Summary Report for the First NSF Workshop on Mobile Community Measurement Infrastructure

David Choffnes
Northeastern University

Z. Morley Mao
University of Michigan

ABSTRACT

The first NSF Workshop on Mobile Community Measurement Infrastructure was held on November 12, 2014, in Washington, D.C. This goal of this workshop was to respond to the increasing needs for increased visibility into network behavior to enable cutting-edge research in mobile computing. The workshop brought together the top researchers in both academia and industry, as well as policy makers from FCC, to discuss the requirements for such an infrastructure and identify the concrete steps to make progress towards the identified goals. Given the diverse research topics within mobile computing, we ensured that different research areas and all the important network layers (e.g., physical, MAC, network, transport, and application) were represented at the workshop to ensure a productive outcome from the discussions. We report on the talks, panels and discussions during the workshop, and make recommendations for future directions to address key challenges toward building critical community measurement infrastructure.

1. MOTIVATION

Flexible measurement infrastructure support for experimentation are critical for enabling and supporting research work with strong experimental components to help validate ideas and evaluate the design in practice. Mobile computing research work must be evaluated in as realistic network settings as possible to help researchers understand how various factors such as signal strength, network load affect the performance and energy metrics of interest. Currently, there is no environment that enables network visibility from an end-to-end perspective of the cellular network protocol stack: from the end-device all the way to the network server traversing through various network elements. The lack of such an integrated measurement infrastructure greatly hinders the innovation in this important research field. Because it is not trivial to design instrumentation to support diverse needs of researchers in mobile computing, the workshop helps define the goals and requirements of a mobile measurement infrastructure that is designed to last at least 10 years to support cutting-edge research in this space.

On November 12, 2014, we held the first NSF workshop on Mobile Community Measurement Infrastructure. The goal of this workshop was to identify the key requirements for designing and developing a mobile community measurement infrastructure to support cutting-edge research in mobile computing, going beyond the current research interests to support experimentation in next-generation mobile networks such as 5G. Today's mobile systems are deployed with ad-hoc measurements providing limited visibility into network behavior; however, we need integrated, cross-layer scalable, and flexible measurement infrastructure to enable transformative and bold research endeavors. In particular, most commercially deployed mobile network infrastructure are completely closed to researchers, making it difficult to experiment with or understand the impact of today's systems or future technologies. We propose that a *knowledge plane* (as proposed by David Clark's vision [3]) for the mobile wireless network must be designed from ground up, rather than as an after-thought after the protocol and infrastructures have been built and deployed.

A goal of this report is to continue the conversation on how the community will conduct measurement in an effective, sustainable, and collaborative manner. Our workshop mission aligns with previous and concurrent workshops on similar topics, namely AIMS 2014 [2] and the NSF Workshop on Future Research Infrastructure held jointly with this workshop. However, our mission is more directly focused on measurement and monitoring support, instead of broadly infrastructure development. In fact, the specialized purpose of a critical part of the infrastructure should support measurement effectively, to enable real-time debugging and diagnosis, in order to evolve the wireless network for future application and user demands.

The broader impacts of this proposed workshop includes the following: (1) an opportunity to bring together researchers in different areas within mobile computing to discuss the requirements for a common measurement infrastructure to support cross-area research and collaboration. (2) the outcome of the workshop is a concrete plan developing a mobile measurement infrastructure which will enable research in key areas such as mobile cloud computing, network protocol design, etc. (3) we expect the discussions at the workshop to also generate new research focus for the mobile computing community.

The list of attendees is: David Choffnes (Northeastern); Z. Morley Mao (Michigan); Thyaga Nandagopal (NSF); Aruna Balasubramanian (Stony Brook); Geoff Challen (Buffalo); Ranveer Chandra (MSR); Romit Choudhury (UIUC); Deepak Ganesan (UMass); Nada T. Golmie (NIST); Anoop Gupta (Microsoft); Walter Johnston (FCC). David Kotz (Dartmouth); Padma Krishnaswamy (FCC); Erran Li (Lucent); Kobus van der Merwe (Utah); Chunyi Peng (OSU); Haiyang Qian (ChinaMobile); Lili Qiu (UT); Gyan Ranjan (Narus); Aaron Schulman (Stanford); Srini Seshan (CMU); Rangam Subranamian (NTIA); Andreas Terzis (Google); Changbo Wen (T-Mobile); Mike Wittie (Montana); Guoliang Xing (MSU); Ying Zhang (Ericsson Research); Lin Zhong (Rice); Xia Zhou (Dartmouth);

