Karlsruher Institut für Technologie

Karlsruhe Reports in Informatics 2016,8
Edited by Karlsruhe Institute of Technology,
Faculty of Informatics
ISSN 2190-4782

Ubiquitäre Systeme (Seminar) und Mobile Computing (Proseminar) WS 2015/16

Mobile und Verteilte Systeme
Ubiquitous Computing

Teil XIII

Herausgeber:
Martin Alexander Neumann, Anja Bachmann,
Matthias Berning

2016

KIT – Die Forschungsuniversität in der Helmholtz-Gemeinschaft

Fakultät für **Informatik**

Please note:
This Report has been published on the Internet under the following Creative Commons License:
http://creativecommons.org/licenses/by-nc-nd/3.0/de.

Ubiquitäre Systeme (Seminar)
und
Mobile Computing (Proseminar)
WS 2015/16

Mobile und Verteilte Systeme
Ubiquitous Computing

Teil XIII

Herausgeber

Martin Alexander Neumann

Anja Bachmann

Matthias Berning

Karlsruhe Institute of Technology (KIT)
Fakultät für Informatik
Lehrstuhl für Pervasive Computing Systems (PCS) und TECO

Interner Bericht 2016-08
ISSN 2190-4782

Vorwort

Die Seminarreihe Mobile Computing und Ubiquitäre Systeme existiert seit dem Wintersemester 2013/2014. Seit diesem Semester findet das Proseminar Mobile Computing am Lehrstuhl für Pervasive Computing System statt. Die Arbeiten des Proseminars werden seit dem mit den Arbeiten des zweiten Seminars des Lehrstuhls, dem Seminar Ubiquitäre Systeme, zusammengefasst und gemeinsam veröffentlicht.

Die Seminarreihe Ubiquitäre Systeme hat eine lange Tradition in der Forschungsgruppe TECO. Im Wintersemester 2010/2011 wurde die Gruppe Teil des Lehrstuhls für Pervasive Computing Systems. Seit dem findet das Seminar Ubiquitäre Systeme in jedem Semester statt. Ebenso wird das Proseminar Mobile Computing seit dem Wintersemester 2013/2014 in jedem Semester durchgeführt. Seit dem Wintersemester 2003/2004 werden die Seminararbeiten als KIT-Berichte veröffentlicht. Ziel der gemeinsamen Seminarreihe ist die Aufarbeitung und Diskussion aktueller Forschungsfragen in den Bereichen Mobile und Ubiquitous Computing.

Dieser Seminarband fasst die Arbeiten der Seminare des Wintersemesters 2015/16 zusammen. Die Themen der hier zusammengefassten Aufsätze umfasst Mobile Sensing und Augmented Reality. Wir danken den Studierenden für ihren besonderen Einsatz, sowohl während des Seminars als auch bei der Fertigstellung dieses Bandes.

Karlsruhe, den 01. April 2016

Martin Alexander Neumann
Anja Bachmann
Matthias Berning

Inhaltsverzeichnis

Doppler Effect in Mobile Sensing

Ingo Sobik*

Advisor: Long Wang†

Karlsruhe Institute of Technology (KIT)
Pervasive Computing Systems – TECO
*ingo.sobik@student.kit.edu
†wanglong@teco.edu

Abstract. Mobile sensing is an emerging area of interest for researchers all around the world. As smart devices like smartphones, smartwatches, etc. continuously increase their computational power and the amount of built-in sensors, the amount of new application cases for smart devices in everybody's everyday life increases as well. The Doppler Effect as physical principal can leverage commodity hardware of these devices to improve interaction with digital interfaces or systems like smart-home environments. This paper provides an overview of a variety of applications that are build on top of the Doppler Effect and that are able make device interaction more comfortable and provide a rich human-computer interaction. It highlights the benefits that results from Doppler Effect based applications as well as possible challenges in the future.

Keywords: doppler effect, activity recognition, gesture recognition, tracking, localization, sound, radio frequency, mobile seinsing

1 Introduction

The development of mobile computing has made large leaps since the beginning of the 21st Century. Shrinking device size and ongoing growth of processing power have leveraged computers to integrate more and more in everybody's everyday life. But not only the technology itself, and the hence resulting change in the way we communicate, work and live, has further evolved. The growth in computational power, as well as the amount of built-in sensors and actuators in *Smartphones*, *SmartTVs*, *Smartwatches*, etc., enable a new way of collecting and processing information and data from devices and their users. This developement kick-started the scientific field of *Mobile Sensing*, which in turn leads to a continous improvement of the way we use and interact with our digital environment. Data, collected by GPS-Sensors, accelerometers, gyroscopes, microphones, cameras, WiFi, etc. can be used to leverage smart devices to perform tasks like navigation, tracking, health monitoring, gesture recognition and many more. All of them continuously improve the usability and user experience of smart devices, without the necessity of expensive additional custom hardware.

To enable such technological progress, new algorithms and principles, suitable for mobile devices, need to be found and evaluated. One of the latter is the *Doppler Effect* or *Doppler Shift*.

The Doppler Effect is a phenomenon, science frequently has profited from. Application cases range from medical engineering and radar tracking to astrophysics. In medical environment e.g. the Doppler Effect is used to improve image quality of sonography. Also many radar systems (i.e. the Doppler Radar systems) use the Doppler Effect and thus enable flight control, in-battlefield surveillance, radar speed traps in traffic, or meterological weather forecasts and tornado surveillance. Whereas astrophysicists, for example, use the Doppler Effect in gravity fields to detect exoplanets or to measure velocity curves of galaxies.

Additionally, this physical phenomenon is recently used to leverage commodity speakers and microphones, as well as WiFi modules of smart devices to monitor moving targets. Applications can be built on top of the Doppler Effect, to perform more complex tasks like gesture recognition [1, 3, 5, 15], activity recognition [4, 9, 12], localization or tracking [16, 18].

This seminar paper provides an overview of the principle itself and a variety of applications that make use of the Doppler Effect in Mobile Sensing. It shows that the Doppler Effect enables a lot of new possibilities for smart devices to collect and process user data. Especially in close range scenarios Doppler Effect based sensing achieves a very high accuracy compared to the relatively small amount of necessary computational costs. On the other hand, this paper also shows the deficits of Doppler Effect based sensing and its limitations like the absent robustness in long range scenarios and the lack of security features.

The rest of the paper is organized as follows: Section 2 introduces the physical principle of the Doppler Effect, Section 3 shows the application of the Doppler Effect in the field of mobile sensing and introduces a variety of applications. Section 4 compares the characteristics of the sound based and the radio frequency approach. Section 5 concludes and sums up the result of this paper.

2 Physical Principle of the Doppler Effect

The Doppler Effect is a well known physical phenomenon. It describes the change in frequency of a wave as its source and an observer move relatively to each other. A canonical example of the occurrence of the Doppler Effect is the audible change in the pitch of car's sound as it approaches and departs from an observer. Figure 1 shows this example: On the left-hand side of the picture both, the car and the observing persons are at rest. The sound waves of the car propagate equally in all directions. Thus, Person A and B observe the same frequency of the car's sound. The right-hand side of the picture shows the moving car while the observing persons A and B remain at rest. Since the propagation speed of sound remains constant, the motion of the car towards the person B induces a shift of the frequency in the car's direction. This results in a higher tone from the perspective of person B. Simultaneously person A observes a lower tone than in the beginning. In the moment the car passes person B, he observes a Doppler shift. In this case this is a down-pitch in frequency. Person A and B now observe a tone with an identical frequency.

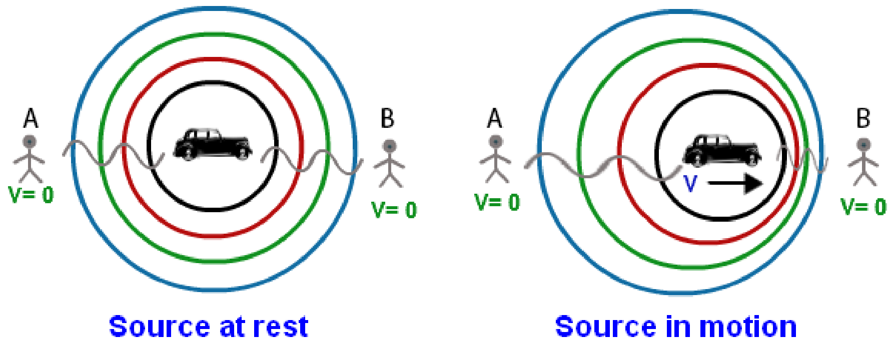

Fig. 1. Frequency shift (Doppler Effect) of a car's sound wave as it passes a resting observer.

This is an example of the Doppler Effect with a moving source and a resting observer. However in general, the Doppler Effect can occure in different manners:

1. Source moves while the observer remains still (see example figure 1)
2. Observer moves while the source remains still
3. Source and observer move relatively to each other with different speeds
4. Source and observer remain still, while an object (e.g. a user's hand) moves relatively to both of them

Item 4 of this enumeration is a special form of item 2. In this case the moving object reflects the source's wave, thus inducing the Doppler Effect. The moving object can now be seen as a virtual emitter of the reflected wave. Therefore this case is identical to the one shown by item 2.

Generally, the Doppler Effect allows us to monitor moving targets. Based on the strength and characteristics of each Doppler shift, it is possible to deduce detailed information about the original movement, i.e. speed and duration of a motion, as well as a motion-specific *Doppler pattern*. The latter is the result of observing multiple Doppler shifts over a period of time.

Equation 2 shows the relation between an emitted frequency f_0 and a perceived one f. The parameters c, v_r and v_s denote the propagation speed of the wave in the medium, the velocity of the observer relative to the medium and the velocity of the source relative to the medium, respectively.

$$f = \frac{c + v_r}{c + v_s} * f_0 \tag{1}$$

Generally, the Doppler Effect occures independently of the type of wave and the medium it travels trough. For example, it is possible to detect Doppler shifts in sound, light or radio frequency waves. However, the amount of the resulting Doppler shift varies with the speed of motion, the propagation speed of the wave, its frequency, as well as the movement of the medium (if any). Therefore the effort necessary to extract the Doppler shifts from the observed wave heavily depends on the wave's characteristics and the speed a motion is performed with.

This gets clearer by the following example: Assuming that it is possible to detect 1Hz frequency shifts in both, WiFi and acoustic signals, the accuracy in speed estimation is still higher in acoustic signals than in WiFi due to the slower popagation speed of sound. Using a frequency of 17Hz and the given propagation speed of sound at 346m/s, the resulting speed resolution is $\frac{346}{17000} \approx 2cm/s$. Assuming a pilot frequency of a WiFi signal at 2.4GHz, the resulting resolution is $\frac{3*10^8}{2.4*10^9} \approx 12.5cm/s$. This shows that for the same motion, the resulting Doppler shift of the acoustic signal is 6 times higher than the one of the WiFi signal [18]. Therefore, most of the Doppler Effect sensing is done with sound or radio frequency. Since they propagate at lower speeds or at least with a lower frequency. Furthermore, all of the papers that are examined in the following sections use the medium air that does not move.

3 Doppler Effect in Mobile Sensing

Before using smart devices like smartphones, there have already been applications that used the Doppler Effect to detect human activity. Paradiso et al., for example, used the Doppler Effect to measure upper-body kinematics like velocity and direction of motion [11]. Later, applications on commodity devices started on notebooks that came up with the required performance. Tarzia et al., for example, detect the presence of computer users by using the Doppler Effect on soundwaves [16]. Leveraging the Doppler Effect to enable gesture recognition or tracking functionality has required custom build hardware for a long time. This is because the performance of the devices was exceeded by the requirements of the applications.

Today, mobile devices have a lot of computational power, and togehter with the amount of their built-in sensors, the number of their application cases has coninuously increased. Mobile devices enable new functionalities, for example tracking the user's heart rate or sleep [4]. This all is possible without the need of custom hardware.

Data deduced from the Doppler Effect: The Doppler Effect is already used in a variety of ways in science and technology to monitor moving targets. The amount of a Doppler shift varies with the speed and direction of the moving object. If the object approaches a sensing device (source as well as observer), a positive Doppler shift is induced. Whereas, if the object departs from a sensing device the induced Doppler shift is negative. Figure 2 illustrates this more clearly. As an example, it shows a moving hand above a sensing device and the thus resulting spectogram of Doppler shifts.

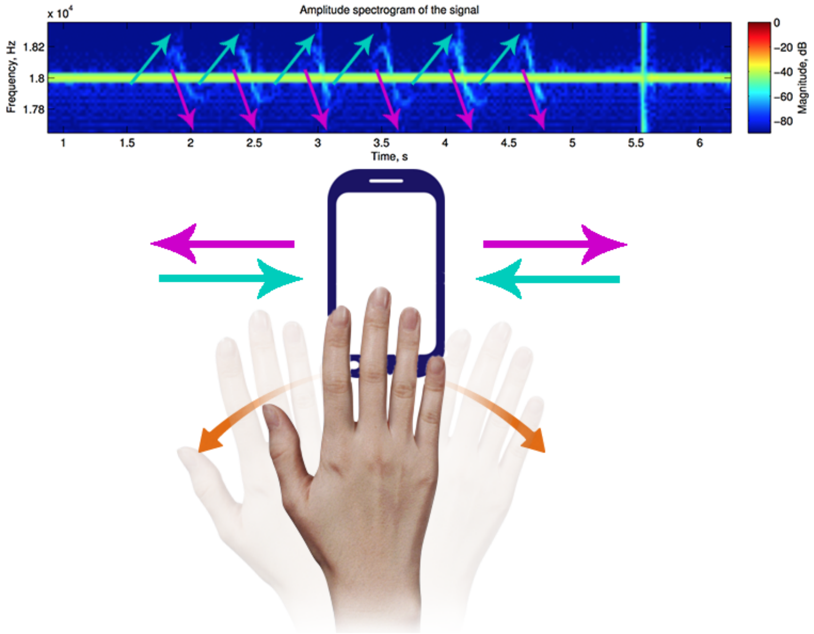

Fig. 2. Moving Hand above a sensing device and the resulting spectogram of Doppler shifts.

As shown in section 1, the Doppler Effect can be used to deduce information like speed and direction (approach/depart) of a moving object relative to a sensing device. If we increase the number of observers and/or sources, it is possible to also increase the amount of information we can get about a motion. Thus, it is possible to collect additional information like the distance an object moves or the angle of the motion relative to the source(s) or observer(s). If we, for example, use multiple smartphones to independently sense the Doppler Effects induced by a motion, the resulting Doppler patterns are going to slightly vary from each other. This is because their relative position to the moving object is slightly different. Therefore, it is possible to aggregate the data collected by the different phones and get more detailed information on the performed motion. The thus collected data can be used on mobile devices to enable new features. This includes person-to-person interaction (e.g. data exchange, or device pairing), person-to-device interaction like gesture control, or interacting with a smarthome environment.

Categories of Doppler Effect sensing: Basically, there are two common ways to leverage the Doppler Effect for sensing purposes: using acoustic sound waves, and using radio frequency waves (especially WiFi). Both of them have their own advantages and disadvantages, as further described in section 3.3. Independent from the type of wave that is used, the usecases of the Doppler Effect in mobile sensing can roughly be devided into three categories:

1. Gesture Recognition: e.g. filesharing, device pairing, multi display synchronization, device control
2. Activity Recognition: e.g. monitoring personal health, detecting falls, determining user precense, recognizing speech and breathing
3. Tracking and Localization: e.g. smartTV mouse, surveillance, person tracking

Gesture recognition is used to interact with digital interfaces or to remote control digital devices. Activity recognition, however, focuses on passive supporting tasks. Thus it might, for example, be possible to automatically change the settings of a smarthome environement, based on a user's activity. Tracking and localization uses the motion-induced Doppler shifts to keep track of a device or person.

When it comes to implementation, there are a couple of challenges all of these categories have in common. These are:

- How to accurately detect the Doppler shifts
- How to match certain activities or gestures with the Doppler shift patterns
- How to differentiate between multiple users
- How to deal with noise and other disturbing environmental influences

However, each category has its own additional challenges. The following sections give an overview of several published works from each of the three categories. They discribe the different approaches and highlight their individual challenges and contributions.

3.1 Gesture Recognition

Gesture recognition focuses on improving the interaction with a mobile device or a digital interface in general. Beside the known input methodes like keyboard, mouse, touch or spech, gesture recognition can make device interaction more comfortable (e.g. while having dirty hands), more safe (e.g. while driving a car), or can just be used to bring a user rich Human Computer Interaction (HCI) through remote control.

Examples of existing sound-based approaches are:

- **Airlink [3]: gesture-based filesharing, device pairing**
- **Doplink [1]: gesture-based filesharing, device control**
- **SoundWave [5]: gesture-based device control**
- **Spartacus [15]: gesture-based device pairing**

– **Dolphin [13]: gesture-based device control**

The basic functional principle of these approaches is very similar. All of them use a continuous pilot tone soundwave with a frequency between 18KHz and 22KHz that is emitted by a mobile device. This frequency is used because almost every commodity smart device is capable of emitting a tone with that frequency. Nevertheless, the authors of [5] state that that their system can work with frequencies down to 6KHz. Although it is technically possible, it is rather impracticable, since sounds of 6KHz can be perceived by humans. If multiple devices are used, there are three different roles a device can have:

1. Transmit the soundwave
2. Receive the soundwave
3. Transmit and receive the soundwave

If a device is just the sound emitter, it is either completely passive, or it is involved in the gesture. If it is the sound receiver, it is a device that waits to detect doppler shifts on a wave emitted by an other source. And if a device performs both actions, it emits a soundwave and perceives possible reflections of it (Doppler Effect).

Basically, there are two different ways to perform a gesture. Either while holding a device, or without holding a device. Performing a gesture without device, requires one or more stationary device(s) to transmit and/or receive the soundwave. In this case, the Doppler shift is induced by the hand reflecting the soundwave. If a gesture is performed while holding a device, it must either receive, or transmit the soundwave. In this case, the Doppler shift is induced by transmitter and receiver moving relatively to each other.

