UNIVERSITY OF OULU

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Vappu Schroderus

CROWDSENSING STUDY TO UNDERSTAND SLEEP AND SMARTPHONE USE

Bachelor's Thesis
Degree Programme in Computer Science and Engineering
February 2021

Schroderus V. (2021) Crowdsensing Study to Understand Sleep and Smartphone Use. University of Oulu, Degree Programme in Computer Science and Engineering, 20 p.

ABSTRACT

Smartphones have become an essential ubiquitous device for people worldwide. There has been a wide range of studies about the negative impact of smartphones as a daily life tool, but the HCI context has not been popular perspective for this topic. In this thesis we are presenting the preliminary results of the relation between different sleeping habits or problems, and the different application usage as a crowdsensed manner. The sleep data is collected with Patient Health Questionnaire-8 and the mobile phone data is collected with Carat. The preliminary results show that there are relations between high frequency of nightly phone use and problems with sleep. In addition, daily music listening was positively related to good sleep and overall alertness of the participants. These results give us an interesting perspective for future work to understand better the relation between mobile phone use and overall human living habits.

Schroderus V. (2021) Unen ja älypuhelimen suhteen ymmärtäminen joukkoaistimisen avulla. Oulun yliopisto, Tietotekniikan tutkinto-ohjelma, 20 s.

TIIVISTELMÄ

Älypuhelimista on tullut maailmanlaajuisesti käytetyimpiä jokapäiväisiä laitteita. Älypuhelimista on tehty useita tutkimuksia liittyen niiden negatiivisiin vaikutuksiin päivittäisessä käytössä, mutta ihmisen ja puhelimen suhdetta ihmisen ja tietokoneen vuorovaikutuksen kontekstissa on tutkittu vähän. Tässä tutkielmassa esittelemme alustavia tuloksia erilaisten nukkumisongelmien sekä erilaisten puhelinapplikaatioiden käyttämisen suhdetta tarkastellen dataa joukkoaistimisen avulla. Unidata on kerätty PHQ-8-terveyskyselyn avulla ja älypuhelimen data on kerätty käyttäen hyödyksi puhelimen datankeräysalustaa Caratia. Alustavat tulokset osoittavat, että runsaalla yöllisellä puhelimen käytöllä ja unionglmilla on yhteys. Lisäksi tulokset osoittavat, että päivittäisellä musiikinkuuntelulla on yhteys hyvään yöuneen ja yleisesti päivittäiseen vireystasoon. Nämä alustavat tulokset antavat meille mielenkiintoisen näkökulman tulevaisuuden tutkimusta varten, jotta pystyttäisiin yhä paremmin ymmärtämään puhelimien ja erilaisten elintapojen välistä suhdetta.

TABLE OF CONTENTS

1. INTRODUCTION

Smartphones have become significantly personal and essential everyday devices for people all around the globe [1]. People have a lot of reliance on their smartphones as they offer useful everyday solutions from basic communication functions like calling and texting to more advanced functions such as calculators, alarm clocks, navigators and overall access to the internet. In addition, easy access to many social media and communication applications such as Facebook or WhatsApp via smartphone offers an easy way to keep in touch with your friends and relatives. Many international and global news sites have implemented their own news mobile applications making it indeed easy to keep in touch with the news from all around the world. Additionally, smartphone games have become popular among people from all ages [2]. Because of different games and for example video streaming applications such as Netflix, smartphones are used for entertainment reasons too rather than only for communicating. It is no wonder that the smartphones have become such a popular ubiquitous device worldwide. However, there is a negative side for this. Multiple studies [2, 3] have shown that people might develop addictive behaviour towards smartphones. High smartphone use is also known for causing problems with sleeping [4] and Lepp et al. [5] has shown that high smartphone use is positively related to anxiety.

The relation between smartphones and sleep has been a popular theme in studies of different viewpoints but the HCI context is still understudied. In this thesis we are viewing the relationship between smartphone use and sleep quality with crowdsensing approach. We are mapping the hourly application category usage against the different PHQ-8 questions response levels. The main aim for this is to identify hourly application usage from different application categories and examine if the different types of application use have relations with sleeping problems or feeling tired. The final data consists of single PHQ-8 question's response levels from 0 to 3 ("Not at all" to "Nearly every day") against the hourly crowdsensed application usage of the users from the separate response levels.

This thesis is based on the previous research [6] made for our WellComp-conference that was held on September 2020. The used raw data and the results of this thesis are the same as in the Sharmila et al. This thesis is an extension, that considers the data structuring and the methods for reaching the final results with more detailed manner.

The smartphone data was collected directly from the voluntarily participants devices via open-source phone data collecting platform called Carat. The data about the participant's sleeping problems were collected with Patient Health Questionnaire - 8 (PHQ-8) which has been used widely for different types of research. In the final results we are viewing the hourly smartphone category use against the PHQ-8 question 3, "Trouble in falling asleep or staying asleep, or sleeping too much" and examine whether there are differences in the application usage between the people who sleep differently. The results of this work could give us new insights for the future work, as there are not many researches about the hourly application use and its relations to the various problems in life, such as lack of motivation, tiredness or eating problems in the HCI context.

