The Design and Evaluation of a Task-Centered Battery Interface

Khai N. Truong\textsuperscript{1}, Julie A. Kientz\textsuperscript{2}, Timothy Sohn\textsuperscript{3}, Alyssa Rosenzweig\textsuperscript{1}, Amanda Fonville\textsuperscript{2}, Tim Smith\textsuperscript{1}

\textsuperscript{1}University of Toronto  
Toronto, Ontario, Canada  
\{khai, alyssacr, tsmith\}@cs.toronto.edu

\textsuperscript{2}University of Washington  
Seattle, Washington, USA  
\{jkientz, ajf32\}@uw.edu

\textsuperscript{3}Nokia Research Center  
Palo Alto, California, USA  
tim.sohn@nokia.com

Presented by Kevin Lo
Motivation

• Battery interfaces provide detailed information on how much usage time left.

• Default interface only gives a high-level overview; not always reliable or accurate.

• User needs to create her own mental model of how much longer she can use this application

• Most interfaces only display time for 1 app.
Design of TCBI

- Existing indicators inadequate. Goal:
  “To provide users with an estimated amount of time remaining for applications running independently or in combination to eliminate guess work”

- Conducted a survey with 104 participants to understand users’ battery usage habits

- 4-week user study to evaluate effectiveness of TCBI
Related Work

• Various research on energy conservation
  – Power management
  – Energy efficient programs
  – Improving human-battery interaction (TCBI)

• Commercial Applications:
  – EZ Battery Life
  – Battery Time
  – Battery Magic
Battery Usage Survey

• Online survey consisted of 26 questions
  – demographics, different portable devices used, use of the battery interface and its perceived accuracy and usefulness, etc.

• 52 male and 52 female participants with diverse backgrounds through email mailing lists and online advertisements

• Incentive: Chance to win 1 of 4 $25 GCs
Device Usage

- 95% own a smartphone/cellphone/PDA and 40% own 2 or more devices.
- 93% own a laptop; 43% own 2 or more.
- 72% own an Internet tablet or media player; 7% own something else (e.g. Amazon Kindle)
- Average of 4.26 devices/person (SD: 2.22)
Device Usage (p2)

• Most use their devices on a daily basis; 6% smartphone & 8% laptop users on a weekly/monthly basis

• Only 57% of Internet tablet/media player users use those devices on a daily basis
Battery Death

- Phones die more frequent while idle than other mobile devices
- 21 participants’ phones die at least once a week; 49 once or more per month
- Less frequent battery deaths while in use:
  - 16 participants reported at least once per week; 29 at least once per month
Low Battery Behavior

- B1: I limit my use to emergency/absolutely necessary situations
- B2: I adjust device settings (e.g., brightness, display time) to conserve power
- B3: I change which applications I am using (e.g., only use text messaging, stop Internet use, or phone conversations)
- B4: I turn off wireless capabilities
- B5: I take measures to preserve my current activities and/or saved documents
- B6: I do not take actions to conserve power and I continue to use it as if I did not receive the message
- B7: Other behavior

Figure 1: Breakdown of responses for low-battery behavior when user does not have access to an external power source.
Better Battery Interface

“The laptop's battery indicator is so unreliable. It says something like 30% remaining and then it starts warning 3% level at the next moment.” –P58

Figure 2: Participant agreement levels when asked “I think the current battery interface on my portable device would be more useful if it showed information that relates specifically to the application I am using.”
Implementation

• Prototype built on Nokia N810

Figure 3: The Task-Centered Battery Interface. The left figure (a) shows the battery time remaining when none of the common applications is running. The right figure (b) shows the battery time remaining when the user is listening to music and estimates.
Implementation (p2)

• Python status bar applet

• Obtains 2 piece of data when launched:
  – Battery’s current voltage reading
  – List of active processes

• Record time-stamped voltage readings every 10 sec while an application runs alone.
Implementation (p3)

- 6 traces for each application; brightness and volume set to 50%, with brand new battery

- For each set of application traces:
  - Average the logged time remaining for each voltage capacity reading
  - Perform a smoothing function
  - Cross-validate each set of traces against their respective battery usage profile
Results – Single Application

• Predicted battery time within an average of 738.37 sec (12.7%) of actual battery time

Table 1: The average error for each application’s battery usage profile. Magnitude error is the average difference between the predicted battery time remaining with the actual battery time remaining. Percentage error is the average percentage difference between the predicted battery time remaining with the actual battery time remaining.