Airlink: [3] In Airlink, Chen et al. use multiple devices that do both, emitting and receiveing the signal. They use the gesture recognition to enable filesharing between two devices. To share a file, the user performs a swiping gesture starting at a specific device and ending at the device the data is meant to share with. All devices emit the same pilot tone at 18-19KHz. The emitted wave is reflected by the hand, inducing a Doppler shift that can be perceived by the mobile phones.

The challenge is, especially if there are more than 2 devices, to specify where the gesture starts and where it ends. This can be solved due to the fact that the Doppler shift varies with the moving direction of the hand. If the hand approaches a device, a positive shift is induces. If it departs a device, a negative one is induced. Due to each device monitoring the Doppler shifts of their own pilot tone, all of them register a different Doppler shift, depending on their position relative to the gesture. The device the gesture starts at only detects a negative Doppler shift, whereas the device the gesture ends at only detects a positive Doppler shift. All other devices that are positioned in between detect both, a positive, as well as a negative Doppler shift.

Figure 3 illustrates this more clearly. If shows the spectograms of a swipe gesture above three phones (A, B, C) that are positioned in a line. A user performs the swipe gesture starting at phone A, passing phone B and ending at phone C. Phone A detects a negative Doppler shift, since the user's hand departs from it. Phone B detects both, a positive as well as a negative Doppler shift, since the user's hand first approaches phone B and then departs from it. In contrast, phone C just detects a positive Doppler shift, since the user's hand approaches phone C where the gesture ends.

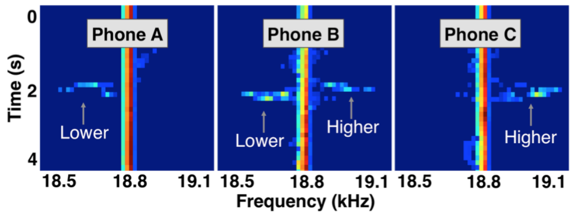

Fig. 3. Spectograms of a swipe gesture above three phones positioned in a line. Phone A detects just a negative Doppler shift, phone B detects a negative as well as a positive Doppler shift, and phone C detects just a positive Doppler shift.

i) Detection Process:
To detect the Doppler shifts each device samples incoming signals at 44.1kHz and processes them in segments of 4096 samples. For each segment it computes a *Fast Fourier Transformation* (FFT) to calculate a 2048-point FFT vector with 10.77 Hz bins. That way, the system creates a stream of FFT vectors that are processed to recognize the gestures. If a device detects a Doppler shift, it sends its processed data to a central server. The server then compares the individual Doppler Effect profiles and determines the source and destination device in order to enable the filesharing.

Besides device-to-device data transfer, Airlink also supports broadcasting. This is realized by performing a specific "patting" gesture above the source device. In this case, the originating device detects the corresponding Doppler shift signature and sends it to the central server. Again, the server enables the broadacast to the neigbouring devices.

ii) Evaluation:
Chen et al. evaluated Airlink in multiple ways. Due to the fact that the Doppler effect varies with the angle the gesture approaches the phones, they varied the topology of the phones. Besides placing them on straight line, they also placed them in a triangle topology consisting of three phones. They evaluated how well a user can select a specific phone by performing a gesture from phone A to B as shown in Figure 4. Although the average accuracy in a straight line (with a distance of 25cm in between) topology is about 96.8%, the accuracy drastically decreases, if angle α is decreased. However, Airlink performs reliably for angles

from 180° to 120°. This is due to the fact that two phones being closer to each other is making their individually detected Doppler Effect signature more similar. How Airlink handles larger distance between the individual phones wasn't further evaluated. However, the authors make clear that the overall accuracy highly depends on the accuracy of the detection process, as well as the speed a gesture is performed with. This means that the accuracy also decreases when a gesture is performed too slow or too far beyond the target device. Therefore it is hard to use Airlink in a real remote scenario. This results in a rather rough cut concerning the usability, and leaves space for further research. Especially to improve intuitivity and robustness.

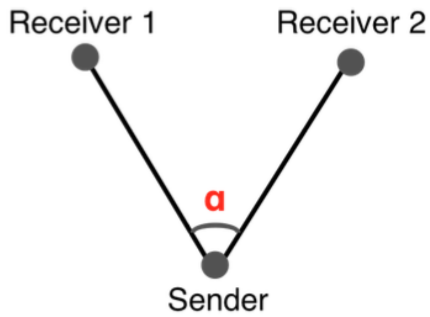

Fig. 4. Triangle topology. [3]

Besides topology issues, Chen et al. state that background noise is also a challenge for Airlink [3]. To cope with backround noise and wrong detections, the authors chose a pilot frequency above 18kHz that does not interfer with environment noise. They implemented an empirically chosen threshold value, that prevents the system from detecting unintentional Doppler shifts. The system does not detect any Doppler shifts that are below 2.5% of the pilot tone's magnitude. This noise cancelling is applied after the FFT of the incoming signal. To evaluate their threshold value, they placed the system for one hour in a noisy indoor environment without performing any gestures. During this period the system did not detect any false alarms. This shows that Airlink achieves a decent robustness to background noise, at least if it is used in an indoor environment.

Prior to the filesharing between multiple devices, the Airlink's central server needs to know the available devices (pairing). Chen et al. suggest to enable this by adding another dimension of interaction. They state that the pairing process could be realized by either performing a special swipe gesture between the devices, or by performing a special pointing gesture towards a device while holding another one in the hand. For this purpose it is mandatory that at least the still standing device is sampling incoming signals, and that the moving device emits a soundwave. This scenario is similar to the Spartacus system [15]. Nonetheless, this feature is not further investigated by the authors of Airlink. Neither is the energy consumption of the system, which is crucial, since other publications like [15] state that the continuous listening process consumes a large amount of energy.

The authors of Soundwave [5] for example, states that energy consumption can be reduced by simply deactivating the system e.g. during keyboard typing or during other activities that clearly make the detection process needless.

Another approach is made by Kellog et al. called *AllSee*. They focus on a low-power approach for gesture recognition, based on WiFi signals. Their key to energy efficiency is to decode the gesture information directly, using analog components intead of performing an analog-to-digital conversion. However, their approach requires custom hardware to be added to the smart devices [8]. This is clearly an disadvantage and does not fit the basic idea of only using commodity hardware.

Another aspect worth mentioning is the performance of Airlink. Although the authors do not evaluate the performance of Airlink itself, the authors of Dolphin [13] indirectly evaluate their system by using it to control interactive games. Since Dolphin is also a system for gesture recognition that works very similar to Airlink, this can be taken to indicate a trend. Another interesting fact about Dolphin is that they increase the accuracy of their system by including data from the gravity sensor, to optimize the sensing ability. This is reasonable, since they also found out that the phone's position influences sensing quality [13].

Spartacus: [15] In this approach Sun et al. also use commodity mobile device to enable device-to-device interaction like pairing or filesharing. The idea is to perform a pointing gesture that enables e.g. device pairing without prior configuration. Spartacus is deployed as an ordinary app.

Just like Airlink [3], Spartacus leverages the Doppler Effect on sound waves to enable gesture recognition. The active device in the pointing gesture emits a pilot tone soundwave while the passive devices perform the detection process.

To enable the device-to-device interaction, a pointing gesture is performed towards a dedicated mobile device while holding another one in the hand. Figure 5 shows a schematically usecase of spartacus: User A wants to interact with another mobile device (e.g. Device B). To enable interaction with device B, user A performs the pointing gesture towards B while holding his own device in its hand. Device B perceives the resulting doppler shift thus enabling the requested interaction (e.g. pairing or file sharing). A similar approach called DopLink is made in [1] by Aumi et al. They also introduce a device selection (or pairing) mechanism based on the DE. Contrary to Spartacus [15], they use notebooks as passive devices instead of smartphones. However, the performed pointing gesture is the same.

While developing Spartacus, Sun et al. were confronted with three major challenges. The first one addresses the resolution of the doppler shift detection. The target device is selected by comparing perceived doppler shifts among nearby

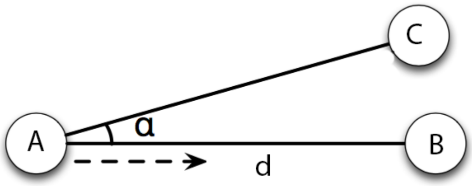

Fig. 5. Pointing gesture Spartacus from device A towards device B.

devices and chosing the one with the maximum peak. As already mentioned in section 3.1, pointing gestures of average users are usually slow with a short duration. This leads to limited Doppler shifts, and makes it even more difficult to accurately detect the peak Doppler shifts.

Using the Airlink approach, meaning to run a audio signal analysis process over a series of audio frames and applying a FFT on each of the audio frames, would result in an angular resolution of $\alpha = 26.7°$. Suppose the distance (d) between A and B in Figure 5 is 5m, this results in a spartial resolution of 2.3m. This means that the distance between device B and C has to be at least 2.3m to be able to clearly select on of them with a gesture. This is definetly too large for an in-field scenario [15].

i) Increasing Angular Resolution:
The spartial resolution hast to be increased, therefore Spartacus utilizes a *undersampling technique*. This means to decrease the sampling rate of the audio signal contrary to the Nyquist theorem [10]. Undersampling is a technique already applied in RF communication or image processing [6, 17] and is feasible because the detected frequency shifts are only a few hundred Hertz compared to the pilot frequency (about 18 - 22KHz) [15]. This means that only a small amount of the bandwith is relevant for the detection process.

Due to the undersampling, the angular resolution of Spartacus can be improved to $\alpha = 10°$. This is twice as high as the traditional FFT approach (like the one used in [3]).

The use of the undersampling process results in a further challenge: Due to the undersampling the spectrum of the original audio samples between 19KHz and 21KHz is mapped to a much lower spectrum between 0.58KHz and 2.58KHz, and a pilot frequency of 1.58KHz. Therefore aliasing arises, and the audio signal interfers with ambient noise. This makes the detection of frequency shifts impossible. To address this issue, Sun et al. designed and implemented a *bandpass audio signal processing pipeline* to recover the frequency shifts. In a nutshell, this means to split consecutive audio samples into overlapping analysis windows to improve the time-domain resolution. Afterwards a 10-order Butterworth bandpass filter is used to eliminate frequencies below 19KHz or above 21KHz. Then, the undersampling process, as well as the FFT is applied. The resulting data is an energy level value that is proportional to the amount of the detected Doppler shift. This energy value then is used to compare the perceived shifts among all nearby devices.

In contrast to Airlink 3.1, there is no additional server involved that takes care of determining the target device. The comparison takes place on the initiating device. Therefore, each device that detects a Doppler Effect reports its computed energy level, together with its unique device ID to the gesture device. The target device is determined by comparing all the received energy levels and picking the maximum.

ii) Energy Efficiency:
One of the major points of criticism at Airlink in section 3.1 is the lack of energy efficency. The continuously listening/detection process consumes a high amount of power, which already is a valuable good of energy limited mobile devices. Therefore, Spartacus additionally focuses on a low-power approach by introducing a *low-power audio listening protocol*. Unlike WiFi or Bluetooth, Spartacus reduces its action scope to devices that are in close proximity to the sender. The limited range of audio signals helps, to only focus on neighbouring devices within the same space. By chosing audio signals, Spartacus automatically seperates from interaction sessions that take place in neighbouring rooms, or even neigbouring buildings.

Furthermore, communication protocols like Bluetooth or WiFi are a lot more complex, and offer much more features than eventually needed for the purpose of Spartacus, or any other DE-based approach analyzed in this seminar paper.

Spartacus utilizes this fact by eschewing additional communication protocols for the gesture recognition process. All the information necessary to detect a gesture is transmitted via audio signals.

To reduce the energy consumption of the detection process, all devices are put in a standby condition. They periodically wake up to listen for a audio beacon.

Prior to a gesture, the sender device emits that beacon to make neighbouring devices aware of the upcoming gesture. The sender device also encodes its device ID and sends it in a transmission succeeding the beacon transmission.

This leads to an energy consumption 4X lower than scanning approaches using *WiFi Direct* and even 5.5X lower than using *Bluethooth 4.0* [15].

iii) Evaluation:
Concerning the evaluation of Spartacus, it achieves an accuracy of 80% - 90% within a distance of 4m to 5m between the devices. These values improve if the distance is reduced, making Spartacus very robust in close-range scenarios. If the interaction range is within one meter, Spartacus also achieves an accuracy above 90%, even if the directional difference between the devices is lower than $\alpha = 20°$. The device selection mechanism DopLink also achieves an average accuracy above 90%. DopLink even evaluated how obstacles between source and receivers influence the accuracy. They were able to detect 77.2% of the performed

gestures from behind a monitor or a whiteboard. To further improve these results, they suggest to apply additional multipath analysis in the future [1]. Once again, since Spartacus and Doplink are very simliar to each other, the evaluation results concerning the obstacles indicate a possible trend.

However, as distances keep increasing, the performance of both, Spartacus as well as DopLink drops gradually, making them only suitable for close-range scenarios. This evaluation results also correspond with those achieved by the Airlink approach [3].

In contrast to Airlink, Spartacus also adresses security issues. They state that the system is vulnerable to security attacks. A malicious device could continuously pretend to have detected higher Dopple Effects, thus disturbing or even blocking device interactions. They suggest to use a secure connection mechanism, so that only trusted and authenticated devices are allowed to report their Doppler shifts. Although these security approaches are yet to be further developed, they show the sensitivity of DE-based applications. Therefore the threat of an malicous user or attaker can be generalized for all DE-based applications.

It is also worth noting that Sun et al. bring up a potential downside of using frequencies above 15KHz on commodity microphones and speakers. Frequencies above 15KHz on commodity hardware result in a strong degradation of sound energy. Microphones and speakers of commodity mobile devices are designed to fulfill the needs of human conversation and media playing. This means, that even though the hardware is technically capable of working with sound frequencies beyond human acoustic perception, the ability to detect the Doppler shifts is significantly influenced by the hardware quality. Although the results in close-proximity scenarios are satisfying, to improve accuracy and increase interaction range, it is recommended to use audio sounds with lower frequencies, or better hardware [15].

3.2 Activity Recognition

Besides gesture recognition, activity recognition is the second application filed of DE-based mobile sensing. However both, activity recognition as well as gesture recognition, have a lot in common. Both are based on the Doppler sensing, therefore they face the same difficulties regarding the detection of the shifts. This allows to exchange insights between both fields. The main difference between activity recognition and gesture recognition is basically that the latter is used to interact with a digital interface or to control a device. Whereas activity recognition rather acts passively, as a supporting system. This can be used to monitor health, support independent living, to provide information about a user's location and direction, or to improve smart home environments.

Opportunities for Activity Recognition: [4] One major disadvantage of e.g. camera-based approaches to enable activity detection, is the induced privacy issues. DE-based applications protect the privacy and are easy and cheap to deploy. The Doppler Effect can be leveraged to detect falls of elderly people, or to detect diseases like sleeping apnea [9]. This all by placing a smartphone on the desk next to a person. Fu et al. show in [4] some opportunities for the Doppler Effect in the field of activity recognition and outline possible usage scenarios for this sensing technique.

They are using custom software to realize a prototype deployed on an unmodified smartphone to evaluate limitations and chances of activity recognition based on the DE.

Just like Spartacus [15] and Airlink [3], Fu et al. rely on commodity hardware to detect the audio Doppler signatures of various activities. They use a continuous wave with a frequency of 20KHz emitted by the smartphone's speakers. The built-in microphpone listens to the pilot tone and detects the tone's reflection at moving objects.

In their generic approach they try to figure out how different environmental conditions influence the functionality and usability of the DE-based sensing.

Although they primarily focus on activity recognition, they start by evaluating the Doppler signatures of various gestures to transfer their insights afterwards.

One of the first things [4] evaluates is the application range of the sensing. As mentioned in section 3.1 the maximum application range of Spartacus is about 5m, without a significant loss of accuracy. Spartacus however uses a pointing gesture that has a larger range of motion than e.g. a person resting. Hence, it is difficult to outline one generic maximum distance that fits all application cases. The maximum distance clearly depends on the specific activity's range of motion and the strength of the thus induced Doppler shift. Additionally, it depends on the setup of the sensing device.

If Doppler Effect sensing with a stationary phone is used to monitor sleep, the maximum distance differs widely compared to the detection of a person running around while holding the sensing phone. In the specific test environment of Fu et al. the maximum distance of their application is 2m [4].

Furthermore, they examined three major setups for audio doppler sensing, i.e stationary, holding the phone and carrying the phone on the body. Each one of these reacts differently to environmental influences.

i) Stationary Phone:
Keeping the phone stationary, they tried to distinguish different types of motion. I.e. simple hand movements (gesture recognition), walking by the phone, sleeping next to the phone and working on a computer next to the phone.

While performing hand motions like swyping, they noted that a major limitation of activity and gesture recognition is the fact that it is only possible to detect the absolute change in distance of an object relative to the device. This means that technically a gesture or activity only induces a Doppler shift if it aproaches or departs the phone. Thus it should be indistinguishable if a user performs a left-to-right or a right-to-left swipe motion.

However, not only the hand is involved in the swiping gesture. What also carries weight is the movement of the arm during the swipe gesture. Because the motion of the arm slightly varies depending on the direction of the swipe motion, it is yet possible to distinguish the two gestures. As opposed to this, Airlink (3.1) managed the distinction between left-to-right and right-to-left gesture by comparing the perceived shifts of multiple devices.

Besides recognizing hand gestures, [4] also shows that it is possible to recognize a person walking by the phone. Although the detected doppler shifts are simliar to those of the swyping gesture, it is possible to differentiate these two activities by including the duration of the movement, since a swiping gesture is executed way faster than walking by the phone. Figure 6 shows the detected Doppler shift signatures of both, the swipe gesture and the person walking by the phone.

However, they also state that arm and upper body movements induce Doppler shifts as well. This can significantly influence the Doppler shift signature of an activity. On the one hand this allows to differentiate more activities, whereas on the other hand it complicates the finding of a doppler shift signature that is capable of clearly identifying a certain activity. This means that the doppler shift signature can get blurred by individual movements of body parts.

