). Smartphones were mainly seen as a tool to improve work quality and efficiency, particularly in health care (Kuziemsky, Laul & Leung, 2005; Lu, Xiao, Sears & Jacko, 2005). During the early stages of smartphone adoption, some effort was made to understand how and why end users used smartphones, but no conclusive answer was discovered. Potential use cases were often related to work, especially in business and in healthcare. Qualitative works (e.g. Kim, 2003) explored the factors involved in using a smartphone. In the end, as we now know, the smartphone was adopted by the general public and became a ubiquitous device.

Similarly, the use of a smartwatch is currently being explored and investigated to understand why and how the devices are used *in-the-wild*. Article IV investigated the use and *adoption of the use* of the smartwatch during 2016 to understand *how* end users adopted their use to obtain the mentioned *tangible* benefits from their devices and whether any conclusions can be drawn of these beneficial usage styles.

From the same data source as Article III, we collected an extended longitudinal data set between December 2015 and March 2017 and identified 81 individuals who had logged data for a minimum period of four months. In addition to data from notification sensor and screen sensor, Article IV included battery-related information, as battery constraints are often cited as critical features in smartwatch adoption (Min et al., 2015) most valued by users (Hosio et al., 2016). From the sensor data, 98 different types of usage characteristics (features) were derived (see the last article at the end of the thesis for a full list and description of these characteristics). These features explore:

a) **Notification-related behaviour** (e.g. ratio of promptly observed notifications).
b) **Session-related behaviour** (e.g. ratio of short duration, less than 30 seconds, or isolated, two consecutive sessions at least 10 minutes apart). The concepts of 'short, isolated, or reward-based' (SIRB) use was originally introduced in Oulasvirta et al. (2012).
c) **Battery-related behaviour** (e.g. the battery level at the beginning of a charging event).
d) **Usage consistency-behaviour** (e.g. the longest streak of consecutive usage days). The concepts of usage consistency was originally introduced in Meyer et al. (2017).

From each user's daily data, daily statistics were generated for each feature in feature groups 1–3 (notification, session and battery-related behaviour), and the statistics were further averaged for every 30 days (one month), with the feature group 4 (usage consistency) included. This process resulted in a total of 486 unique 30-day long usage style periods.

4.4.1 Exploratory and accepted usage behaviours

Based on these 30-day periods of averaged usage statistics, the *k*-means clustering method was used as an approach to group similar usage behaviours. The process of grouping behaviours together according to different number of unique behaviours, ranging from 5 to 100, benchmarked using Dunn index (Maulik & Bandyopadhyay, 2002) (measuring the distance between each unique group) and Shannon's Entropy (Pincus, 1991) (measuring the entropy in distribution of behaviours into different groups), resulted in an optimal set of 33 unique types of behaviours. As each 30-day was mandated to fit into a group, some of them could be placed sub-optimally.

Thus, we discarded any outlier behaviours within groups by measuring the mean distance (Euclidean distance of all the combined features) from the behaviour to the *centre* of that behaviour and discard those that were deemed too *different* from the core behaviour (centre).

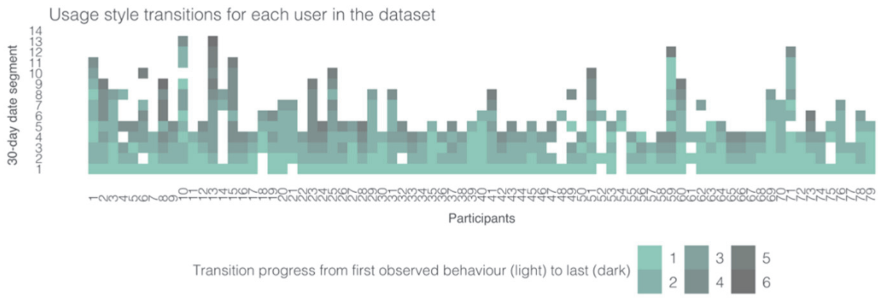

Fig. 11. Usage style behaviour transitions (Reprinted by permission from Paper IV © 2019 Authors).

For each user, their behaviours were then ordered based on their timing within the data collection period, thus creating a longitudinal set of user behaviours. Figure 11 highlights the transitions from one behaviour to the next for each of the participants. From the time series data set, we can then observe whether users returned to any past behaviours, and how long they spent within the same behaviour type. By assessing whether (and when) a behaviour recurs during the data set, we can categorise the usage behaviours into *exploratory* and *accepted* behaviours. Each behaviour that occurred once only ($N = 6$ out of 33) is discarded as irrelevant for further analysis.

– Total of 10 behaviours recurred and decreased in popularity over time indicating that these usage styles were prominent during early device use but were habitually abandoned as the adopted usage style. These 10 behaviours are classified as **exploratory behaviours**—behaviours that the users tried but deemed unviable for long-term device use.
– Total of 16 behaviours recurred and became increasingly popular, indicating that the users collectively adapted their device usage to the characteristics of this behaviour. These 16 behaviours are thus classified as **accepted behaviours**—behaviours that benefit the user according to their goals and expectations for owning and using a smartwatch.

Another evidence found from the data to support these exploratory and accepted behaviours was derived from looking at the transitions and transition frequency from one usage behaviour to another. For example, annotating one participant's (P71) transitions from one usage behaviour (annotated by Bxx) to another with an arrow notation resulted in the following:

B14→B17→B21→B21→B21→B21→B32→B32→B32→B32→B32

In this example, P71 starts with exploratory behaviours (B14 and B17), then transitions to another behaviour (B21) and does not return to the previous behaviour. The article also revealed that outside of the select 16 accepted behaviours, users would not return to their previously abandoned exploratory behaviours and tend to remain within an accepted behaviour for longer periods. These results highlight the effectiveness of the used methodology and reveal user's proneness to adapt their use away from unbeneficial usage styles.

Lastly, as the exploratory and accepted behaviours were analysed in terms of the individual usage features (e.g. the number of usage sessions during the day or number of arriving notifications), three key findings were revealed to distinguish between the two types of usage behaviours:

- **Notification-related use**: Accepted behaviours elicit fewer daily notifications (M = -64.8, $p < .01$), less promptly observed notifications (difference of -.30 in the ratio, $p < .05$) and more notifications from 'Other' (+.04, $p < 0.01$) and 'Games' (-.001, $p < .05$) categories. Lastly, the notification-initiated sessions are less likely for accepted behaviours (difference of -.03 in the ratio, $p < 0.01$), albeit by a small margin and with more variance within the exploratory behaviours. The differences are illustrated in Figure 12.