The rest of the summary is organized according to the sessions and panels during the workshop [1]. In the following section, we summarize existing efforts discussed at the workshop. We then present a summary of the issues and challenges raised during discussion and feedback from surveys during the workshop. Last, we present our recommendations for future efforts in this space.

2. EXISTING PROGRAMS AND EFFORTS

2.1 Measurements and Application Needs

2.1.1 Wireless Spectrum

Rangam Subramanian (NTIA) discussed the topic of federal initiatives for wireless innovation measurement. The government controls 43% of the of the spectrum (below 3.5GHz) and is interested in studying how we can share this spectrum. The NTIA is responsible for federal spectrum management (the FCC is the commercial side), and there are initiatives of interest to researchers in the space of spectrum innovation and spectrum sharing. For instance, the Center for Advanced Communication (CAC) and National Advanced Spectrum Communications Test Network (NASTCN) are inititiatives aimed at facilitating collaborative research and testing. The Wireless Spectrum R&D Senior Senior Steering group coordinates spectrum-related R&D, and its next workshop (planned for March 19, 2015) is on federal incentives for sharing spectrum. The WSRD Test-bed Portal[1] provides a list of the major test beds in the US, and facilitates information exchange across academia and government. Spectrum.gov provides a compendium of federal agency spectrum usage in 225 MHz to 5 GHz.

The Model City public notice describes a program that seeks to create cities for experimentation. The notice seeks public comment to establish a public-private partnership facilitating the creation of an urban test city that would support rapid experimentation and development of policies, underlying technologies, and system capabilities for advanced, dynamic spectrum sharing. The idea is for government and industry to comes up with devices to study bands of interest. Importantly, we need to address how to enable collaboration and define measurement standards for a Model City. Researchers could use the Model city to focus on specific "important" spectra and understand properties such as range, bandwidth and the like.

Ranveer Chandra (MSR) discussed the Microsoft Spectrum Observatory. It is a cloud service where different observatories (USRP based spectrum analyzers) collect spectrum data and upload it to the cloud. The use case for the spectrum observatory is to help policy makers, to detect rogue transmitters, to populate white space database, and to model the real world. There are around O(10) universities/sites where these analyzers are deployed. Key challenges are how to reduce the cost of the boxes (currently $5k), how to more efficiently/effectively collect and summarize the data, how to make spectrum scans more dynamic, get space-time spectrum utilization, make inferences from the raw data, and detect transmitters.

Aaron Schulman (Standford) discussed cellular base station PHY measurement and smartphone power measurement. As a motivating example, there is a need for measurements at base stations to quantify PHY faults, which would allow stations to adapt to PHY conditions and help inform the design of 5G. Leveraging the observation that cellular providers already use programmable DSPs in their base stations that allow for such measurements, the ATOM project builds a programmable base station by composing it from processing blocks with fixed processing time constraints.

2.1.2 Measurements up the stack

Andreas Terzis (Google) described measurements to improve YouTube. 20% of Internet traffic is YouTube, out of which 40% comes from mobile. A key problem is that there are many (middle)boxes between the user and the YouTube cache, and their impact on performance was unknown. To test this, Google worked with T-Mobile, where half the traffic was sent over the proxy and half the data bypassed the proxy. The results are that without the proxy, there are fewer retransmissions and higher throughput (because the proxy throttles the connection), decreased energy consumption, reduced buffer bloat, and better QoE. An important mea-

[1]`http://www.nitrd.gov/Subcommittee/wsrd/Testbeds/map.aspx`

surement question for Google is understanding latency in 2G networks, given that large populations (generally in developing regions) access the Internet over 2G networks. While proxies may have many roles and performance impact may be intentional, an important issue is that interactions between the proxy and other parts of the network can have unintended consequences.

