Large-Scale Assessment of Mobile Notifications

Alireza Sahami Shirazi, Niels Henze, Tilman Dingler,
Martin Pielot[†], Dominik Weber, Albrecht Schmidt

VIS, University of Stuttgart
Stuttgart, Germany
{firstname.lastname}@vis.uni-stuttgart.de
{dominik.weber}@studi.informatik.uni-stuttgart.de

[†]Telefonica Research Barcelona
Barcelona, Spain
pielot@tid.es

= 重点内容

ABSTRACT

Notifications are a core feature of mobile phones. They inform users about a variety of events. Users may take immediate action or ignore them depending on the importance of a notification as well as their current context. The nature of notifications is manifold, applications use them both sparsely and frequently. In this paper we present the first large-scale analysis of mobile notifications with a focus on users' subjective perceptions. We derive a holistic picture of notifications on mobile phones by collecting close to 200 million notifications from more than 40,000 users. Using a data-driven approach, we break down what users like and dislike about notifications. Our results reveal differences in importance of notifications and how users value notifications from messaging apps as well as notifications that include information about people and events. Based on these results we derive a number of findings about the nature of notifications and guidelines to effectively use them.

Author Keywords

notification; mobile phone; large-scale; in the wild; apps; mobile HCI

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Notifications on mobile phones inform users about a variety of events, such as the arrival of a message, a new comment on one of their social network posts, or the availability of an application update. Some notifications, such as application updates, appear silently, whereas, others (e.g., incoming short messages) can create brief visual, auditory, and/or tactile signals to capture the user's attention. While in some cases immediate action is taken by the users, in other cases notifications are widely ignored depending on their importance and the user's current context. A unified method to present notifications is used on current iPhone and Android

smartphones. The notification is briefly presented in a notification bar at the top of the screen where other notifications may be listed. Through the unified presentation mechanism, notifications are widely used by a broad range of mobile applications. The nature of notifications is manifold, thus some applications use them sparsely while others literally flood the notification bar with their messages.

The effect of notifications created by applications in the desktop workplace, such as email clients or instant messengers applications, has been studied thoroughly in previous work [4, 5, 13, 16]. With mobile phones, however, notifications enter a new space, as they (1) are delivered to a highly unified mechanism, (2) inform about a much larger variety of events, ranging from messages to system events, and (3) became pervasive due to the omnipresent nature of current smartphones that are virtually always with the user [12, 21, 19].

Previous work on mobile notifications informs about the behavior of specific mobile applications or the notifications from specific application categories [8]. In contrast, we try to derive a holistic picture of mobile notifications. We report about a study, in which we collected information about the notifications from 40,191 people over the course of 8 months. Using a research-in-the-large approach [11, 10, 17], we developed an application called *Desktop Notification* and published it on Google Play. *Desktop Notification* forwards notifications from the phone to the PC's browser, where they are displayed as notification popups. After phone and browser have been paired, each notification appearing on the phone is shown in the browser. We not only record the type and arrival timestamp of each notification but also use the app to collect subjective feedback and qualitative comments on the importance of individual notifications in situ.

The core contribution of this work includes: (1) the first large-scale analysis of mobile notifications including users' subjective perceptions, (2) a detailed break down of what kind of notifications users like and dislike (3) the derivation of guidelines for designers about how to make effective use of notifications in mobile apps.

This paper is structured as follows: after providing an overview on the state of related work, we present the apparatus developed for data acquisition. We then describe the dataset collected and used for the analysis. Later, we present the objective and subjective analysis and report the findings. Finally, we conclude the paper with implications derived from the analysis.

RELATED WORK

Notification has been defined as a visual cue, auditory signal, or haptic alert generated by an application or service that relays information to a user outside of the current focus of attention [13]. Current major smartphone platforms offer a unified mechanism for notifications. They show all notifications in the notification bar that is located in the top of the screen. On modern smartphones, notifications play a key role in informing the user about a wide range of system activities, including the arrival of a message, a new update available, or an alarm by the calendar about an upcoming meeting.

As shown by previous work, a core problem with notifications is that users tend to drop their current task to check the notification instead. Studies by Iqbal *et al.* [13] and Mark et al. [16] show that information workers engage less in opportunistic email checking, multitask less, spend more consecutive time on tasks, and were are stressed, when disabling email notifications at work.

In the work context, this can be explained by the fact that notifications interrupt workers [1, 5, 4, 6]. Information workers find it difficult to return to a previous task after having been interrupted by *e.g.,* a notification [6], and they report higher subjective workloads when receiving notifications during phases of focused work [1]. The more cognitively demanding the work is, the more pronounced the effect [4]. In the context of phone usage, Leiva et al. [15] found that participants of their study spent significantly more time on an app, when they were interrupted by a phone call. Consequently, users report about being annoyed when receiving too many notifications on their phones [3].

Nevertheless, previous work has also shown that people value notifications. While participants in the study of Iqbal *et al.* [13] and Mark et al. [16] were aware of the notifications' disruptive effects, they appreciated the awareness that they provided. Sometimes, people even seem to be eager to receive notifications, as they keep frequently checking the phone [18]. Hence, most people would not want to disable notifications.

However, previous work also suggests that notifications vary in level of importance. According to a field study with 11 co-workers by Fischer *et al.* [9], the subjective importance of SMS-related notifications depends on how interesting, entertaining, relevant, and actionable a message is. If apps, which are not perceived as useful, keep sending notifications, users become annoyed and consider deleting those apps [7]. This suggests that learning about the importance of notifications for the user may be a feasible approach to deliver them in more appropriate ways.