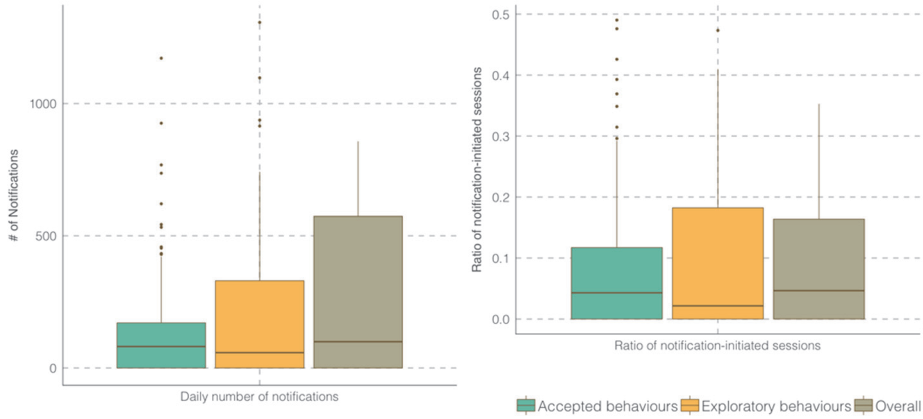

Fig. 12. Differences in notification-related usage characteristics between accepted and exploratory behaviours (Reprinted by permission from Paper IV © 2019 Authors).

- **Battery-related use**: There are noticeable differences in both daytime charging and draining behaviours between the accepted and exploratory behaviours. The accepted behaviours generally elicit higher charging likelihood (+.06, $p < .05$) during the evening hours (21 p.m.–23 p.m.) and higher hourly drain rate (ranging between +0.4 and 1.4, $p <.05$) during the night (12 a.m.– 2 a.m.) and the afternoon (14 p.m.–17 p.m.). The differences are illustrated in Figure 13.

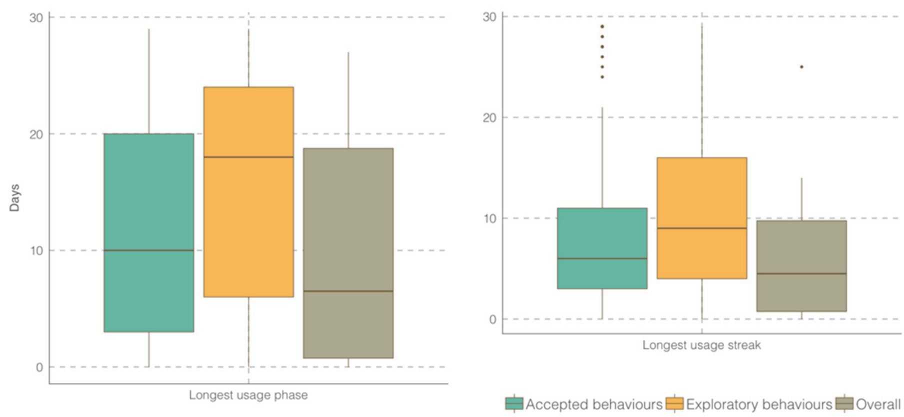

Fig. 13. Differences in battery charging and battery drain features between the accepted and exploratory behaviours. X-axis denotes hours and Y-axis the frequency of charging

events (right) and the draining ratio (left) (Reprinted by permission from Paper IV © 2019 Authors).

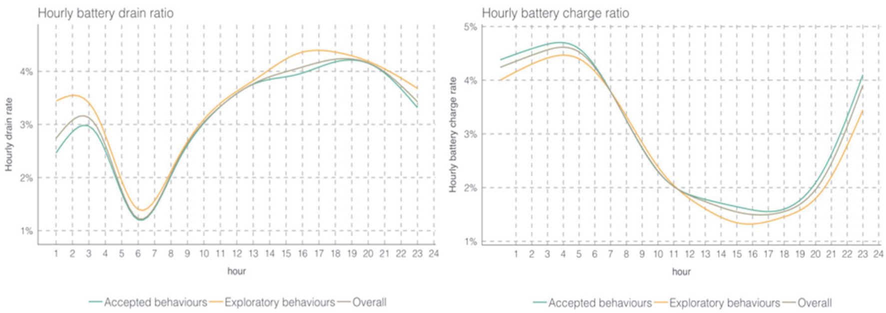

Fig. 14. Differences in usage frequency between the accepted and exploratory behaviours (Reprinted by permission from Paper IV © 2019 Authors).

- **Frequency of use**: Accepted behaviours show decreases both in terms of daily usage sessions **(**-25.6, $p < 0.01$), frequency of use (-.01, $p < -.05$) during the morning hours (7 a.m.–9 a.m.) as well as shorter usage streaks (-2.7 days, $p < .01$) and usage phases (-3.7 days, $p < .01$) (see Figure 14 for details).

As described in the conclusion of the previous article, we hypothesised that smartwatch users self-regulated their engagement with the smartwatch. Article IV explored this idea further by revealing several key metrics that are preferred by the smartwatch users. The use evolves to become significantly selective, and smartwatches are used in chosen scenarios rather than being ubiquitous in use.

4.4.2 Limitations

One significant limitation in the exploratory analysis of Article IV is the lack of certainty about when a user first installs the logging application. Thus, some users might have been using it on their smartwatches for a long time, while others might install it right after purchasing a new smartwatch. This article examined numerous revisions, and this was part of the reviewer's critique of the article. The finalised approach takes this limitation into account, as it only attempts to identify two types of behaviour without commenting on when such behaviours would be exhibited.

4.5 Discussion

Due to the emergence of increased connectivity in our smartphones, compared with earlier used terms such as 'cellular phone' for similar form-factor, the research community has been provided with efficient tools for measuring end users' technology use. While the 2000s saw a share of this type of HCI research, the possibilities provided by mobile sensing took it further. For other smart devices (e.g. the re-emergent smartwatch), this type of sensing technology did not exist either. This thesis and its presented articles related to technology use among individuals explored ways to understand further human-computer interaction as well as methods for leveraging that understanding.