Additionally, Fu et al. analyzed the Doppler shifts induced by more complex activities, i.e. a person sleeping and a person working on a computer, each time with the sensing phone next to the the person. In all test cases they were able to distinguish the activities with an stationary phone. Although movements like typing on a keyboard can influence the detection process. They bring to mind that the more complex an acitivity gets, the more difficult it gets to identify a clear Doppler shift signature [4].

ii) Holding the Phone:
Similar to Spartacus [15], Fu et al. also tested a couple of gestures and activities while holding the phone in one hand. This includes upward or downward motions, e.g. to regulate volume of a sound system, as well as swiping motions, e.g. to sort through a set of digital photographies. Although it is possible to differentiate these gestures on the basis of their doppler shift signatures, [4] shows that moving the phone itself results in an increased amount of "frequency shift noise" and hinders the detection process a lot.

It is stated that by Fu et al. that it might be possible to perform some sort of noise reduction by substracting an estimated amount of background and "walking noise" from the perceived signal. Anyways, this suggestion is not further discussed. Furthermore, they show that holding the phone is only suitable while walking in wide areas. Certain rooms like narrow corridors may cause many unintentional frequency shifts.

iii) Wearing the Phone on the Body:
The activity recognition based on audio Doppler shifts reaches its limits when it come to wearing the phone on the body. If the phone is worn inside the clothing, [4] states that there is too much noise covering the echoed signal. Therefore, it is almost impossible to clearly differentiate between certain activities.

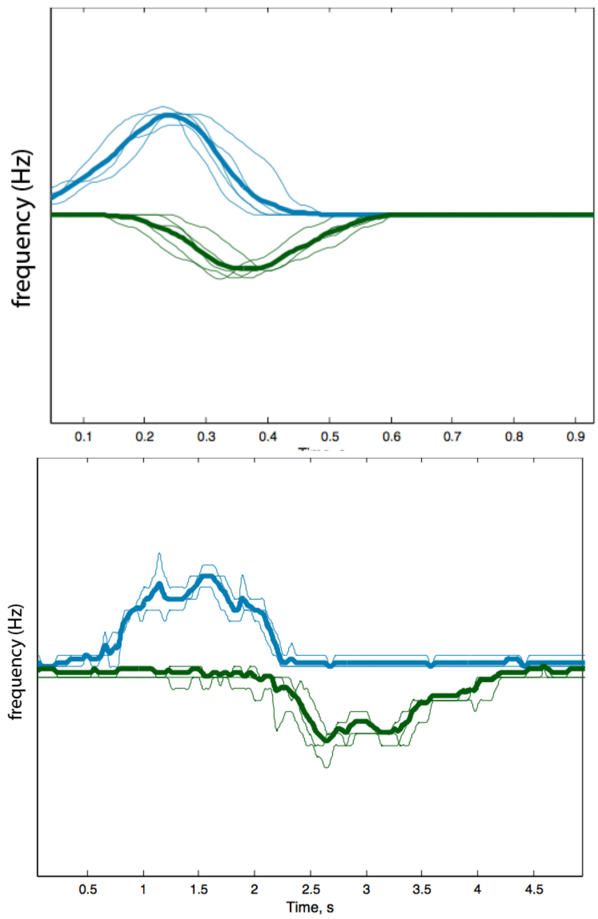

Fig. 6. Spectograms of swipe (top) gesture in comparison with the one of a walking activity (bottom). (Blue lines and green lines denote positive and negative Doppler shifts, respectively.) [4]

Fu et al. even started an approach to identify the surroundings of a jogger wearing the phone at his arm. They hoped to distinguish between e.g. urban, wood or fields, based on the reflections of the audio signal. Unfortunatley there is too much noise that hinders the extraction of a clear signal [4]. They suggest to work on an appropriate noise estimation based on the proximity of certain entities simliar to researches shown in [14], to clean the incoming signal.

In [4], Fu et al. note a couple of problems concerning the audio based Doppler Effect sensing that are also mentioned by Sun et al. in [15]. This includes the vulnerability of sound waves to noise, as well as the low quality of the devices' hardware. Furthermore, they leave out the security issues and the power consumption as well, although both of them can make the system impractical.

3.3 Tracking and Localization

The most ambitious application field of DE-based mobile sensing is tracking and localization. Target tracking based on GPS sensors, or body tracking by wearing nodes is widely spread by now. However, these approaches require a lot of additional hardware or lack accuracy in some cases. Yun et al. show in [18] that the Doppler Effect can also be leveraged to enable spatial device tracking by using the hardware already available in mobile devices. They introduce a mechanism that tracks the position of a mobile device to use it as a mouse in the air to interact with a smart TV. They call it the *AAMouse*.

AAMouse : [18] Tracking based on the Doppler Effect is a lot more challenging than gesture or activity recognition, since the most challenging part of the latter two is to detect the Doppler shift signatures induced by a motion. Basically, once the Doppler shift signature of a gesture or activity is known, the only thing left to do, is to apply pattern matching to identify the performed motion.

Whereas tracking faces some more challenges. Besides accurately detecting the Doppler Effect itself, the estimated frequency shift has to be used to derive the change in the position of the device. Additionally, the initial position of the device has to be estimated at first.

To track the mobile device, [18] uses a smartTV with two speakers that emit inaudible acoustic sound pulses. The device perceivces the audio signals, detects the doppler shifts and feeds the measured data back to the smartTV. There, the data is used to estimate the new position of the mouse pointer.

Refering to section 1, this means that the source of the signal remains static while the receiver moves. Contrary to Spartacus [15] or Airlink [3], introduced in section 3, the smartTV in [18] emits 17Hz pilot sound pulses instead of a continuous wave.

The mobile device samples the incoming signal at 44.1Hz. The FFT that is used by other approaches like Airlink does not allow tracking in real time. The time needed to store all the FFT samples exceeds one second what would make AAMouse impractical.

Therefore Yun et al. use an undersampling technique called *Short-term Fourier Transformation*(STFT) and a Hamming window to reduce the occuring aliasing [18]. This narrows down the frequency domain and makes it easier to detect the shifts.

To estimate the change of the device's position, equation 2 can be rearranged to get the speed of the movement. This allows to calculate the distance by integration. This leads to a lot more accuracy compared to acceleration-based tracking techniques that require a double integration to calculate the distance.

The two speakers emit different frequencies. This is necessary since Yun et al. want to realize a 2D-tracking technique. Therefore two different anchor points are needed. It is even possible to realize a 3D tracking mechanism. In order to do this, the only thing further needed is a third wave emitter [18]. Something similar to this is shown in [7], where the Doppler Effect is used to enable one-handed 3D gesture recognition. However, the authors use custom hardware that consists of one transmitter and three receiver microphones.

i) Estimating the new Device Position:
At first Yun et al. assume that the initial device position as well as the distance between the two speakers are known. The device's initial position is the combination of the relative distance between the device and speaker L, and the one between the device and speaker R.

The device's new position now is estimated by calculating the two travelled distances relative to the speakers based on their distinct Doppler shift. This is done by integrating over the speed of the motion estimated from the perceived Doppler shifts.

The updated distance from the speakers can be used as the radius of a circle around each speaker. The device's new position can be estimated by calculating the intersection of the two circles. Figure 7 shows how the new position is estimated.

The accuracy of the Doppler shift detection in part depends on the *Signal-Noise-Ratio*(SNR). The SNR varies across frequencies and can significantly influence the detection process of the Doppler shitft. To enhance robustness, Yun et al. send 1Hz sound tones at multiple pilot frequencies that are 200Hz apart and use all of them in combination to estimate the Doppler shift [18].

ii) Estimating the Distance between Speakers:
Further challenges are to esimtate the distance between the two speakers and to estimate the device's initial position. To estimate the distance between the speakers Yun et al. introduce a simple, yet effective calibration mecahnism.

To calibrate the smart TV the user holds the mobile device infront of the TV and starts at the left side of the TV and moves towards the right side. The TV emits a pilot tone that is perceived by the device. It measures the Doppler shifts and sends the data back to the TV. This process can be repeated a couple of times to increase accuracy of the calibration.

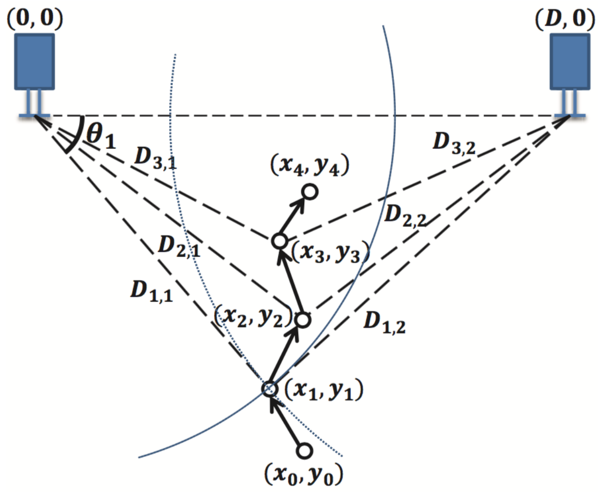

Fig. 7. Estimating new position based on the two independent Doppler shifts of the speakers [18].

iii) Estimating the initial Device Position:
A more difficult challenge is to estimate the device's initial position. To solve this, Yun et al. use a *particle filter*.

Basically, the particle filter generates many particles unfiromly distributed in an area. Each particle corresponds to a possible initial position of the device. In the following sampling interval the device movement is determined from the current particles. The particle filter filters those particles that are not consistent with the device movement.

As already mentioned in section 3.2 the device's new position is estimated by the intersection of the circles around the speakers. The result can be one, two, or no intersection points. If there is one or two intersection points, the corresponding particle is called feasible. If there is no intersection, the particle is regarded infeasible and filtered out.

Afterwards, the device's intitial position is estimated as the centroid of the remaining particles [18].

iv) Using AAMouse on a single-speaker System:
Yun et al. also introduced a technique to enable the AAMouse on a system with only a single speaker.

Generally, the more speakers are available, the higher the tracking accuracy, hence more anchor points are available. Nonetheless, there are also systems with only one speaker.

To enable the tracking of the mobile device another anchor point is needed. Yun et al. have extended their approach to use an additional WiFi wave as second anchor point. The only requirement is that the single-speaker device offers a WiFi interface.

Due to the fact that the Doppler shifts occur independently from the wave, the tracking mechanism can be used for the WiFi signal, too. To estimate the distance between the speaker and the WiFi device on the TV, they track the WiFi signal's phase change over time and use a known relationship from the phase rotation [18]. This example shows that both, sound-based and RF-based Doppler sensing can be combined to adapt varying circumstances.

v) Evaluation of AAMouse:
The evaluation of AAMouse shows that it is much better than e.g. accelerometer-based tracking system. The average error of accelerometers is 17.9cm, whereas the AAMouse can reduce this to only 1.4cm. These values deteriorate to 2.4cm if the AAMouse is used with a WiFi source as second anchor point.

Even if the speakers generate both, music and the pilot tone, the average error is not effected. This is because the maximum frequency of music hardly exceeds 15KHz [18].

AAMouse works flawless at a distance of 2-3m. Yun et al. state that if the distance between the speakers is increased, the accuracy can be increased as well. Since large TVs are getting more popular, this can lead to a high accuracy beyond a distance of 4m.

4 The Doppler Effect in Acoustics compared to Radio Frequency

As mentioned in section 1 the Doppler Effect is observed in any wave, including acoustic and RF signals. Both areas are closely related to each other. And both techniques can be used to realize similar applications.

The following table provides an overview of the basic differences between Doppler Effect based sensing with sound waves and radio frequency.

	Sound	**Radio Frequency**
Propagation Speed	slow (speed of sound)	fast (speed of light)
Bandwidth	narrow	wide
Frequency	18 - 22 KHz	2.4 GHz / 5 GHz (WiFi)
Line of sight	required	not required
Required Hardware	Microphone & Speaker	WiFi Module
Individual Strengths	- Waves are easy to generate and manipulate, - Detection process relatively simple	- Can penetrate walls, - relatively robust against noise
Individual Weaknesses	- Not robust against background noise, - Smartphone hardware quality should be improved	- Detection process complex, - Transport protocols often hide information from upper layers, - RF interfaces are not designed for sensing purposes

i) Other types of Radio Frequency besides WiFi:
Although, the previous sections focus on WiFi when it comes to radio frequency, this does not mean that this is the only form of radio frequency suitable for Doppler Effect based sensing. Zhao et al., for example, introduced *SideSwipe*, a mechanism to detect in-air gestured based on the Doppler Effect in GSM signals [19].

ii) Combining Sound and Radio Frequency:
Depending on the specific scenario, one of both types of wave might be more appropriate than the other one. Although application cases like the AAMouse [18] have shown that both techniques can also be combined to achieve a more accurate result, or to extend the applicability of a system. Another application called *WALRUS* verifies that, too. Although WALRUS is no pure mobile application, the authors introduce a localization mechanism that uses both, acoustic and radio frequency waves to localize a person within various rooms. The WiFi offers accuracy on room-level, whereas the acoustic waves are used to locate a person within the boundaries of a specific room [2].

iii) Line of Sight:
This also underlines one of the key differences between acoustic and radio frequency Doppler Effect sensing. Most radio frequency signals can easily penetrate walls, whereas acoustic waves are mostly bound by them. This can be both, an advantage as well as an disadvantage. Whole-home gesture recognition like the one introduced in [12], would be very difficult to achieve, just by using soundwaves. The previous sections show that most of the applications in the field of mobile sensing are based on the acoustic Doppler Effect. This is because

some of the key advantages compared to the radio frequency Doppler Effect are the narrower bandwith and the slower propagation speed of the wave. Both of them make it easier for acoustic applications to differentiate the relatively small Doppler shifts from the pilot frequency. Furthermore, they directly influence the achievable frequency resolution. This is encouraged by the papers [4, 12, 18]. All of them use diverse techniques to reduce or narrow the bandwith and therefore increase the accuracy of their systems.

iv) Wave Generation and Manipulation:
Another advantage of acoustic waves is that they are a lot easier to gernerate. Almost every smart device, whether it is a smartphone, smartwatch or a smartTV, has built-in speakers. Furthermore, most of them also have a built-in microphone to record the incoming sound signal. Although a lot of them do have a WiFi interface as well, it is more difficult to generate the WiFi signal. Often an additional source, e.g. an AccessPoint is necessary to fulfill this task. Furthermore, WiFi or Bluetooth interfaces are not designed for sensing purposes, neither are the protocols. Transport protocols often hide information from upper layers. Therefore it is very difficult to access the data on the lower layers.

v) Robustness:
A sound signal, however, is more likely to be disturbed by a too high amount of background noise. Although sound frequencies above 15KHz are rather uncommon in a home environment, they still exist and can significantly influence the detection process. However, WiFi signals are more vulnerable to Doppler shifts induced by random human activity. This means that e.g. tracking becomes more difficult if there are too many people around that could probably disturb the detection process.

5 Summary

This seminar paper gives an overview of the Doppler Effect in mobile sensing. The Doppler Effect can be used to sense gestures and interact with digital interfaces, to detect human activities like sleeping or working, or to track humans or devices. All the examined applications can be assigned to three major application fileds, i.e. gesture recognition, activity recognition and tracking. Although all of them do have a lot in common, they in part face different challenges. The major ones are to correctly estimate the Doppler shift, create precise Doppler shift signatures of gestures or activities, how to deal with power consumption and how to increase the accuracy of the applications.

Most of the applications are sound-based, like Spartacus [15] or Airlink [3]. This is because Doppler shifts are easier to detect in sound waves than they are in RF waves. Nevertheless, there are a couple of applications that rely on radio frequency, like WiSee [12] that enables Whole-home gesture recognition based on WiFi. Furthermore, there are also applications that combine both techniques like WALRUS or the AAMouse [18]. The combination of both techniques can be used to extend the applicability of a system or to improve its robustness.

Although most of the papers achieved an acuracy higher than 90%. It is worth mentioning that this accuracy depends heavily on environmental influences. Sound-based applications, for example, are sensitive to obstacles that are between transmitter and receiver of the signal. Whereas RF-based applications are more likely to be disturbed by too much human activity in close proximity. This means that the reliability of the various systems heaviliy depends on the environment in which a system is used. However, if the environmental influences are rather low, systems like Airlink [3], Spartacus [15] or the AAMouse [18] can have a lot of benefit for one's everyday life. Especially the latter example offers a high amount of usability and is relatively robust when it comes to background noise.

To further improve usability of DE-based sensing applications, the systems in general need to become more robust. On the one hand, this means to improve the noise handling mechanisms, and on the other hand to improve the power consumption. Both points of are crutial for the applicability of the systems. Furthermore, besides Spartacus [15] no one really addresses security concerns. However, this is very important, especially for pairing or filesharing mechanisms based on the DE.

Nonetheless, all the examined papers show that the Doppler Effect can be leveraged to fulfill a variety of tasks. Since the computational power, as well as the quality of smart devices improves gradually, the applications themselves can be further improved. Furthermore, all the different advantages of the unique systems should be combined to form an ubiquitous and robust system.