2. RELATED WORK

Mobile phones have become one of the most popular and used ubiquitous consumer devices all around the world [7]. People are having their mobile devices with them all the time of the day, even at night in bed. Because smartphones provide users many useful everyday solutions from communication tools to different applications to make their lives easier it is no wonder that people use a lot of time on their mobile phones. Although mobile phones provide many useful daily life's solutions there is always the other side of the aspect. There has been multiple studies about negative side effects of mobile phone usage such as mobile phone addiction [3] and high mobile phone usage is associated with sleep disturbance and depression [8]. Mobile phones, being a device that is carried to anywhere, is a great target of sensing people's daily life's habits such as movements through GPS-sensors [7] and sleep. With mobile phone sensing we can flexibly collect information about our participant's mobile usage and analyse for example the relation between participant's nightly mobile phone usage and sleeping.

There has been a wide range of different definitions for high and problematic mobile phone usage [9]. In this related work we are going to refer high mobile phone usage as highly increased mobile phone usage from the average mobile phone usage, and problematic mobile phone usage as more of a kind of a mobile phone addiction, that has actual impact to one's well-being and health. In the final results of the work, rather than focusing to a single user's smart phone use habits we are viewing the data as crowdsensed data.

Although high mobile phone usage is associated with different kinds of mental health problems such as depression, anxiety and increased stress, Katevas et al. [9] shows that high mobile phone usage alone does not predict one's negative wellbeing. Katevas et al. claimed that rather focusing on the intensive mobile phone use the focus should be in the way the mobile phone is being used. The study clustered five different types of mobile phone use. Cluster1 -limited use was the cluster that had the least mobile phone use, and it was used as baseline for other clusters. Cluster2 -business use was a cluster where the mobile phone users had frequent phone calls but low nighttime use. Cluster3 -power use was a cluster where users had increased duration of daytime use and nightly email application use. Cluster4 -personality-induced problematic phone use is a cluster where participants had increased nightly use and session length and Cluster5 -externally-included problematic phone use was basically the same as cluster 4 but use of messaging apps is greatly higher. Although participants from clusters 2 and 3 had rather high mobile phone usage at daytime and nighttime, and scored highest in amount of phone calls and session durability, the participant's PHQ-8 depression questionnaire answers showed no bias to the participant's negative well-being. Clusters 4 and 5 participants had less mobile phone use than cluster 3 participants, but they scored the highest scores in the PHQ-8 depression questionnaire. The main difference to the cluster 3 was that clusters 4 and 5 participants has highly increased nightly mobile phone usage. All together the study indicates that high mobile phone usage does not necessarily imply that the user has problems with mental health however participants with strongly increased nightly use of mobile phone had overall lack of interest in doing things and they felt tenser and worse than participants from clusters 1, 2 and 3.

Although high mobile phone use does not directly predict one's negative well-being, an explanatory study has been conducted that explains the relation between high frequency mobile phone and multiple problems in daily life. Lepp et al. [5] shows that problematic mobile phone usage was negatively related to academic performance (GPA) and positively related to anxiety. The same study shows that problematic mobile phone use has a negative impact to overall satisfaction with life. The mobile phone usage data was collected through questionnaires and the participants were asked to assess their daily mobile phone use hours and the amount they send and receive text messages throughout day. The satisfaction with life was assessed with SWLS (satisfaction with life scale), which was a questionnaire that contained questions about general satisfaction with life through 7-point scale from "Disagree strongly" (Coded as 1) to "Agree strongly" (Coded as 7). Higher score indicates more satisfaction with life.

Multiple studies [8, 10, 11, 4] has showed, that both high and problematic mobile phone usage has been associated with sleeping problems such as sleep disturbances and weakened sleep quality. Thomée et al [8] discovered a relation between increased sleep disturbances and high mobile phone use among both women and men between ages 20-24 for a 1-year follow up. The study pointed out, that one of the biggest reasons of the sleep disturbances was the feeling, that the person was expected to be reachable all around the clock. However, most of the participants did not consider the accessibility to be stressful. Foerster et al. [10] discovered with one-year follow up that adolescents who had at least one nocturnal awakenings from their own smartphone developed restless sleeping and troubles with falling asleep compared to the participants that did not have nocturnal awakenings. Hysing et al. [11] found out that among adolescents mobile phone use at evening affected to the later bed time and it generally affected the sleep latency (the time to fall asleep). After assessing adolescent's screen time and sleep habits for over month Perrault et al. [4] not only found out, that intensive mobile phone use among adolescents at evening and nighttime reduces the amount of sleep and sleep latency but it also has a negative influence to school performance, body weight, and overall mood of the adolescents. The study also reported that controlling the screen use of adolescents in the evening has a positive effect in sleep quality and overall daytime functioning.

The relations between smartphone use, depression, and sleeping problems has been a popular field of study as there has been various researches about the phone usage habits and health problems of different age groups. Smartphones can cause addiction in certain groups of users and when used as a compulsive habit, it can lower overall satisfaction with life. When smartphones used with an adequate manner, smartphone can become a reliable device that can have a positive impact to daily life.

3. METHODS

The main aim of this work is to examine participant's hourly application usage from different application categories and identify whether different manners with application uses have some relations with different types of sleeping habits or problems. Rather than viewing the participants as individuals, we are taking a crowdsensed approach.

The data about the sleeping habits or problems were collected through Patient Health Questionnaire-8. Every two weeks, the participants answered to the self-reported depression assessment questionnaire called PHQ-8 that contains eight questions about the feelings and moods of the participants during the last two-week period. The mobile phone data is collected two weeks prior to the day the questionnaire was answered meaning that we have mobile usage data about the two week's period the participants themselves assessed in the PHQ-8 questionnaire. With these components we can evaluate typical mobile phone usage of various groups of users that answered differently in the PHQ-8 questionnaire.