In summary, previous work has shown that notifications can be interrupting and frustrating. Notifications are still valued as long as the content is interesting. Previous work, however, mainly focused on specific applications and cannot inform about general interaction with notification. What is missing is a deeper understanding of what notifications people are dealing with on their phones and the importance of different types of notifications.
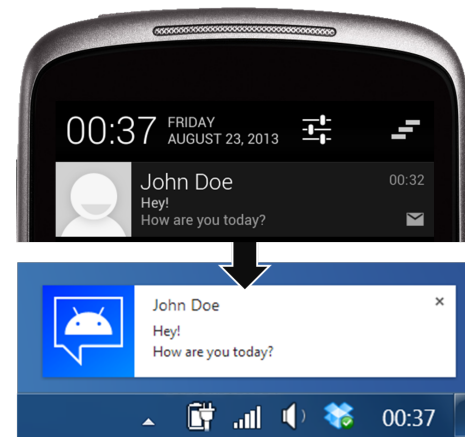

Figure 1: The *Desktop Notification* app automatically pushes notifications from the phone to the desktop computer.

DESKTOP NOTIFICATION

To collect information about notifications users receive on their phone and their interactions with them, we developed a system called *Desktop Notification*. The system bridges between the mobile phone and the desktop computer. It allows users to receive notifications from their phone on their desktop computer (see Figure 1). It consists of two parts: an Android application and a web browser plug-in. The Android application runs on an Android mobile phone (Android OS 2.2 and higher) and detects notifications appearing on the phone. By *notification* we refer to any message which is shown in the notification area or drawer of Android phones. This can include silent notifications, such as a notification about an available update for an installed app. After detecting new notifications, the app uploads their content to a central online database with the timestamp of its reception, the app's name, and the contained text.

To receive notifications on the PC, we developed a plug-in for Chrome and Firefox browsers. The plug-in periodically polls new notifications from the database that occurred over the most recent 2 minutes and displays them to the user. Generally, notifications are polled every 10 seconds. By default, a notification is shown for 7 seconds. However, the user is able to set the interval to 10 or 15 seconds. In case multiple notifications are retrieved simultaneously, they are displayed in parallel.

The app and the plug-in need to be paired for each user. Therefore, a unique private key is generated after installing the app on the phone (Figure 2). When users install the plug-in in the browser they need to enter that key in order to receive notifications on their PC. Notifications can be pushed to any operating system with a Chrome or a Firefox browser (e.g., Linux, Windows, OS X), hence the system works platform-independent. In the current implementation, the plug-in supports up to three browser instances and is able to poll notifications from a maximum of three smartphones. In that case a separate key is used for each device.

As the number of notifications received from apps on the mobile phone vary and have different importance for different
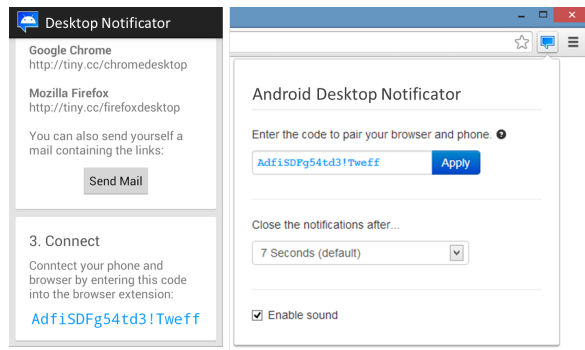
Figure 2: A secret code is generated after a user installed the *Desktop Notification* on the phone. The code must be entered in the browser plug-in to pair it with the phone.

users, we integrated a blacklist in the mobile app. The blacklist allows users to prevent notifications of certain apps from showing up in the browser by adding the apps to that list. If an app is in a user's blacklist, its notifications are ignored and not pushed to the connected computer. This enables users to administrate which apps' notifications they want to see on the desktop computer.

Participants Recruitment
We aimed to distribute the system widely to be able to derive a holistic picture of notifications on current smartphones. Therefore, we released the app in the mid of December 2012 in the Google Play marketplace for free. In the app's description in the marketplace we informed users that the app is part of a research project and that anonymous information is collected for research purposes.

According to Google Play's Developer Console, the app was installed 58,592 times up until the end of July 2013, and had 23,032 active installations at that time. The top three devices were Google Nexus 4 (11.11%), Samsung Galaxy S III (8.55%), and Samsung Galaxy S II (5.15%). The app was rated 684 times (Mean=4.44) and received 600 comments. The comments included new feature suggestions, reporting bugs, and qualitative feedback. Further, the app was listed in the top 200 apps in the Communication category of Google Play and was featured in several popular blogs.

Data Acquisition
Using the *Desktop Notification* app, different information can be collected from notifications, which users received on their phones. We recorded when a notification was shown and to which application it belonged. Users could explicitly click a notification message in the notification drawer. When the user clicked on a notification, we logged the time. Thus, we were able to calculate the time from when a notification was shown until the user explicitly clicked on it. We call this time *click time* from here on.

Furthermore, we collected subjective feedback from users. During the last two weeks of July 2013, we randomly added a button to notifications shown from certain apps in the browser. To avoid rare apps, we only chose apps which had been used by at least 50 unique users. By pressing the button,
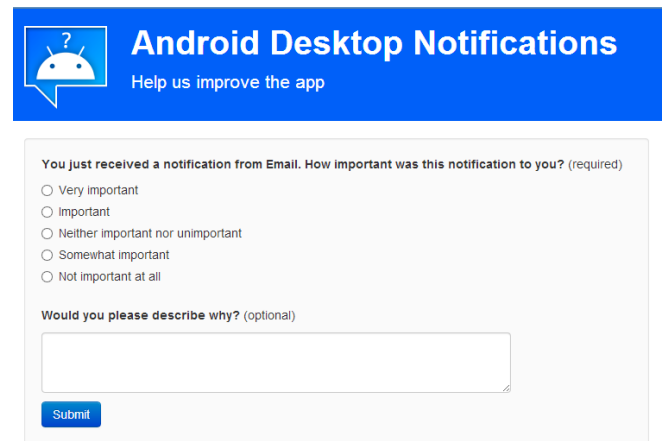

Figure 3: Questionnaire to collect subjective feedback.

a new tab was opened in the browser and a question was displayed (Figure 3). We asked the users to rate how important was the message received from this app using 5-point Likert scale (1=not important at all, 5= very important). Additionally, we asked users to optionally provide us with feedback on their rating using free text.