References

1. Aumi, M.T.I., Gupta, S., Goel, M., Larson, E., Patel, S.: Doplink: Using the doppler effect for multi-device interaction. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 583–586. UbiComp '13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2493432.2493515
2. Borriello, G., Liu, A., Offer, T., Palistrant, C., Sharp, R.: Walrus: Wireless acoustic location with room-level resolution using ultrasound. In: Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services. pp. 191–203. MobiSys '05, ACM, New York, NY, USA (2005), http://doi.acm.org/10.1145/1067170.1067191
3. Chen, K.Y., Ashbrook, D., Goel, M., Lee, S.H., Patel, S.: Airlink: Sharing files between multiple devices using in-air gestures. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 565–569. UbiComp '14, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2632048.2632090
4. Fu, B., Karolus, J., Grosse-Puppendahl, T., Hermann, J., Kuijper, A.: Opportunities for activity recognition using ultrasound doppler sensing on unmodified mobile phones. In: Proceedings of the 2Nd International Workshop on Sensor-based Activ-

ity Recognition and Interaction. pp. 8:1–8:10. WOAR '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2790044.2790046
5. Gupta, S., Morris, D., Patel, S., Tan, D.: Soundwave: Using the doppler effect to sense gestures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1911–1914. CHI '12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2207676.2208331
6. Harada, H., Prasad, R.: Simulation and Software Radio for Mobile Communications. Artech House, Inc., Norwood, MA, USA (2002)
7. Kalgaonkar, K., Raj, B.: One-handed gesture recognition using ultrasonic doppler sonar. In: Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. pp. 1889–1892 (April 2009)
8. Kellogg, B., Talla, V., Gollakota, S.: Bringing gesture recognition to all devices. In: Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation. pp. 303–316. NSDI'14, USENIX Association, Berkeley, CA, USA (2014), http://dl.acm.org/citation.cfm?id=2616448.2616477
9. Nandakumar, R., Gollakota, S., Watson, N.: Contactless sleep apnea detection on smartphones. GetMobile: Mobile Comp. and Comm. 19(3), 22–24 (Dec 2015), http://doi.acm.org/10.1145/2867070.2867078
10. Oppenheim, A.V., Schafer, R.W., Yuen, C.K.: Digital signal processing. Systems, Man and Cybernetics, IEEE Transactions on 8(2), 146–146 (Feb 1978)
11. Paradiso, J., Abler, C., Hsiao, K.y., Reynolds, M.: The magic carpet: Physical sensing for immersive environments. In: CHI '97 Extended Abstracts on Human Factors in Computing Systems. pp. 277–278. CHI EA '97, ACM, New York, NY, USA (1997), http://doi.acm.org/10.1145/1120212.1120391
12. Pu, Q., Gupta, S., Gollakota, S., Patel, S.: Whole-home gesture recognition using wireless signals. In: Proceedings of the 19th Annual International Conference on Mobile Computing & Networking. pp. 27–38. MobiCom '13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2500423.2500436
13. Qifan, Y., Hao, T., Xuebing, Z., Yin, L., Sanfeng, Z.: Dolphin: Ultrasonic-based gesture recognition on smartphone platform. In: Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on. pp. 1461–1468 (Dec 2014)
14. Schulz, A., Karolus, J., Janssen, F., Schweizer, I.: Accurate pollutant modeling and mapping: Applying machine learning to participatory sensing and urban topology data. In: Networked Systems (NetSys), 2015 International Conference and Workshops on. pp. 1–8 (March 2015)
15. Sun, Z., Bose, R., Zhang, P.: Spartacus: Spatially-aware interaction for mobile devices through energy-efficient audio sensing. GetMobile: Mobile Comp. and Comm. 18(4), 11–14 (Jan 2015), http://doi.acm.org/10.1145/2721914.2721918
16. Tarzia, S.P., Dick, R.P., Dinda, P.A., Memik, G.: Sonar-based measurement of user presence and attention. In: Proceedings of the 11th International Conference on Ubiquitous Computing. pp. 89–92. UbiComp '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1620545.1620559
17. Vandewalle, P., Sbaiz, L., Vandewalle, J., Vetterli, M.: How to take advantage of aliasing in bandlimited signals. In: Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on. vol. 3, pp. iii–948–51 vol.3 (May 2004)
18. Yun, S., Chen, Y.C., Qiu, L.: Turning a mobile device into a mouse in the air. In: Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services. pp. 15–29. MobiSys '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2742647.2742662

19. Zhao, C., Chen, K.Y., Aumi, M.T.I., Patel, S., Reynolds, M.S.: Sideswipe: Detecting in-air gestures around mobile devices using actual gsm signal. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. pp. 527–534. UIST '14, ACM, New York, NY, USA (2014), `http://doi.acm.org/10.1145/2642918.2647380`

Entwurfswerkzeuge für Augmented Reality Systeme

Stelian Stoev*

Betreuer: Matthias Berning†

Karlsruher Institut für Technologie (KIT)
Pervasive Computing Systems – TECO
*stelian.stoev@student.kit.edu
†berning@teco.edu

Zusammenfassung. In Rahmen dieser Seminararbeit untersuche ich, inwieweit sich die bestehenden Toolkits zur Erstellung von Augmented Reality entwickelt haben und welche neuen Anwendungsbereiche sie auf dem Markt ermöglichen. Im Folgenden werden einige Toolkits basierend auf verschiedenen Kriterien im Detail beschrieben und verglichen. Betrachtet werden Layar, Wikitude, DART, ARToolKit und Metaio.

Schlüsselwörter: augmented reality, Layar, Wikitude, Metaio, DART, ARToolKit

1 Einleitung

Augmented Reality (AR) hat zusammen mit Virtual Reality (VR) einen immer weiter wachsenden Einfluss auf unser Leben. Die Technologie bietet fast unbegrenzte Möglichkeiten und hat das Potenzial, nicht nur die Computerspielbranche stark zu beeinflussen, sondern auch unser Sozialleben, die Werbung, das Marketing, die Industrie und die Lehre können massiv davon profitieren. Das englische Wort „augment" bedeutet „erweitert" oder „vergrößert". Dabei wird die Realität durch die Anzeige zusätzlicher Informationen ergänzt. Auf die Basis der reellen Welt wird mit Hilfe von AR eine virtuelle Schicht gelegt. Abhängig von der Position und Orientierung des Nutzers können auf dieser Schicht virtuelle Objekte jeglicher Art platziert werden. Diese virtuellen Objekte werden entweder mit einem bestimmten Gegenstand oder mit einer bestimmten Position in der reellen Welt verknüpft.

Das menschliche Auge benötigt ein Gerät in beliebiger Form, um die virtuellen Inhalte zu sehen. Handys, interaktive Brillen wie z.B. Google Glass und Datenhelme (Head-mounted Display) sind nur einige Beispiele aus diesem Gebiet. Während die Datenhelme zur spezifischen Hardware gehören, die die meisten Menschen nicht immer dabei haben und die Brillen, die Informationen direkt auf das Auge projizieren, nach wie vor soziale Probleme verursachen, sind Handys sehr weit verbreitet und begleiten fast jeden Menschen im Alltag. Die Technologie, die die Augmented Reality ermöglicht, beruht auf einfachen Sensoren wie

Kameras oder Accelerometer und dem Bildschirm. Diese Tatsache verstärkt die Möglichkeit zur Nutzung des Handys als AR Hardware, da moderne Handys sowohl alle nötigen Sensoren aufweisen, als auch die nötige Rechenleistung besitzen. Aus diesen Gründen kommen mehr und mehr Entwicklungsumgebungen für mobile AR-Anwendungen auf den Markt.

Je mehr Interesse das Thema Augmented Reality auf sich zieht, desto mehr Sinn macht es, die Entwicklungsmöglichkeiten zu erweitern. Insbesondere die Zahl der an der Technologie interessierten Designer hat in den letzten Jahren stark zugenommen, und das in einem Themengebiet, in dem sich bisher beinahe ausschließlich Anwender mit guten Programmierkenntnissen aufgehalten haben [6].

Abhängig vom Anwendungsszenario und den gestellten Anforderungen ist für jedes Projekt eine passende Software zu bestimmen. Das erste Kriterium bei der Auswahl betrifft in den meisten Fällen die Betriebssysteme und Plattformen, die vom Produkt unterstützt werden sollen. Da die verbreiteten mobilen Plattformen fast von jedem Softwareprodukt unterstützt werden, stellen solche Betriebssysteme eine gute initiale Basis für die Auswahl dar. Es gibt keine Software, die alle Techniken zur Erstellung von Augmented Reality unterstützt. Ein gut definierter Plan und eine konsistente Sicht auf das Projekt mit einer darausfolgenden klaren Sicht auf die Anforderungen an eine Entwicklungsumgebung sind absolut kritisch für den Erfolg eines AR-Projektes. So kann z.B. eine geschickte Kombination von Tracking Möglichkeiten und unterstützten Inhalt eine Anwendung sehr hervorheben. Wie bei jeder Software, spielen auch die Dokumentation und die Kosten eine entscheidende Rolle. Im Rahmen dieser Arbeit werden die einzelnen Werkzeuge bezüglich der folgenden Kategorien verglichen:

- Lizenz
- Unterstützte Plattformen
- Dokumentation
- Zielgruppe
- Inhalt
- Tracking-Möglichkeiten

2 Terminologie

Ein grundlegender Aspekt von den Entwurfswerkzeugen ist die Art und Weise, wie die Umgebung erfasst werden kann. Im Wesentlichen wird versucht, die Daten der Sensoren geschickt zu interpretieren, so z.B. bei der Suche nach bestimmten vordefinierten Mustern in den Kameradaten. Diese Methode wird bei markerbasiertem Tracking benutzt. Die Marker können wiederum ganz verschiedene Formen besitzen, von QR- und Bar-Codes bis hin zu einem ganz normalen Bild. Wichtig dabei ist, dass sie hart in die Anwendung hineinegecodet wurden. Auf der anderen Seite gibt es die nicht-markerbasierten Varianten des Trackings, die sich auf das Erkennen von größeren Mustern, wie zusammenhängenden Farbflächen, oder ganzer Objekte, wie z.B. einem Hund, konzentrieren. Um bessere

Ergebnisse zu erziehen, werden hier meist Daten verschiedener Sensoren gleichzeitig ausgewertet. Ein standort-bezogenes Tracking nutzt die geographischen Koordinaten des Anwenders (GPS), die Ausrichtung des Geräts (Kompass) und sein Accelerometer, um eine möglichst präzise Positionierung des virtuellen Objekts relativ zur reellen Welt zu erreichen.

Die durch zusätzlichen Inhalt erweiterte Sicht wird durch optische bzw. videobasierte See-Through-Technik erreicht. Bei dem Video See-Through werden virtuelle Objekte auf einem laufenden Videostream (meist in Echtzeit) platziert. Im Unterschied dazu benötigt man beim Optical See-Through eine beliebige semitransparente Oberfläche, die vor den Augen des Benutzers angebracht wird, wie auf Abb. 1 dargestellt. Auf diese Oberfläche werden gerenderte Inhalte projiziert. Es gibt sowohl Stereo-Oberflächen, mit einer Oberfläche pro Auge, welche die Möglichkeit bieten, stereoskopische Inhalte darzustellen, als auch monokulare Oberflächen [13].

Abb. 1. Beispiel für Optical See-Through [7].

3 Beschreibung der einzelnen Werkzeuge

Für jedes Projekt spielen die Anforderungen und eine gute Planung eine wichtige Rolle. Kommt Augmented Reality ins Spiel, müssen für ein Projekt besondere Überlegungen angestellt werden, da kein universelles Tool für die Erstellung von Augmented-Reality-Inhalten existiert. Die angebotenen Möglichkeiten sind sehr unterschiedlich und sie überlappen sich kaum. Dieser Abschnitt dient dazu, mehr über die einzelnen Werkzeugen zu erfahren. Ein Vergleich findet sich im nächsten Kapitel.

Der Vollständigkeit wegen müsste das Unity Augmented Reality Plugin Teil dieses Abschnittes sein. Unity ist eine sehr beliebte Entwicklungsumgebung für 3D-Spiele. Genau solche Spiele können das volle Potenzial von Augmented Reality ausschöpfen, da in dieser Umgebung sowohl die einzelnen Szenen im 3D-Editor erstellt, als auch die Logik mittels eigener Skripte und der Ton vordefiniert werden muss. Sie beinhaltet desweiteren auch eine Grafik-Engine und eine PhysX-Engine, die die physikalischen Kräfte auf ein Objekt darstellen kann. Dementsprechend ist der Einbau eines AR-Plugins in diese Umgebung eine komplexe Aufgabe. Aus Komplexitätsgründen wäre es ineffektiv,dieses Plugin in der vorliegenden Arbeit zu betrachten. Das Unity Augmented Reality Plugin basiert jedoch auf ARToolKit, welches später näher betrachtet wird.

3.1 Layar

Layar wurde im Sommer 2009 gegründet und war einer der ersten mobilen AR-Browser auf dem Markt. Die Layar Plattform beinhaltet ein SDK und ermöglicht es auf sehr einfache Art und Weise, durch den Browser eine AR-Anwendung zu erstellen. Das SDK stellt eine statische Bibliothek mit der Layar-Vision-Funktionalität und der Geo-Lokalisierung-Funktionalität direkt in einer iOS- und/oder Android-Anwendung zur Verfügung. Ein erheblicher Vorteil besteht darin, den interaktiven Inhalt unter eigener Handelsmarke in den eigenen Anwendungen einbauen zu können.

Obwohl der Browser leicht zur verwenden ist, verwendet er eine relativ komplexe Architektur. Die Layar Plattform [4] beinhaltet 5 Komponenten:

- **The Layar Reality Browser:** Klient, der auf dem mobilen Gerät des Benutzers läuft.
- **The Layar Server:** Server, der die Schnittstellen zum Reality Browser, the Layar Publishing site und den externen Layar Service Providers zur Verfügung stellt.
- **The Layar Publishing Website:** Eine Webseite, in welcher die Entwickler neue Schichten von Inhalten registrieren, sowie vorhandene Schichten und Konten verwalten können.
- **The Layar Service Providers:** Dienstleistungen, die durch Dritt-Entwickler erstellt werden.
- **The Layer Content Sources:** Stellt den Inhalt zur Verfügung, der dann im Layar Reality Browser dargestellt wird. Die Grenze zwischen Service Providers und Content Sources ist in manchen Fällen nicht ganz eindeutig, jedoch sollten sie dennoch als unterschiedliche logische Einheiten dargestellt werden.

Der typische Ablauf der Kommunikation beginnt mit dem Start des Layar Reality Browsers auf einem mobilen Gerät. Damit verschickt der Klient eine Anfrage an den Server. Darauf basierend kann der Server eine Liste mit den für diesen Benutzer vorhandenen Schichten zusammenstellen und diese dem Benutzer präsentieren. Nachdem der Benutzer sich für eine Schicht entschieden hat, werden alle Inhalte, die dieser Schicht angehören, über den Layer Service Provider zur Verfügung gestellt [4].

Die SDK ist nicht kostenlos, jedoch gibt es eine 30-tägige Probezeit, während der man ohne jegliche Begrenzungen Inhalte erstellen kann. Nach dieser Zeit steht nur noch die kostenpflichtige Variante zur Verfügung. Der Probeschlüssel

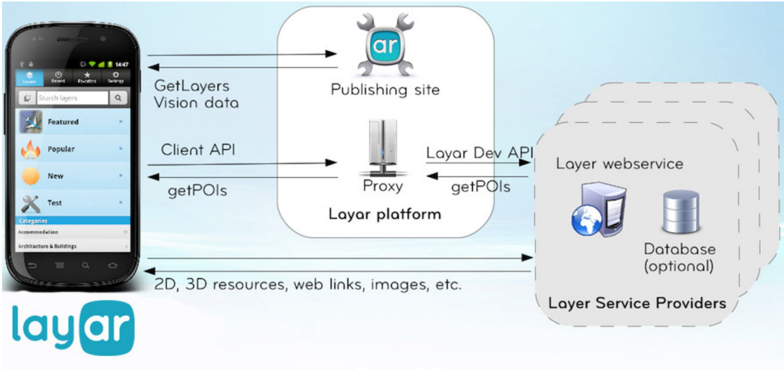

Abb. 2. Layar Platform Architektur [4].

läuft nach 30 Tage ab, entscheidet man sich nicht für den Kauf werden alle bis dahin erstellen Inhalte gelöscht.

Layar setzt auf nicht-markerbasiertes Tracking: QR- oder Barcodes werden nicht als Marker verwendet, was natürlich auch Multimarker Tracking ausschließt. Die Vielfalt der Sensordaten werden als ganzheitliches Bild betrachtet und darin wird versucht, Muster zu erkennen. Eine Gesichtserkennung wird jedoch immer noch nicht unterstützt.

Beim Einsatz der SDK sind gute Programmierkenntnisse in der gewünschten Plattform (sowohl bei Android als auch bei iOS) erforderlich. Um die Lernkurve möglichst steil zu halten, steht neben einer ausführlichen Dokumentation auch eine Entwickler-Webseite zur Verfügung. Des weiteren ist umfangreiche Unterstützung durch die Community im Internet zu finden.

3.2 Wikitude

Wikitude ist eine Firma, die die AR-Technologie seit 2008 entwickelt. Sie wurde in Salzburg gegründet, wobei sie bis 2012 die standort-bezogene AR-App Wikitude World Browser anbot. Das bedeutet, dass die Positionierung des virtuellen Objektes über die Position des Anwenders (GPS), seine Ausrichtung (Kompass) und seinem Accelerometer berechnet wird. Im Jahr 2012 veröffentlichte die Firma das Wikitude SDK. Dieses SDK unterstützt nicht nur standort-bezogene AR, sondern auch die Bilderkennung in 2D und in 3D sowie die Bildverfolgung in 2D und 3D [8]. Besonders beliebt ist sie, weil sie sich sehr leicht durch vorhandene Plugins für Unity, Apache Cordova, Titanium Mobile und Xamarin in eine bestehende Entwicklungsumgebung integrieren lässt [8]. Daher ist das Hinzufügen von Augmented-Text, -Bild, -Ton, -Video und -3D-Modellen sowohl statisch als auch animiert gut möglich und auch ganze HTML Widgets können in einer eigenen Anwendung sehr schnell und leicht realisiert werden. Auf Abb. 3 ist die Oberfläche der webbasierten Version von Wikitude dargestellt.

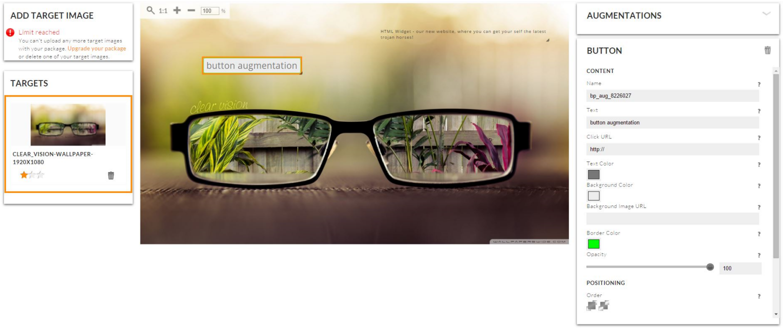

Abb. 3. Oberfläche von Wikitude Studio

Wikitude setzt ebenfalls stark auf nicht-markerbasiertes Tracking, indem das SDK Möglichkeiten sowohl für 2D-Erkennung und -Verfolgung von Bildern, als auch eine 3D-Variante im Betastadium anbietet. Trotzdem bietet das Werkzeug auch optional QR- und Barcodes für das Tracking an.

