PhoneLab: A Large Programmable Smartphone Testbed

Anandatirtha Nandugudi, Anudipa Maiti, Taeyeon Ki, Fatih Bulut
Murat Demirbas, Tevfik Kosar, Chunming Qiao, Steven Y. Ko and Geoffrey Challen
Department of Computer Science and Engineering
University at Buffalo
team@phone-lab.org
www.phone-lab.org

Abstract
As smartphones have emerged as the most widely deployed mobile computing platform, the scale of smartphone experimentation has lagged behind. New facilities enabling large-scale experiments are needed to ensure that research discoveries translate to the billions of smartphones in use today. To meet this challenge, we introduce PhoneLab, a 288-device smartphone testbed deployed at the University at Buffalo. PhoneLab provides access to smartphone users incentivized to participate in experiments while simplifying experiment data collection. The testbed will open for public experimentation in October, 2013, and continue to expand in 2014. To demonstrate the power of PhoneLab, we present three selected results from a usage characterization experiment run on 115 phones for 21 days. We use each result to motivate a future PhoneLab experiment, demonstrating how PhoneLab will enable mobile systems research.

Categories and Subject Descriptors
D.4 [**Operating Systems**]: Distributed systems

General Terms
Design, Experimentation, Measurement, Performance

Keywords
Smartphones, testbed, mobile devices

1. INTRODUCTION
Smartphones have quickly become the most popular computing platform. Google announced in September, 2013, that it had activated over 1 billion Android devices in only six years, with 500 million in the last year alone [5]. The International Data Corporation (IDC) projects that 224 M smartphone units will ship worldwide in 2013 Q4, a 40% increase over 2012 Q4 [7]. Taken as a whole, the growing network of smartphone devices represents the largest and most pervasive distributed system ever built.

SENSEMINE'13, November 14 2013, Roma, Italy.
Copyright 2013 ACM 978-1-4503-2430-4/13/11 ...$15.00
http://dx.doi.org/10.1145/2536714.2536718

Unfortunately, the scale of smartphone experimentation is not keeping pace. A quick survey of MobiSys'12, MobiSys'13, and SenSys'13 papers reveals that frequently smartphone evaluations on real devices use small numbers of phones—for example, 2, 3, or 7 [14, 15, 6]. Many other experiments use simulations driven by small, old, or synthesized data sets [18, 9, 8]. In either case, large-scale results from real users would be more compelling. While user recruitment, human subjects compliance, and reliable data collection make large-scale smartphone experimentation challenging, harnessing the growth of smartphones requires evaluating ideas at scale.

In this paper, we present PhoneLab, a large public smartphone testbed hosted by the University at Buffalo that enable research at scales currently impractical. PhoneLab provides the features necessary for smartphone research:

- **Scale:** PhoneLab currently has 288 participants incentivized to join experiments. Unlike applications distributed on the Play Store, PhoneLab experiments do not need to benefit or even interact with users, facilitating research unlikely to be popular on application marketplaces.
- **Power:** By utilizing the Android open-source smartphone platform, PhoneLab allows interactive experiments, non-interactive data collection, and changes to the Linux kernel; Android middleware and libraries; and Dalvik virtual machine. Passive experimentation is difficult to perform without platform support, and platform experiments are impossible to distribute on application marketplaces.
- **Realism:** Participants use the phones as their primary device.
- **Relevance:** PhoneLab allows researchers to stop relying on out-of-date datasets. Instead, new data can be collected in the most appropriate way for the experiment.

Section 2 describes PhoneLab, including its design, implementation, and data collection mechanism. Next, we demonstrate that PhoneLab is powerful and usable in Section 3 by giving three example data analysis results. For this purpose, we have conducted a usage measurement experiment run by 115 PhoneLab participants for six months. Rather than attempting a comprehensive analysis of the dataset, we use it to highlight the power of PhoneLab and breadth of research it supports by presenting three results:

- overall energy breakdown (Section 3.1),
- opportunistic charging (Section 3.2),
- 3G to Wifi transitions (Section 3.3).

These examples are intended to encourage others in the mobile systems community to use PhoneLab.

