Designing and Evaluating Mobile User Interfaces

Info 152 Spring 2010
Guest Lecturer: Jeffrey Nichols
February 24, 2010

Last Time

• User Interface Design Process

• Mobile Design Guidelines
Today

- Mobile Design User Study
- Mobile UI Evaluation

Mobile Design User Study

- Study of 21 iPhone and Android G1 users at Almaden Research Center
- Before and after interviews
  - Impact of smart phone acquisition
- IBM Research Report
  [http://tinyurl.com/yatzdqw](http://tinyurl.com/yatzdqw)
Lessons Learned: Interaction Patterns

• Typical mobile interaction is very short
  – < 1 minute for most interactions (making vs. filling time)
  – Consumption predominates

• Context shapes and constrains use
  – What people do and how long they have to do it

• Interruptions are common

• Common interaction patterns:
  – Monitoring
  – Information seeking
  – Information availing / entertainment

Lessons Learned: Access, Initialize, Work times
Example: check email while out

Example: reply to email while out
Example: reply to email at your desk

Lessons Learned: Minimize Initialization Time

• Interactions < 60 seconds, so use time productively
  – Little control over access time, so focus on initialization time

• Native vs. web services
  – Native typically initialize much more quickly

• Need for improved mobile usable security
  – Passcode entry is a hit on initialization time
  – 10 seconds to enter an 8 character alphanumeric password is 15% of interaction time
  – IBM Security policy a “showstopper” for ~85% of our user pool
    • Won’t use iPhone for work if that’s the necessary cost
    • Most transactions are NOT work-related, and don’t want to take hit for those non-work tasks (e.g., check weather, Twitter)
Usable mobile security

• Partial authentication for partial access
  – # new messages vs. contents

• Graphical authentication mechanisms

• Lots more work needed
  – e.g., Separate work/personal apps

Lesson Learned: Travel Shifts Interaction Patterns

• Impacts access time
  – People are more likely to be away from their laptops/desktops for hours / days
  – People engage in work tasks on their phones that they would otherwise defer to their laptops / desktops
    • e.g., writing email messages, blogging

• Information needs change
  – More likely to need directions, restaurant recommendations, store hours, etc.
  – More coordination with friends / colleagues
  – More sharing of experiences
Lessons Learned: Activities Span Devices

• Actions deferred from one device to another

• Example: E-mail
  – Monitor and triage on mobile devices
    • Users spend most of their time in the inbox
      – Form intended actions for carrying out on laptop/desktop
  – Coping behaviors
    • Explicit marking of read messages as unread
    • Open a message just to mark it read
  – Deferred sending from desktop devices

Evaluating Mobile User Interfaces
Evaluation Methods

**Expert Evaluation**
- Cognitive Walkthrough
- Heuristic Evaluation

**User Evaluation in the Lab**
- Think Aloud Study
- Performance Evaluation

**User Evaluation in the Field**
- Diary Study
- Experience Sampling Method (ESM)
- Logging Study

Expert Evaluation
Cognitive Walkthrough

Formalized technique for imagining user’s thoughts and actions when using an interface

• Given detailed description of interface
• Select task
• Tell a story motivating user actions required to do task
• Interface should give motivations via prompts/feedback
• Breakdown in motivations imply problem with interface

Walkthroughs are difficult to do when tasks are ill defined and can be accomplished in many ways

Heuristic Evaluation

Developed by Jakob Nielsen (1994)

Can be performed on working UI or sketches

Small set (3-5) of evaluators (experts) examine UI
  – Check compliance with usability heuristics
  – Different evaluators will find different problems
  – Evaluators only communicate afterwards to aggregate findings
  – Use violations to redesign/fix problems
Heuristics

H2-1: Visibility of system status
H2-2: Match system and real world
H2-3: User control and freedom
H2-4: Consistency and standards
H2-5: Error prevention
H2-6: Recognition rather than recall
H2-7: Flexibility and efficiency of use
H2-8: Aesthetic and minimalist design
H2-9: Help users recognize, diagnose and recover from errors
H2-10: Help and documentation

Examples

Can’t copy info from one window to another
   – Violates “Minimize the users’ memory load” (H1-3)
   – Fix: allow copying

Typography uses mix of upper/lower case formats and fonts
   – Violates “Consistency and standards” (H2-4)
   – Slows users down
   – Fix: pick a single format for entire interface

   – Probably wouldn’t be found by user testing
Why Multiple Evaluators?

- Every evaluator doesn’t find every problem
- Good evaluators find both easy & hard ones

Expert vs. User Evaluation

Expert evaluation is much faster
  - 1-2 hours each evaluator vs. days-weeks

Doesn’t require interpreting user’s actions

User testing is far more accurate (by def.)
  - Takes into account actual users and tasks
  - HE may miss problems & find “false positives”

Good to alternate between expert & user-based testing
  - Find different problems
  - Don’t waste participants
User Evaluation in the Lab

Lab Testing

• Bring real people into the lab
• Participants perform tasks with user interface
• Often pre- and post-questionnaires to elicit extra feedback
Two Kinds of Lab Studies

“Think Aloud” Study
- Focus is on qualitative data
- Find confusing and difficult elements of UI
- Typically used early in the design process

Performance Study
- Focus is on quantitative data
- Measure completion time, errors
- Typically used later in the design process
- Requires a point of comparison
  (earlier or competitor’s designs)

“Think Aloud” Study

Key: Ask users to say everything they are thinking aloud
- Remind them when they forget

As with any other user study, be careful not to introduce bias
- Let users work through any confusion,
  don’t help them
- Avoid laughing, making facial expressions, etc.
Note: Can Run Study At Any Time

Don’t need a complete implementation (or any implementation at all!)

Wizard of Oz Technique

Human operator mimics advanced computational functionality

– Speech recognition, gesture recognition, vision, etc.

Allows for testing advanced functionality without full implementation
The Problem with Lab Studies

Lab studies cannot perfectly simulate the real world

Inherently artificial
  – Artificial setting
  – Artificial motivation
  – (in the case of mobile) Artificial events
  – Etc.

User Evaluation in the Field
Three Field Study Methods

1. Diary Study
   - Capture activities from users’ real mobile environments
   - Users jot down responses to questions given beforehand in a journal
   - Relies on self-reporting, so may miss some data points

Adapted from Tim Sohn, Nokia Research Center
Experience Sampling Method (ESM)

• Developed in the mid 70’s by Csikszentmihalyi, Larson, and Prescott

• Useful for studying mobile users because it gathers data in their environment

• Beep users several times a day to answer questions
  – At random
  – After a relevant event

Adapted from Tim Sohn, Nokia Research Center

Logging Study

• Install program on users’ devices to capture data

• Infer user activity or intent from log data

• Capture data that may be missed through self-reporting methods

• Obviously, many privacy concerns!

Adapted from Tim Sohn, Nokia Research Center
Summary

Once you’ve started building your mobile interface, there are many ways to test it for usability

– Expert evaluation methods should be used early and often to find basic problems

– User evaluations can be conducted in the lab, and are useful for identifying domain-specific problems

– Once you have a deployable application, field evaluations can help you understand how your application is used in context