Several previous or parallel works have either leveraged sensing or qualitative approaches to understanding technology use in smartphones or watches. An early landmark work was conducted by Falaki et al. (2010) who explored diversity in smartphone user base. The work revealed ever-present diversity within the users with no two identical users in a data set of 225 users, analysed from the perspective of usage sessions, battery and application choices. Soikkeli, Karikoski, and Hämmäinen (2011) also focused on the topic of user diversity in usage sessions according to varying contexts. The paper described more extended usage sessions within Home-context vis-à-vis Office-context (location-based); it concluded by recommending adapting the design to accommodate for context-awareness.

Both Oulasvirta et al. (2012) and Jones et al. (2015) explored application choices and usage and showcased distinct differences between individuals, how they revisited their applications (e.g. recurring use of communication applications) and how this behaviour made smartphone usage *pervasive*. Several factors can influence application choices, and many of them remain unexplored. The variance in the subject's affect can be one of them, which was discussed in Article I. Naturally, as technology and application space evolve over time, so does the underlying reasoning for user choices and interactions.

Another factor in technology use is understanding not only the influence of the user on technology (user-driven interaction choices) but how technology can be used to influence user decisions. Within the scope of this thesis, we addressed the *reminder mechanisms* and application-initiated data collection methods (i.e. ESM - experience sampling methods) (Larson & Csikszentmihalyi, 1983). Prior work in enhancing data collection frequency, quality and accuracy via reminders used both SMS messages (Patrick et al., 2009) and notifications (Bentley & Tollmar, 2013). Both showcase benefits; however, the problem of continuously interrupting the user can persist. Article II explored the use of a more covert intervention mechanism via the alert dialogue and revealed advancements over prior work.

Many works also elicit the use of qualitative methods for both labelling and enhancing the collected sensor data or simply gathering similar information sets via surveys, interviews or focus groups. In many cases, it is absolutely necessary to collect qualitative information to capture and label sufficient training data for modelling human behaviour. Investigating fragmented but continuous smartphone usage sessions (van Berkel et al., 2016), understanding people's receptivity to mobile notifications (Mehrotra, Pejovic, Vermeulen, Hendley & Musolesi, 2016) or exploring the value that end users attach to their existing battery charge levels (Hosio et al., 2016) may not be conducted based solely on sensor data. Similarly, both Articles I and II collected self-reported information of human affect for data collection and labelling purposes as well as for modelling human behaviour. This approach was continued in Article V and addressed in Chapter 5.

The existence of qualitative information is not always necessary, as indicated by both Articles III and IV that investigated the use of smartwatches solely from the sensor data perspective. The quantitative approach functions well as long as the agenda of the analysis is not the *exact* understanding of the user's motivations for actions but merely the actions themselves. However, in understanding the use of smartwatches, authors such as Schirra & Bentley (2015), Cecchinato et al. (2015) and Pizza et al. (2016) have taken a purely qualitative approach, using surveys, interviews or observations, for example, via lifelogging (continuous video recordings of an individual's daily life). Both Weber (Weber, Voit, Kratzer & Henze, 2016) and Min et al. (2015) combine the approaches to understand specific parts of smartwatch use, namely notifications in multi-device environments and how smartwatch users place value and manage their smartwatch batteries—one of the critical points of critique for the end-user adoption of smartwatches. Tables 7 and 8 summarise some of the recent research contributions in understanding technology use in both smartphones and smartwatches.

Table 7. Summary of recent research in smartphone use, and how the presented articles fit into the landscape.

Author	Year	Device		Quantitative					Qualitative			
		SP	SW	Sessions (screen)	Notifications	Battery	Location	Other	Survey & interviews	Focus groups	Self-reports	Lifelogging
(Falaki et al., 2010)	2010	X		X		X		X				
(Soikkeli et al., 2011)	2011	X		X								
(Denzil Ferreira et al., 2011)	2011	X				X						
(Oulasvirta et al., 2012)	2012	X		X				X			X	
(Rahmati & Zhong, 2013)	2013	X		X				X	X	X		
(Lee et al., 2014)	2014	X							X			
(Ferreira et al., 2014)	2014	X		X				X	X		X	
(Pielot et al., 2014)	2014	X		X	X				X			
(Jones et al., 2015)	2015	X						X				
(Van Deursen et al., 2015)	2015	X							X			
(Hosio et al., 2016)	2016	X		X	X	X	X	X	X			
(van Berkel et al., 2016)	2016	X		X					X		X	
(Mehrotra et al., 2016b)	2016	X			X		X	X			X	
Article I	**2016**	**X**		**X**							**X**	
Article II	**2016**	**X**		**X**							**X**	

Table 8. Summary of recent research in smartwatch use, and how the presented articles fit into the landscape.

Author	Year	Device		Quantitative					Qualitative			
		SP	SW	Sessions (screen)	Notifications	Battery	Location	Other	Survey & interviews	Focus groups	Self-reports	Lifelogging
(Schirra & Bentley, 2015)	2015		X						X			
(Min et al., 2015)	2015		X			X			X			
(Cecchinato et al., 2015)	2015		X						X			
(Pizza et al., 2016)	2016		X									X
(Cecchinato et al., 2017)	2017		X						X			
Article III	2017	X	X	X	X							
Article IV	2018		X	X	X	X		X				

5 Interruptibility and human attentiveness

The last chapter of this thesis is devoted to the field of measuring human attention via sensing. A recent comprehensive survey on attention management systems by Anderson et al. (2018) provided several key takeaways in designing such systems as well as understanding human attention. Attention is either captured or steered by external and internal stimuli, while the stimulus properties impact the user's reaction to the stimulus.

> "Attention is often considered as selective processing of incoming sensory information, with limited capacity and reactive and deliberate processes. Attention is also referred to as the ability to ignore irrelevant information."
>
> (Anderson, 2018, p. 3)

Ashcraft (2006) summarise the definition of attention via cognitive psychology as the mental process of concentrating effort on a stimulus or mental event.

Thus, the lack of attention can often lead to *interruptions*, and similar niche of research is described as research in *interruptibility*. Both Chang and Tang (2015) and Dingler and Pielot (2015) investigated attentiveness to mobile messaging, while Pielot (2014) approached the same scope from the perspective of availability to calls. The field of interruptibility research aims to determine a user's availability, readiness and interest in a given content element (Fogarty et al., 2005). Social and behavioural cues allow humans to assess a person's level of interruptibility (Barker, 1968; Hatch, 1987). In 2005, Fogarty et al. (2005) found that relatively simple external sensors could be used to successfully construct a model on a person's interruptibility.