3.1. Smartphone Data Collecting

Application usage data was collected from 743 unique users using an open-source platform called Carat [12]. Carat is a downloadable mobile phone platform used to gather mobile usage data from the participants of the study. Carat collects data about currently running applications and whether the applications are running in the background or foreground. Carat does not collect data every second the mobile phone is used but rather every time the mobile phone's battery level drops by 1%, meaning that the samples of the data are applications that were running at the time when battery level dropped [13]. Each participant is given a unique user id and the application usage data for example the process id of the application is gathered along with the user's id's and the time and the date the data was collected. In this study rather than viewing all the applications as separate, we decided to categorise all the similar unique applications into different big categories.

3.2. Application Categorising

Participants were gathered around the globe and because of that we had a lot of different types of mobile application usage. Participants from different countries uses different social media and communication applications and therefore there is a possibility that we are only viewing the results from a certain group of participants in the final results. This is the reason it is not effective to view the results application-wise. Application categorising allows us to view the similar applications as whole one category and we can have the data from participants that might use applications that are not globally popular.

The used applications are categorised by Google Play-store's application categories. The process ids of the applications are as well fetched from the Google Play-store making it easy to categorise them by Google's categorisation. Google Play although has multiple categories to different gaming applications such as trivia games, strategy games and word games. All these gaming categories were merged to one big category called "games" in this study because viewing them as unique categories is not relevant and we can get a better overall picture of the mobile game's users. The final data consists of 24 different application categories such as weather, productivity, and transportation. In this work the main categories to be examined are communication, social, music, games, and productivity categories.

The application category usage data was binary data, that had the information of the time and the day the participants used their smartphones and which category's application they were using at that time. When the application was used, the binary number 1 was saved in the data file's corresponding column of the application's category. Other categories were saved as 0 at that time in the other application categories columns, because only one application can be run in the foreground at a time. Every row of the final data consists of the participant's unique user id, date and hour the data sample was collected, and the binary data about the application that was used at that moment. We will refer this data file later by file 'A'. The number of total amount of data collecting samples of different categories is calculated simply by summing all the data samples by categories together. In this work rather than summing all the samples together, we want to view the samples by users and all the 24 hours and find out which categories' applications were used in the previous two weeks period that the users assessed in the PHQ-8 questionnaire. With this information we can find out which time of the day people were using certain application categories and whether it correlates or not with the PHQ-8 questionnaire answers.

3.3. PHQ-8

PHQ-8 (Patient Health Questionnaire -8) is a self-reported questionnaire used mostly in population-based studies [14]. PHQ-8 assesses patient's possible depression level through 8 questions about their last two weeks performance of daily life. The questionnaire asks questions about the issues in daily life's habits such as eating, sleeping and the overall mood of the patient. The patient rates each question with the response level that best describes the last two weeks performance. The response levels are 0 (Not at all), 1 (Several days), 2 (More than half the days), or 3 (Nearly every day). The final PHQ-8 score is the sum of all the response levels of all 8 questions. The final score 10 or above is considered major depression and the score over 20 is considered severe major depression. In this work, rather than focusing to the final PHQ-8 score, we are focusing on the individual questions and their response level's relation to the application category usage with a crowd-sensing approach. We will study separately each question all the different response levels and categorise the users with that parameter for each PHQ-8 question. The main PHQ-8 questions we are focusing on in this work are the lack of interest in doing things, problems with sleeping such as sleeping too little or too much and feeling tired.

The PHQ-8 questionnaire was brought to the users through the questionnaire tool of the Android version of Carat. The platform provides the questionnaire to the participants and it maps the same unique user ids to the PHQ-8 answers as to the mobile phone usage data. The demographic information that was collected had single choice questions about gender, age, current occupation, highest completed education, household situation, yearly income compared to the living country's average, debt as percent to their income, savings and current location. With this information, we can bring together different demographic information about single users and compare them to other users [15].

3.4. Participants

The application category usage data and demographic data was collected from 743 individual participants around the world. The participants were asked to answer to a demographic questionnaire that contained questions about their age, gender, household situation, education, occupation, and country they are living in. The most important data in this demographic questionnaire is the gender and the age of the participants as with this information we can get a view of the main statistics about the participants and whether there is something to take account in the final results.

As there are a lot of users from different countries and continents, there is a big range of different applications and mobile phone users. Data categorising provides us a way to view all the similar types of applications that are used by different users.

The participant's gender distribution is clearly unevenly divided as the percent of male participants is 86.3% and the percent of female participants is 11.4%. The remaining 2.3% are the participants that answered, "Other or rather not to tell". This is although an aspect we have no control over as the study is conducted as a crowdsensed study and the participants are voluntary people across the globe.

The participants age distribution is not as unevenly divided as the gender distribution. The participants were divided into different age groups. 12.1% of the participants are 18-24 years old, about 32.6% are 25-33 years old, 25.0% belongs to the age group of 35-44 years, 26.2% belongs to the age group of 45-64 years and the remaining 4.0% belongs to the group of over 64 years old meaning that majority of the participants are people in student age or working age.