2. THE PHONELAB TESTBED

PhoneLab began operating in 2012 with 191 participants[1] using Nexus S 4G smartphones running Android 4.1.1. The first year of operating was a beta test allowing us to develop the software and expertise necessary to manage a large testbed. No external experiments were solicited, but several internal experiments were performed including a usage characterization study that produced the results described in Section 3. In 2013, PhoneLab grew to 288 participants using Samsung Galaxy Nexus devices running Android 4.2.2. Participants receive discounted voice, data, and messaging from Sprint in exchange for their participation, and are required to use their PhoneLab phone as their primary device.

PhoneLab experiments are either distributed through the Play Store or as platform over-the-air (OTA) updates. Participants are notified of new experiments and choose whether to participate after reviewing what information will be collected about them. PhoneLab participants are *required* to participate in experimentation but *not required* to participate in any particular experiment. They may remove experiments that they deem too intrusive or that negatively affect their device. Some experiments may run in the foreground and interact with users like typical applications, while others may gather data silently in the background.

PhoneLab users must provide human subjects review documentation, a list of log tags to capture that identify their log messages, and their experimental software—either a link to the Play Store or a patch against the current PhoneLab platform source. Experiments generate data through the standard Android logging interface. Log messages generated by PhoneLab experiments are captured and uploaded to a central server while the device is plugged in and charging. When experimentation completes, the user receives an archive containing every log message matching their tags generated by all participating devices.

2.1 Platform and Device

PhoneLab phones run the popular Google Android Open-source Smartphone Platform (AOSP). Using an open-source platform for PhoneLab was an obvious choice for obvious and less-obvious reasons. The obvious reason is that the AOSP allows PhoneLab users to experiment with any software component, meeting our goal of providing a powerful testbed. Modifications to Android services that provide location, access networks, and manage energy can be benchmarked alongside unmodified devices. Of course, power also creates problems: faulty experiments can render phones inoperable and threaten participation. As a result, experimentation at the platform level will require additional pre-deployment testing and interaction with the PhoneLab team when compared with experiments that only distribute novel applications or collect data at the application level. We plan to support external platform experimentation in 2014.

We have also found that using an open-source platform has less obvious benefits. First, the availability of the Android source makes PhoneLab instrumentation easier even when collecting data from the application level because it gives a visibility into hidden APIs. For example, our usage characterization experiment, described in Section 3, uses Java reflection to access hidden battery usage APIs. Second, using the AOSP allows us to sign the platform image used by our participants. When the same key is used to sign a software package, that application may run as the system user with root privileges. Using this feature allows us to distribute and update core PhoneLab experimental management software via the Play Store while retaining the privileges necessary to collect logs and perform platform updates.

[1] We refer to people carrying participating in PhoneLab experiments as PhoneLab *participants*, to differentiate them from researchers running PhoneLab experiments who we call *users*.

	2012–2013	2013–2014
Total Participants	191	288
Survey Responses	191	249
By Gender		
Female	51	127
Male	140	122
By Age		
< 18	12	0
18–20	74	12
21–24	34	21
25–29	29	34
30–34	15	35
35–39	6	28
40–49	13	52
50–59	7	34
> 60	1	9

Table 1: **PhoneLab demographic breakdown.**

We distributed Samsung Nexus S 4G smartphones to our first year of participants and Samsung Galaxy Nexus smartphones to our second year. Both were official AOSP development phones and are well-supported. While we expect to receive yearly phone upgrades and distribute more up-to-date devices each year, we anticipate that the prohibitive cost of the newest smartphone models will prevent us from deploying them on PhoneLab.

2.2 Participants

Recruiting a large number of PhoneLab participants requires effective incentives. In their first year of PhoneLab participation, voice, data and messaging are free with funding provided by the National Science Foundation (NSF). This free year of service plays a major role in our recruiting efforts. In subsequent years, participants pay a deeply discounted $45 per month rate for unlimited data and messaging through a deal negotiated with Sprint. Sprint has proved to be an ideal partner for the PhoneLab project, both helpful with testbed logistics and still providing unlimited data plans, negating any economic impact of our data collection.

Because participants may leave at any time, the front-loaded cost structure of our incentives makes it most efficient to recruit participants who will be able to continue as part of PhoneLab for multiple years. While we anticipate that some participants will leave after the first free year, interviews with them will help us identify long-term participants during subsequent years. Long-term participants allow us to amortize the first free year and provide a stable group comfortable being a part of PhoneLab experimentation.