Sensor-based modelling approaches leverage machine learning classifiers traditionally to create predictions (or evaluate prediction accuracy when *in-the-wild* deployments are not conducted). Similarly, machine learning has been used to model human affect based on sensor data (e.g. different mental states, coined as *affective computing*) (Mohr, Zhang & Schueller, 2017) or simpler conditions such as boredom (Pielot et al., 2015). The use of machine learning in this scope has its drawbacks; however, as the target of the modelling is *human*, who tends to have individual behavioural habits. Additionally, collecting sufficient contextual information is challenging (if not impossible), as a vast number of different predictors (or *features)* influence human decision-making and activities. Simply put, while machine learning is a reliable method for modelling scenarios with perfect (or near perfect) information, introducing the human subject creates uncertainty in the predictions. Including the individual differences introduces another challenge, as it is not necessarily appropriate to include person A's data into person B's model. These issues are part of the reason why the gold standard for sensor selection and prediction accuracy is yet to be reached in tandem with mobile sensing, and frequently a prediction accuracy of roughly 80% can be seen as sufficient.

5.1 Training data collection for human interruptibility

Addressing the challenges of training data collection in machine learning is difficult, especially the so-called *cold start* problem; if the model has no prior knowledge of a person's behaviour, it cannot offer any predictions. Traditionally, this is addressed by generating a *general model* that comprises the data of all available subjects (users). However, as previously described, this can be troublesome as individual differences can account for substantial prediction inaccuracies. On the other hand, relying solely on *personal models* collected solely from a person's training data over time introduces a time restraint to predictions, as it will be necessary to wait until sufficient amount of training data have been collected prior to receiving any predictions.

Article V aimed to mitigate this challenge by using *group-based models*. In essence, through the use of external factors to assign subjects to similarly behaving groups we can alleviate the cold start problem by assigning a new user into the best fitting group and use the groups' combined model for predicting individual user's behaviour.

Similar to Article II, the data for Article V were collected via the LifeTracker application (Visuri, 2016) but with an extended number of features, somewhat similar to those in Article IV. The analysis slightly considered processed raw data from the screen sensor (processed into usage sessions), network sensor (type, availability), battery sensor (charging state and level), proximity sensor, the physical activity sensor, LifeTracker dialogue generation details (delay and timing) and interactions with the said LifeTracker application. Each event of an alert dialogue that was presented to the user was then tagged with this contextual sensor data, and the label (prediction) for each event was extracted from the interactions with the application. We used a 2 x 2 matrix to infer the label classification according to whether the user contributed data using the alert dialogue or opted to classify the dialogue as interrupting (clicked 'Do not bother me') or not (clicked 'Ok'). These resulted in four classifications (A, B, C and D) for each event, as described in Table 9.

Table 9. Training data label classification in Article V (Reprinted by permission from Paper V © 2017 Authors).

	Dialogue accepted (non-interrupting)	Dialogue rejected (interrupting)
Data contributed	A. Non-interrupting, data contributed (N = 8,099)	C. Interrupting, data contributed (N = 374)
No data contributed	B. Non-interrupting, no data contributed (N = 7,490)	D. Interrupting, no data contributed (N = 3,277)

5.2 Grouping similar users

Our grouping approach aimed to construct distinct user group profiles. In Article V, we utilised demographic information (age, gender), the pre-study survey (Q1–Q4 below) and each user's device daily usage patterns (usage frequency during different hours of the day) to form user groups with a total of 40 dimensions, using the k means clustering algorithm. Since none of this information was included in the data used to create the machine learning classifiers, the grouping procedure did not bias the predictions (e.g. by appointing higher importance to features more prominent in one group over another).

- Q1. 'Do you often read the arriving notifications immediately?' On a 4-point scale (Never, Sometimes, Usually, Always).
- Q2. 'What kind of applications (categories) do you use on your smartphone?' According to a selection from Play Store categories, including the 'Other' category, which allows the user to be more specific using free text.
- Q3. 'Would you describe your smartphone use as active (frequent short periods), passive (only check when you are prompted by notification, for example) or mixed?'.
- Q4. 'Would you describe yourself as a technology enthusiast?' On a 4-point scale (Definitely not, Not really, Somewhat, Definitely).

Using this approach, we identified four groups of users with distinguishable (and human understandable) characteristics: 'Casual users', 'Social chatterers', 'Work on-the-go' and 'Night owls'.

5.3 Discussion

The approach in Article V turns out to be an efficient method for alleviating the cold start problem and deployable in a real-world setting. The best performing classifier out of the benchmarked ones, the Random Forest, offers a 79.0% prediction accuracy across the four classes as a general model. The leave-one-out validation method—each user is used as testing data and the *remaining users* as training data—to simulate the cold start challenge offers somewhat lower accuracy (78.7%). The user models—each model trained solely on the user's data—increase the accuracy slightly (80.4%).

The group-based modelling, however, shows significantly ($p < .05$, t = 3.20, df = 151.35,) higher mean accuracy (81.9%, SD = 2.8%) when compared with any other models. The changes are small, albeit statistically significant. The reasons for such small differences are likely the validity of the base classification approach and sensor selection as well as the quality of the training data. Many prior works have used similar sensor sets when classifying either interruptibility (willingness to address arriving notifications, messages or calls) or for modelling and predicting human behaviour (e.g. stress or boredom). Notably omitted sensors in Article V were arriving notifications, details of application choices (namely application names) and location information, which were at the time omitted in order to retain user's privacy during the experiment. The applications were collected; however, numerical identifiers (first used application as 1, second as 2, etc.) were used for each separate user, which made it infeasible to collectively know which application was assigned to which number (as the numbering was different for each user).

Another positive indicator of the approach was revealed when we looked at the *feature rankings* in each distinct group. To understand how and why our users interacted with the dialogues in specific ways, we used feature extraction to gain insights into each factor. MDA showed the impact of each feature on the accuracy of the classifier if the feature was removed, and MDI was used as the impurity function.

Overall, the classifiers that used user clustering understood changes in physical activity and proximity in more detail, and individual clusters put weight on features such as network type and hour ('Casual Users', Cluster 1), Wi-Fi and Internet availability ('Night Owls', Cluster 2) and session duration ('Work On-the-go', Cluster 5). Also, while the general model used the dialogue delay as the most important feature, four out of five clusters found dialogue delay to be less impactful. The same applied to session type, new or continuing.