We were also interested in how many users used the blacklist feature and which type of apps were added to the blacklist. On August 12th we released an update for the app and collected the apps added to the blacklist from the users.

The text that accompanies notifications can include very sensitive information, including SMS bodies, email titles, and the user's contacts. To avoid ethical issues and preserve the users' privacy we did not collect any information about the content of notifications due to the high sensitivity of the information that can be included in notifications. Uploaded content of notifications is automatically and periodically deleted to preserve users' privacy and anonymity.

DATASET
Between the 10th of January 2013 and the 19th of July 2013 we collected 197,515,366 notifications from 40,191 unique users (253 unique locales) using *Desktop Notification*. In total 19.06% of the notifications stored in the database were polled by the plug-in and shown on the desktop computer. The three most frequent locales are en_US (29.7%), es_ES (10.35%), and en_GB (9.85%). The users used 20,014 unique apps from 30 different categories that we retrieved from the Google Play marketplace.

A large number of apps in the dataset were only used by very few users. Hence, we selected a subset of the data for the analysis. In the first step we only included apps with at least 400 users in a new dataset. This resulted in 173 apps from 23 categories (70.9% of the total number of notifications). Further, we realized the categories retrieved from the Google Play marketplace were too generic. For example, email apps (e.g., Gmail or Outlook), text messaging apps (e.g., Whatsapp or SMS apps), and Voice messaging apps (e.g., Google Hangout and Skype) were all in the same Communication category. However, previous work shows that they are used differently

	Number of			
Category	**Users**	**Apps**	**Click times**	**Ratings**
messenger	29,627	10	2,508,203	883
voice & messenger	27,768	9	335,530	470
mail	26,120	8	781,502	610
social	22,173	12	501,159	382
calendar	12,292	5	18,990	305
alarm clock	11,849	7	18,678	50
music	15,366	16	47,786	114
game	3,046	6	43,206	41
market	29,326	6	192,357	156
reader/news	3,958	7	11,036	50
utility	13,202	27	29,990	105
tool	11,494	18	33,352	42
system	27,269	21	229,067	304
other	13,511	21	44,370	124
overall	37,233	173	4,795,226	3,636

Table 1: The 14 categories of applications that we derived. It shows the number of unique users that used an app from the category, the number of apps, how often we observed that a user clicked on a notification, and the number of ratings we received.

[14]. Beside, prior work categorization is not publicly available or do not include all apps in our dataset [2]. This encouraged us to develop a more specific categorization scheme. To achieve this, two researchers independently reviewed the selected apps and each derived new categories. These new categories then were discussed and finalized, which resulted in 14 categories total. Table 1 shows the derived categories and provides an overview of the number of users, the apps in the respective category, the amount of click time data points, and the total number of ratings we collected through our questionnaire. In the following we use these categories for further analysis.

Constraints

One of the interesting aspects in the dataset is the number of notifications that the different apps show a user per day. In the Android OS design guideline it says: "*Think of notifications as a news channel that alerts the user to important events as they happen or a log that chronicles events while the user is not paying attention*". Apps use this feature based on their internal logic. When a user is voice chatting using the Skype app, for example, the app shows and updates the current duration of the ongoing call in the notification bar. The app achieves this by generating a new notification each second until the call is ended. Technically, this results in a very large number of notifications. On the other hand, the Kakao Talk app, for example, shows only a single notification when a user voice chats.

Due to this constraint, comparing apps by the number of notifications they show, is difficult. One approach to detect such a flooding is the analysis of the notifications' texts. However, as discussed above, we do not collect any information about the notifications' content due to the severe privacy implications. Thus, we cannot detect such flooding systematically.
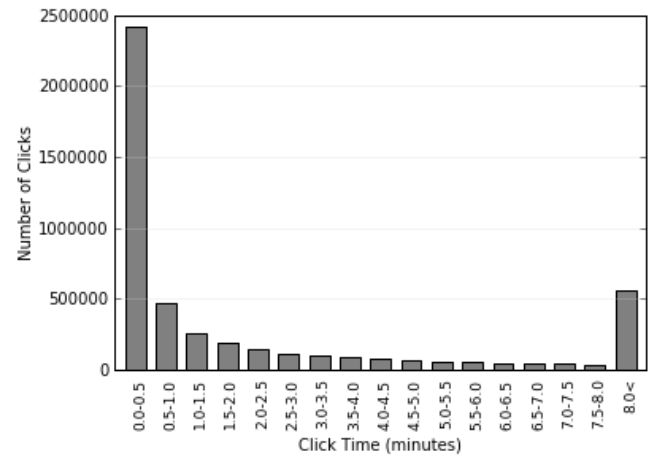

Figure 4: The histogram of the click time for all click data points. 50% of the interaction with notifications happened in the first 30 seconds.

In the result section, we assured to only consider apps that do not create a flood of notifications that result from updating existing ones. To achieve this, we downloaded all apps and tested their behavior manually. We ensured that the apps did not cause notification flooding due to continued actions on the phone, e.g., uploading/downloading a content, scanning the file system, or play music, etc.

OBJECTIVE MEASURES

In the following we analyze the data collected for 173 apps with at least 400 users each. Figure 1 provides an overview of the 14 categories that we use. A category had on average 17,627.64 users (SD=9,251.54) and 12.35 apps (SD=7). The utility category had the maximum number of apps (27 apps) and the calendar category had the minimum one (5 apps). A category had on average 342,516.14 reaction time data points (SD=663415.46). The messenger category had the maximum number of data points (2,508,203) followed by the voice & messenger (335,530) and mail categories (781,502).