Chunyi Peng (OSU) talked about protocol verification in cellular networks. A key problem is that the control plane protocol in recent 3GPP standards is layered and spans different domains such as circuit/packet-switched, and there are different versions of 3G/4G with different protocol specifications. The operations are also distributed, some operations reside on the base station, the phone, gateways, etc. In fact, interactions among protocols (*e.g.,* 3G and 4G) are a big problem in operational networks. To make progress in this research area, we need better access to network cores (currently closed to researchers) and shared datasets to understand protocol behavior and configurations in practice.

2.1.3 Putting the mobile in mobile infrastructure

Srini Seshan (CMU) described a key challenge in wireless measurement: existing studies either too stale or they lack sufficient context to be useful to others. To solve this problem, we need continuous measurements like the spectrum observatory. But of course, such measurements are either done manually or require dense deployment that are impractical. To address this, they are using robots to collect measurements because they are programmable, repeatable and will be ubiquitous (eventually). A deployment at CMU currently measures daily and weekly WiFi RSSIs.

Romit Choudhury (UIUC) discussed several types of mobility for wireless infrastructure, including small antenna mobility, routers under a table moving at a micro level, WiFi routers on the ceiling moving within and between rooms for macro mobility, and even quad-copters flying out of the cell towers according to client location for even larger mobility. Measurement results indicate that even micro mobility improves performance 1.7x. because fading and interference goes down as the routers move around. Areas of future work include how to coordinate these routers/robots from the cloud to do topology management based on network needs, and using quad-copters to change the quality of the link.

2.1.4 Sensing

Deepak Ganesan (UMass) described how to use backscatter communication to help enable an Internet of Things (loT). Trends indicate sensing is no longer the bottleneck; rather it is the communication in terms of power consumption. Backscatter is a way to achieve this low power communication using RFID. A key challenge is how to do this for other domains, given that the available energy is tin and the signal quality is extremely dynamic. Open questions include how well backscatter systems will work at larger (human, house) scales, and what are the fundamental limitations on how much information we can communicate via backscatter.

Lili Qiu (UT–Austin) discussed robust network compressive sensing, *i.e.,* coping with measurements that have missing values, errors, and anomalies. This work breaks down measurements into matrices, and decompose the matrix into a low rank matrix, a sparse anomaly matrix and noise matrix. This provides a way to remove the noise from measurement data.

2.1.5 Energy measurement

Guoliang Xing (MSU) talked about non-invasive power metering for mobile and embedded systems. The goal is for in-situ power meters to measure power for embedded systems without requiring them to be soldered to each embedded system. The Nemo system is a new power meter that snaps on the battery pack on the embedded node and does not require hardware changes. Nemo has been adapted to work with smartphones apart from embedded systems. This allows non-invasive power measurement for different platforms. Nemo has >5Khz sampling rate and <1 micro amp resolution.

Aaron Schulman's BattOR smartphone power monitor is a hardware device that performs fine-grained battery consumption measurement and modeling. Importantly, it is highly accurate, small, and can be easily integrated with phones.

2.2 Testbeds

Spectrum testing effort: Walter Johnston from FCC discussed the topic on city-wide spectrum test and posed questions such as what structures would serve the public and research community to establish principles that would make spectrum available ubiquitously where there are incumbents in the spectrum, but don't use it effectively. There has been an opportunity in search of a plan — PCAST announced $200million dollars, but wasn't authorized. It is important to get service providers on board in exploring how to gain spectrum on shared basis. There is a long wait before the next opportunity before 30Ghz and 5G. What has worked well by FCC with carriers has been talking with operators to tell them what they will do before they do it. There has been significant discussions about Turbo Boost, though as speeds go up not a bit deal. Operators tend to be unhappy about comparative studies, but can support characterization studies.

Mobile network testbed: Professor Kobus van der Merwe (Utah) presented the PhantomNet testbed, which is a mobile network testbed - configurable network over an experimental license spectrum, based on Emulab and OpenEPC stack. phantomnet.org open for business, can be used for class projects. He is actively looking for users. There is support for emulated traffic but with somewhat limited in scope. The audience asked several questions about the utility of the testbed, because of the limited scope. It is interesting to note that there is right to experiment which can't be unreasonably withheld. FCC is also looking on a program license, where the right is to experiment and someone would have to prove that the experiments interfere with other work.