Die SDK stellt fast keine Begrenzungen, was die unterstützten Plattformen angeht. Unterstützt werden iOS, Android und Smart Glasses. Eine Voraussetzung für eine gut geschriebene Software ist eine ausführliche und gut strukturierte Dokumentation [15]. Wikitude hat eine führende Position in diesem Kriterium.

3.3 Metaio

Metaio ist eine deutsche Firma, die im Jahr 2003 in München gegründet wurde. Zusammen mit DART und ARToolKit hat Metaio die Entwicklung von Grundlagen der AR vorangetrieben. Sie hat über 500 Projekte erfolgreich abgeschlossen und mittlerweile 10 Jahre Erfahrung gesammelt. Ein weiteres Zeichen dafür, dass die reine AR-Firma erfolgreich war, ist auch der Aufkauf durch Apple im Jahr 2015. Seitdem werden das SDK und das Creator-Programm nicht mehr verkauft.

Das Metaio-SDK ist modular aufgebaut und beinhaltet mehrere Komponenten: die Renderingkomponente, die Aufnahmekomponente, die Sensorkomponente, die Verfolgungskomponente und die Metaio-SDK-Schnittstelle. Sie stellt die Interaktion zwischen der Implementierung der Anwendung und den anderen vier Komponenten zur Verfügung. Dadurch werden die Implementierungsdetails gekapselt und der Benutzer muss sich um das Aufnehmen, Rendern usw. nicht kümmern. Die wichtigsten Funktionalitäten sind durch die SDK APIs realisiert worden, welche eine leichte Implementierung von AR-Anwendungen gewährleisten. Auf Abb. 4 ist die Aufgabe der Metaio SDK als Schnittstelle zwischen dem Betriebsystem und den Anwendungen dargestellt.

Das Metaio SDK ist mit Android, iOS, Unity3D und Windows kompatibel. Dafür ist die Schicht „Plattform-specific programming interface“ zuständig.

Der Metaio Creator hat den Vorteil, dass er wesentlich weniger kostet und dass er auch für Leute ohne Programmiererfahrung geeignet ist. Es lassen sich

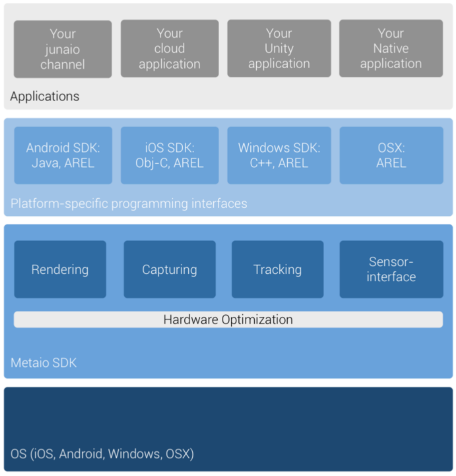

Abb. 4. Metaio SDK Overview [5].

mittels Drag & Drop interaktive AR-Anwendungen erstellen. Der Nachteil besteht drin, dass es nicht mehr möglich ist geobasiertes AR zu nutzen und dass er weder Unity noch die native Entwicklung unterstützt [5].

3.4 DART

DART (The Designer‘s Augmented Reality Tool-Kit) wurde 2003 von Blair MacIntyre, Meribeth Gandy, Steven Dow und Jay David Bolter am Georgia‘s Institute of Technology vorgestellt. DART ist ein Plug-In für die populäre Software Macromedia Director. Die Entwickler von DART haben eng mit Designern zusammengearbeitet, um deren größten Probleme im Zusammenhang bei der AR-Entwicklung zu finden. Das Plugin ist dafür entwickelt, diese zu lösen und soll die direkte und effektive Arbeit mit AR gewährleisten.

Macromedia Director ist eine sehr beliebte Umgebung mit zahlreichen Funktionen, unter anderem auch einer Skriptsprache namens Lingo. Als Plug-In für diese Umgebung kann DART sehr großen Nutzen bieten. Die meisten AR-Tools sind von Softwareentwicklern für Softwareentwickler gemacht. Designer haben jedoch im Allgemeinen die notwendige Erfahrung mit dem Programmieren nicht, um diese komplexeren Tools zu benutzen [9]. Es gibt viele Bücher und Informationen darüber, wie man mit der Umgebung arbeitet. Zusätzlich gibt es andere Plug-Ins und eine offizielle Dokumentation. Damit sind Designer gut versorgt, was Information bezüglich der Nutzung des Macromedia Directors angeht. Sie können auch davon profitieren, dass das der Macromedia Director kein unbekanntes Tool ist. Die meisten Designer sind mit ihm schon vertraut und können weiterhin dieselbe Software zur Erstellung von AR benutzen.

Designer sind durch die AR-Tools meistens gezwungen, sowohl auf eine höheren Ebene zu arbeiten (das Design zu entwerfen), als auch auf einer sehr

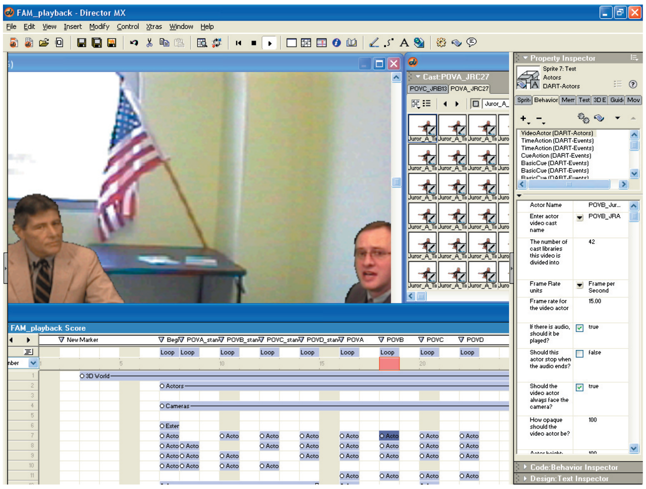

Abb. 5. Screenshot von der DART-Oberfläche [3].

niedrigen Ebene zu skripten. Das vermischt die Konzepte und macht das Austauschen von einzelnen Technologien schwieriger. DART ist modular aufgebaut und seine Entwickler haben diesen Aspekt bei der Entwicklung berücksichtigt.

3.5 ARToolKit

Das ARToolKit ist die wohl am meisten verwendete Software für Augmented Reality [14]. Diese Bibliothek wurde 1999 von Hirokazu Kato von Nara Institute of Science and Technology entwickelt. Sie wurde dann von der Universität Washington HIT Lab als Open-Source Projekt verfügbar gemacht. Später wurde auch die erste mobile Bibliothek für AR mit Symbian im Jahr 2005 vorgestellt, welche 2008 für iOS und 2010 für Android kompatibel gemacht wurde. Auf Grund der Tatsache, dass das ARToolKit Open-Source ist, ist der dazugehörige Quelltext in Github [2] zu finden. Es ist aber auch möglich eine kompilierte SDK direkt von der offiziellen Webseite herunterzuladen [10]. 2015 wurde ARToolKit von der Firma DARQI übernommen und neu veröffentlicht, ist aber dennoch Open-Source geblieben.

ARToolKit unterstützt Windows, Linux, Mac OS X, iOS und Android und bietet die gleiche Funktionalität auf alle Plattformen, wobei die Leistung der verschiedenen Hardwarekomponenten sehr großen Einfluss auf die gesamte Leistung des System haben kann. Die Software ist in C/C++ geschrieben und kann somit sehr leicht für andere, auch experimentelle, Plattformen erweitert werden.

Eine der größten Schwierigkeiten der AR-Entwicklung ist die genaue Kalkulation des Blinkwinkels des Benutzers in Echtzeit, damit die virtuellen Objekte genau auf die reellen Objekte ausgerichtet werden können. ARToolKit benutzt eine Bilderkennungs-Technologie, um die Kameraposition und die -orientierung relativ zu quadratischen Objekte oder glatte Oberflächen zu berechnen. So wird

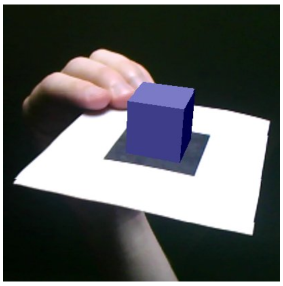

Abb. 6. Mit ARToolKit erstelltem AR [1].

ermöglicht, virtuelle Objekte im Bild zu positionieren. ARToolKit unterstützt zurzeit die klassischen quadratischen Marker, 2D-Barcode und Multimarker. Außerdem unterstützt ARToolKit jede beliebige Kombination dieser Marker. Die schnelle und präzise Erkennung dieser Marker hat eine rasche Entwicklung von tausenden von Anwendungen ermöglicht.

ARToolKit unterstützt sowohl Video See-Through als auch Optical See-Through. Ein sehr großer Vorteil des ARToolKits gegenüber fast alle anderen Kits ist dessen umfassende Dokumentation, die durch ein Community-Forum ergänzt wird. Im Internet lässt sich dadurch schneller Unterstützung finden.

4 Vergleich

Beim Entwurf einer Anwendung definieren die Entwickler möglichst viele Anforderungen, die für ein gutes Ergebnis nötig sind. Sobald diese zur Verfügung stehen, wird entschieden, mit welchen Werkzeugen am besten gearbeitet wird, um das Projekt zu realisieren. Neben naheliegenden Kriterien wie zum Beispiel, ob markerbasiertes Tracking unterstützt werden soll, kommen auch viele andere, weniger technische Punkte hinzu, die erheblichen Einfluss auf die endgültige Entscheidung, welches Tool eingesetzt wird, haben können. Ein gutes Beispiel hierfür wäre eine gut strukturierte Dokumentation. Je mehr Zeit ein Entwickler braucht, um auftretende Probleme aufzulösen, desto teurer wird das Projekt.

4.1 Lizenz

Grundsätzlich ist in dieser Kategorie zwischen kostenlos und kostenpflichtig zu unterscheiden. Es gibt aber auch große Unterschiede, was die kostenpflichtigen Modelle betrifft. Zum Einen könnte man ein Subscription-Modell abonnieren, in dem eine Lizenz für eine bestimmte Zeit wie zum Beispiel einen Monat oder ein Jahr gekauft wird. Zum Anderen könnte man sich für ein Projekt pro Plattform eine Lizenz kaufen. Oft besteht auch die Möglichkeit, einmalig eine Gebühr zu zahlen. Beim ARToolKit handelt es sich um eine komplett frei verfügbare Software. Sowohl das SDK ist kostenlos zu erhalten, als auch der gesamte Quelltext im GitHub zu finden. Metaio verfolgte ein Subscription Model, jedoch sind das SDK und der Creator seit dem 15. Dezember 2015 nicht mehr im Angebot. Dennoch wird die Firma die laufenden Verträge nicht kündigen.

Lizenz	Layar	Wikitude	Metaio	DART	ARToolKit
Kostenlos	-	x	-	x	x
Open Source	-	-	-	-	x
Kommerziell	x	x	-	-	-

Tabelle 1. Vergleich Lizenzen

4.2 Unterstützte Plattformen

Anhand der Tabelle ist sehr deutlich zu erkennen, dass die Entscheidung, mit welchen Tool man arbeiten möchte, immer noch sehr stark davon abhängt, für welche Plattformen das Projekt zu realisieren ist. Da iOS und Android sehr verbreitete AR-Plattformen sind, unterstützen diese auch die meisten Toolkits. Möchte man aber Projekte für andere Betriebssysteme entwickeln, ist die Auswahl stark begrenzt.

Plattformen	Layar	Wikitude	Metaio	DART	ARToolKit
iOS	x	x	x	-	x
Android	x	x	x	-	x
Windows	-	-	x	x	x
Linux	-	-	-	-	x
OS X	-	-	x	x	x
Smart Glasses	-	x	-	-	x

Tabelle 2. Übersicht der unterstützten Plattformen.

4.3 Dokumentation

Wie von großen Produkten zu erwarten, bringen alle eine ausführliche Dokumentation mit sich. Da DART und ARToolKit schon längere Zeit existieren, gibt es für diese auch eine gute Unterstützung von Entwicklern für Entwickler (Community).

4.4 Zielgruppe

Die Zielgruppen wurden in drei Kategorien unterteilt. Grund dafür ist, dass das Angebot eines ToolKits sehr unterschiedlich sein kann. Wikitude hat zum

Dokumentation	Layar	Wikitude	Metaio	DART	ARToolKit
Eigene	x	x	x	x	x
Community	-	-	-	x	x

Tabelle 3. Dokumentation

Beispiel eine webbasierte Version, mit der man ohne jegliche Programmierkenntnisse eine eigene AR-Anwendung erstellen kann. Wenn man sich gut auskennt, kann man sehr gute Ergebnisse mit dieser Webversion erzielen. Zusätzlich ist ein SDK vorhanden, das es einem erlaubt, native Anwendungen mit AR-Inhalten zu versehen. Metaio SDK ist ebenso dafür da, native Anwendungen zu erstellen, im Gegensatz zum Metaio Creator, der eher Drag & Drop nutzt und für Nichtprogrammierer geeignet ist.

Zielgruppe	Layar	Wikitude	Metaio	DART	ARToolKit
Nichtprogrammierer	x	x	x	-	-
wenig Programmierkenntnisse	-	x	-	x	-
Programmierer	x	x	x	x	x

Tabelle 4. Zielgruppen.

4.5 Inhalt

Um sich für eine Entwicklungsumgebung zu entscheiden, ist neben der gewünschten Plattform auch der unterstützte Inhalt ein wichtiger Entscheidungsfaktor. Durch die Tabelle 4.5 bekommt man eine gute Übersicht, was für Möglichkeiten welche Toolkits momentan bieten. Wikitude und ARToolKit zeichnen sich durch ihre Vielfalt an möglichen Inhalten aus, da sie neben Text, Ton, Video und Bild auch 3D-Modelle unterstützen. Obwohl Layar auch viele Freiheiten bietet, kann die Plattform immer noch nicht mit 3D-Modellen umgehen. Metaio und DART bieten ähnliche Inhalte zur Darstellung an.

Inhalt	Layar	Wikitude	Metaio	DART	ARToolKit
Text	x	x	x	-	x
Bild	x	x	x	x	x
Video	x	x	x	x	x
HTML Widgets	x	x	-	-	x
Ton	x	x	-	-	x
animierte Bilder	x	x	x	x	x
3D-Modelle	-	x	x	x	x

Tabelle 5. Inhalt

4.6 Tracking Möglichkeiten

Vor dem Schreiben einer AR-Anwendung muss eine wichtige konzeptionelle Entscheidung getroffen werden. Diese hängt mit der Art der Erkennung der Umwelt zusammen. In diesem Aspekt sind die beschriebenen Werkzeuge sehr unterschiedlich. Während die Möglichkeiten von Layar und Wikitude sich fast überlappen, entwickelt sich Metaio weiter und zeichnet sich mit einer einzigartigen Eigenschaft aus - Face-Tracking. DART kann QR- und Barcode-Tracking durchführen, wie fast alle anderen auch, es unterstützt aber auch Online-Video-Streams, was sonst nur noch von ARToolKit angeboten wird. ARToolKit erzielt mit seinem Multimarker-Tracking einen höheren Nutzen.

Tracking	**Layar**	**Wikitude**	**Metaio**	**DART**	**ARToolKit**
2D Markerless Image Recognition	x	x	x	-	-
2D Markerless Image Tracking	x	x	x	-	-
Geolocation Tracking	x	x	x	-	-
QR und Barcode Tracking	-	x	x	x	x
Face Tracking	-	-	x	-	-
Multimarker Tracking	-	-	-	-	x
Tracker von online video streams	-	-	-	x	x

Tabelle 6. Tracking Vergleich

5 Fazit

Augmented Reality ist ein sehr spannendes und innovatives Themenfeld, in dem viele große Firmen wie Google, Apple, Tesla, Facebook und Microsoft aktiv forschen [11,12]. Man findet immer mehr Augmented-Reality-Anwendungen, die immer interaktiver mit der Welt agieren. Diese kommen meist kleinen und mittelständigen Unternehmen und sind nicht im Portfolio der großen IT-Firmen. Facebook hat mit dem Erwerben von Oculus Rift und Apple mit Metaio klare Zeichen dafür gesetzt, dass diese Firmen stark in dieser Richtung arbeiten. Dadurch wird die Konkurrenz in der Branche viel größer und das verspricht ein bedeutendes Wachstum für alle Entwicklungsumgebungen für Augmented Reality und Virtual Reality. Aufgrund dessen, dass die Werkzeuge für Augmented Reality derzeit sehr unterschiedliche Funktionalitäten auf verschiedenen Plattformen anbieten, muss man sehr präzise Anforderungen für sein Projekt definieren, um sich für das passende Werkzeug entscheiden zu können.