Click time

To gather insights into how users act on notifications we tracked users' clicks on their mobile phones. So whenever a user tapped a notification, we derived the time that elapsed between presentation of the notification and the user's click. Figure 4 shows a histogram of the click time. If a user clicked on a notification, 50% happened in the first 30 seconds. If a user did not click on a notification in the first five minutes the probability that she will ever click on it is 17%.

We statically compared how long it took users to click on a notification for the 14 categories. Figure 5 shows the click time for all categories. After removing the outliers [20], we conducted a one-way ANOVA. Levene's test indicated that the assumption of homogeneity of variance had been violated, $F(1,13)=209,661.326$, $p<.001$. We compared the click time between the categories. The ANOVA show that there is a statistically significant difference between the categories
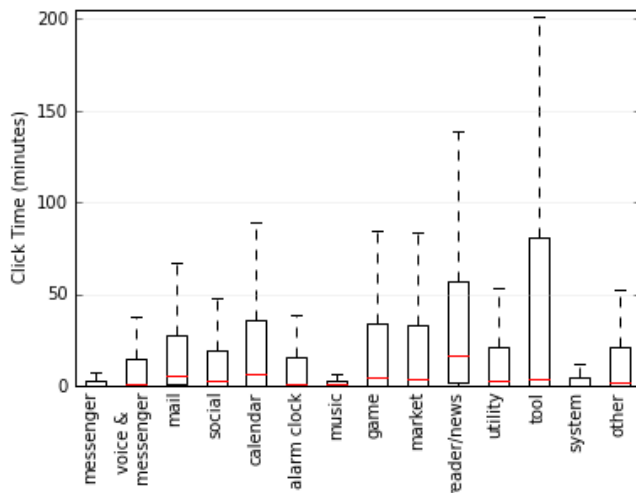
Figure 5: Click time for apps from 14 categories.

(F(1,13)= 79426.85, p<.001). As the categories have different variances and sample sizes, a Games-Howell post-hoc test was used for pairwise comparison. The post-hoc test revealed significant differences between all categories (p<0.01) except between calendar and game (p=0.82) as well as between social and Utility (p=0.67) categories.

The analysis of the click time between the categories showed that the system category had the shortest click time (Mdn=0.21 minutes) followed by the messenger category (Mdn=0.25 minutes). The native Android SMS (Mdn=0.24 minutes), WhatsApp (Mdn=0.25 minutes), and the native Motorola SMS (Mdn=0.26 minutes) apps had the shortest click time in the messenger category. Apps from the reader/news category had the longest click time (Mdn=16.30 minutes). In particular, Pulse News (Mdn=52.86 minutes), Google Reader (Mdn=18.61 minutes), and Appy Geek (Mdn=14.50 minutes) had the longest click time in this category.

Apps in Blacklist

Users can exclude notifications from showing up on the desktop by adding apps to a blacklist. Looking at these disabled apps we wanted to find out which kind of apps users appreciate notifications from and which kind is mostly found in the blacklist. We initialized the blacklist with a list of in total 164 apps and phone features that we had previously reviewed and found to be producing a lot of unnecessary notifications (flooding) which polluted our service. Examples include the aforementioned Skype app, which updates its notification bar with every second a call is held. Further examples include Mozilla's Firefox, Download, Viber, Dropsync–a Dropbox synchronization tool–, multiple weather apps, antivirus software, various battery indicators and power management tools. Users can add more apps to the blacklist as well as remove them. Apart from these flagged apps, users on average actively disabled 5.51% (SD=13.35) of all apps that raise notifications on their phone.
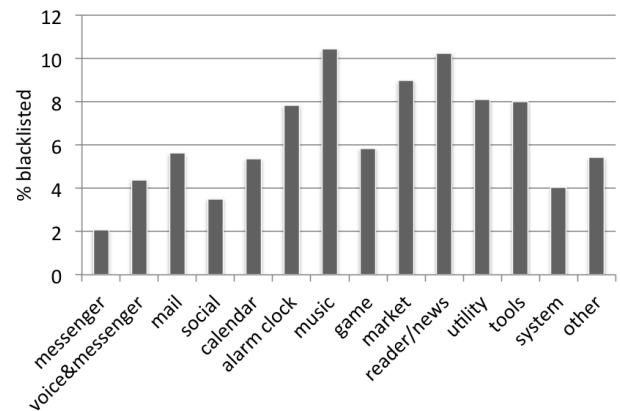

Figure 6: Percentage of apps per category that are actively blacklisted by users.

Figure 6 shows the percentage of apps that users actively added to the blacklist. Messengers and social apps are rarely disabled whereas apps in categories like music and reader/news seem to be blacklisted more frequently. In fact, 10.4% of blacklisted apps are music related. They seem to frequently notify users about changed titles, which may be perceived as annoying. Apps from the category reader/news (10.2% of blacklisted apps) update the notification bar quite often as well. Hence, frequency of notifications when the information given is not perceived as important, is a crucial factor that leads to increased blacklisting. Similarly, we observed apps in the categories alarm clock (7.8%), Utility (8.1%) and Tools (8.0%) to be blacklisted relatively often which confirms that internal system information and user-triggered actions are not of perceived importance.

Notifications per Day

In order to get a better understanding of the prevalence of different applications despite the flooding, we filtered out all flooding applications from the data set. In addition to many system services we could, most notably, not include notifications from the native telephone app, *Skype*, and *Viber*.

Table 2 shows the top 10 non-flooding applications that created most notifications for an average user at an average day. Only applications related to communication (messengers, email, social network) and reminders (todo list, calendar) made it on the list.

WhatsApp, by far, dominates the list, suggesting that this application is used most widely and most frequently amongst our participants. The second most frequent type of notifications comes from emails. The third most frequent type of notifications is SMS. These numbers show that facilitating communication is still one of the most importance function of a mobile phone.