From our work, it is evident that personal applications should not rely on generalised models, as differences in smartphone use between users have been brought up repeatedly in the literature. Different user types are more active during different times of the day, have different usage styles in terms of usage session frequency and duration (van Berkel et al., 2016; Visuri et al., 2017), prefer different types of applications (Welke et al., 2016; Zhao et al., 2016) and interact with their devices differently (Falaki et al., 2010). Applications can leverage this approach to use historical data from their user base as training data for new users by matching the characteristics of new users to existing user groups. From the perspective of the presented research question, there are increases in the prediction accuracy when the data was separated and filtered during pre-processing. A more important result, however, could be that data separation in Article V could mitigate the larger cold start problem.

5.4 Limitations

In many cases, the approaches proposed in Article V cannot be universally applied. The approach requires external details inputted by the user or collected via sensing approaches, and not all applications support this type of information collection. Users can also opt to deny access to such information on both personal and sensor level.

6 Conclusion and future work

This thesis explored the field of mobile sensing via increasing and adding to our combined knowledge of how smart technologies are used in-the-wild. Throughout the thesis, we presented methodologies and insights on how end users interact with their devices. This knowledge can be leveraged to effectively create intelligent and less disruptive mobile applications and solutions. A plethora of other research exists within the same scope; thus one of the limitations of this thesis is its lack of coverage of the entire scope. To address this limitation, we included a range of other works in the discussion (and within the presented articles) to highlight our contribution. We believe that this thesis in tandem with the other conducted research offers a solid basis for understanding the fundamentals of mobile sensing, deploying application and experiments that leverage mobile sensing and drawing useful insights from the sensed data.

While the works presented in this manuscript all address the topic of attention, it should be noted that any influence on attention is left mostly unanalysed. Measuring human attention, especially a longitudinal analysis during a long-term study performed in-the-wild can be challenging. Some research has been conducted in modelling attention, e.g. (Grillon, Robinson, Mathur & Ernst, 2016); however, the work is still in its early stage, and models and methods for assessing changes in attention are not in extensive use. The area itself is both exciting and interdisciplinary and could prove to offer good grounds for future research.

Article I demonstrated that mood affected smartphone application use. However, it is difficult to encourage granular logging of such mood. Thus, Article II explored different methods to improve the quality and quantity of self-reported data by analysing usage traits when data were willingly contributed by end users. In Article III, we wanted to know how smartwatches were used in-the-wild and how they compare with smartphones' use. Article IV explored the smartwatch-related usage over time to understand how the usage of such novel technology evolves. Article V continued the overall theme of mitigating interruptions, which are key to both self-reported data on any device type and the reduction of interruptions caused by notifications and information overload—one of the leading motivations for obtaining smartwatches, although it ends up amplifying the problem instead of alleviating it.

An added benefit of this wide-ranging field of research within mobile sensing is the fact that much of it is still open for future work. This thesis forms a research arc, starting from a fundamental understanding of technology use in a particular case

(Article I) to creating a more high-level understanding of users' available attentiveness to interruptions (Article V). The study of attention, how our available attention is influenced by the existence of technology around us and how we use it pose an important and interesting research agenda that could be investigated in detail via mobile sensing. Physiological measurements can offer longitudinal insights into the user's affect and well-being *in-the-wild*, which is something that has been somewhat lacking, as most experiments are focused on either using lab studies or survey methods or relying solely on *modelling* affect instead of measuring it.

In summary, we are very optimistic about the future of mobile sensing and its applications in research scope. The limitations discussed within this thesis can efficiently be solved by applying new technologies and by the development of innovative methods. Consumer interests in areas such as personal health and activity tracking are increasing, which offer both an agenda for researchers and a group of participants with inherent interest. Wearable and mobile technologies to measure activity and health are becoming ubiquitous and could be the next step in mobile sensing and applying sensing to understand human decision-making and everyday life in general.

List of references

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999). Towards a better understanding of context and context-awareness. *Handheld and Ubiquitous Computing (HUC'99)*, *1707,* 304-307. doi:10.1007/3-540-48157-5_29

Anderson, C., Hübener, I., Seipp, A. K., Ohly, S., David, K., & Pejovic, V. (2018). A survey of attention management systems in ubiquitous computing environments., 2(2), 58.

Ashcraft, M. (2006). *Cognition* (4. suppl.). New Jersey: Pearson Education.

Barker, R. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford, California: Stanford University Press.

Bentley, F., & Tollmar, K. (2013, April). The power of mobile notifications to increase wellbeing logging behavior. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1095-1098). ACM.

Böhmer, M., Lander, C., Gehring, S., Brumby, D. P., & Krüger, A. (2014, April). Interrupted by a phone call: exploring designs for lowering the impact of call notifications for smartphone users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3045-3054). ACM. Breiman, L. (2015). Random forests. *Machine Learning October 2001*, *45*(1), 5-32. doi:10.1023/A:1010933404324

Brown, B., McGregor, M., & McMillan, D. (2014, September). 100 days of iPhone use: understanding the details of mobile device use. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 223-232). ACM.

Cecchinato, M. E., Cox, A. L., & Bird, J. (2015, April). Smartwatches: the Good, the Bad and the Ugly?. In *Proceedings of the 33rd Annual ACM Conference extended abstracts on human factors in computing systems* (pp. 2133-2138). ACM.

Cecchinato, M. E., Cox, A. L., & Bird, J. (2017, May). Always on (line)?: user experience of smartwatches and their role within multi-device ecologies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3557-3568). ACM.

Chang, Y. J., & Tang, J. C. (2015, August). Investigating Mobile Users' Ringer Mode Usage and Attentiveness and Responsiveness to Communication. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 6-15). ACM.

Chase, M. E., & Herrod, M. (2005). College student behaviors and attitudes toward technology on campus. *Computer*, *94*, 98.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). ACM.

Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing*, *5*(1), 4-7. doi:10.1007/s007790170019

Dingler, T., & Pielot, M. (2015, August). I'll be there for you: Quantifying Attentiveness towards Mobile Messaging. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1-5). ACM.

Dirican, A. C., & Aksoy, S. (2017). Step counting using smartphone accelerometer and fast Fourier transform. *Sigma Journal of Engineering and Natural Sciences*, *8*, 175-182.

Dong, Y., Scisco, J., Wilson, M., Muth, E., & Hoover, A. (2014). Detecting periods of eating during free-living by tracking wrist motion. *IEEE Journal of Biomedical and Health Informatics*, *18*(4), 1253-1260. doi:10.1109/JBHI.2013.2282471

Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., & Estrin, D. (2010, June). Diversity in smartphone usage. In *Proceedings of the 8th international conference on Mobile systems, applications, and services* (pp. 179-194). ACM.