When recruiting our first batch of 2012–2013 participants, we targeted freshman, sophomores, and new PhD students. The University at Buffalo has a large international graduate student community, and many of these students arrive on campus without phones or phone contracts, making them ideal PhoneLab participants. Unfortunately, retention after our first year was poor: only 43 of 191 first-year participants chose to continue after the free year ended. We believe multiple factors are at play. First, we now consider our decision to recruit undergraduates to be a mistake, since many of these students still have the option of continuing on family plans paid for by their parents. Second, multiple participants have complained about poor Sprint coverage, which, while out of our control, is a concern in our area.

As a response to our low first-year retention, during the second year we targeted staff and faculty along with incoming PhD students and did not recruit or accept undergraduates. We hope that adult staff members are more accustomed to paying for phone service and less likely to have a free family plan to return to. This change to our recruitment focus has also had the benefit of significantly improving both the gender and age diversity of our participant pool. As shown in Table 1, while our first-year participants were primarily young men, our second-year participants are a good mix of ages and genders, making results more representative.

2.3 Testbed Software

PhoneLab smartphones are distributed with a small piece of testbed management software integrated into the Android platform. This heartbeat service uploads periodic reports including information about device location, battery levels, and the installation status of other required PhoneLab software. This information is only used for testbed management and never released to researchers.

On boot, the heartbeat service starts the PhoneLab Conductor, the primary PhoneLab configuration and data collection tool. Experimental configuration, log collection, data upload and platform updates are performed by the Conductor, which is installed and updated through the Google Play Store. We sign it with the platform build key, which allows it to run with the root privileges necessary to collect logs from all applications, perform platform updates, and start experiments that lack a foreground activity. Periodically, the Conductor retrieves an XML configuration from the PhoneLab server. The configuration specifies what background experiments to start or stop, which log tags to collect, and the upload policy and endpoints. The Conductor also uploads status information to the server during the configuration exchange, including version numbers of PhoneLab software, what experiments are currently running and how much data is queued for upload.

PhoneLab logging and data collection must be unintrusive. If it is not, either our participants will quit, or their usage patterns will be affected. We believe that we have achieved this goal. Measured battery usage of PhoneLab software is low. A conservative estimate that includes all of the applications that run as the shared system user comes to a per-participant average of 2.4%. Given that this estimate includes many non-PhoneLab applications that also run as the system user, it should be considered a strict maximum. Our default policy of only uploading while the device is charging reduces the overhead of the most power-hungry task. In addition, we have received no complaints about our tools, even from participants that initially used their phones without them.

2.4 Safety and Privacy

There is an important difference between PhoneLab and other testbeds, such as Emulab [17], PlanetLab [10], MoteLab [16], or OpenCirrus [3]: our experiments involve real people. There are two core requirements regarding our participants. First, they should use their phone normally, which motivated the design of unintrusive testbed management software. Second, and more importantly, they must feel safe while participating in PhoneLab experiments.

To accomplish this, we leverage several existing safety mechanisms when possible. First, we require an Institutional Review Board (IRB) to review each PhoneLab experiment for human subjects compliance. IRB approval or exemption is required before any PhoneLab experiment can begin.

Second, we distribute experimental applications to a group of developers prior to broader release, allowing us to identify any significant problems before they reach our participants. This step is particularly important for platform experiments, which must be established as stable before being distributed.

Finally, we utilize Android's existing safety and privacy mechanisms. Participants are presented with the typical Android privacy dialog during experiment installation. Rather than building an alternate distribution channel or privacy mechanism, we felt it was sufficient and probably better to use a process that participants are familiar with. After installation, if a participant discovers that an experiment malfunctions or wastes power, they can uninstall it. If we notice patterns of experimental removal, we will flag the experiment and notify the researcher.

2.5 Experimental Procedures

To conclude, we review PhoneLab experimentation from a researcher's perspective. First, develop your application locally. Any information logged through the standard Android logging library can be recorded. In addition, the platform may already be logging useful information for you. Keep track of all the log tags you want PhoneLab to capture. Approach your local IRB and receive experimental approval and upload your application to the Play Store.