Discussion

Two cases can be taken into account when considering notifications generated on the phone. A notification is received either due to an action the user explicitly started (e.g., downloading an app) or due to an external action (e.g., receiving

Category	Apps	Mean	SD
messenger	WhatsApp	19.9	64.5
mail	Gmail	7.0	15.8
messenger	SMS/MMS (Stock)	4.2	16.7
social	Facebook	3.7	23.1
messenger	Facebook Messenger	3.4	16.4
mail	Email (Stock)	1.7	17.8
calendar	Google Calendar	1.6	33.0
calendar	Any.do To-do List	1.5	15.5
social	Google Plus	1.0	8.9
calendar	Calendar (Stock)	0.9	19.1

Table 2: Top-10 Applications in terms of notifications generated per day. "Stock" refers to applications that are preinstalled on the Android OS.

an SMS). In the first case, the user actively uses the phone and (immediately) observers and clicks the notification. This leads to a short reaction time. This behavior can be observed for apps from system, utility, and tool categories. In the second case, the user may not actively use the phone. Even though the phone notifies users about incoming notifications using various modalities, such as audio and vibration feedback, the reaction can depend on the importance of the notification received as well as other factors such as user's context. The higher the importance of the notification, the shorter is the click time. Such a trend, for example, is observed for categories used for the communication purposes, i.e., Messenger, Voice& Messenger, and Mail.

The analysis of the reaction time reveals that the interaction with notifications generated on the phone are significantly different between categories. The probability that the user clicks the notification within the first five minutes is 83%. Surprisingly, with the 50% probability the interaction is taking place in the first 30 second. This reveals that notifications on phones are in general of great importance for users. However, this is highly dependent on the app's category. While click time for apps from the Messenger and System categories are the shortest, Reader/News apps have the longest click time.

The blacklist is a way to collect subjective feedback on the kind of apps from which notifications are generally appreciated or avoided by users. Besides the pre-selected list of apps which we blacklisted for flooding reasons, users have made active use of this feature in order to prevent notifications from certain applications to be pushed to the Desktop. Spamming apps which update the notification bar too frequently (music, reader/news) or often state the obvious (utilities, tools) are most likely to be blocked by users. They may make sense being displayed on the mobile device, but often loose their context when pushed to the desktop.

SUBJECTIVE FEEDBACK
We used the short questionnaire described above to ask users about the last notification they received. The questionnaire asked users to provide the importance of the notification using a five-point scale (1=unimportant to 5=important). In ad-

dition, a free-form text field asked to describe why the notification was important or unimportant (see Figure 3). In total, we collected 4,964 quantitative ratings and 796 qualitative answers from 4,816 different users.

Quantitative Feedback
We summarized the notifications' importance using the 14 categories described above. As the number of applications in the categories differ and the applications generate different numbers of notifications on the phone, the number of ratings also largely differs for the 14 categories (see Table 1). We received the smallest number of ratings for games and the most ratings for messengers. We conducted an ANOVA to determine if the importance of notifications from different categories significantly differs. We use a Games-Howell post-hoc test for pairwise comparison due to the varying sample sizes. The ANOVA shows that the importance of notifications importance rated significantly different between the categories ($F=(1,13)=117.26$, $p<0.001$).

Figure 7 shows the average rating for each category. Post-hoc tests revealed that the 14 categories can be separated in four groups. With an average importance of 4.43 (SD=1.04), users rated notifications from applications in messaging category as the most important ones. The notifications of several popular messaging applications received ratings above 4.5, including Sony Ericsson's SMS app (M=4.92), Handcent SMS (M=4.74), the native Android SMS app (M=4.68), and WhatsApp (M=4.53). Notifications with the lowest ratings were created by the Facebook Messenger (M=3.55) and Snapchat (M=3.93). The post-hoc test revealed that notifications generated by messaging applications are rated significantly more important than notifications from all other categories (p<0.001).

Notifications from applications in the other three communication categories receives similar ratings (voice & messenger M=3.66, SD=1.63; mail M=3.56, SD=1.50; social M=3.45, SD=1.51). Post-hoc tests revealed no significant differences between the three categories and no significant difference with calendar category but with all other categories (p<0.001). In the voice & messaging category the notifications from Line (M=4.67) and Google Voice (M=4.67) apps received the highest rating, whereas, Skype (M=2.58) received the lowest rating. In the mail category, the Outlook app received the highest ratings (M=4.21) and the K-9 Mail app received the lowest (M=2.54). Notifications from the Instagram (M=4.10) and Twitter (M=4.05) apps received the highest ratings in the social category, while Foursquare (M=2.00) clearly received the lowest.

The importance ratings for the calendars (M=3.50, SD=1.57) and the alarm clocks (M=2.07, SD=1.61) clearly differ. The ratings of notifications' importance from the calendar do not significantly differ from the importance of notifications from the three communication categories described above but the difference is significant for all other categories (p<0.001). The alarm clock category only significantly differs from all communication categories and the calendar (p<0.001). Apart from system notifications, the notifications from the remaining categories (music M=2.61, SD=1.58; game M=2.34,
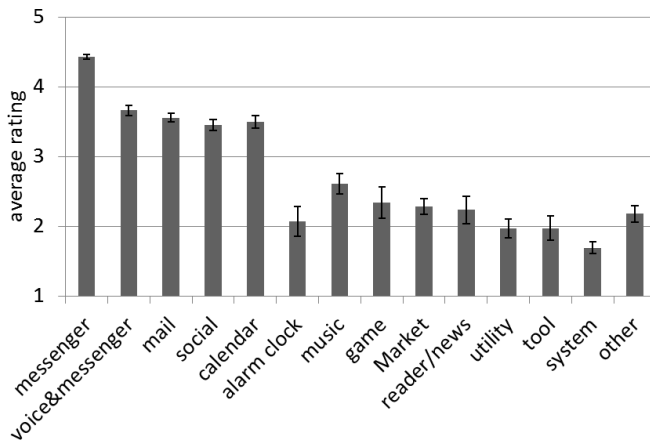
Figure 7: Rated importance of the notification generated by apps from the 14 categories. Error bars show the standard error.