Ferreira, D. (2013). *AWARE: A mobile context instrumentation middleware to collaboratively understand human behavior* (Doctoral dissertation, University of Oulu). Retrieved from http://herkules.oulu.fi/isbn9789526201900/isbn9789526201900.pdf

Ferreira, D., Dey, A. K., & Kostakos, V. (2011, June). Understanding human-smartphone concerns: a study of battery life. In *International Conference on Pervasive Computing* (pp. 19-33). Springer, Berlin, Heidelberg.

Ferreira, D., Goncalves, J., Kostakos, V., Barkhuus, L., & Dey, A. K. (2014, September). Contextual experience sampling of mobile application micro-usage. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 91-100). ACM.

Ferreira, D., Kostakos, V., & Dey, A. K. (2015). AWARE: Mobile context instrumentation framework. *Frontiers in ICT*, *2*(6), 1-9. doi:10.3389/fict.2015.00006

Fogarty, J., Hudson, S. E., Atkeson, C. G., Avrahami, D., Forlizzi, J., Kiesler, S., . . . Yang, J. (2005). Predicting human interruptibility with sensors. *ACM Transactions in Computer-Human Interaction*, *12*(1), 119-146. doi:10.1145/1057237.1057243

Gong, J., Yang, X. D., & Irani, P. (2016, October). Wristwhirl: One-handed continuous smartwatch input using wrist gestures. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 861-872). ACM.

Gouveia, R., Karapanos, E., & Hassenzahl, M. (2015, September). How do we engage with activity trackers?: a longitudinal study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1305-1316). ACM.

Gouveia, R., Pereira, F., Karapanos, E., Munson, S. A., & Hassenzahl, M. (2016, September). Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 144-155). ACM.

Grillon, C., Robinson, O. J., Mathur, A., & Ernst, M. (2016). Effect of attention control on sustained attention during induced anxiety. *Cognition and emotion*, *30*(4), 700-712.

Hatch, M. J. (1987). Physical barriers, task characteristics, and interaction activity in research and development firms. *Administrative Science Quarterly*, *32*(3), 387-399. doi:10.2307/2392911

Ho, J., & Intille, S. S. (2005, April). Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 909-918). ACM.

Horvitz, E., Apacible, J., & Subramani, M. (2005, July). Balancing awareness and interruption: Investigation of notification deferral policies. In *International Conference on User Modeling* (pp. 433-437). Springer, Berlin, Heidelberg.

Hosio, S., Ferreira, D., Goncalves, J., van Berkel, N., Luo, C., Ahmed, M., ... & Kostakos, V. (2016, May). Monetary assessment of battery life on smartphones. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1869-1880). ACM.

Iqbal, S. T., & Bailey, B. P. (2008, April). Effects of intelligent notification management on users and their tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 93-102). ACM.

Jones, S. L., Ferreira, D., Hosio, S., Goncalves, J., & Kostakos, V. (2015, September). Revisitation analysis of smartphone app use. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1197-1208). ACM.

Kern, N., Schiele, B., & Schmidt, A. (2007). Recognizing context for annotating a live life recording. *Personal and Ubiquitous Computing*, *11*(4), 251-263. doi:10.1007/s00779-006-0086-3

Kim, S. (2003). *Exploring factors influencing personal digital assistant (pda) adoption.* University of Florida.

Kracheel, M., Bronzi, W., & Kazemi, H. (2014). A wearable revolution: Is the smartwatch the next small big thing? *IT ONE Magazine 2014*, *7*(December), 18-19.

Kushlev, K., Cardoso, B., & Pielot, M. (2017, September). Too tense for candy crush: affect influences user engagement with proactively suggested content. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (p. 13). ACM.

Kuziemsky, C. E., Laul, F., & Leung, R. C. (2005). A review on diffusion of personal digital assistants in healthcare. *Journal of medical systems*, *29*(4), 335-342.

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 41-56). Wiley Jossey-Bass.

Lee, Y.-K., Chang, C.-T., Lin, Y., & Cheng, Z.-H. (2014). The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. *Computers in Human Behavior*, *31*, 373-383.

Leiva, L., Böhmer, M., Gehring, S., & Krüger, A. (2012, September). Back to the app: the costs of mobile application interruptions. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services* (pp. 291-294). ACM.

LiKamWa, R., Liu, Y., Lane, N. D., & Zhong, L. (2013, June). Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services* (pp. 389-402). ACM.

Lin, L. Y., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., Colditz, J. B., . . . Primack, B. A. (2016). Association between social media use and depression among US young adults. *Depression and Anxiety*, *33*(4), 323-331.

Liu, M. (2013). A study of mobile sensing using smartphones. *International Journal of Distributed Sensor Networks*, *9*(3), 272916.

Lu, Y.-C., Xiao, Y., Sears, A., & Jacko, J. A. (2005). A review and a framework of handheld computer adoption in healthcare. *International Journal of Medical Informatics*, *74*(5), 409-422.

Lukoff, K., Yu, C., Kientz, J., & Hiniker, A. (2018). What makes smartphone use meaningful or meaningless? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 22.

Maier, J., & Wörndl, W. (2015, September). Smartwatch Interaction-More than just Notifications. In *Mensch & Computer Workshopband* (pp. 299-303).

Matic, A., Pielot, M., & Oliver, N. (2015, September). Boredom-computer interaction: Boredom proneness and the use of smartphone. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 837-841). ACM.

Maulik, U., & Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(12), 1650-1654.

McClernon, F. J., & Roy Choudhury, R. (2013). I am your smartphone, and I know you are about to smoke: The application of mobile sensing and computing approaches to smoking research and treatment. *Nicotine & tobacco research*, *15*(10), 1651-1654.

Mehrotra, A., Hendley, R., & Musolesi, M. (2016). *PrefMiner: Mining user's preferences for intelligent mobile notification management*. Paper presented at the ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Mehrotra, A., Musolesi, M., Hendley, R., & Pejovic, V. (2015, September). Designing content-driven intelligent notification mechanisms for mobile applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 813-824). ACM.

Mehrotra, A., Pejovic, V., Vermeulen, J., Hendley, R., & Musolesi, M. (2016, May). My phone and me: understanding people's receptivity to mobile notifications. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 1021-1032). ACM.

Mehrotra, A., Tsapeli, F., Hendley, R., & Musolesi, M. (2017). MyTraces: Investigating correlation and causation between users' emotional states and mobile phone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 83.