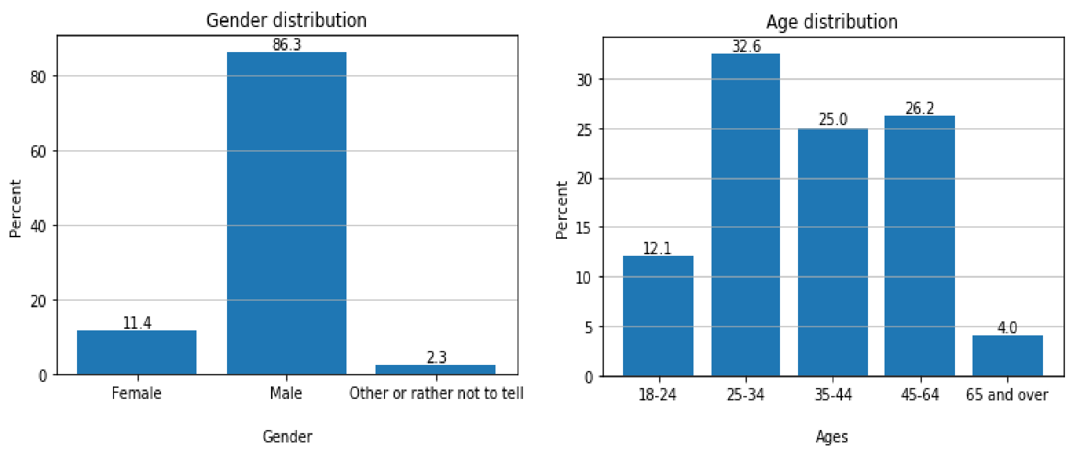

Figure 1. Participant's gender and age distributions

3.5. Data Structuring

In order to make further analyses about the application data, we need to give the raw data a new structure. At first, we had two different files. The file A contained the information about all the binary application category usage data and the date and the time when the data was collected. The file B contained the data about the dates when the single users answered to the PHQ-8 questionnaire and each response levels for each question and the total score of the PHQ-8. The data A is a raw data and needs to be structured in order to reach the information about the relation between application usage and PHQ-8 answers.

Firstly, the raw data (The file A) needs to be structured into two weeks periods. For every unique user, we take the data samples of the user that has been collected two weeks prior to their personal PHQ-8 answering date that are listed in the file B. Next, we take the users each two week's period's data and sort the data according to which hour the data was collected. For example, we take all the two week's data samples that were collected at hour 17. Lastly, we sum separately the samples of each application usage together by each hour and turn the sum into binary meaning that if the application has been used in that hour, the value is set to 1 and if it has not been used, the value will be set to 0. In the final data we have each user's two week hourly data, that tells if the application category has been used during that two-week period at that certain hour.

It is not productive to examine all the 24 different application categories as there are multiple categories that were not used regularly by participants. We decided to examine the categories that were the most popular among the participants or that were particularly interesting according to the theme of this research. The categories that are viewed in this work are communication, social, music, games, and productivity categories.

3.6. Data Visualisation

Data visualisation is one of the best ways to examine possible relations between the application usage data and the PHQ-8 questionnaire's answers. In this research we decided to compare and visualise the different usage of applications according to the PHQ-8 questions response levels. We are going to compare participants data depending on what PHQ-8 response level (0, 1, 2, or 3) they answered to the question in hand. The data visualisation is done for only one PHQ-8 questionnaire's question at a time. In this work we are focusing on the questions about problems with sleeping, lack of interest in doing things, and feeling tired or having a little energy.

In order to properly visualise the data, we need to divide the wanted application categories from the earlier structured data frame. We created own csv-files for all the wanted categories. The new csv-files include the participants user ids, the ending date of each two weeks period and the PHQ-8 answers for each 8 questions the participant assessed in the questionnaire. It also includes the binary hourly usage data separately from hours 0 to 23 of the application in question.

In order to make clear visualisations of the data, we decided to use heatmaps. Heatmaps are mainly used for graphical representation and analysing of data and heatmaps uses colour coding to represent different values [16]. In these heatmaps we are representing the differences of hourly application usage depending on the participants response level in certain PHQ-8 questionnaire's questions. It is not reasonable to compare the total values of application category usages between different response levels as the amount of response level answers differ a lot from each other. For example total of response level in question about lack of interest in doing things: 719 questionnaire response level answers were 0 (Not at all), and 115 response level answers were 3 (Nearly every day). If we compared these response level answered participants total application usage data, the data would be biased to the participants that answered response level 0, as there is simply more data and users, which would also affect the total application usage. To avoid this, we decided to make percentual analyses for the data.

To make percentual analyses for the data files, we need to first convert the total usage values to percents. At first, we have the total hourly application category usages among the different response values usage. To convert the total amount of usages into percents, we need to divide the hourly total usage values by the amount of users that answered the certain response value for the PHQ-8 question in hand. For example, let us view the PHQ-8 questionnaire's question number one (little interest or pleasure in doing things) and the communication application category: There were 34 times of total usage of communication category at hour 13 among the participants who answered response level 3 to the PHQ-8 question number one. There were 115 participants who assessed response level 3 to the question number one. Next, we divide the total usage (34) by the number of participants that answered the response level 3 (115) we can get the answer of 30%, which is the percent of participants who answered response level 3 and used the communication application category at the hour 13. We will do the same process for all the response levels and all the hours and we can compare the different percentages of usage. We lastly implemented a vector named 'all' that contained all the response levels of hourly application category usage divided by the amount of all the taken PHQ-8 questionnaires. We are presenting the hourly percentual vectors 0-3 and 'all' by application categories with heatmaps. This way we can evaluate and compare the graphical representations and search differences from for example nightly mobile phone usage among participants who answered differently in the question about problems with sleeping.