<table>
<thead>
<tr>
<th>Application</th>
<th>Magnitude Error</th>
<th>Percentage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS</td>
<td>615.88 sec</td>
<td>13.0%</td>
</tr>
<tr>
<td>Music</td>
<td>859.55 sec</td>
<td>14.4%</td>
</tr>
<tr>
<td>Skype</td>
<td>685.77 sec</td>
<td>26.7%</td>
</tr>
<tr>
<td>Video</td>
<td>649.25 sec</td>
<td>13.8%</td>
</tr>
<tr>
<td>Web</td>
<td>1400.41 sec</td>
<td>21.6%</td>
</tr>
<tr>
<td>Nothing</td>
<td>1696.12 sec</td>
<td>11.9%</td>
</tr>
</tbody>
</table>
Results – Multiple Applications

• Predicted battery time within an average of 769.3 sec (17.1%) of actual battery time

\[
t_{predicted} = C \times \frac{1}{(1/t_1 + \cdots + 1/t_n)}, \text{ where } \\
C = \frac{(n + 1)}{n}, \text{ and } \\
n = \text{the number of applications running.}
\]

Table 2: The average error for our multi-application battery time remaining estimation function.

<table>
<thead>
<tr>
<th>Application</th>
<th>Magnitude Error</th>
<th>Percentage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS + Music</td>
<td>388.30 sec</td>
<td>9.4%</td>
</tr>
<tr>
<td>GPS + Skype</td>
<td>500.21 sec</td>
<td>18.1%</td>
</tr>
<tr>
<td>GPS + Video</td>
<td>378.08 sec</td>
<td>8.9%</td>
</tr>
<tr>
<td>Web + Music</td>
<td>981.17 sec</td>
<td>21.9%</td>
</tr>
<tr>
<td>Web + Skype</td>
<td>895.76 sec</td>
<td>21.4%</td>
</tr>
<tr>
<td>Web + GPS + Music</td>
<td>1472.42 sec</td>
<td>23.1%</td>
</tr>
</tbody>
</table>
Deployment Study

- 2 phases, 1 interface (standard vs. TCBI) per phase

Table 3: Timeline of user study across the 4-week deployment.

<table>
<thead>
<tr>
<th>Day</th>
<th>Study Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>First Meeting: Study explanation, initial interview, first battery interface deployed</td>
</tr>
<tr>
<td>1</td>
<td>Task on web browsing, survey #1-1</td>
</tr>
<tr>
<td>4</td>
<td>Task on Skype, survey #1-2</td>
</tr>
<tr>
<td>7</td>
<td>Task on listening to music, survey #1-3</td>
</tr>
<tr>
<td>10</td>
<td>Task on watching videos, survey #1-4</td>
</tr>
<tr>
<td>13</td>
<td>Task on both web browsing and listening to music, survey #1-5</td>
</tr>
<tr>
<td>14</td>
<td>Second Meeting: Interview on first interface, logs copied, switch to second interface</td>
</tr>
<tr>
<td>15</td>
<td>Task on web browsing, survey #2-1</td>
</tr>
<tr>
<td>18</td>
<td>Task on Skype, survey #2-2</td>
</tr>
<tr>
<td>21</td>
<td>Task on listening to music, survey #2-3</td>
</tr>
<tr>
<td>24</td>
<td>Task on watching videos, survey #2-4</td>
</tr>
<tr>
<td>27</td>
<td>Task on both web browsing and listening to music, survey #2-5</td>
</tr>
<tr>
<td>28</td>
<td>Final Meeting: Interview on second interface, logs copied, devices and journals collected</td>
</tr>
</tbody>
</table>
Deployment Study (p2)

- Send emails on specific day with specific task

- Participants answer set of 7 questions before & after task:
  - Current Time & Battery remaining
  - Prediction on whether the task could be completed; Likert scale used to rate predictions

- Logging software used to track participation
Deployment Study (p3)

- 8 final participants; $60 compensation for completing the study

Table 4: Participants and demographics for the 8 final participants from the TCBI deployment study.