SD=1.42; market M=2.28, SD=1.39; reader/news M=2.24, SD=1.39; utility M=1.97, SD=1.37; tool M=1.97, SD=1.43) do not significantly differ from each other. Their notifications are all rated significantly less important than notifications from the communication categories and more important than the system notifications. Finally, notifications from the system (M=1.69, SD=1.26) are rated significantly less important than all other categories (p<0.001).

Qualitative Feedback

We analyzed the qualitative data by manually assessing the users' statements in the questionnaire. First, two researchers went independently through 795 qualitative answers to derive a classification. Afterwards, they discussed the classes and agreed on a consistent set of classes. They agreed to classify the comments along two dimensions: the sentiment of the comment (annoyed, happy, effective, empowered, or neutral) as well as nine reasons for this sentiment (see Table 3). Afterwards, the researchers went again through the answers and assigned one or more classes to each answer without considering the application that created the notification or the rating of the notification's importance. The average quantitative rating for all questionnaires with a qualitative answer is 2.99 (SD=1.70) on the five-point scale.

Notification's frequency
We classified 91 comments that addressed the frequency of notifications. This includes comments that criticized to receive a particular notification at all. Exemplary comments are *"Daily update, i know that it happens so the notification is not really important."*, *"email notifications is important, but i also get it via other ways. the way the phone is handling it now is annoying because it keeps showing it to me everytime the email count changes"*, and *"too often"*. Notifications in this class are mainly rated of little importance (M=1.45, SD=0.91).

Notification's usefulness
In total 174 comments mentioned that the received notification is useful or not and often also provide reasons for the usefulness. Examples include *"Although useful to know, it is*

Class	Count	Rating	Rating SD
Frequency	91	1.45	0.91
Usefulness	174	3.16	1.60
Relevance on computer	143	2.37	1.61
Self-initiated actions	23	1.43	0.84
System internals	101	1.78	1.23
About an event	27	4.07	1.27
Context or contacts	72	4.50	0.98
Description of app	418	3.16	1.69
Apparatus related	60	3.32	1.71

Table 3: The classes of qualitative comments we received through the questionnaire.

not critical.", *"happening all the time, no relevant info"*, and *"It's not that useful and quite unnecessary (to me)."*. Notifications in this class are rated somewhat important on average (M=3.16, SD=1.60).

Notification's relevance on computer
Users get the notification also on their computer due to the nature of the apparatus. As a result, 143 comments addressed the relevance of the notification on their computer. Comments include *"gmail already has his own notification system"*, *"I have fb chat on my computer"*, and *"I couldn't hear someone called me, but I saw it on my computer"*. Notifications in this class are rated fairly important on average (M=2.37, SD=1.61).

Notification about self-initiated actions
Twenty three comments explained that the received notification was about a self-initiated action. One user, for example, commented *"I used computer to send to my phone. Why would i want it pop up here again?*, and another *"i activated it, so I know that it happened. i don't need a notification telling me that it happened"*. With an average importance of 1.43 (SD=0.84) comments about self-initiated actions are least important compared to the other classes.

Notification about system events
Notifications about system events are addressed by 101 comments. Such comments include *"I don't want to be bothered with Google Play Store App Update Notifications or Installation Notifications"* and *"I hate this popup. I don't want to be notified about my keyboard opening/closing."*. On average, notification in this class seem of little importance (M=1.78, SD=1.23).

Notification about an event
In total 27 comments address a real life event, such as a meeting or an upcoming task. Examples include *"Appointment to meet my parents for dinner. Very important not to forget :)"*, *"Notification of my dental appointment later in the day."*, and *"Repetitive action on a set schedule. I only need to know about unique events."*. Notifications in this class are rated important on average (M=4.07, SD=1.27).

Description of context and contacts
We classified 72 comments that describe the context that makes the notification important or not import (e.g. *"I am in I.T. at work and use these notifications quite a bit to keep in*

contact with everything I do.") as well as comments that provide information about the person or group of persons that the notification is related to (e.g. *"Text from my mother"*). Further examples include *"My Princess!"* and *"Because it came from someone that posts approx. once every half hour"*. Notification in this class are considered very important on average (M=4.50, SD=0.98).

Description of the notification's cause

With 418 comments, almost half of the analyzed comments stated the application that created the notification (e.g. just *"gmail"* or *"Task from ANY DO"*) or described the type of application (e.g. *"Text message"* or *"Weather info that is displayed in my notifications bar."*). On average, notification in this class seem fairly important (M=3.16, SD=1.69).

Comment on the apparatus

Sixty comments were related to the apparatus. In particular, users used this channel for feature requests (e.g. *"It'll perfect if I can answer for whatsapp and sms like in krome app"* or *"Nice to know, btw it would be nice to see app icons from android and to be able to ignore apps chrome side"*) or general comments such as *"I'm using your app to replace windows 8 notification system that was bogging down my computer."*. On average, notification in this class have been rated fairly important (M=3.16, SD=1.69).

Correlations with Importance

We analyzed the correlation between click time and importance of notifications. Hereby, we considered the apps in categories which had both rating and click time information. Pearson's correlation was computed to assess the relation between reaction time and rating. The result showed a negative correlation between the two variables r=-0.27, p=0.02. Thus, the click time is shorter when a notification is more important.

Blacklisted apps tend to receive lower ratings in our questionnaire with regards to notifications' relevance. Hereby, Pearson's correlation is negative, r=-0.62, n=63, p<0.001, which shows that low ratings are in line with blacklisted apps since the perceived relevance is low. The qualitative analysis aforementioned confirms this as well.