3.7. Statistical Analysis

Lastly we are going to analyse the coefficient of determination of the hourly usage vectors $X = x_0$ to x_{23} and the 'all' (Y) vector with linear regression. The linear regression is a linear approach to model relationships between dependent variables (Y) and independent variables (X). R^2-value is a measurement value that is used to measure the goodness of fit of the model. R^2-value determines how well the model fits the data on a scale 0-100%. In this study a high R^2-value indicates that participants that answered certain way to the PHQ-8 questionnaire are statistically using the application category similarly in comparison to 'all' participants together. Correspondingly if the R^2-value is low, it indicates that the participants that answered certain way to the PHQ-8 are statistically using applications differently to 'all' participants.

4. RESULTS

In this work we have gathered data from total of 743 participants and we are going to focus on the questions about lack of interest in doing things, problems in sleeping, and feeling tired or having a little energy. It should be noted that participants answered the PHQ-8 multiple times as the data collecting lasted longer than 2 weeks.

In the PHQ-8 questionnaire's question (question number 1) about lack of interest or pleasure in doing things, 719 of the answers was marked as the response level 0 (Not at all), 473 of the answers were the response level 1 (Several days), 138 was marked as response level 2 (More than half the days) and 115 were 3 (Nearly every day) the question number 1 distribution can be seen in the figure 2a)

The question about troubles falling asleep (question number 3) was marked 655 times as the response level 0 (Not at all), 485 as level 1 (Several days), 178 as level 2 (More than half the days) and lastly 127 answered 3 (Nearly every day). This distribution has similar trend as the previous distribution. The lower response level is directly proportional to the high number of answers meaning that generally there are more people that do not have problems with sleeping. The distribution for the question number 3 is shown in the figure 2b).

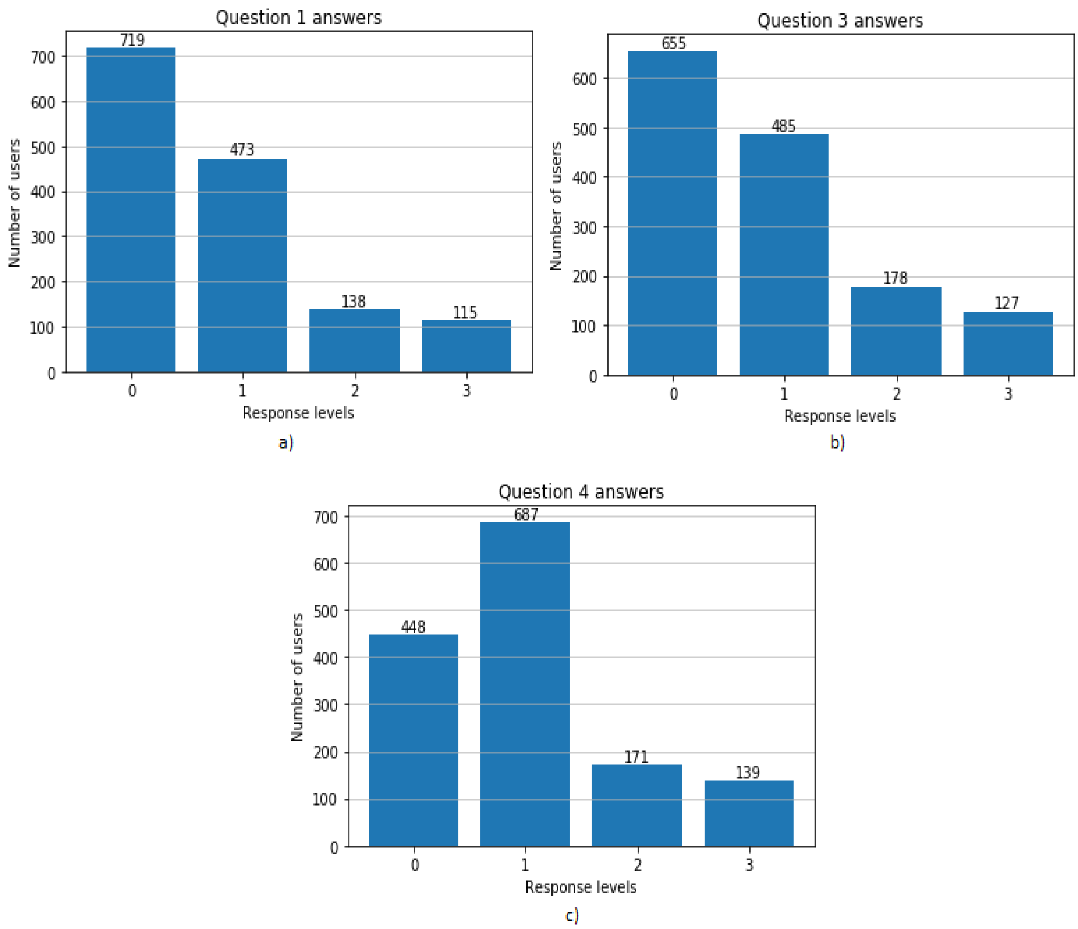

Figure 2. Different questions response level distributions

Lastly the question about feeling tired or having a little energy (question number 4) was marked 448 times as the response level 0 (Not at all), 687 times as level 1 (Several days), 171 times as 2 (More than half the days) and 139 times 3 (Nearly every day). In this question the response level 1 was clearly the most answered response level as also seen in the figure 2c). According to these numbers, generally people are feeling tired several days in the two weeks period than not at all.