Mehrotra, A., Vermeulen, J., Pejovic, V., & Musolesi, M. (2015, September). Ask, but don't interrupt: the case for interruptibility-aware mobile experience sampling. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (pp. 723-732). ACM.

Meyer, J., Wasmann, M., Heuten, W., El Ali, A., & Boll, S. C. (2017, May). Identification and classification of usage patterns in long-term activity tracking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 667-678). ACM.

Min, C., Kang, S., Yoo, C., Cha, J., Choi, S., Oh, Y., & Song, J. (2015, September). Exploring current practices for battery use and management of smartwatches. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*(pp. 11-18). ACM.

Mishina, Y., Murata, R., Yamauchi, Y., Yamashita, T., & Fujiyoshi, H. (2015). Boosted random forest. *IEICE TRANSACTIONS on Information and Systems, 98*(9), 1630-1636.

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, *13*, 23-47.

Mortazavi, B. J., Pourhomayoun, M., Alsheikh, G., Alshurafa, N., Lee, S. I., & Sarrafzadeh, M. (2014, June). Determining the single best axis for exercise repetition recognition and counting on smartwatches. In *2014 11th International Conference on Wearable and Implantable Body Sensor Networks* (pp. 33-38). IEEE.

Niforatos, E., & Karapanos, E. (2014). EmoSnaps: A mobile application for emotion recall from facial expressions. *Personal and Ubiquitous Computing*, *19*(2), 425-444. doi:10.1007/s00779-014-0777-0

Obuchi, M., Sasaki, W., Okoshi, T., Nakazawa, J., & Tokuda, H. (2016, September). Investigating interruptibility at activity breakpoints using smartphone activity recognition API. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (pp. 1602-1607). ACM.

Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K., & Tokuda, H. (2015, March). Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (pp. 96-104). IEEE.

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105-114. doi:10.1007/s00779-011-0412-2

Palen, L., Salzman, M., & Youngs, E. (2000, December). Going wireless: behavior & practice of new mobile phone users. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 201-210). ACM.

Park, C., Lim, J., Kim, J., Lee, S. J., & Lee, D. (2017, February). Don't Bother Me. I'm Socializing!: A Breakpoint-Based Smartphone Notification System. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 541-554). ACM.

Patrick, K., Raab, F., Adams, M. A., Dillon, L., Zabinski, M., Rock, C. L., . . . Norman, G. J. (2009). A text message-based intervention for weight loss: Randomized controlled trial. *Journal of Medical Internet Research*, *11*(1), e1. doi:10.2196/jmir.1100

Pielot, M. (2014, September). Large-scale evaluation of call-availability prediction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 933-937). ACM.

Pielot, M., Cardoso, B., Katevas, K., Serrà, J., Matic, A., & Oliver, N. (2017). Beyond interruptibility: Predicting opportune moments to engage mobile phone users. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 91.

Pielot, M., Church, K., & De Oliveira, R. (2014, September). An in-situ study of mobile phone notifications. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 233-242). ACM.

Pielot, M., De Oliveira, R., Kwak, H., & Oliver, N. (2014, April). Didn't you see my message?: predicting attentiveness to mobile instant messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3319-3328). ACM.

Pielot, M., Dingler, T., Pedro, J. S., & Oliver, N. (2015, September). When attention is not scarce-detecting boredom from mobile phone usage. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing* (pp. 825-836). ACM.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, *88*(6), 2297-2301.

Pizza, S., Brown, B., McMillan, D., & Lampinen, A. (2016, May). Smartwatch in vivo. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5456-5469). ACM.

Poppinga, B., Heuten, W., & Boll, S. (2014). Sensor-based identification of opportune moments for triggering notifications. *Pervasive Computing, IEEE*, *13*(1), 22-29. doi:10.1109/MPRV.2014.15

Puro, J.-P. (2002). Finland: a mobile culture. *Perpetual contact: Mobile communication, private talk, public performance*, 19-29.

Rahmati, A., & Zhong, L. (2013). Studying smartphone usage: Lessons from a four-month field study. *IEEE Transactions on Mobile Computing*, *12*(7), 1417-1427.

Rawassizadeh, R., Price, B. A., & Petre, M. (2015). Wearables: Has the age of smartwatches finally arrived? *Communications of the ACM*, *58*(1), 45-47.

Rawassizadeh, R., Tomitsch, M., Nourizadeh, M., Momeni, E., Peery, A., Ulanova, L., & Pazzani, M. (2015). Energy-efficient integration of continuous context sensing and prediction into smartwatches. *Sensors*, *15*(9), 22616-22645.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161. doi:10.1037/h0077714

Sahami Shirazi, A., Henze, N., Dingler, T., Pielot, M., Weber, D., & Schmidt, A. (2014, April). Large-scale assessment of mobile notifications. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3055-3064). ACM.

Salovaara, A., Lindqvist, A., Hasu, T., & Häkkilä, J. (2011, August). The phone rings but the user doesn't answer: unavailability in mobile communication. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (pp. 503-512). ACM.

Sano, A., & Picard, R. W. (2013, September). Stress recognition using wearable sensors and mobile phones. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 671-676). IEEE.

Sarsenbayeva, Z., Ferreira, D., van Berkel, N., Luo, C., Vaisanen, M., Kostakos, V., & Goncalves, J. (2017, September). Vision-based happiness inference: a feasibility case-study. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (pp. 494-499). ACM.

Schilit, B. N. N., & Theimer, M. M. M. (1994). Disseminating active map information to mobile hosts. *IEEE Network*, *8*(5), 22-32. doi:10.1109/65.313011

Schirra, S., & Bentley, F. R. (2015, April). It's kind of like an extra screen for my phone: Understanding Everyday Uses of Consumer Smart Watches. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2151-2156). ACM.

Scholl, P. M., & Van Laerhoven, K. (2012, July). A feasibility study of wrist-worn accelerometer based detection of smoking habits. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing* (pp. 886-891). IEEE.

Servia-Rodríguez, S., Rachuri, K. K., Mascolo, C., Rentfrow, P. J., Lathia, N., & Sandstrom, G. M. (2017, April). Mobile sensing at the service of mental well-being: a large-scale longitudinal study. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 103-112). International World Wide Web Conferences Steering Committee.

Shoaib, M., Bosch, S., Scholten, H., Havinga, P. J., & Incel, O. D. (2015, March). Towards detection of bad habits by fusing smartphone and smartwatch sensors. In *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)* (pp. 591-596). IEEE.