Discussion

From the ratings alone, the 14 categories can be split into four groups. With a rated importance of 4.43 the messenger applications clearly stand out. They are more important than notifications from all other categories. Notification from the three other communication categories and from the calendar also receive fairly high ratings with averages between 3.66 and 3.45. With an importance of 1.68, notification from the system clearly received the lowest rating. Ratings for notifications from the other categories provide a diffuse picture with music applications receiving a rating of 2.61 and with 1.98 utilities receiving the lowest rating in this group.

The classification of the qualitative comments provides insights for the (un)importance of notifications. Main aspects that lead to a low importance are: notifications being received too frequently (M=1.45), the notifications about actions initiated by the user (M=1.43), and notifications about system

events (M=1.78). For these three classes, there is a negative correlation with the notifications' rating of importance (r=-0.33, r=-0.16, and r=-0.27). If a comment contains one of the three aspects the rating is significantly lower (p<0.001). Comments that provide the notification's context (e.g. that it is work related) or information about a user's contacts are of high importance (M=4.50). Similarly, if a real-life event is mentioned in the comment the notification is rated more important. For these two aspects, there is a significant positive (p<0.001) correlation (r=0.28, r=0.12) with the notifications' rating of importance.

The messenger, mail, voice & messenger, and social categories received the clear majority of comments that provided information about the user's contacts or the real-life context of the notification's cause. This seems to be the main factor for their importance. The calendar received almost all comments about events but only four comments provided information about context or contacts. It may be concluded that notifications are perceived as important if they notify about communication with other persons, inform about other persons' actions, or about events. Main factors why notifications are considered unimportant are the frequency with which they are created, if they inform about actions that have been self-initiated by the user, or if they provide information about the phone's internals. This can explain the very low importance of notifications from the system. They are created too frequently, and users find notifications about the phone's internals not important in general. We assume that the alarm clocks are a special case. Their notifications also received a very low rating. From the comments, we assume that the most likely explanation is that users do not need the notification as they already have been informed through other means. One user, for example, commented: *"Nice to see what my next alarm is. However, the nature of an alarm is that it reminds and alerts. I dont' really need a notification in addition to an alarm."*.

FINDINGS

Users receive a truly large number of notifications on their smartphones. There is a large number of apps that create notifications and their nature is diverse. Thus, the notifications serve diverse purposes. Previous work points out the disruptive nature of notifications and, in line with our qualitative results, also shows that users value notifications nonetheless. Thus, it is necessary to find a balance between disrupting the user and providing valuable information. In the following we list the major findings derived from the analysis.

Nature of notifications is disruptive. It has been shown that notifications about incoming mails on desktop computers reduce users' overall performance because they distract from other tasks [4, 5, 13, 16]. Our results show that if a user clicks on a notification this typically happens in less than 30 seconds. It is not only true for messaging apps but it is true for notifications created by apps from most categories. Thus, if a user clicks on a notification, the time between seeing it and clicking on it is short. Thus, our results support the assumption that mobile notification are similarly disruptive like notifications shown on desktop computers. This might be even

worse than for desktop applications as smartphones accompany the users throughout the day.

Important notifications do not necessarily cause immediate attention. A significant correlation between notification's rated importance and click time shows that the time until a user engages with notifications that are important is shorter than for notifications that are not important. Thus, our results support the assumption that important notification cause immediate attention. The correlation is, however, weak. It can be concluded that there is a difference between being important and requiring immediate attention.

Notifications are for messaging. Our results clearly show how much users value notifications from messaging apps. The notifications are considered highly important, users rarely want to block them and the click time after a notification is created, is typically very short. The clear difference for all measures between emails apps and messaging apps also shows that they are in fact different even though the two categories provide almost the same functionalities. Previous work assessed notifications on the desktop might not be applicable to mobile notifications in general. The clear difference between mail and messaging apps that we found, however, further shows that these are in fact different application categories.

Important notifications are about people and events. We observed a diverse user behavior in terms of click time and blacklisting for apps from the different categories. The subjective results are, however, clear. Notifications from apps that can be used for communication with others as well as calendars are rated significantly more important than notifications from all other categories. This is in line with the results derived from the qualitative data. Notifications are considered important if they are about events or provide information about the user's context or contacts.

Not all notifications are important. It is rather obvious that not all notifications are similarly important. The subjective results, however, provide reasons for the unimportance of notifications: Notifying the user about an action that she/he initiated herself/himself makes little sense, providing a large number of notifications is not appreciated by users, and notifications about system events are also not valued. None of these factors is utterly surprising. What is surprising is the large number of notifications that are created that fit these criteria.

Recommendations for Developers

Our data indicates that mobile phone users have to deal with plenty of notifications each day. Subjective feedback shows that some of them are highly valued but others are not appreciated. Previous work on notifications for desktop computers shows that notifications are disruptive and our data suggests that this is also true for the mobile domain. For users' benefit, a system should therefore present only important notifications. The user is however not the only stakeholder. App developers can use notifications to capture users' attention and encourage repeated use. Furthermore, current systems

provide users with little support when it comes to managing notifications.

With current operating systems for smartphones, every app is free to create notifications. While, for example, the Android platform has a rather detailed permission system it does not include permissions for notifications. Users can disable notifications for specific apps but this feature is rather hidden and requires multiple steps. Thus, operating systems should enable users to easily manage notifications from selected apps. In addition, the system should enable app developers to assign priorities to notifications so that users can decide to only receive notifications of a selected importance.

Developers can use notifications to provoke repeated use of their app. As we showed, not all notifications are appreciated and thus creating them can have an inverse effect. Therefore, app developers need to be careful about the notifications they create. If an app creates a notification it should be ensured that the content is important and relevant for the user. This is typically the case if the notification reports incoming messages, real-life events, or is related to the user's contacts. For a game, for example, it would be better to report that a friend installed the game rather than reporting an unknown person achieved a new high score.