Literatur

1. Artoolkit erzeugte graphik, http://america.pink/images/2/0/5/3/5/4/en/3-artoolkit.jpg
2. Artoolkit github website, https://github.com/artoolkit
3. Dart screenshot from interface, http://i0.wp.com/ael.gatech.edu/lab/files/2007/07/DART-image.jpg
4. Layar platform website, https://www.layar.com/documentation/browser/layar-platform-overview/
5. Metaio website from the internet archive, https://web.archive.org/web/20150504224834/http://www.metaio.com/products/creator/compare-products/
6. New york times partners with google for virtual reality films - fortune, http://fortune.com/2015/10/20/nyt-google-cardboard/
7. Optical see-through beispiel, https://www.google.com/imgres?imgurl=http%3A%2F%2Frenmar.ru%2Fshop%2Fimg%2Fmotoshlem-boss-520-61740-small.jpg&imgrefurl=http%3A%2F%2Frenmar.ru%2Fshop%2Ftermobele%2Fmotoshlem-boss-520&docid=QXdCl3FXsXzCyM&tbnid=rOMZwMzvQMI4DM%3A&w=1000&h=731&hl=en&ei=Ju7yVtSwFsqR6AT42p2wDg
8. Wikitude sdk feature-comparision, http://www.wikitude.com/products/wikitude-sdk-features/
9. UIST 2004: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology, Vancouver, October 24-27, 2004, Santa Fe, NM, CHI letters, vol. vol. 6, issue 2. Association for Computing Machinery, New York, N.Y. (2004)
10. ARToolKit official website.: http://artoolkit.org/
11. Benner, K., Wingfield, N.: Apple sets its sights on virtual reality, http://www.nytimes.com/2016/01/30/technology/apple-sets-its-sights-on-virtual-reality.html
12. Lomas, N.: What google sees in augmented reality, http://techcrunch.com/2015/12/12/lifelike-adverts/
13. Rolland, C.: A comparison of optical and video see-through head-mounted displays. Information Systems 19(3), 197 (1994)
14. Wagner, M.: Building wide-area applications with the artoolkit. In: First IEEE International Augmented Reality Toolkit Workshop. p. 7 (29 Sept 2002)
15. Wikitude Documentation: (2016), http://www.wikitude.com/documentation/

What smartphone interaction reveals about your personality, sociability, and sentiment

Max Vogler*

Advisor: Anja Bachmann†

Karlsruhe Institute of Technology (KIT)
Pervasive Computing Systems – TECO
*max.vogler@student.kit.edu
†bachmann@teco.edu

Abstract. During the last two decades, the increasing technological possibilities of mobile computing devices and widespread global distribution of smartphones enabled researches to automatically detect reliable and precise information about the user's personality characteristics. This work summarizes a variety of studies, displaying how these characteristics correlate with smartphone usage patterns and trying to explain the reasons for them. While the analysis contains a variety of apps, it focuses on messaging, telephony, and social media usage. Before the detailed analysis, this work explains psychological and technical fundamentals and compares historical and current mobile device usage. The last chapters contain critical discussions, limitations, and possible future work.

Keywords: Big Five, Personality, Sentiment, Mood, Smartphone, Messaging, Telephony, Apps, Social Media, Facebook, WhatsApp, Contacts

1 Introduction

When the first SMS – simply containing "Merry Christmas" – was sent in December 1992, who would have thought that this was one of the first steps to completely revolutionize the global computing and communications industry. About two decades later, a staggering 6.1 trillion SMS were sent in the year 2010. This was not an isolated case: Among many other services, Facebook, a popular social network, experienced a similar explosive growth. Its monthly active member count has grown to 1.4 billion (one fifth of the world population) since its foundation in 2004.

Strongly correlated to this is the growing use of smartphones: Starting with 122 million smartphones sold worldwide in 2007, when the first iPhone was released, this number increased tenfold in only 7 years, to 1.2 billion in 2014[15]. Since then, smartphone penetration has increased steadily, ranging from 55% in the USA and Germany up to 83% in Sweden in 2015 [13,10,19]. Deprecating other computing, communication, and media devices, a 2015 study reported a daily usage mean of 2.7 hours per person. More than one third even use their smartphone directly before going to sleep or after waking up, making it truly an essential part of everyday life. [33]

Based on collected smartphone data including the usage of apps, games messaging, social media, and more, South Korean researchers found strong correlations between sociographics and smartphone use: Young, educated, and wealthy subjects used their smartphone generally more, females used e-commerce and relational apps more than males and other findings indicate a strong connection between demographics, personality, and smartphone use. [22]

Furthermore, current developments are so intense, researchers even coined the term psychoinformatics, a new field of study combining psychology and computer science. Due to the vast global spread of smartphones, the steadily increasing usage, and the ongoing innovation of technical possibilities, they predict that experience sampling via mobile applications will assume a central role in psychological data collection and lay the ground for truly new computing paradigms. [42]

Researchers, psychologists, psychiatrists, and smartphone users themselves may be able to detect personality traits and sentiment more quickly and reliably and find out previous unknown correlations. Software developers may be able to write affective apps which react better and more sensible to a user's current emotional state. Finally, the whole marketing industry may possibly use the new-found emotional information for better targeted ads, services, and products. Especially recommender systems, e.g. suggested movies on Netflix or products you may like on Amazon, could benefit greatly. Instead of relying only on superficial information, they could provide suggestions based on your emotional state and personality.

This work is structured in five main parts: After this introduction, in Chapter 2, the fundamentals of personality, sociability, and sentiment are broadly covered. Standardized assessment methods are explained and historical mobile phone use and current smartphone use described and compared.

In Chapter 3, typical usage patterns of messaging, telephony, social media, and other apps are analyzed and correlations to personality characteristics explained. Furthermore, a state of the art machine learning approach is shown. Chapters 4 and 5 contain critical discussion, conclusion, and future work.

2 Fundamentals

2.1 Personality

Personality is defined as stable set of characteristics that describe how one's thoughts, feelings, and actions are common or different to others [25]. Personality traits are related to school attendance, gambling behavior, leadership behavior, job performance, participation in sports, and more [24]. They have even been shown to predict job satisfaction, professional, and romantic relationship success and many more parts of life [16].

Studies have shown that all personality measures can be categorized under the same five basic personality traits [21]: extraversion, agreeableness, conscientiousness, neuroticism, openness – called the **Big Five personality traits**. Consequently, this work uses Goldberg's Five Factor Model [18], the most widespread and acknowledged model that describes the human personality based on those five traits [37]. The model has been successfully validated across cultures using a diverse set of methods and has been proven to stay stable over time [21].

Trait	Typically described as ...
Extraversion	active, assertive, energetic, enthusiastic, outgoing, talkative
Agreeableness	appreciative, forgiving, generous, kind, sympathetic
Conscientiousness	efficient, organized, planful, reliable, responsible, thorough
Neuroticism	anxious, self-pitying, tense, touchy, unstable, worrying
Openness	artistic, curious, imaginative, insightful, original, widely interested

Table 1. Typical descriptions of the Big Five traits [9]

Extraversion, the first of the five personality traits, describes one's desire to be with other people [23]. Extraverted people are characterized to be active, assertive, energetic, enthusiastic, outgoing, and talkative [30]. This trait additionally correlates with warmth, gregariousness, excitement seeking, and general positive emotions [30]. High extraversion can be perceived as attention-seeking and domineering. On the other hand, low extraversion may appear as self-absorbed [23].

Agreeableness describes how friendly one is perceived. Agreeable persons appear appreciative, forgiving, generous, kind, sympathetic, and trusting [30]. This trait additionally correlates with straightforwardness, altruism, compliance, and modesty [30]. While high agreeableness can be perceived as naive and submissive, low agreeableness appears argumentative and untrustworthy [23].

Conscientiousness refers to one's work ethic, orderliness and thoroughness [11]. Conscientious people are described as efficient, organized, planful, reliable, responsive, and thorough [30]. This trait additionally correlates with competence, order, dutifulness, achievement striving, self-discipline, and deliberation. High conscientiousness may appear as stubborn and obsessive, while low conscientiousness can appear sloppy and unreliable [23].

Neuroticism refers to one's lack of emotional control. Neurotic people are described as anxious, self-pitying, tense, touchy, unstable, and worrying. This trait additionally correlates with hostility, depression, impulsiveness and vulnerability. Further, it is negatively correlated to self-consciousness [30]. Low levels suggest a good level of control over emotions. High neuroticism can indicate a need for stability. [11]

High **openness to experience** indicates that one is broadly interested and seeks novelty. Open people seem to be artistic, curious, imaginative, insightful, and original [30]. This trait additionally correlates with a high amount of fantasy [30]. People with very high openness may appear unpredictable and unfocused. Low openness on the other hand can indicate a closed-minded, more dogmatic person preferring familiarity and convention. [20]

Narcissism - even though not captured in the Big Five - is another personality characteristic which is of great importance for this work. Narcissistic individuals tend to possess inflated and unrealistic positive views of themselves. This correlates with strong self-focus, feelings of entitlement, and lack of regard for others [26]. While highly narcissistic people focus on benefits for themselves, without regarding potential harm for others, moderate narcissism can result in taking particular care of the physical appearance, frequent boasting about accomplishments, and other forms of attention seeking [29]. Three typical subtraits of narcissism are need for leadership/authority, grandiose exhibitionism, and entitlement/exploitativeness [29].

Self-esteem describes a person's self-evaluation of their worth. This self-evaluation may occur unconscious and intrinsic (implicit self-esteem) or conscious and reflective (explicit self-esteem). Following a vital need to keep or increase the self-esteem, people strive for reinforcing positive views about themselves. [31].

2.2 Sociability

The **egocentric network density** measures the interconnections between common friends of one person. A high network density means that a high percentage of one's friends know each other and have friendships between each other. In a sparse network (with a low density) one's friends do not know each other. This can be the result if a person has a variety of different interests, hobbies, or activities. [17]

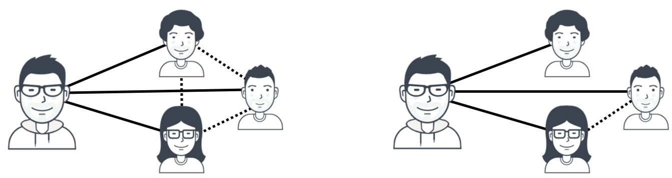

Fig. 1. A dense (left) and sparse (right) network of friends.

2.3 Sentiment

Mood is a mental state, influenced by a variety of intrinsic and external factors. In contrast to **emotions**, which typically do not last very long, different moods can last hours or even days. Mood does not only influence our behavior but is also a very important social signal to guide people on how to interact with each other [27].

Loneliness is a negative experience arising from the absence of satisfying social relationships. Even though loneliness because of a missing romantic partner is a common form, loneliness because of family and social issues are important forms as well. [36].

Social anxiety is the fear of being unable to make a positive impression on others, especially in encounters with strangers [36]. It is often correlated to loneliness, shyness and social isolation.

2.4 Assessment methods

The assessment of the personality traits is commonly done with specialized questionnaires, called *inventories.*These questionnaires contain entries about one's behavior, thoughts, feelings and other characteristics. One inventory to measure the **Big Five** is the Revised NEO Personality Inventory (NEO-PI-R), consisting of 240-items to detect not only the Big Five, but also measure 6 different facets of each trait. In this and most other inventories, participants have to assess personality characteristics and behaviors of themselves on 5-point Likert scales ranging from "Disagree strongly (1)" to "Agree strongly (5)". The reliability of the NEO-PI-R has been tested thoroughly: It ranges from 78 - 91 % according to the traits [2]. A shorter version with only 12 items per trait is the 60-item NEO-FFI. It has a reliability of 74 - 83 %.

Another common type of inventory is the 44-item long BFI-44 [34] where participants have to answer 44 sentences starting with "I see myself as someone who...":

- is talkative
- tends to be lazy
- tends to find fault with others
- can be moody
- does a thorough job
- ...

Even though it seemed radical that *only* 44 items may measure the Big Five accurately, the results were promising. The BFI-44 is now one of the most used inventories. To allow even faster studies, 10-item inventories were created. Both, the Ten Item Personality Inventory (TIPI) and the Ten Item Big Five Inventory (BFI-10, see Table 2), measure the Big Five using exactly 2 items per personality trait. Interestingly, the BFI-10 uses sentences where the TIPI only uses adjectives for the self-assessment. The authors of the BFI-10 have shown that it is a valid inventory and even exceeds the TIPI in accuracy [34]. Nevertheless, the shorter scale is not as reliable as the BFI-44: It only retained 85% of the retest reliability of the BFI-44, especially the detection of agreeableness was worsened. The authors suggest adding a third item for agreeableness "Is considerate and kind to almost everyone" to greatly increase the measured validity of this trait.

To assess **narcissism**, the Narcissistic Personality Inventory was created in the 1980s. The NPI-40 contains 40 items, measuring a variety of sub-traits of

I see myself as someone who...	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
... is reserved	(1)	(2)	(3)	(4)	(5)
... is generally trusting	(1)	(2)	(3)	(4)	(5)
... tends to be lazy	(1)	(2)	(3)	(4)	(5)
... is relaxed, handles stress well	(1)	(2)	(3)	(4)	(5)
... has few artistic interests	(1)	(2)	(3)	(4)	(5)
... is outgoing, sociable	(1)	(2)	(3)	(4)	(5)
... tends to find fault with others	(1)	(2)	(3)	(4)	(5)
... does a thorough job	(1)	(2)	(3)	(4)	(5)
... gets nervous easily	(1)	(2)	(3)	(4)	(5)
... has an active imagination	(1)	(2)	(3)	(4)	(5)

Table 2. Big Five Inventory 10 [34]

narcissism. Keeping with the times, a shorter 13-item inventory, the NPI-13 was created in 2013. It focuses on three main sub-traits: the need for leadership/authority, grandiose exhibitionism, and entitlement/exploitativeness. [29]

To measure **self-esteem**, the Rosenberg Self-Esteem Scale can be used [29]. It contains 10 items measuring the perceived self-worth, e.g. "I feel that I have a number of good qualities" which have to be answered on a scale from "Strongly disagree (1)" to "Strongly agree (5)".

Loneliness is typically assessed, using a 10-item abbreviated form of the UCLA loneliness scale. In this inventory, individuals rate their interpersonal relationships by answering five positive statements and five negative statements on a scale from "Never feel this way (1)" to "Always feel this way (4)". An exemplary question is "How often do you find yourself waiting for people to call or write?". Even though the inventory contains only 10 items, its reliability was trialed successfully: 84 %. [36]

Social anxiety can be measured using the Leary Social Anxiousness scale. It contains 15 items measuring the frequency and intensity of experienced anxiety before and during social interactions. Participants have to answer 11 positive and 4 negative statements on a scale from "Not at all characteristic (1)" to "Extremely characteristic (5)", e.g. "I often feel nervous when calling someone I don't know very well on the telephone.". The reliability of the Leary Social Anxiousness scale is 89 %. [36]

Smartphone usage patterns, can be detected in three ways: Self-reports and interviews are the most time-consuming, effortful, and possibly distorted way. Even though researchers have a direct contact to the subjects, only basic information can be extracted. Interviews are good for ground-laying questions like "Do you own a smartphone?" or "Which messenger do you use the most?", but not good for fine-grained questions like "How often do you look on your smartphone every day?" or even "How many messages do you write per day?". Especially in the last questions, self-perception may be wrong by an order of magnitude.

The second option is to manually collect data from smartphones, e.g. by taking a picture of a conversation on the screen as done in [38]. Even though this method is still very intrusive, it allows access to undistorted data.

The third and most promising method is the fully automatized data collection by installing a dedicated app on the subject's smartphone. This app can not only record real-time usage data, but also ask the user short questions for further insights. Especially for recording the **current mood** and emotions, this is a much better option than interviews or diaries. Furthermore, the subject can be asked to connect to a dedicated Facebook app, which then extracts profile information [17].

2.5 Historical mobile phone use

Even though mobile phone technology use is relatively new (the first SMS was sent in 1992), both - technology and usage - changed completely over the last decade (2005-2015). Typical use cases of traditional mobile phones are limited to telephony, SMS, and simple games, because data plans were not widespread and mobile internet browsing was not consumer friendly enough. Finally, multimedia-capable phones which were able to play MP3s, take pictures, record videos, and send MMS added more usage patterns.

A 2007 study [6] serves as good example for this paradigm change: The study and more recent studies report unisono, that extraverts spend more time calling. This is an example for a behavior that was enabled by mobile phones and is still present. Plus, the study reports that extraverts spend more time adjusting ringtones and wallpapers. This is a great example for a behavior that has radically changed since the introduction of smartphones. In the time of the study, smartphone penetration was minuscule and typical phones were strongly limited in their usage possibilities. That is why the study names "changing ringtones and wallpapers" as "means of stimulation". However, both use cases are not even mentioned once in any other source after 2010, indicating a strong shift in usage.

Furthermore, multiple studies from 2004 - 2008 [36,35,4] divide their test subjects into the categories "Texters" and "Talkers" according to their preference of sending SMS or calling contacts. The studies assign multiple personality characteristics to the groups (especially extroversion, social anxiety, and loneliness). While more up-to-date studies still indicate similar personality characteristics, usage patterns have changed again: Due to the widespread nature and increased versatility of messaging services like WhatsApp, messaging gained much traction across all users.

Today, highly extroverted people may use messaging to reach massive numbers of contacts in short time, in comparison to older studies, indicating a preference of calling for those subjects. This differentiation becomes even more blurred, with the increasing adoption "voice messages", the asynchronous and effortless sending of recorded speech via messengers.

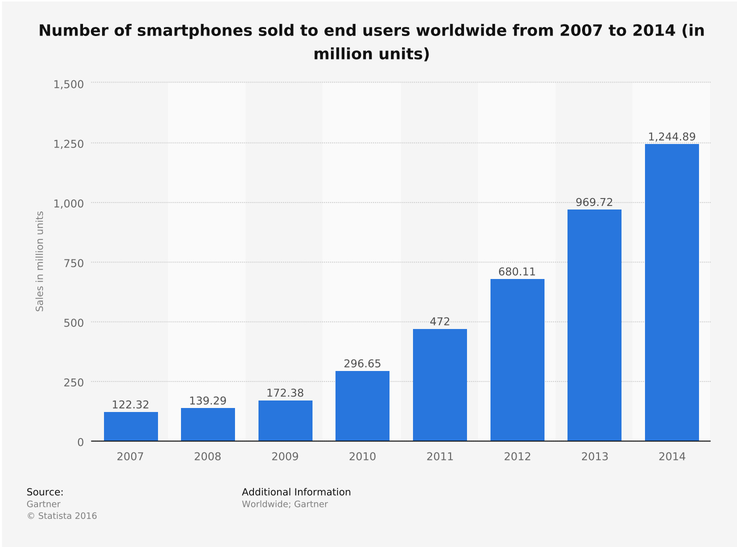

Fig. 2. Number of smartphones sold to end users worldwide from 2007 to 2014 [15]

2.6 Current smartphone use

Since the introduction of the iPhone in 2007, the first globally successful consumer smartphone, smartphone sales, penetration, and usage skyrocketed. Starting with 122 million smartphones sold worldwide in 2007, this number increased tenfold over 7 years, to 1.2 billion in 2014 (see Fig.2) [15]. Strongly correlated to this, smartphone penetration increased steadily, ranging in 2015 from 55 % in the USA and Germany [13,10], to 83 % in Sweden [19].