We are going to view the results of the PHQ-8 and application category usage data from two viewpoints: firstly, we view the frequency of hourly category usage by the participants of different response levels of the question in hand. Secondly, we compare the different response levels to the 'all' category and study how well the frequency of certain response level can explain the overall category usage by all the users. We are going to use the R^2 statistics for the coefficient of determination analysis.

4.1. Communication Category

Communication applications were the most popular applications among the participants. The category covers various communication applications such as Facebook messenger which was one of the most popular among the users. The communication category was viewed from the perspective of two PHQ-8 questions: question 1 -"Little pleasure in doing things" and question 3 -"Trouble falling or staying asleep" or "sleeping too much".

In the figures 3 and 4 the vertical axis represents the PHQ-8 questions different response levels and the 'all' category, that contains all the users from different response levels. The horizontal axis shows hourly the percent of frequency that the communication category's applications were used. The 'all' category consists of all the category use divided with the number of all users. The colour of the segment represents the frequency of the usage. Segment with lighter colour has higher frequency of application category usage at that certain hour.

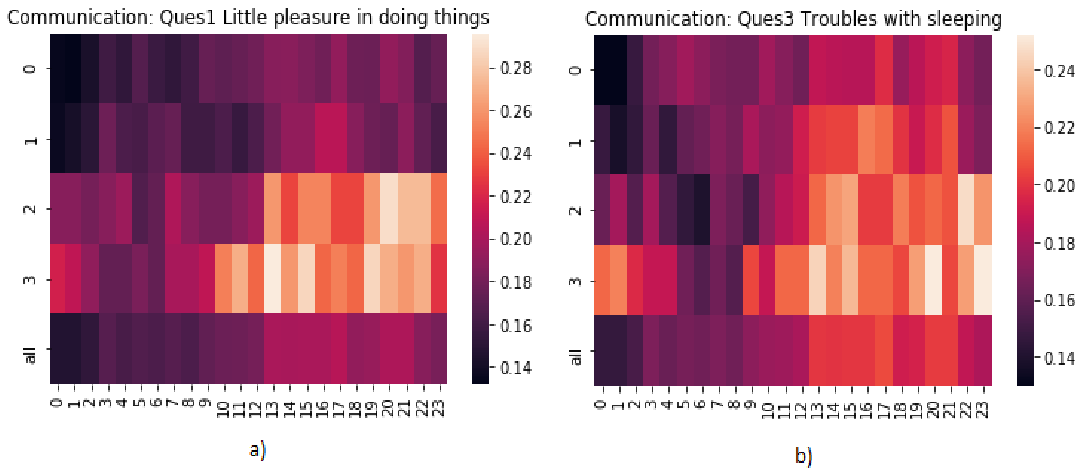

Figure 3. Hourly communication application category usage frequencies against the PHQ-8 question number 1: Little pleasure in doing things and question number 3: Trouble falling asleep or sleeping too much.

Figure 3 a) presents the communication category against the question 1, "Little pleasure in doing things". In the figure is shown that people, who answered the response level 2 "More than the half days" or response level 3 "Nearly every day" had significantly higher frequency of communication application use compared to the levels 0 and 1. This indicates that users that does not feel pleasure in doing things use much more smartphones than users that feel more motivated in general. Elhai et al. [17] found out, that boredom was positively related to the problematic mobile phone use which has a similar trend with the figure 3 a). In the figure in question can also be seen that the phone usage of the people of the response level groups 2 and 3 are using their communication applications at daytime. This might indicate that person who does not have interest in doing things is not only using their smart phones in their free time but at school or work too.

Figure 3 b) presents the Communication category against the PHQ-8 question 3, "Trouble falling or staying asleep or sleeping too much". The overall usage among all response levels is a bit more balanced than the use of the different response levels of the previous heatmap of the communication category and PHQ-8 question number 1. The significant difference between two of these heatmaps is the frequency of the nighttime application category usage. In the figure 3 b) is shown that people who answered to the sleeping problems question with the response 0, "Not at all" or 1, "Several days" has measurably lower frequency of nightly usage. On the contrary users that answered to the same question with the response level 3, "Nearly every day", has clearly the most nightly application usage meaning that users with problems with sleeping have significantly higher frequency of communication application usage at nighttime between 22 pm and 4 am. Communication applications are made for communicating with other people. People who have problems in falling asleep might start conversations via mobile phones with other people who have problems falling asleep. That would give us new viewpoints of the high frequency of the communication application nighttime usage among the people who are have sleeping problems. In addition, it is possible that people who use their smartphones at night because they have problems with falling asleep might have same types of problems with upcoming nights as their natural sleep pattern might switch forward.

We are using the R^2 statistics to determine, how well each response level of the question of the problems with sleep can explain the communication application category usage by the "all" usage. Comparing the response level 3 "Nearly every day" to the "all" category, the R^2-fit is lower (R^2 = 0.52) than the R^2-fit for the people who answered the "More than the half the days" (R^2 = 0.76), "Several days" (R^2 = 0.95) or "Not at all" (R^2 = 0.88). The higher R^2-value of the response level means that the response level in question is statistically similar to the overall usage of the category while the lower R^2-value indicates that the viewed response level is statistically different from the category usage of all users.