Mobile user study testbed: Geoff Challen (University at Buffalo) gave a talk on "My Platform Knows More Than Your App: Android Platform Experimentation on PhoneLab". This support uses logcat for recording information. Phonelab can provide more information than vanilla android API. Researchers are encouraged to do a fork of their platform for other researchers (IRB is required). Many metrics that are not available normally (per app energy usage, file usage, screen usage) are supported. It is a good place to evaluate new proposals are the platform level and there is a possibility of integration with SciWiNet from Clemson. Dave Choffnes (NEU) presented Mobilyzer, a collaboration project with Morley Mao at University of Michigan. He discussed how the platform supports a variety of crowdsourced apps. One of the measurement finding is that all US cellular carriers they measured proxied connections.

Suman Banerjee (Wisconsin) described WiENST, a city-wide edge infrastructure (includes vehicles and base stations). The motivation is that campus wide is too small in terms of scale, and people don't want two devices – citywide would work a little-bit better. The testbed's size is between Orbit and PlanetLab. They plan to use 8-10 base stations from an existing company, involving interested parties and a good management plan between them. Suman also mentioned the AirShark/WiSense tool that use wifi (atheros) cards as a spectrum analyzer: with RF translator there is support to sense licensed spectrum in receive mode as well.

2.3 Operational Challenges

2.3.1 *QoE challenges*

A perspective of QoE challenges is given by Changbo Wen (T-Mobile)'s position on crowdsourcing data collection. Wen works for the performance team. There is significant carrier interest in crowdsourcing data - KPI - key performance indicator. To understand network performance, one source of data is from vendor, third party drive tests (missing negative events - no service, no measurement). To characterize user experience, metrics of interest include coverage, signal quality, speed, application performance. Performance during congestion and service VoLTE is of interest. T-mobile is the first to deploy, but will need to work through lots of bugs, network layer. One challenge is that across different layers, there is non-uniform spectrum. It is difficult to compare new layers to mature layers. Some of the interesting problems include how do different spectrum bands perform in different areas, access failure, drop call, what's the performance in the basement, sport stadium, what's the data service like (Web, video). It is useful to collect a competitive view: inter-carrier (A vs B), intra carrier: markets, network layers (AWS vs PCS bands), vendors, device, MVNO. Customer feedback after a negative effect is collected: surveys are not good enough, need location, how bad it is (missed a call, but phone didn't tell you there was a call).

2.3.2 *SDN-based support*

Dr. Haiyang Qian (China Mobile) discussed the topic of " SDN for Mobile Network Mobility", where he described how to use SDN to measure network performance. Dr. Ying Zhang (Ericsson) also gave her position statement on SDN based measurement for network anomaly detection (SDN adaptive network measurement). There is interest to collect lots of data from different areas to help analyze new business plans (impact of iPhone on verizon when launched), new features/data plans to study how to improve customer loyalty, influence development of future products (offered by the IDPs). To achieve these goals, self-organizing networks (SON) – a big thing in cellular, needs to be carried out on an end-to-end basis in cellular networks. To support scalable measurements and how to control measurements, SDN can be very useful: one of the examples is anomaly detection. An interactive interface is designed between anomaly detector and the measurement module (tell network to collect more data on the anomaly once detected) hard to decide and hard to implement in hardware. SDNs prevent the vendor lock-in problem. Their system is called OpenWatch: adaptive rule update, rule installation (where to install), data preparation (send data back to the anomaly detector).

2.3.3 *Managing cellular networks*

As the final position statement of this panel, Erran Li (Bell Labs) presented his idea on Making Cellular Networks Scalable and Flexible. Given the dense and chaotic deployments, distributed control plane makes it hard to manage inter-cell interference and variable load of different cells. There is a strong need to share cost among operators. He proposes the design of SoftRAN: Big Base Station Abstraction with centralized controller for multiple base-stations to tackle the service chaining problem - the use of different middleboxes for different features, but how to scale the flows efficiently in these networks. The division of spectrum resources into regions leads to poor spectrum utilization (regions are not connected until the Internet, which results in different radio performance profiles in each region).