This is not surprising: Due to their universal usage possibilities, smartphones deprecated traditional mobile phones, MP3 players, compact cameras, and portable gaming consoles. They even overtook the computer as preferred medium for accessing the internet. Common use cases include: telephony, messaging, accessing social networks, browsing the internet, shopping, sending emails, and reading news. Smartphone use is so intense, researchers coined the term psychoinformatics. They predict that experience sampling via mobile applications will assume a central role in psychological data collection [42]. Mobile experience sampling can draw on a vast amount of sensors built into typical smartphones:

- front- and back-facing cameras
- microphones
- proximity sensors
- touchscreens
- pressure sensors
- geolocation sensors (GPS)
- accelerometers
- gyroscopes
- light sensors
- humidity sensors
- fingerprint readers
- heart rate monitors

Even with their limited processing power, typical smartphones in 2015 can sense and detect a variety of features:

- detect walking and running
- detect cycling and other sports
- detect driving
- track eyes
- identify fingerprints, face & voice
- understand spoken language
- recognize complex gestures

Recent studies strengthen this prediction: in 2015, a study over 2500 German smartphone users reported a general daily smartphone usage mean of 2.7 hours. More than one third of smartphone owners even use their smartphone in the last 5 min before going to sleep or in the first 5 min after waking up, making it truly an essential part of the everyday life [33]. Additionally, hundreds of millions of users use social networks to upload personal information to the internet. They upload personal information in the form of texts, images, and videos, enabling extensive studies over millions of datasets which would not have been possible before [42].

3 Analysis

3.1 Messaging

Messaging is a key usage of every smartphone. Sending texts via common messengers like WhatsApp, Facebook, or simply SMS is a integral part of every smartphone user's day. In 2015, the popular WhatsApp messenger alone reported 900 million monthly users, tripling its user base over 2 years. German studies in 2015 found that the app accounted for one fifth of the daily smartphone usage (32 minutes) having deprecated SMS and reached a massive 667 million messages sent very day [33,40] (see Fig. 3).

Contrary to face-to-face conversations, e-mails, and voice calls, text exchanges lack openings and closures. In auditory (spoken) communication, openings like summons, greetings, or simply "How are you?" are needed to secure attention, connection, and mutual identification. To secure attention, one could ask "Hello?", to secure connection "Can you hear me?", and to secure mutual identification "Who am I speaking with?". E-mails on the other hand traditionally contain formal openings and greetings, including long signatures. This is not necessary with text messaging. It exploits affordances of the smartphones: Connection does not need to be secured, because smartphones are nearly constantly connected to the internet and deliver messages automatically and reliably. Attention does not need to be secured verbally, because the smartphone secures it in a non-verbal way by playing a ringtone, vibrating, and displaying a notification. Mutual identification does not need to be secured, because smartphones are assumed to be personal tools and messages are usually exchanged between people who know each other beforehand. Additionally messaging can be used where speaking loudly is forbidden or annoying, e.g. in class. [38]

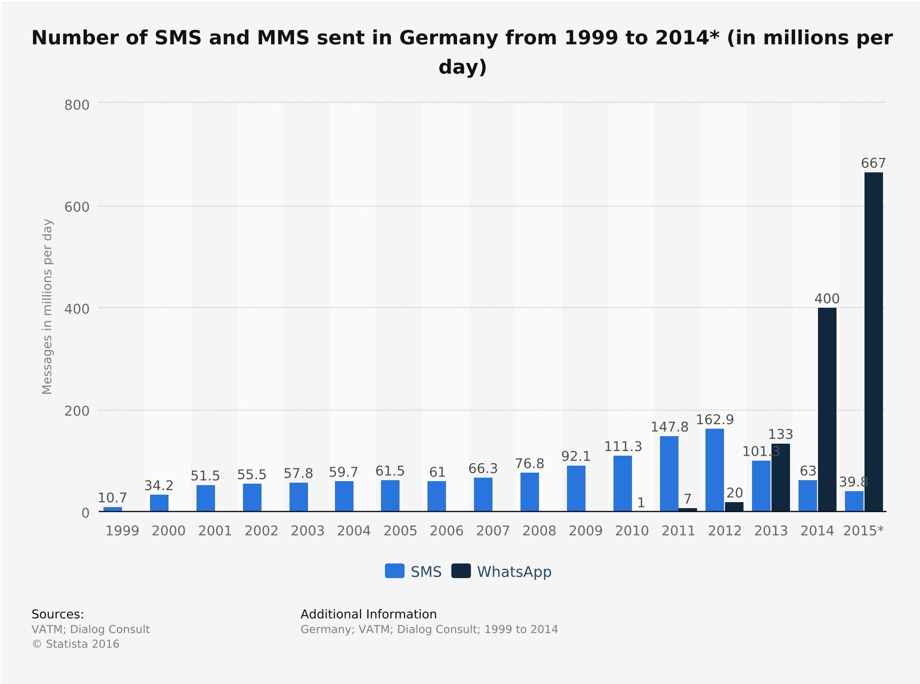

Fig. 3. Number of messages sent in Germany from 1999 to 2015 [40]

Early studies in 2005 [12,4] have already shown that text messaging is generally used in both social and task-oriented contexts. Social contexts include chatting with friends and relatives, while task-oriented contexts include educational and professional discussions. The studies found that "Texters" often prefer the conciseness and asynchronicity of texting over calling and other forms of communication. It has to be noted, that the clear distinction between "Texters" and "Talkers" is not as appropriate today as it was in 2005: Due to the widespread adoption of smartphones, the improvements in on-screen keyboards and the richness of today's messaging applications, many people use messaging a lot more than calling, even though they do not have strong favors for it.

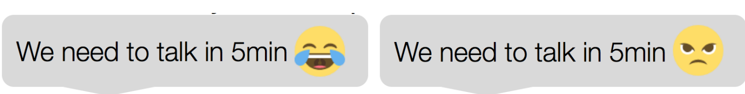

Fig. 4. Visualizing the intent of a message by using emoji (based on [28])

Concerning the use of emoji[1] in both private and professional settings, two studies were performed: The work in [12] analyzed short internet chats of 158 secondary school students in 2007. The study in [28] recorded 20.000 messages in a corporate instant messenger in 2008. Both studies have shown that positive and negative emoji are correlated to positive and negative communication scenarios respectively. In person, positive communication scenarios may be expressed by smiling, a cheerful voice, and soft gestures. A negative communication scenario on the other hand may be expressed by angry mimics, strong or even loud intonation, and intensive gestures. Because text messaging is such a simplistic communication medium which misses that emotional channel, messages become easily ambiguous. A message with an opaque intent like "We have to discuss in 5min" needs to be clarified by the use of emoji. The studies showed that adding happy emoji created a positive emotion in the receiver, while adding sad emoji created a negative emotion (Fig. 4,[28]). Smartphone users learn this very quickly and get used to it. This means that a person sending positive emoji will most likely experience **positive sentiment** and a person sending negative emoji will most likely experience a negative situation.

In the study in [4], 21 participants had to fill out a diary describing sent and received SMS over a 2 week period in 2005. The results suggest that texting is not only used for social grooming (e.g. asking "How was your day?"), but also to overcome shyness, because the user was too awkward to call. The study in [36], done at the University of Plymouth in 2005 supports this: After analyzing 158 submissions to an online questionnaire, it suggests that anxious participants prefer texting over calling for expressive and intimate contact. Thus, an increased usage of texting in comparison to calling may indicate **social anxiety**.

As part of the Lausanne data collection campaign in 2009, the study in [9] analyzed SMS, call, and application logs of 83 participants over a period of 8 months. It suggests that emotionally stable (less neurotic) and open people typically receive less SMS. Controversially people with low openness tend to send more SMS, as do less neurotic people. The study suggests that this may indicate that the personality may be better judged by SMS composed by users than received by them. Analogous to the study's results, it seems natural that neurotic individuals send more SMS, because of their unstable and worrying nature.

Interestingly, a low agreeableness correlates positively with time spent in the SMS app, but not necessarily resulting in more SMS sent or received. This may be, because less agreeable people intensively and repeatedly read and analyze received messages and prepare lengthy arguments for outgoing messages.

The newer, larger German study in [33] recorded the WhatsApp behavior of

[1] Even though emoji and emoticons are inherently different - emoji are graphical icons and emoticons are simple texts like *:-)* - the words are used interchangeably in the literature. Because both trigger similar emotional responses, this work will not differentiate between the two.

2400 participants over a 4 week period in 2015. It agrees with the Swiss study [9] in the fact that neurotic individuals use messengers for a longer time. This effect of the neurotic personality trait seems to be constant, even over technological change.

Furthermore, the study reported that the duration of daily WhatsApp use is positively correlated to extraversion. This could be an indicator that instant messengers like WhatsApp have become useful for highly communicative and extraverted people, which would not have used instant messaging that much in the years before. As last result, conscientiousness is negatively correlated to the duration of WhatsApp use: True to their dutifulness and orderliness, conscientious individuals do not seem to have the motivation for the longer use of instant messengers.

3.2 Telephony

Analyzing telephony meta-data is an important mean of classifying the social behavior ob subjects. This work features three studies. Additionally to [36,9], the study done in [8] analyzed the smartphone interaction of 117 Swiss individuals over a period of 17 months.

The study in [36] suggests that **lonely** people prefer making voice calls. They tend to use texting only as last resort, because of it's reduced intimacy. It additionally showed that increased usage of texting over calling may indicate social anxiety, thus increased usage of telephony on smartphones may indicate a **social confidence**.

This is supported by [8]. They found that telephony is heavily correlated with **extraversion**: Extraverted people receive more calls, have longer incoming calls in average, have longer incoming calls in total and have more unique contacts called. This could be a simple indicator for the vibrant social life of extraverted people. The study also found that incoming calls are a better predictor than outgoing calls, claiming that extraverted people do not only receive more calls because of their larger social network, but also because their peers feel fine when calling them. [8]

Agreeableness correlates with the number of calls received [9] and the number of unique contacts called, confirming their friendly personality trait [8].

Interestingly, people with high **openness** are more likely to miss calls. [9]

Conscientious people were found to talk with fewer unique contacts via telephone, reinforcing their structured and responsible nature [8].

3.3 Social Media

Beside messaging, using social media and accessing social networks like Facebook, Google+, and Twitter is another key usage of Smartphones: A study in 2015 showed, that Facebook alone accounted for 9% (15 minutes per day) of the daily smartphone usage of German test subjects [33]. Americans tend to use it even more, their self-reports averaged at 107 minutes of daily Facebook usage [29].

Beside messaging, posting status updates is one of the most preferred features of Facebook. Even though only simple texts, photos and videos may be posted, Facebook users use this feature to share anecdotes, declare political o pinions, post family videos, brag about achievements or simply post a picture of their last meal [29]. The genuine truth of social media profiles was doubted highly because social networks are commonly used by people to present themselves. Do social media profiles only represent idealized version of the users, not their real personalities? To falsify this theory, the German researchers Back et. al. analyzed 236 social media profiles in the year 2010 [3]. To ensure that participants did not alter their profiles, the researchers saved the profiles before the potential participants were recruited. After comparing the profiles with personality reports made by the users and assessments made by their friends, the researchers concluded that social media profiles reflect the actual personality, not an idealized version.

4 further social media related studies concerning the personality of their subjects are shown in this work: The study in [2] collected Facebook data of 237 Israeli students in 2010, who had to fill out the NEO-PI-R self-report. The study in [17] analyzed linguistic data of 167 Facebook users in 2011, who also had to fill out a 45-item questionnaire about the Big Five. The study in [29] collected data of 555 US Facebook users in 2015, who had to fill out questionnaires about Big Five, self-esteem, and narcissism additionally. Finally, the study in [16] collected linguistic data of 50 Twitter users, again with a 45-question inventory measuring the Big Five.

Because of the widespread usage possibilities of social media for self-presentation and discussion, the usage frequency is positively associated with extraversion, neuroticism, and openness [29].

The studies in [29,16] showed that **extraverts** used Facebook generally more frequently and wrote more words per tweet. Due to their talkative nature, they post status updates about their social activities and everyday life more often. They do not only have a greater number of Facebook friends [29], but their egocentric network density is lower, which means that they have a larger, more open, less interconnected network of friends [17]. This confirms the expectation, that extroverts easily make new friends and use social networks according to their outgoing nature.

Possibly to showcase their extraversion, the "activities" profile field typically contains more text [17]. Surprisingly, the study in [2] found that extroverts tend to fill out profile fields less, suspecting that they prefer their social skills instead of profile fields for sharing other personal information. Future research may clarify, whether the two studies contradict, or the "activities" field is simply an exception.

Linguistic analysis showed that extraverts tend to tweet about life and family, but less about health. They used words like "mate", "child", "husband" more and words like "clinic", "flu", and "pill" less. Possibly due to their general positivity, they avoid negative topics like illnesses and prefer positive topics like family and friends.

Even though **Neurotic** individuals have an increased usage frequency of Facebook - like extraverts - the reasons for this most likely differ: They have been shown to use Facebook to seek attention, social support, and validation. This is further indicated by a high usage frequency of the social network for emotional discourse, e.g. discussing personal dramas [29].

Additionally, linguistic studies have found that neurotic subjects write Facebook posts increasingly about anxiety, using words like “worried”, “fearful”, and “nervous” more frequently. On Twitter, words concerning perceptual processes like “listen”, “hear”, “feel”, “touch” are increased, possible indicators for personal dramas. Furthermore, their tweets contain a higher amount of exclamation marks - a possible indicator for intense, emotionally charged tweets.

Interestingly, the length of the last name correlates with neuroticism [17]. While the authors of [17] believe that this is because common misspelling of long last names may lead to anxiety, another possible reason is that neurotic individuals may use fake last names on Facebook to gain further attention.

Like the previous two groups, people with high **openness** show increased Facebook usage as well, but instead of socializing, they are shown to use Facebook more for finding and sharing information about current events, research, and politics [29]. Because, they are more likely to post status updates about intellectual topics, they may receive less likes and comments [29]. Possibly to present their openness, their “favorite books” profile field typically contains more text and they are more likely to name their personal website address [17]. While the size of their network is not significantly influenced by their openness, the egocentric network density is lower - indicating that hey have a diverse set of friends, following their diverse interests [17].

Despite their open nature, linguistic analysis in [16] only showed that their tweets contain specific words about work and humans (“adult”, “baby”, “boy”) more often, but hashtags even less. Using a General Inquirer dataset, which maps a sentiment from -1 (negative) to +1 (positive) to English words, open individuals were the only group to show a significant correlation to the average sentiment of their tweets: Their average tweet has a positive sentiment.

Furthermore heir tweets contain an increased amount of articles (“a”, “an”, “the”), quantifiers (“few”, “much”, “many”), causations (“because”, “effect”, “hence”), and words of certainty (“always”, “never”), but possible reasons for this are not yet clear.

Agreeableness is positively associated with using Facebook for communication and negatively associated with using Facebook to seek attention and badmouth others [29]. Their Facebook posts contain a higher frequency of affective words (“happy”, “cry”), especially positive emotions (“love”, “nice”, “sweet”) [16]. Despite their friendly and warm nature, the networks of agreeable individuals do not show any signification variations in size or density. Linguistic analysis showed that agreeable subjects tweeted less about achievements and money ("earn", "win", "cash"). Causations like "because" were also found less in their tweets, possibly indicating that they avoid debating and delicate subjects on social media. Instead, they were found to write increased about food and eating, a lighter, less and ultimately more agreeable critical topic [16].

People who are high in **conscientiousness** tend to use Facebook less frequently [29]. They avoid attention seeking and badmouthing, especially perceptual words ("hear", "see", "feel") and swear words are found less in their status updates [17]. On the other hand, they are more likely to update social topics [17], especially about their children [29]. Possibly, they try to demonstrate their conscientiousness, by demonstrating the well-doing of their own children.

The tweeting behavior of conscientious subjects showed many significant correlations: tweets contained less future tenses, negations, negative emotions, sadness, feelings, and words related to death. Unsurprisingly, "will", "gonna", "never", "should", "would", "could", fillers like "blah" and other words that do not match their orderliness were found less in their tweets. Instead, conscientious subjects used more words about work. Furthermore they use colons, exclamation marks, and links increased, undermining their professionalism even more. [16]

Contrary to previous beliefs, people with **low self-esteem** typically do not use Facebook to *present* themselves. They tend to use Facebook only to *express* theirself. Because they have problems with disclosing themselves in person, they tend to use status updates for that [29]. This often results in posts expressing negative sentiment. **Social anxiety** - which often correlates with low self-esteem - results in an increased number of positive posts about one's romantic relationships. Interestingly, this behavior is most likely not intended to boast, brag, or force a good public image of the relationship. The authors [29] found that posts about romantic partners are most likely to be created on days when the author feels the most insecure and fearful of losing their partner: They try to strengthen the relationship when feeling threatened. In average, posts of individuals with low self-esteem receive fewer comments and likes for their Facebook posts. This may be because of the regular negativity of their posts, or simply because posts about romantic partners don't appear likable to others.

Narcissistic people use Facebook for attention-seeking and validation. They are more likely to post status updates about their accomplishments to gain validation for their self view. Additionally, they post frequently about diet plans and exercise routines to express the importance they place on their physical appearance [29]. Reinforcing their narcissistic view, they receive more likes and comments to their status updates. To expand their audience on social media, narcissistic people tend to have high friend counts and are more likely to accept strangers as friends [7]. Narcissism may even create anti-social behavior: Retaliating against bad comments, seeking more social support than one gives, getting angry about a lack of comments on one's status and reading others' status updates to find out if they are talking about oneself are possible results [7] .

3.4 Other Apps

It is no surprise, that the most correlations were reported between personality and social apps (messaging, social media, telephony). Simply because of the reason that social interaction speaks volumes about one's personality. Nevertheless, the usage of other apps may correlate to certain characteristics as well.