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>Age</th>
<th>Occupation</th>
<th>Computer Experience</th>
<th>City</th>
<th>First Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP1</td>
<td>M</td>
<td>20-29</td>
<td>Student</td>
<td>Exp.</td>
<td>Toronto</td>
<td>TCBI</td>
</tr>
<tr>
<td>DP2</td>
<td>M</td>
<td>20-29</td>
<td>Student</td>
<td>Exp.</td>
<td>Toronto</td>
<td>Nokia</td>
</tr>
<tr>
<td>DP3</td>
<td>F</td>
<td>20-29</td>
<td>Student</td>
<td>Inter.</td>
<td>Toronto</td>
<td>TCBI</td>
</tr>
<tr>
<td>DP4</td>
<td>M</td>
<td>30-39</td>
<td>Child care</td>
<td>Inter.</td>
<td>Toronto</td>
<td>Nokia</td>
</tr>
<tr>
<td>DP5</td>
<td>F</td>
<td>30-39</td>
<td>Health clinic worker</td>
<td>Exp.</td>
<td>Seattle</td>
<td>TCBI</td>
</tr>
<tr>
<td>DP6</td>
<td>F</td>
<td>50-59</td>
<td>Unemployed</td>
<td>Basic</td>
<td>Seattle</td>
<td>Nokia</td>
</tr>
<tr>
<td>DP7</td>
<td>M</td>
<td>20-29</td>
<td>Student</td>
<td>Exp.</td>
<td>Seattle</td>
<td>Nokia</td>
</tr>
<tr>
<td>DP8</td>
<td>F</td>
<td>30-39</td>
<td>Homemaker</td>
<td>Inter.</td>
<td>Seattle</td>
<td>Nokia</td>
</tr>
</tbody>
</table>
Results

• Definitely and Probably answers are considered correct for affirmative answers

• Definitely Not and Probably Not are considered correct for negative answers

• Participants answered 94.8% of the questions correctly using TCBI and 73.5% using the Nokia Interface
Results (p2)

“It totally works. I can leave my house without charging the thing, and I don’t have to worry about it dying.” —DP6

“And for the first one [TCBI], it actually felt like when I was checking it... I felt like I was getting more information. At least I was getting a better idea.” —DP3

“It [TCBI] was definitely more informative than the first one, and I felt like I could answer the questions that were being posed with more confidence.” —DP1

“I felt like I could mostly trust it. I tested it out once, just to see if it was really accurate down to the minute. It was, it said in 2 hours the battery would go dead, and in like, 2 hours and 1 minute, the battery went dead.” —DP5
Results (p3)

• Participants were asked to rate their confidence based on the interface they were using

• Average ratings: 1.76 (TCBI) 1.31 (Nokia)
Possible Improvements

- Increase the font size on TCBI
- Only 1 battery meter graphic needed
- Make screen less cluttered
- Ability to explore different application combinations (without actually opening them)
- A more interactive desktop icon
Possible Improvements (p2)

- Display warning (in different color or style) when model is more prone to error.

---

Figure 4: Graphs showing video playback battery traces (left) and the percent error in the predicted value in relation to the actual time remaining (right).
Possible Improvements (p3)

• Use crowdsourcing to gather a large number of battery traces over a wide range of hardware settings

• Per-device training for each application for a more accurate model for that device when there are other background processes running

• Generalizing TCBI to other platforms