Soikkeli, T., Karikoski, J., & Hammainen, H. (2011, September). Diversity and end user context in smartphone usage sessions. In *2011 Fifth International Conference on Next Generation Mobile Applications, Services and Technologies* (pp. 7-12). IEEE.

Sun, L., Zhang, D., Li, B., Guo, B., & Li, S. (2010). Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In *Ubiquitous Intelligence and Computing* (pp. 548-562): Springer Berlin Heidelberg.

van Berkel, N., Luo, C., Anagnostopoulos, T., Ferreira, D., Goncalves, J., Hosio, S., & Kostakos, V. (2016, May). A systematic assessment of smartphone usage gaps. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4711-4721). ACM.

van Berkel, N., Luo, C., Ferreira, D., Goncalves, J., & Kostakos, V. (2015, September). The curse of quantified-self: an endless quest for answers. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (pp. 973-978). ACM.

Van Deursen, A. J., Bolle, C. L., Hegner, S. M., & Kommers, P. A. (2015). Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Computers in Human Behavior*, *45*, 411-420.

Visuri, A. (2016). Smartphone Based Contextual Symptom Tracking and Data Gathering.

Visuri, A., Sarsenbayeva, Z., van Berkel, N., Goncalves, J., Rawassizadeh, R., Kostakos, V., & Ferreira, D. (2017, May). Quantifying sources and types of smartwatch usage sessions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3569-3581). ACM.

Wang, Y., Lin, J., Annavaram, M., Jacobson, Q. A., Hong, J., Krishnamachari, B., & Sadeh, N. (2009, June). A framework of energy efficient mobile sensing for automatic user state recognition. In *Proceedings of the 7th international conference on Mobile systems, applications, and services* (pp. 179-192). ACM.

Weber, D., Voit, A., Kratzer, P., & Henze, N. (2016, September). In-situ investigation of notifications in multi-device environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1259-1264). ACM.

Welke, P., Andone, I., Blaszkiewicz, K., & Markowetz, A. (2016, September). Differentiating smartphone users by app usage. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 519-523). ACM.

Wu, L.-H., Wu, L.-C., & Chang, S.-C. (2016). Exploring consumers' intention to accept smartwatch. *Computers in Human Behavior*, *64*, 383-392.

Xu, C., & Lyons, K. (2015, January). Shimmering smartwatches: Exploring the smartwatch design space. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 69-76). ACM.

Yan, T., Chu, D., Ganesan, D., Kansal, A., & Liu, J. (2012, June). Fast app launching for mobile devices using predictive user context. In *Proceedings of the 10th international conference on Mobile systems, applications, and services* (pp. 113-126). ACM.

Zhao, S., Ramos, J., Tao, J., Jiang, Z., Li, S., Wu, Z., ... & Dey, A. K. (2016, September). Discovering different kinds of smartphone users through their application usage behaviors. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 498-509). ACM.

Original publications

I Visuri, A., Sarsenbayeva, Z., Goncalves, J., Karapanos, E., & Jones, S. (2016, September). Impact of mood changes on application selection. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (pp. 535-540). ACM.

II Visuri, A., van Berkel, N., Luo, C., Goncalves, J., Ferreira, D., & Kostakos, V. (2017, July). Challenges of quantified-self: encouraging self-reported data logging during recurrent smartphone usage. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference* (p. 81). BCS Learning & Development Ltd.

III Visuri, A., Sarsenbayeva, Z., van Berkel, N., Goncalves, J., Rawassizadeh, R., Kostakos, V., & Ferreira, D. (2017, May). Quantifying sources and types of smartwatch usage sessions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3569-3581). ACM.

IV Visuri, A., van Berkel, N., Goncalves, J., Rawassizadeh, R., Kostakos, V., & Ferreira, D. *Understanding usage style transformation during long-term smartwatch use*. Manuscript.

V Visuri, A., van Berkel, N., Luo, C., Goncalves, J., Ferreira, D., & Kostakos, V. (2017, September). Predicting interruptibility for manual data collection: a cluster-based user model. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (p. 12). ACM.

Original publications are not included in the electronic version of the dissertation.

691. Alavesa, Paula (2018) Playful appropriations of hybrid space : combining virtual and physical environments in urban pervasive games

692. Sethi, Jatin (2018) Cellulose nanopapers with improved preparation time, mechanical properties, and water resistance

693. Sanguanpuak, Tachporn (2019) Radio resource sharing with edge caching for multi-operator in large cellular networks

694. Hintikka, Mikko (2019) Integrated CMOS receiver techniques for sub-ns based pulsed time-of-flight laser rangefinding

695. Järvenpää, Antti (2019) Microstructures, mechanical stability and strength of low-temperature reversion-treated AISI 301LN stainless steel under monotonic and dynamic loading

696. Klakegg, Simon (2019) Enabling awareness in nursing homes with mobile health technologies

697. Goldmann Valdés, Werner Marcelo (2019) Valorization of pine kraft lignin by fractionation and partial depolymerization

698. Mekonnen, Tenager (2019) Efficient resource management in Multimedia Internet of Things

699. Liu, Xin (2019) Human motion detection and gesture recognition using computer vision methods

700. Varghese, Jobin (2019) MoO_3, PZ29 and TiO_2 based ultra-low fabrication temperature glass-ceramics for future microelectronic devices

701. Koivupalo, Maarit (2019) Health and safety management in a global steel company and in shared workplaces : Case description and development needs

702. Ojala, Jonna (2019) Functionalized cellulose nanoparticles in the stabilization of oil-in-water emulsions : Bio-based approach to chemical oil spill response

703. Vu, Kien (2019) Integrated access-backhaul for 5G wireless networks

704. Miettinen, Jyrki & Visuri, Ville-Valtteri & Fabritius, Timo (2019) Thermodynamic description of the Fe–Al–Mn–Si–C system for modelling solidification of steels

705. Karvinen, Tuulikki (2019) Ultra high consistency forming

706. Nguyen, Kien-Giang (2019) Energy-Efficient Transmission Strategies for Multiantenna Systems

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/

UNIVERSITY OF OULU P.O. Box 8000 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

UNIVERSITY of OULU
OULUN YLIOPISTO

ISBN 978-952-62-2270-7 (Paperback)
ISBN 978-952-62-2271-4 (PDF)
ISSN 0355-3213 (Print)
ISSN 1796-2226 (Online)