The relation between communication application use, and pleasure in doing this or sleeping problems is clear as seen in the figure 3. However, from these results it is not fully possible to determine, whether people are using communication applications because they are bored or not able to fall asleep, or they are bored and have sleeping problems because they are using mobile phone throughout the day. Finding an answer to this question would be rather eye-opening and it would give us directions towards new studies.

4.2. Music Category

Music has been studied to improve one's sleeping quality [18]. Harmat et al. found out that relaxing music such as classical music could work against insomnia and other sleeping problems. We viewed the music category from the perspective of two PHQ-8 questions: Question number 3 "Trouble falling asleep or staying asleep or sleeping too much" and question number 4 "Feeling tired or having a little energy".

The figure 4 a) shows the relation of the hourly music category use and the PHQ-8 question number 3. It can be seen that the music category is used with higher frequency throughout the day by the users that answered the response level 0, "Not at all" or response level 1, "Several days". However, is it important to note that the music application use at hours 10 pm to 5 am is clearly the highest among the users that answered the response level 3, "Nearly every day". The lowest music application use frequency is among the users who answered the response level 2, "More than the half days". The high music application use at daytime among the people who answered, "Not at all" or "Several days" might have a positive impact to the nightly sleeping quality and reducing the time between going to bed and falling asleep as the users claimed not to have any bigger problems with sleeping. Correspondingly the users that have problems with sleeping might be using the music as a medication at night. The R^2-fits of the users who answered, "More than the half days" (R^2 = 0.33) and "nearly every day" (R^2 = 0.13) were much lower than the users who answered "Several days" ((R^2 = 0.78) or "Not at all" (R^2 = 0.85). This means that the people who have problems with sleeping are using the music applications statistically in a different manner compared to the general use of the "all" category.

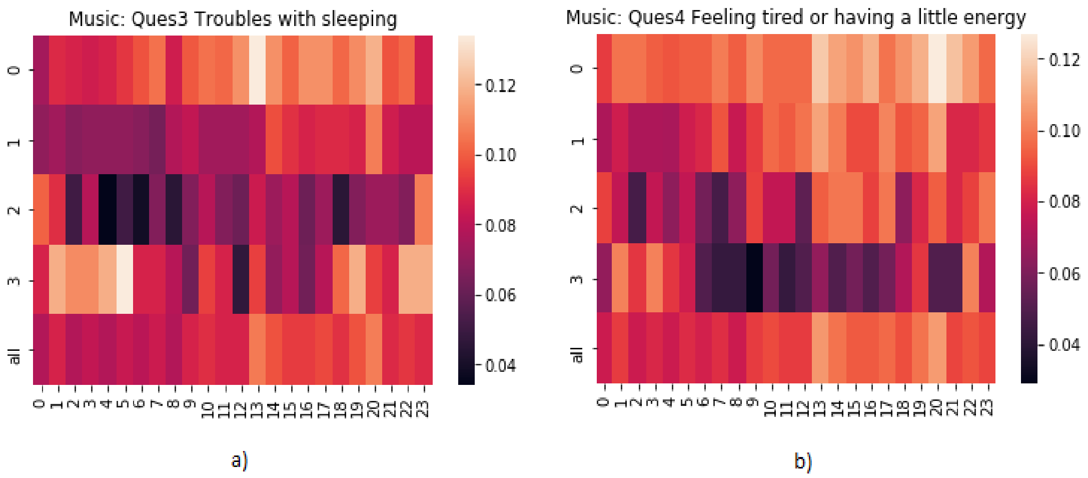

Figure 4. Hourly music application category usage frequencies against the PHQ-8 question number 3: "Trouble falling asleep or sleeping too much" and question number 4: "Feeling tired or having a little energy"

The figure 4 b) represents the relation between music category use and the PHQ-8 question number 4, "Feeling tired or having a little energy". The pattern of the heatmap 4 b) has similarities with the figure 4 a). The music application usage of the users that answered, "Not at all" and "Several days" are quite similar to the previous figure as the music application use is seemingly high throughout the full 24 hours of the day. The biggest difference between the two figures is the response levels 2 and 3. In the figure 4 b) participants that answered "More than the half days" are clearly using more music applications than the participants that answered "Nearly every day" but in the previous figure 4 a) the situation is different. The similarities between the two heatmaps can be seen from the R^2-values too. People who answered the response level 0 or 1 has seemingly high R^2-fit (0.84 and 0.87) and people from the response level 3 have low R^2-fit (R^2=0.12) meaning that the highest and lowest R^2-values are seemingly similar to the figure 4 a).

People who are listening to music throughout the day seem to be feeling less tired and having less problems with sleeping than the people who are listening less music. Both heatmaps in the figure 4 has similar patterns of the daily music application use of the response levels 0 and 1. There is also similarities in the nightly music application use of the response level 3 as in the both heatmaps is seen that the frequency of the music application use is highest at night time. One reason for this could be that people who are tired or have troubles with sleeping are using the music as a method to relax and fall asleep at night. People who are not feeling tired or do not have troubles with sleeping have high frequency of daily music application use. The reason could be either people listening to music because they are not feeling tired and slept well, or people sleeping well and feeling less tired because they are listening to music throughout day. Either way according to these result listening to music has positive impact to one's nightly sleeping and alertness.

5. SUMMARY

In this study total of 743 participants allowed the Carat-platform to collect smartphone usage data. The same participants answered the Patient Health Questionnaire-8, which allowed us to view the hourly application use and the participant's two weeks moods and habits together. The participant gender distribution was biased towards men, which might have some type of effect to the overall application use and the overall PHQ-8 scores. For the rest demographics participants were more evenly divided from different living conditions and household situations.