Second, upload your list of log tags, IRB letter, and link to your application on the Play Store through the PhoneLab website. We will contact you when we begin beta testing and again once your experiment is ready for the testbed. During beta testing you will be provided with PhoneLab log output to ensure that your experiment is running properly.

Finally, your experiment will be scheduled. Our goal is to maintain a medium-sized list of active experiments for our participants: large enough to make good use of the testbed, but small enough to ensure that each experiment is picked up by many participants. When your experiment completes, you will receive a archive with messages matching the tags you selected.

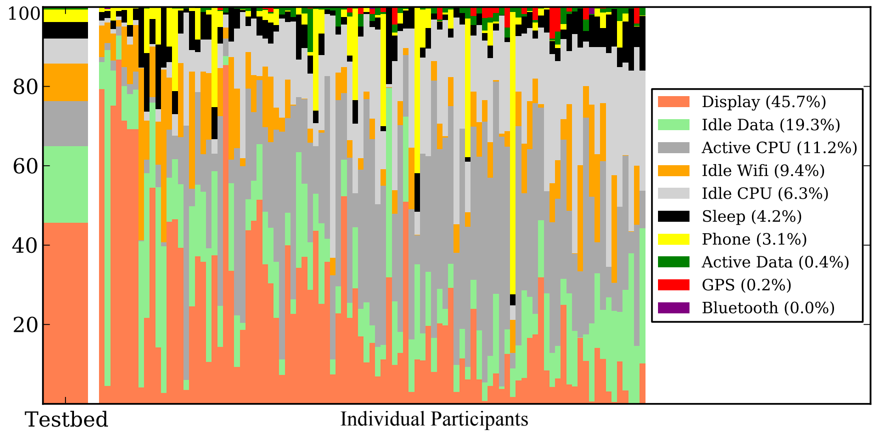

Figure 1: **Power usage by component.** The large bar at left shows an aggregated breakdown for all participants. The participant bars are scaled against the participant with the most energy usage.

3. EXAMPLE EXPERIMENTS

During our first beta test year we performed a usage measurement study. 115 participants joined the study which ran for over six months. For this purpose, we developed a measurement application that collects multiple salient features of smartphone usage: networking, mobility, power consumption, and application usage. This section presents selected results on energy usage, wireless network transitions, and battery charging behavior. Our goal is to demonstrate the types of experiments that can be performed on PhoneLab and the insights they can achieve.

3.1 Energy Breakdown

A single-day component-by-component breakdown is shown in Figure 1. Our results are similar to those reported by a previous smaller-scale study [12], and indicate that mobile data (labeled as "Idle data" and "Active data" depending on the state), the screen, and CPU usage are the main sources of power consumption. The per-participant bars also show a great deal of variation, with differences in both the amount and the breakdown of energy consumed by each participant.

One supposedly power-hungry component that has less of an impact than we had expected is the GPS. This is particularly surprising given the large amount of location-monitoring work motivated by GPS power consumption. One of several factors may be at work. First, the Android platform estimates the GPS chipset current consumption at 50 mA. This number is used by the standard "Fuel Gauge" battery monitor and by our calculations. However, it is lower than the data sheet for the Broadcom 4751 GPS receiver [1] and may represent a best-case average. Still, even if the GPS current consumption is off by as much as a factor of five, it does not represent a significant contribution. Other hypotheses are that Android network location is providing location with sufficient accuracy for many applications, eliminating the need for GPS, or participants may be conscious of GPS power consumption and taking steps to control it.

3.2 Opportunistic Charging

One way that users work around the battery limitations of their smartphone devices is by finding new times and places to charge their phones: plugging in at their desk at work, in the car during their commute, or at home before a long night out. We refer to these charging sessions as *opportunistic* to distinguish them from *habitual* nightly charging. Assuming that many smartphone users encounter plug points throughout the day, engaging in opportunistic charging becomes an additional sign of energy awareness, and understanding opportunistic charging becomes necessary to improving energy management on mobile devices. Others have analyzed this behavior before [4] and our goal is to examine the battery charging behavior of PhoneLab participants.