Conscientiousness has the most distinct correlations: Conscientious subjects use office, mail, and calendar apps more, but audio, video, music, and e-commerce apps less [9] [22]. The first three may be categorized as *productivity* apps, while the last four may be used for entertainment. To use the smartphone for productivity purposes instead of entertainment clearly correlates to the efficient and planful nature of conscientious individuals.

Extraversion is negatively correlated with literacy applications [22] and Internet use [8]. Fitting to the nature of introverts, the studies provided clear indicators that they not only have less social interaction in person, but also using their smartphones. Instead they prefer reading books and articles in literacy apps and browsing the internet with their browser.

People with high **openness to experience** were found to use video, audio, and music apps more, possibly as sources for their inspiration, or due to their curiosity and wide interests. Furthermore, unlike their conscientious counterparts, they used office and calendar apps less. [8]

Behavior	Extra.	Agree.	Cons.	Neuro.	Open.
Preferring texting over calling	-				
Miss incoming calls					+
Number of unique contacts called	+				
Duration of calls	+				
Number of calls received	+				
Number of messages received				+	-
Number of messages sent				-	-
Time spent in messaging app	+	-	-	+	
Usage frequency of Social Media	+		-	+	+
Number of Facebook friends	+				
Information in Facebook profile	-				
Seek attention & emotional drama		-	-	+	-
Sharing research, politics, ...					+
Self presentation & boasting	+	-	-	+	
Usage of Office apps			+		-
Usage of Video/Audio/music			-		+
Browsing the Internet	-	-			

Table 3. Overview over behaviors and their correlation to characteristics

3.5 Automatic prediction of personality

The work in [8] did not only perform a 17 month long study over the smartphone interaction of 117 subjects, but also implemented and tested a supervised machine learning method for the automatic prediction of personality. They defined a binary classification for each trait of the Big Five, e.g. users could be either classified as "more extroverted" or "less extroverted" than the median. The best results were achieved with a hybrid model that uses different models for male and female participant.

4 Discussion

In contrast to detecting personality characteristics, studies found it difficult to detect the **current mood** of their users: The only substantial results were achieved when analyzing the relation between emoji and emotions in laboratory settings [12,28]. It is still unclear how the usage of emoji expresses sentiment in intimate, real life messaging conversations. Concerning social networks, the only finding related to mood was made in [7]: people use social networks to get social support when they feel upset, or to feel better when feeling distressed. Again, no precise usage patterns could be extracted. Further undermining the problem, the study in [32] reported a low mood detection accuracy by analyzing tweets, due to the skewness of the data. It seems that detecting mood is more difficult than detecting the personality, possibly because mood changes are quick and frequent. Furthermore, it is possible that the mood may not even have a remarkable impact on many parts of the smartphone interaction at all.

An overall limitation of smartphone-based detection is the steadily ongoing change of both the technology and the human interaction with it. First of all, a constant change of popular apps limits the app-based detection. Even though apps may be grouped into similar categories (messaging, social networks, games, productivity, ...), each app has important differences: In contrast to Twitter, where tweets are mainly public, Facebook posts may be sent to friends only, to groups, or publicly. Additionally, the two social networks have such a large market share over all ages, that teenagers may prefer other apps to experience digital freedom from their parents. Second, the interaction with the smartphone changes constantly. Findings from the last decade that extroverts prefer calling over messaging may not be applicable anymore. Highly extroverted people use messaging to reach many contacts in a fast way, in comparison to older studies, indicating a preference of calling for extroverts.

A strong positive outcome of smartphone-based studies is the increased reliability in comparison to self-reported metrics: While self-reports most likely contain small errors, huge misperceptions of oneself are common, especially when it comes to daily smartphone usage times and frequencies. This empiric problem can be abolished by automatized, large-scale smartphone studies.

5 Conclusion

Personality is defined as stable set of characteristics that describe how one's thoughts, feelings, and actions are common or different to others. It is commonly characterized into the Big Five traits extraversion, agreeableness, conscientiousness, neuroticism, and openness. Further, we looked at other characteristics, including narcissism, self-esteem, loneliness, social anxiety, sociability, and the current mood.

We showed that these characteristics are highly relevant in many parts of life, including social interaction, school, job, sport, and relationships. Because they even can predict job satisfaction, professional and romantic relationship success, and much more there exists high interests in detecting the personality. Measuring the personality characteristics is commonly done using questionnaires, in which test subjects have to assess their thoughts behaviors, feelings. Typical inventories measuring the Big Five are the BFI-44, NEO-PI-R, and TIPI. Because it is effortful to use these inventories, we showed that there exists interest in the automatized detection of personality characteristics using smartphones.

Smartphone usage is a globally spread and steadily increasing everyday part of life for all demographics. The devices are used for both corporate and private purposes, ranging from sending business e-mails to sharing intimate moments. Because smartphones fulfill such a vital role, many studies have analyzed the correlations between smartphone and internet usage and the user's personality. Even though there is room for improvement in this new field, called psychoinformatics, the studies yielded accurate and significant results. Personality characteristics - especially extraversion, agreeableness, conscientiousness, neuroticism, openness, and narcissism can be detected reliably by analyzing app usage frequencies, message and telephony behaviors, and social media usage patterns. It is important to note that both meta data analysis and linguistic content analysis lead to significant results:

The amount and diversity of contacts and frequency of social media usage provided strong indicators for extraversion, openness, and narcissism. Linguistic analysis of social media posts illustrated conscientiousness and neuroticism particularly well. Usage of productivity, mail, and calendar apps indicated high conscientiousness and low openness. Further, the usage of games, music, and video apps provided more indicators for high openness and reduced conscientiousness. Finally, literacy apps showed correlation with introversion.

While personality characteristics can be detected reliably through smartphone interaction, other emotional data is much more problematic to detect. Emotions, moods, and the general sentiment can only be detected on a shallow level by analyzing superficial information, e.g. the usage of positive and negative emoji in messages. Loneliness and social anxiety can be detected on a basic level as well: By analyzing linguistic features of social media posts and whether users prefer to call or to text, simple assumptions can be made. Likewise, there exist only basic indicators for low self-esteem as well.

Conclusive, we showed that it is already possible to use machine learning to create a fully automatized classifier for personality traits.

5.1 Future work

To improve well-being and health, smartphones may support and warn their users when detecting psychological **stress**. This is a feeling of strain, pressure, and anxiety which may be caused by external and or internal factors, which can even result in depression, heart attacks, and strokes. While stress may increase the heart rate or make an individual sweaty and nervous, the most implications appear in the individuals mind, unnoticeable by traditional sensors. Today, sensing stress is limited to special sensors. Smartphone-based stress detection has not yet achieved significant results [5]. In the future though, smartphones could act as sensing devices for stress. The smartphone could not only automatically sense heart rate, sweat, and trembling, but also analyze even the tiniest interactions to detect stress. Once high stress levels are detected, the smartphone could assist its user at regulating the emotions or even at finding help.

Most studies focus on the Big Five, so social anxiety, low-self esteem, loneliness, and other feelings may be analyzed in a more fine-grained way in future studies. Beyond the scope of basic indicators covered by today's studies, more small, subtle and indirect indicators may be found.

Studies indicate that people try to regulate their **emotional state** by purposefully listening to specific music. Because the mood is a important factor of music listening, the selected songs could be a viable indicators for the current mood of the user [41]. Current studies try to explore how a song's emotions can be classified and how users try to regulate their emotions through music [14]. Even though current studies had only limited success at correlating music and mood [39], through advantages in those fields, three use cases emerge: First, music players could detect the user's emotions and enable the environment to react better to them. Additionally, automatically created playlists could react to the user's mood better, by making use of specialized recommender systems based on the user's mood (called CAMR - context-aware music recommendation) [39]. Third, the user's personality could be predicted even better, because it is shown to correlate to the music taste [16].

Very extreme emotional may be caused by **medical conditions**, for example bipolar disorder or deliberate self-harming. First studies are already being performed to detect the mood of bipolar disorder patients [1]. To support the patients and protect them from possible harm, further studies need to be performed on how to detect those extreme emotional states with smartphones. Following the example of Google, which provides users with phone numbers of advice services when users search for "suicide", smartphones could provide support on more fine-grained and subtle levels.

There is more work to be done concerning **personality characteristics**. Current studies only focused on one demographic group (e.g. students) or did not take demographics into account at all. This may result in problems: Teenagers may possibly appear very neurotic, simply because of their age and development, even if they are average in comparison to their peer group. Future studies comparing the results of detected personality characteristics between different groups, could further improve the detection. Interesting comparisons could be made across genders, cultures, and ages.

Finally, **new detection algorithms** and information sources will most likely bring new breakthroughs in the field of psychoinformatics. Concerning messaging, most studies only analyze the meta data of messages, not the content. Linguistic text analysis could show up new coherences. Furthermore, no work has analyzed group chats, a possible source for information about a users network.

When regarding the device data, more research has to be done about movement history (using geolocation). Movement patterns may be a viable source of information, especially for detecting mood and sentiment. Using the front-camera to detect emotions in the face could bring a breakthrough at automatic mood sensing.

Replacing traditional data analysis with machine learning, could significantly improve detection accuracy. Additionally, it may show up previously unknown correlations between smartphone usage times, events, and the user's sentiment and personality. Particularly concerning telephony, only meta-data (caller, receiver, time, and duration) is currently analyzed. Further progress in the analysis of speech patterns, intonation, vocal range, and speech recognition could improve telephony as data source greatly.

In terms of social media usage, automatized semantic image analysis could bring new insights. In current studies, only the frequency and amount of posted photos was analyzed. Future studies could analyze the contents of photos and videos to imply further facts about the personality. Beside detecting new coherences, detecting photos about food, sports, family, and politics could further prove today's findings.

References

1. Alvarez-lozano, J., Frost, M., Osmani, V., Bardram, J., Kessing, L.V., Mayora, O., Faurholt-Jepsen, M.: Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. 7th International Conference on Pervasive Technologies Related to Assistive Envotinments (PETRA) 1(212) (2014)
2. Amichai-Hamburger, Y., Vinitzky, G.: Social network use and personality. Computers in Human Behavior 26(6), 1289–1295 (2010), http://dx.doi.org/10.1016/j.chb.2010.03.018
3. Back, M.D., Stopfer, J.M., Vazire, S., Gaddis, S., Schmukle, S.C., Egloff, B., Gosling, S.D.: Facebook Profiles Reflect Actual Personality, Not Self-Idealization. Psychological Science 21(3), 372–374 (2010), http://pss.sagepub.com/lookup/doi/10.1177/0956797609360756
4. Barkhuus, L.: Why everyone loves to text message: social management with SMS. Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work pp. 324–325 (2005)

5. Bauer, G., Lukowicz, P.: Can smartphones detect stress-related changes in the behaviour of individuals? 2012 IEEE International Conference on Pervasive Computing and Communications Workshops, PERCOM Workshops 2012 (March), 423–426 (2012)
6. Butt, S., Phillips, J.G.: Personality and self reported mobile phone use. Computers in Human Behavior 24(2), 346–360 (2008), http://linkinghub.elsevier.com/retrieve/pii/S0747563207000295
7. Carpenter, C.J.: Narcissism on Facebook: Self-promotional and anti-social behavior. Personality and Individual Differences 52(4), 482–486 (2012), http://dx.doi.org/10.1016/j.paid.2011.11.011
8. Chittaranjan, G., Blom, J., Gatica-Perez, D.: Mining large-scale smartphone data for personality studies. Personal and Ubiquitous ... (September), 1–18 (2013), http://link.springer.com/article/10.1007/s00779-011-0490-1
9. Chittaranjan, G., Blom, J., Gatica-perez, D.: Who ' s Who with Big-Five : Analyzing and Classifying Personality Traits with Smartphones (2011)
10. ComScore: Number of smartphone users in Germany from January 2009 to July 2015 (in millions) (2016), http://www.statista.com/statistics/461801/number-of-smartphone-users-in-germany/
11. Costa, P.T., McCrae, R.R.: Four ways five factors are basic. Personality and Individual Differences 13(6), 667–673 (1992)
12. Derks, D., Bos, A.E., von Grumbkow, J.: Emoticons and social interaction on the Internet: the importance of social context. Computers in Human Behavior 23(1), 842–849 (2007), http://linkinghub.elsevier.com/retrieve/pii/S0747563204002079
13. EMarketer: Smartphone penetration in the United States from 2010 to 2017 (2016), http://www.statista.com/statistics/201183/forecast-of-smartphone-penetration-in-the-us/
14. Ferwerda, B., Schedl, M.: Enhancing music recommender systems with personality information and emotional states: A proposal. CEUR Workshop Proceedings 1181, 36–44 (2014)
15. Gartner: Number of smartphones sold to end users worldwide from 2007 to 2014 (in million units) (2016), http://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/
16. Golbeck, J., Robles, C., Edmondson, M., Turner, K.: Predicting personality from twitter. Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011 pp. 149–156 (2011)
17. Golbeck, J., Robles, C., Turner, K.: Predicting personality with social media. Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11 p. 253 (2011), http://portal.acm.org/citation.cfm?doid=1979742.1979614
18. Goldberg, L.R.: The Structure of Phenotypic Personality Traits 48(I), 26–34 (1993)
19. Google: Percentage of people who use a smartphone in Sweden from 2012 to 2015 (2016)
20. Hughes, D.J., Rowe, M., Batey, M., Lee, A.: A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. Computers in Human Behavior 28(2), 561–569 (2012), http://dx.doi.org/10.1016/j.chb.2011.11.001
21. Judge, T.A., Higgins, C.A., Thoresen, C.J., Barrick, M.R.: The big five personality traits , general mental ability , and career success a ... (1999)

22. Kim, Y., Briley, D.A., Ocepek, M.G.: Differential innovation of smartphone and application use by sociodemographics and personality. Computers in Human Behavior 44, 141–147 (2015), http://www.sciencedirect.com/science/article/pii/S0747563214006694
23. Kron, T.: How to become a better leader. Nursing 6(10), 67–70, 72 (1976)
24. Landers, R.N., Lounsbury, J.W.: An investigation of Big Five and narrow personality traits in relation to Internet usage. Computers in Human Behavior 22(2), 283–293 (2006), http://linkinghub.elsevier.com/retrieve/pii/S0747563204001128
25. Lane, W., Manner, C.: The Impact of Personality Traits on Smartphone Ownership and Use. International Journal of Business & Social Science 2(17), 22 (2011)
26. Leung, L.: Generational differences in content generation in social media: The roles of the gratifications sought and of narcissism. Computers in Human Behavior 29(3), 997–1006 (2013), http://linkinghub.elsevier.com/retrieve/pii/S0747563212003706
27. LiKamWa, R., Liu, Y., Lane, N., Zhong, L.: Can your smartphone infer your mood. PhoneSense workshop pp. 1–5 (2011)
28. Luor, T., Wu, L.L., Lu, H.P., Tao, Y.H.: The effect of emoticons in simplex and complex task-oriented communication: An empirical study of instant messaging. Computers in Human Behavior 26(5), 889–895 (2010), http://dx.doi.org/10.1016/j.chb.2010.02.003
29. Marshall, T.C., Lefringhausen, K., Ferenczi, N.: The Big Five, self-esteem, and narcissism as predictors of the topics people write about in Facebook status updates. Personality and Individual Differences 85, 35–40 (2015), http://linkinghub.elsevier.com/retrieve/pii/S0191886915003025
30. McCrae, R.R., John, O.P.: An introduction to the five-factor model and its applications. Journal of personality 60(2), 175–215 (1992)
31. Mehdizadeh, S.: Self-presentation 2.0: narcissism and self-esteem on Facebook. Cyberpsychology, behavior and social networking 13(4), 357–364 (2010), http://www.fyiliving.com/wp-content/uploads/2010/08/facebooknarcissist1.pdf
32. Mogadala, A., Varma, V.: Twitter user behavior understanding with mood transition prediction. Proceedings of the 2012 workshop on Data-driven user behavioral modelling and mining from social media - DUBMMSM '12 (October), 31 (2012), http://dl.acm.org/citation.cfm?doid=2390131.2390145
33. Montag, C., Błaszkiewicz, K., Sariyska, R., Lachmann, B., Andone, I., Trendafilov, B., Eibes, M., Markowetz, A.: Smartphone usage in the 21st century: who is active on WhatsApp? BMC Research Notes 8(1), 331 (2015), http://www.biomedcentral.com/1756-0500/8/331
34. Rammstedt, B., John, O.P.: Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41(1), 203–212 (2007), http://linkinghub.elsevier.com/retrieve/pii/S0092656606000195
35. Reid, D., Reid, F.: Insights into the Social and Psychological Effects of SMS Text Messaging. The Social and Psychological Effects of Text 2005(February), 1–11 (2004), http://www.160characters.org/documents/SocialEffectsOfTextMessaging.pdf
36. Reid, D., Reid, F.: Text or talk? Social anxiety, loneliness, and differential preferences for cell phone use. CyberPsychology and Behavior, 10(3), 424–435 (2007)
37. Ryan, T., Xenos, S.: Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. Com-

puters in Human Behavior 27(5), 1658–1664 (2011), http://dx.doi.org/10.1016/j.chb.2011.02.004
38. Spagnolli, A., Gamberini, L.: Interacting via SMS: practices of social closeness and reciprocation. The British journal of social psychology / the British Psychological Society 46(Pt 2), 343–364 (2007)
39. Teng, Y.C., Kuo, Y.S., Yang, Y.H.: A large in-situ dataset for context-aware music recommendation on smartphones. Electronic Proceedings of the 2013 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2013 (2013)
40. VATM: Number of SMS and MMS sent in Germany from 1999 to 2014* (in millions per day) (2015), http://www.statista.com/statistics/461700/number-of-sms-and-mms-sent-per-day-germany/
41. Yang, Y.H., Teng, Y.C.: Quantitative Study of Music Listening Behavior in a Smartphone Context. ACM Trans. Interact. Intell. Syst. 5(3), 14:1—-14:30 (2015), http://doi.acm.org/10.1145/2738220
42. Yarkoni, T.: Psychoinformatics: New Horizons at the Interface of the Psychological and Computing Sciences. Current Directions in Psychological Science 21(6), 391–397 (2012), http://cdp.sagepub.com/lookup/doi/10.1177/0963721412457362