Users should make active use of mechanisms that prevent notifications. For most apps, notifications can only be completely enabled or disabled. Some apps such as the messenger *WhatsApp* provide users with more sophisticated means to adjust which notifications are created. In general, users need to be aware that notifications are an important tool to stay connected through their phone, but can also be a significant attention-grabbing distractor.

Limitations

The findings of this paper are mainly limited by three factors: First, the application *Desktop Notification* that we used as the apparatus to collect the analyzed data certainly influenced the results. Using *Desktop Notification* can, for example, influence the time until a user clicks on the notifications. Having a notification already received and read on the computer could reduce the motivation to click on it to open an application on the phone. Yet, with a total median click time of 30 seconds, and a median click time of a few minutes for most categories, we believe that this situation happened only infrequently, and hence, did not bias the reported findings significantly. Furthermore, only 19% of all collected notifications have been shown on the connected computer and even for these notifications not all were actually seen by the user.

Second, users might have specifically considered the importance of notifications on the computer when rating and blacklisting them. Due to the nature of the *Desktop Notification* apparatus users might have blacklisted apps because they do not want the notifications to their desktop computer even though the notifications are valuable on the phone. We tried to minimize this effect by looking at a variety of measures and asking about the importance of the specific notification. In addition, we assume that asking about specific notifications in-situ is in general more valid than using other means of determining their importance (e.g. through surveys). Finally, we assessed

the value and importance of notifications through a number of objective and subjective data and assume that through these different means we are able to generalize beyond users of our apparatus.

Third, the application biased the user base that we recruited as participants. Users interested in our application certainly differ from the general population. Thus, the participants might have different interests, behaviors, and opinions compared to the average population. However, smartphone users in general also differ from the general population. In addition, we recruited a very large sample with 40,191 participants using 253 different locales. As such our sample is more diverse than those in typical controlled studies.

CONCLUSION
In this paper, we conducted a large-scale assessment of mobile notifications. Through an application that enables users to connect their smartphones with their desktop computer we collected almost 200 million notifications generated on the mobile phones of more than 40,000 users. We further collected 4,964 quantitative ratings of notifications and 796 qualitative comments.

We showed through objective and subjective measures that notifications created by apps from different categories are valued differently by users. Results revealed, for example, a difference between messaging and email apps in terms of all analyzed measures. Overall, users value notifications from messengers, other communication apps, and calendars while they do not value notifications from system applications. By analyzing users' subjective feedback we show why notifications are considered important by users.

Future work should investigate if the perceived importance of notifications can be assessed automatically. This will enable users to more easily specify which notifications they want to receive. Combined with information about the user's context it would further be possible to deliver notifications when they were most appreciated and had the least negative impact on users.

REFERENCES
1. Adamczyk, P. D., and Bailey, B. P. If not now, when?: the effects of interruption at different moments within task execution. In *Proc. CHI* (2004).

2. Böhmer, M., Hecht, B., Schöning, J., Krüger, A., and Bauer, G. Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In *Proc. of MobileHCI* (2011).

3. Church, K., and de Oliveira, R. What's up with whatsapp? comparing mobile instant messaging behaviors with traditional sms. In *Proc. MobileHCI* (2013).

4. Cutrell, E., Czerwinski, M., and Horvitz, E. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In *Proc. Interact* (2001).

5. Czerwinski, M., Cutrell, E., and Horvitz, E. Instant messaging and interruption: Influence of task type on performance. In *Proc. OZCHI* (2000).

6. Czerwinski, M., Horvitz, E., and Wilhite, S. A diary study of task switching and interruptions. In *Proc. CHI* (2004).

7. Felt, A. P., Egelman, S., and Wagner, D. I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. In *Proc. SPSM* (2012).

8. Fischer, J. E., Greenhalgh, C., and Benford, S. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. In *Proc. MobileHCI* (2011).

9. Fischer, J. E., Yee, N., Bellotti, V., Good, N., Benford, S., and Greenhalgh, C. Effects of content and time of delivery on receptivity to mobile interruptions. In *Proc. MobileHCI* (2010).

10. Henze, N., and Pielot, M. App stores: external validity for mobile hci. *interactions 20*, 2 (2013).

11. Henze, N., Pielot, M., Poppinga, B., Schinke, T., and Boll, S. My app is an experiment: Experience from user studies in mobile app stores. *IJMHCI* (2011).

12. Ichikawa, F., Chipchase, J., and Grignani, R. Where's the phone? a study of mobile phone location in public spaces. In *Proc. Mobility* (2005).

13. Iqbal, S. T., and Horvitz, E. Notifications and awareness: a field study of alert usage and preferences. In *Proc. CSCW* (2010).

14. Kim, H., Kim, G. J., Park, H. W., and Rice, R. E. Configurations of relationships in different media: Ftf, email, instant messenger, mobile phone, and sms. *Computer-Mediated Communication* (2007).

15. Leiva, L., Böhmer, M., Gehring, S., and Krüger, A. Back to the app: the costs of mobile application interruptions. In *Proc. MobileHCI* (2012).

16. Mark, G., Voida, S., and Cardello, A. "a pace not dictated by electrons": an empirical study of work without email. In *Proc. CHI* (2012).

17. McMillan, D., Morrison, A., Brown, O., Hall, M., and Chalmers, M. Further into the wild: Running worldwide trials of mobile systems. In *Proc. Pervasive* (2010).

18. Oulasvirta, A., Rattenbury, T., Ma, L., and Raita, E. Habits make smartphone use more pervasive. *Personal Ubiquitous Computing* (2012).

19. Pielot, M., de Oliveira, R., Kwak, H., and Oliver, N. Didn't you see my message? predicting attentiveness to mobile instant messages. In *Proc. CHI* (2014).

20. Tukey, J. W. *Exploratory data analysis*. 1977.

21. Wiese, J., Saponas, T. S., and Brush, A. Phonedprioception: enabling mobile phones to infer where they are kept. In *Proc. CHI* (2013).