2.4 Data Collection and Privacy

David Kotz discussed privacy factors to consider when doing mobile measurement and data collection. These include IRB permission if human subjects are involved, informed consent of subject, the need to plan ahead for rights to share data, securing data on mobile devices, in transit, in repository. To address some of these issues, researchers can use single-study cloud servers for receiving, storing, processing, archiving data. Given that data is hard to collect, researchers should plan to share it.

The CRAWDAD public archive of wireless/mobile data contains 116 datasets and tools used over 1,500 academic papers, currently with 6.7 thousand users. Several lessons learned from collecting CRAWDAD data include: (1) metadata is important, (2) understand privacy regulations: wiretap law, HIPAA, FERPA, EU Privacy, etc, (3) gain permission of users & operators, (4) plan for data sanitization, (5) when sharing, demand that recipients respect privacy of human subjects (e.g., disallow re-identifying subjects).

Gyan Ranjan talked about security, privacy and encryption in mobile networks. In mobile, HTTP still dominates the app-layer protocol and the HTTP header contains much information, but HTTPS prevents the visibility to the contextual data. A key challenge is how to allow operators and other parties to access information currently hidden in HTTPS? Put another way, from the network management perspective, what type of measurement data is really needed and can we make this available without compromising data privacy? Do we have to do this for millions of apps or can we apply sampling techniques?

Anoop Gupta focused on openness and collaboration in the Spectrum Observatory. The spectrum observatory platform contains a measurement station attached to RF sensors, which puts all the collected data in a single data format. The actual file format is extensible and readable, including the time, configuration and the spectrum information. Each file also includes the context information, such as location, sensor, connectors, scanning algorithm, cabling, and antenna. To handle privacy concerns, they use aggregation. The open question is how to balance between providing real research value out of data while still guaranteeing privacy. The current state of the art is to use data aggregation to alleviate these issues, but we need new approaches for future to allow more sharing.

Lin Zhong pointed out that managing platforms collecting participant data is a pain, and sharing any data with location is impossible if you want to guarantee user privacy. Instead, since industry already has all this sensitive location data (*e.g.,* , from Google Map and Carrier IQ), there should be a way for academics to access it instead of trying to reproduce it at a small scale in their own testbeds. Academic researchers need to do more "crazy things" that industry isn't going to do, and take risks, even if it means a significant chance of failure.

3. ISSUES AND CHALLENGES

3.1 Measurements and Application Needs

What are the key questions we should be asking? How can we structure efforts to support the maximum amount of research? What measurements do we want and where can we get them? What are the most pressing research challenges in enabling accurate, fine-grained, scalable measurement of mobile systems? What information do we really want to collect? How might this data be used in both protocols or applications? Is there a coherent architecture that makes this data available to different applications/groups/etc? How can we perform large numbers of wireless measurements quickly and easily? How do we perform measurements at many different locations? Can spectrum parking measurements be crowd-sourced? What practical use cases can be motivated for mobile AP's as opposed to multiple repeaters? What are new degrees of freedom in networks that we are not seeing? Network mobility is an example. Human users is an example. What else? What application is the measurement going to benefit? What hypotheses are we testing with the measurement infrastructure?

What are the key challenges for developing measurements to meet application needs? One of the key challenges highlighted by participants is openness and ground truth: we need access to mobile systems and platforms, representative datasets for user/device populations, and infrastructure support for measurement. Other key challenges include privacy, incentives for supporting/reporting measurements and covering higher-frequency spectrum.

3.2 Testbeds

The following issues came up when discussing existing testbeds for mobile measurement.

- **Limited scope**: most of the focus of testbed is on network protocols, very little on layer 2. There is a lack of a focus on end-to-end measurement with more common focus on link layer measurements by most carriers. Testbed and measuring tools are available but did not see many studies to address user experience with new technology-LTE.

