A Computational Method for Evaluating UI Patterns

Bardia Doosti[1], Tao Dong[2], Biplab Deka[3], Jeffrey Nichols[2]

[1]Indiana University Bloomington
bdoosti@Indiana.edu

[2]Google Inc.
{taodong, jwnichols}@google.com

[3]University of Illinois at Urbana-Champaign
deka2@illinois.edu

Scan me

Purpose

We introduce a method for measuring the impact of UI design languages and informing design debates. In particular, we demonstrate this method using two common UI patterns in Material Design:

- Floating Action Button
- Navigation Drawer (Hamburger Menu)

Dataset

We used the Rico dataset [1], which includes over 72,000 Android UI screens from over 9,700 popular apps across 27 categories on the Google Play Store in early 2017.

For each UI screen, the dataset contains the following:

- Screenshot as a JPEG file.
- View Hierarchy as a JSON file.

We also used every app's average rating and number of downloads from Google Play Store as proxies of their quality.

Methodology

Step 1:

- Used a generic keyword to find each material design element via JSON file.
- But some of the items found this way was irrelevant.

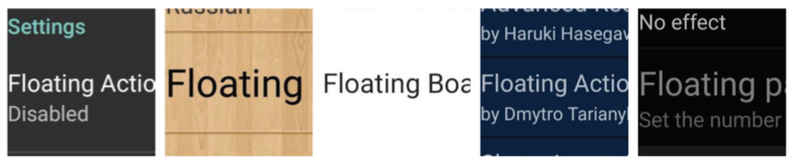


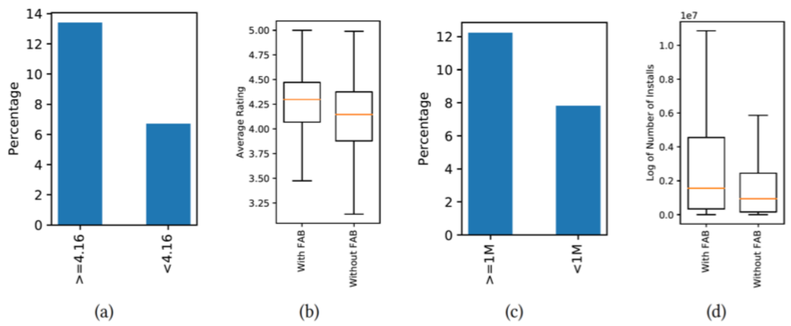
Figure 1: Items found with keyword "Float", irrelevant to "Floating Action Button"

Step 2:

- Detecting false positives using their JPEG file.
- Trained a Convolutional Neural Network for each Material Design which is responsible to detect that element.
- Got >95% accuracy for each Material Design element.

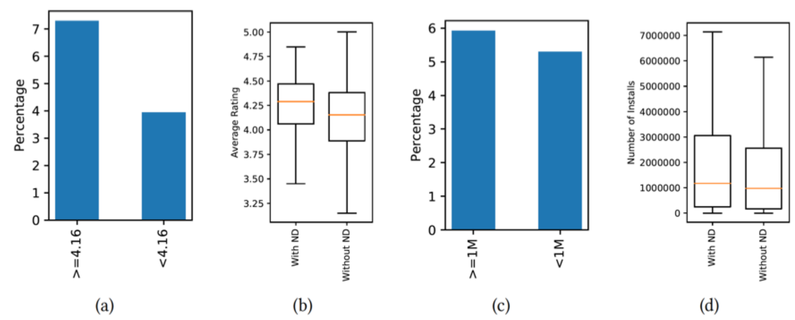
Results

Usage of Floating Action Buttons



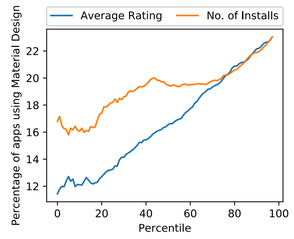
(a) The percentage of apps using the Floating Action Button (FAB) in the high-rating group vs. the low-rating group.
(b) Box plots of the average ratings of apps using the FAB vs. those not using the FAB.
(c) The percentage of apps using the FAB in the more-installed group vs. the less-installed group.
(d) Box plots of the number of installs of apps using the FAB vs. those not using the FAB.

Usage of Navigation Drawers



(a) The percentage of apps using the Navigation Drawer in the high-rating group vs. the low-rating group.
(b) Box plots of the average ratings of apps using the Navigation Drawer vs. those not using the Navigation Drawer.
(c) The percentage of apps using the Navigation Drawer in the more-installed group vs. the less-installed groups.
(d) Box plots of the number of installs of apps using the Navigation Drawer vs. those not using the Navigation Drawer.

Material Design and App Quality



Distribution of the percentage of apps using at least one of the six common Material Design elements over percentiles in average rating (blue) and number of installs (orange)

The usage of Material Design was highly correlated with the average rating percentile ($\rho = 0.99$ and p-value = 3.1×10^{-91}) and number of installs ($\rho = 0.94$ and p-value = 2.3×10^{-47})

Conclusion

- We developed a computational method to examine the use of UI design pattern languages in the wild by combining app hierarchy, screenshots, and marketplace metadata.
- Our data analysis showed that this method was effective in demonstrating the overall impact of a pattern language, Material Design, on the quality of apps in the Android ecosystem.

References

[1] Biplab Deka, et al. Rico: A Mobile App Dataset for Building Data-Driven Design Applications. In 30th Annual Symposium on User Interface Software and Technology (UIST '17)