The results of the hourly frequencies against the PHQ-8 response levels were visualised by forming them into heatmaps, that shows the results of different frequencies with different colour codes. The statistical analysis for the data was implemented with the coefficient of determination (R^2-value) that was used to determine, whether a certain response level user's application use could explain the application use of the "all" category.

The heatmap visualisations and the R^2-values gave us rather interesting new information about the different mobile phone application use. The preliminary results showed us, that the users that felt little pleasure in doing things were using a lot of more communication applications throughout the day, even at nighttime. The communication applications were used significantly less by the people who were motivated and felt pleasure in doing things. In addition, the results gave us important information about the relation between sleep and smartphone use. The results showed that users who are feeling tired or have problems with sleep tend to have higher frequency of nightly music and communication application usage than the people who do not have problems with sleep or feel tired. Correspondingly people who does not feel tired or have problems in sleep are using their music applications more frequently and in a more balanced manner. These results give us a direction for new points of view as it would be eye-opening to find out, whether there is deeper causality between these variables.

For future work we are going to focus more to the relation between the PHQ-8 questions and examine whether the same users answered in a similar manner to different questions and whether the users have some types of similarities with smartphone use. In addition, there might be something to take account from the final PHQ-8 scores and the depression levels that were not reviewed in this paper. The direction is to find out more interesting results from the crowdsensed data and the new results can lead us to further sleep and mobile phone related research.

6. REFERENCES

[1] Peltonen E., Sharmila P., Asare K.O., Visuri A., Lagerspetz E. & Ferreira D. (2020) When phones get personal: Predicting big five personality traits from application usage. Pervasive and Mobile Computing 69, p. 101269.

[2] Chen C., Zhang K.Z., Gong X. & Lee M.K. (2019) Failure to decrease the addictive usage of information technologies: A theoretical model and empirical examination of smartphone game users. Computers in Human Behavior 92, pp. 256–265.

[3] Sahin S., Ozdemir K., Unsal A. & Temiz N. (2013) Evaluation of mobile phone addiction level and sleep quality in university students. Pakistan journal of medical sciences 29, p. 913.

[4] Perrault A.A., Bayer L., Peuvrier M., Afyouni A., Ghisletta P., Brockmann C., Spiridon M., Hulo Vesely S., Haller D.M., Pichon S. et al. (2019) Reducing the use of screen electronic devices in the evening is associated with improved sleep and daytime vigilance in adolescents. Sleep 42, p. zsz125.

[5] Lepp A., Barkley J.E. & Karpinski A.C. (2014) The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. Computers in human behavior 31, pp. 343–350.

[6] Sharmila P., Schroderus V., Lagerspetz E. & Peltonen E. (2020) Towards understanding smartphone usage and sleep with a crowdsensing approach. In: Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, pp. 700–703.

[7] Saeb S., Zhang M., Karr C.J., Schueller S.M., Corden M.E., Kording K.P. & Mohr D.C. (2015) Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. Journal of medical Internet research 17, p. e175.

[8] Thomée, Härenstam & Hagberg (2011) Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study. BMC Public Health 11, p. 66.

[9] Katevas K., Arapakis I. & Pielot M. (2018) Typical phone use habits: Intense use does not predict negative well-being. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 1–13.

[10] Foerster M., Henneke A., Chetty-Mhlanga S. & Röösli M. (2019) Impact of adolescents' screen time and nocturnal mobile phone-related awakenings on sleep and general health symptoms: a prospective cohort study. International journal of environmental research and public health 16, p. 518.

[11] Hysing M., Pallesen S., Stormark K.M., Jakobsen R., Lundervold A.J. & Sivertsen B. (2015) Sleep and use of electronic devices in adolescence: results from a large population-based study. BMJ open 5, p. e006748.

[12] Oliner A.J., Iyer A.P., Stoica I., Lagerspetz E. & Tarkoma S. (2013) Carat: Collaborative energy diagnosis for mobile devices. In: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, pp. 1–14.

[13] Peltonen E., Lagerspetz E., Nurmi P. & Tarkoma S. (2016) Too big to mail: On the way to publish large-scale mobile analytics data. In: 2016 IEEE International Conference on Big Data (Big Data), IEEE, pp. 2374–2377.

[14] Kroenke K., Strine T.W., Spitzer R.L., Williams J.B., Berry J.T. & Mokdad A.H. (2009) The phq-8 as a measure of current depression in the general population. Journal of affective disorders 114, pp. 163–173.

[15] Peltonen E., Lagerspetz E., Hamberg J., Mehrotra A., Musolesi M., Nurmi P. & Tarkoma S. (2018) The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 1–12.

[16] Pryke A., Mostaghim S. & Nazemi A. (2007) Heatmap visualization of population based multi objective algorithms. In: International Conference on Evolutionary Multi-Criterion Optimization, Springer, pp. 361–375.

[17] Elhai J.D., Vasquez J.K., Lustgarten S.D., Levine J.C. & Hall B.J. (2018) Proneness to boredom mediates relationships between problematic smartphone use with depression and anxiety severity. Social Science Computer Review 36, pp. 707–720.

[18] Harmat L., Takács J. & Bódizs R. (2008) Music improves sleep quality in students. Journal of advanced nursing 62, pp. 327–335.