Figure 2 shows that many users engage in opportunistic charging. We define a charging session as opportunistic if is longer than 10 minutes but does not fully charge the battery, indicating that the device was disconnected before charging could finish. On a representative day during our experiment, of the 245 charging sessions we observed that day, 96 (39%) were opportunistic by this definition. 50 of 95 active participants engaged in opportunistic charging at some point during our experiment an average of once per day.

Opportunistic charging may be a response to an anticipated need for more smartphone battery power, as when a student plugs her smartphone in to charge before a night out. Our data also allowed us to examine how many of these opportunistic charging sessions were necessary to bridge the gap to the next full charge. We found that only 24 of the 96 opportunistic charges we observed were necessary. We believe that this indicates that participants have responded to their smartphones' battery limitations by engaging in conservative charging behavior, grabbing power whenever possible even if they do not anticipate needing it later.

Opportunistic charging combined with the varied rhythms of our participants creates a second interesting effect: at any given point there is a wide disparity in the amount of power available on different phones. Figure 3 displays the top, bottom, and middle (median) quartiles for a single day on PhoneLab. Only phones that are discharging are shown, which explains the sharp increase between 6 and 10AM as participants end nightly charging cycles. As the graph indicates, it is likely that when two smartphones meet they have very different battery levels.

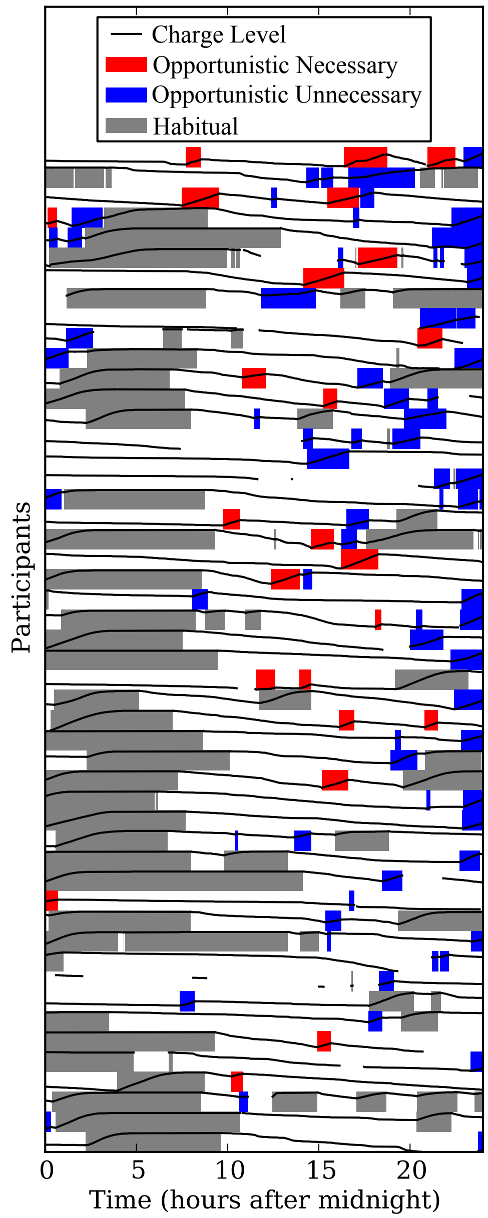

Figure 2: **Charging patterns.** Many users perform opportunistic charging during the day, with habitual charging occurring at night.

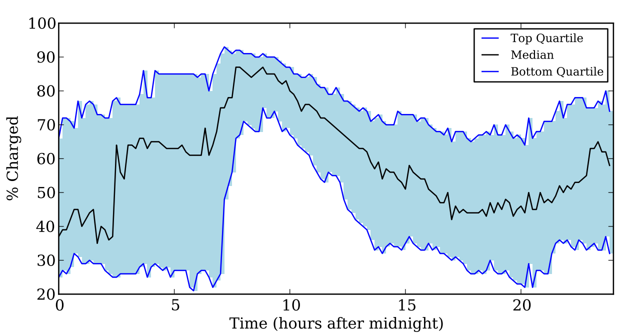

Figure 3: **Charge difference between participants during one day.** The graph plots the top and bottom quartiles and median. A significant spread is present at all times.

