Master Thesis: Deep Activity Tracking on Android Smartphones

Smartphones are supporting us in our everyday lives, e.g. by providing tailored information on demand anywhere and anytime. However the smartphone mostly acts as a "passive servant", meaning that users have to actively request help or information. Smartphones can only rarely act proactively themselves yet.

The basis for proactive support is intent prediction: The smartphone needs to know what the user might want without the user explicitly stating it. This might be impossible to achieve for any arbitrary situation, however for context dependent, repetitive tasks smartphones could learn from past behaviors to predict user desires in the future.

Therefore a broad picture of the users activities (e.g. Which apps are used? What is the user doing in the app?), state (e.g. cognitive load, emotion, stress level), and context (e.g. Where is the user? What is he currently doing in the real world?) is needed. While context aware applications are meanwhile supported by device APIs [e.g. the Google Snapshot API[1]] and assessing user's (mental) state is an active research topic with many proposed solutions, work on collecting smartphone usage is superficial: I.e. most work satisfies with knowledge on what apps are used, but do not regard what users are actually doing inside an app. However "deep app tracking" could empower intent prediction systems to not just prompt the next used app, but instead offer a deep link that directly performs a user action.

Example: Tom commutes to his office daily by public transport. He has multiple connection options to choose of, depending on when exactly he is leaving from home and whether there are any service issues on one of the connections. Thus he checks the public transport company's app every morning, querying for the fastest connection from home to work by entering his home- and office location. Current intelligent devices would learn that Tom uses the public transport company app every morning, and thus place its icon prominently on the screen on mornings. However with the ability of deep app tracking, the device could learn that Tom is querying the fastest connection from his home to office every morning. The device could offer a deep link, i.e. an icon that directly performs the desired action.

Existing approaches for deep app tracking evaluate screenshots [e.g. the Screenomics project[2]]. Although that approach is comprehensive, it is very privacy invasive as raw screen contents are transferred to a server. Furthermore it's resulting data is hard to use, as features describing higher level actions have to be extracted with computer vision algorithms. A different approach is enabled by Android's accessibility service[3]: An app can retrieve the window's XML structure and subscribe to interaction events. However this approach does not work for all views, as for example the content of web views and prerendered content is not represented in the XML structure. Furthermore this vast amount of unstructured data is hard to gain valuable information from.

The research question on this topic could be: How can we collect deep app usage data on Android smartphones in the wild, in a privacy friendly, usable and scalable manner?

- How can the data resulting from the Android accessibility service be processed, to extract user-level actions and intents?
- Can we use the accessibility service's data to train an image recognition model that can preprocess screenshots? Can the trained screenshot processing model be used to extract features about apps that cannot be accessed via the accessibility service?

---

[1] https://developers.google.com/awareness/android-api/snapshot-api-overview
[2] http://clgiles.ist.psu.edu/pubs/HCI2019.pdf
[3] https://developer.android.com/reference/android/accessibilityservice/AccessibilityService#retrieving-window-content

You are a good fit for this topic if you…

- Have advanced Android programming skills and are confident with data analysis (e.g. in R or Python)
- Love to dive deep into unstructured data and are eager to extract insights from it
- Have a creative and proactive mindset to come up with ideas and approaches to tackle problems
- Are in your third (or higher) semester of your Master studies in media informatics, computer science, or a related field

If you are interested, please drop me a message including some information why you feel you are a good fit for this topic and a transcript of records to florian.bemmann@ifi.lmu.de !