3.3 Mobile Network Transitions

Mobile smartphones move through a complex network environment. Providing the illusion of seamless connectivity requires negotiating hand-offs both between Wifi access points and between Wifi and 3G radios. We were interested in observing hand-offs between 3G (provided by Sprint, PhoneLab's operational partner) and Wifi and found many in the dataset collected by our usage experiment. Since the Android `ConnectivityService` frequently switches network interfaces for exploration purposes, we have defined a transition as two one-minute or longer sessions on different interfaces separated by less than one minute. We further limit ourselves to cases where we received a location update during the transition. Data for the first 21 days of the usage experiment is shown.

Figure 4 plots the location of transitions that occurred on or near the University at Buffalo North Campus. We notice many clusters in expected locations: near the entrance and exits of buildings where participants are likely to be moving from campus Wifi to 3G.

4. RELATED WORK

Most similar to PhoneLab are the NetSense [13] project at Notre Dame and the LiveLabs [2] testbed at Singapore Management University. NetSense has many similarities to PhoneLab: it distributed instrumented smartphones to several hundred incoming freshman undergraduate students. In contrast to PhoneLab, however, NetSense was built to support a single study—on how use of digital technologies impacts friendship formation—and was never designed or operated as a public testbed. LiveLabs is a city-scale research testbed designed to allow companies to run large-scale consumer trials and experiment with novel services. It aims to recruit thousands of participants, potentially providing scale exceeding that of PhoneLab, but is also not public.

Other computer science testbeds meet domain-specific needs. PlanetLab [10] operates more than 1,000 machines worldwide to facilitate large-scale, realistic Internet research. Emulab [17] provides emulated network environments to enable controlled, repeatable network experiments. MoteLab [16] provided access to 200 sensor network nodes deployed in a multi-story office building. ORBIT [11] takes a two-tier approach allowing emulated experiments as well as real deployments, targeting reproducibility and realism at the same time. OpenCirrus [3] is a geographically distributed cluster designed to support cloud computing research.

5. CONCLUSIONS

We have introduced PhoneLab, a new large-scale programmable smartphone testbed at the University at Buffalo supporting experimentation both above and below the application-platform interface. Through three example experiments we have demonstrated the power of PhoneLab to enable the next generation of mobile systems research. We look forward to working with researchers interested in using PhoneLab once the testbed opens to the public in October, 2013.

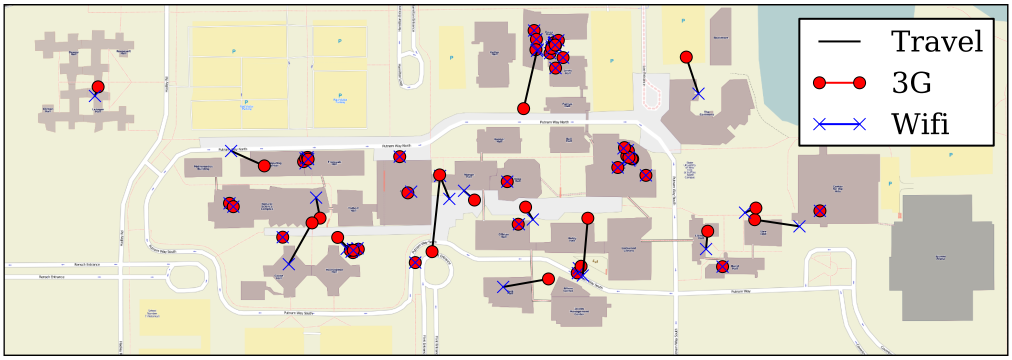

Figure 4: **3G to Wifi transition locations.** The map indicates that there are several common areas where network hand-offs occur.

Acknowledgments

Many students have worked on the PhoneLab testbed, including Sonali Batra, Micheal Benedict, Vinu Charanya, Tong Guan, Rishi Baldawa, Jay Inamdar, Manoj Mylapore, Mitchell Nguyen, and Sean Zawicki. Maulik Dave is currently employed as the testbed administrator and helped prepare this paper. PhoneLab is funded by the National Science Foundation's Computing Research Infrastructure program under award number CI-ADDO-NEW-1205656.

6. REFERENCES

[1] Broadcom BCM4751 Integrated Monolithic GPS Receiver. `http://www.broadcom.com/products/GPS/GPS-Silicon-Solutions/BCM4751`.

[2] Livelabs. `http://centres.smu.edu.sg/livelabs/`.

[3] A. I. Avetisyan, R. Campbell, I. Gupta, M. T. Heath, S. Y. Ko, G. R. Ganger, M. A. Kozuch, D. O'Hallaron, M. Kunze, T. T. Kwan, K. Lai, M. Lyons, D. S. Milojicic, H. Y. Lee, Y. C. Soh, N. K. Ming, J.-Y. Luke, and H. Namgoong. Open Cirrus: A Global Cloud Computing Testbed. IEEE Computer, 43(4):35–43, Apr. 2010.

[4] N. Banerjee, A. Rahmati, M. D. Corner, S. Rollins, and L. Zhong. Users and Batteries: Interactions and Adaptive Energy Management in Mobile Systems. In Proceedings of the 9th International Conference on Ubiquitous Computing (UbiComp), 2007.

[5] BGR Media, LLC. Google announces 1 billion Android activations, teases Android 4.4. `http://goo.gl/u8m4HA`, September 2013.

[6] T. Hao, G. Xing, and G. Zhou. iSleep: Unobtrusive Sleep Quality Monitoring using Smartphones. In Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys), 2013.

[7] International Data Corporation. Worldwide Mobile Phone Growth Expected to Drop to 1.4Despite Continued Growth Of Smartphones. `http://goo.gl/ROLY0`.

[8] S. Isaacman, R. Becker, R. Cáceres, M. Martonosi, J. Rowland, A. Varshavsky, and W. Willinger. Human mobility modeling at metropolitan scales. In Proceedings of the 10th international conference on Mobile systems, applications, and services, MobiSys '12, pages 239–252, New York, NY, USA, 2012. ACM.

[9] S. Nath. Ace: exploiting correlation for energy-efficient and continuous context sensing. In Proceedings of the 10th international conference on Mobile systems, applications, and services, MobiSys '12, pages 29–42, New York, NY, USA, 2012. ACM.

[10] L. Peterson, T. Anderson, D. Culler, and T. Roscoe. A Blueprint for Introducing Disruptive Technology into the Internet. SIGCOMM Comput. Commun. Rev., 33(1):59–64, Jan. 2003.

[11] D. Raychaudhuri, M. Ott, and I. Secker. ORBIT Radio Grid Tested for Evaluation of Next-Generation Wireless Network Protocols. In Proceedings of the First International Conference on Testbeds and Research Infrastructures for the DEvelopment of NeTworks andA COMmunities (TRIDENTCOM), 2005.

[12] A. Shye, B. Scholbrock, and G. Memik. Into the Wild: Studying Real User Activity Patterns to Guide Power Optimizations for Mobile Architectures. In Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), 2009.

[13] A. Striegel, S. Liu, L. Meng, C. Poellabauer, D. Hachen, and O. Lizardo. Lessons learned from the netsense smartphone study. In Proceedings of the 5th ACM workshop on HotPlanet, HotPlanet '13, pages 51–56, New York, NY, USA, 2013. ACM.

[14] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. R. Choudhury. No need to war-drive: unsupervised indoor localization. In Proceedings of the 10th international conference on Mobile systems, applications, and services, MobiSys '12, pages 197–210, New York, NY, USA, 2012. ACM.

[15] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin. Sensing Vehicle Dynamics for Determining Driver Phone Use. In Proceeding of the 11th annual international conference on Mobile systems, applications, and services, MobiSys '13, 2013.

[16] G. Werner-Allen, P. Swieskowski, and M. Welsh. MoteLab: a Wireless Sensor Network Testbed. In Proceedings of the 4th International Symposium on Information Processing in Sensor Networks (IPSN), 2005.

[17] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar. An Integrated Experimental Environment for Distributed Systems and Networks. In Proceedings of the Fifth Symposium on Operating Systems Design and Implementation (OSDI), 2002.

[18] T. Yan, D. Chu, D. Ganesan, A. Kansal, and J. Liu. Fast app launching for mobile devices using predictive user context. In Proceedings of the 10th international conference on Mobile systems, applications, and services, MobiSys '12, pages 113–126, New York, NY, USA, 2012. ACM.