- **Privacy issues**: Crowdsourcing is an effective way to collect mobile measurement data, compared to war-driving every 6 months. But integrating these crowdsourcing data, possibly noisy data, is challenging especially when they are under different privacy policies. Because of the opposition from providers, it is hard to distribute data collected. Government has not adopted specific privacy policy and instead adopts policies of whoever they work with). One of the challenge is the difficulty to integrate data with different privacy policies.

- **Scalability challenges**: it is generally difficult to scale measurement.
- **Standardization needs**: there is a strong need to standardize measurements embedded within different equipment.
- **Diverse platforms**: Different platforms have different levels of support, e.g., Android is generally better than Apple's iOS platform due to a lack of the support to schedule measurements on iPhone, but both could still be further improved.
- **Data completeness**: Using crowd-sourced mobile measurement data, one can end up with lots of data, but also faced with a sparse data problem. An alternative approach is drive testing (hired companies). For both, it is hard to extrapolate inside buildings. An important question is how much data is enough in order to make claims to generalize the findings from the data.
- **Relevance to emerging new commercial wireless technology**: the testbeds track the evolution of commercial wireless technology.
- **Data analytics**: large amount of data can turn into sparse data without analyzing the data and revealing deep insights or patterns, privacy concerns when collecting data, quality of collected data. If the data are from users, how do we know users truthfully report the data?
- **Testbed usability**: Testbeds are usually created with a target application in mind. As a result, they are carefully tuned to the PI's research needs. Adapting the testbed for shared use is often surprisingly difficult and creates in sub-optimal results. There is also concern that we are sometimes replicating or performing work that might be done better by industrial partners, so partnership with industry can be helpful.

From the feedback we collected, we also noted the following key questions that we should ask for designing testbeds:

- **Testbed purpose**: What is the purpose of a testbed, How can it be broadened to support the maximum number of projects.
- **Data collection**: What is the structure that would best benefit industry, government, and academia to build integrated testbed or collect measurement data? How to deal with the noise, privacy concerns, and incentives when collecting crowdsourcing data?
- **Integration with existing testbeds**: How can we best utilize existing testbeds? What new infrastructure is needed to perform exciting measurements?
- **Usability of testbeds**: There is an enormous amount of effort expended in creating each testbed. As a result, the PIs should be able to get "first access" privileges. However, I think far too many of these testbeds only see use by one or two groups. It might be necessary to reconsider how we approach testbeds, in terms of how to make it more widely accessible.
- **Industry involvement**: What is the main challenges the Wireless industry facing, do Academic studies provide good feedback to the industry?

3.3 Operational Challenges

We summarize the key issues and challenges identified in this panel here.

- **Difficulties of instrumentation to support measurement**: there is a limited ability to perform instrumentation, e.g., phones, base stations. There are clear tradeoffs between active monitoring and passive measurement: Monitoring baseband information requires more energy, as waking up the phone to do the measurement can be expensive. In general, companies are not really interested in altruistic investment to makes access to data easier, not clear how to generate critical mass to get them to invest into instrumentation.
- **Virtualization-imposed challenges**: virtualization has its own challenges (following a bad experience report, hard to know what happened at the time). It is hard to look at CPU path for all packets due to slow down of the network. It is important to design measurement from scratch in this network.
- **Network complexity**: it adds complexity to measurement challenge. QoE at today's cellular network faces cross-layer challenges and the complex protocol interactions.
- **Lack of data sharing**: Carriers, device manufacturers and chipset providers do not often share data or make their interface for collecting data openly accessible. This is detrimental to improving performance or diagnosing faults. Carriers may

know things but don't seem to be operationalizing that knowledge.

- **Cross-company coordination**: Surprisingly Google does not control what Android information to be available. This implies that coordination across companies is necessary. Academia may be able to provide a bridge between these disparate parties to create incentives for sharing. NSF could broker industry-university consortium that have legal agreements for collaboration and data-sharing, in which the consortium builds shared testbeds, share datasets, and pilot new protocols and algorithms. Academia can act as a bridge.

- **Innovation in areas controlled by industry**: we discussed the need to create an infrastructure allowing us to innovate in areas currently controlled by cellular providers. It also helps to have standardized virtual platform be implemented for a range of measurements and also adapt to specific network characteristics.

- **Realistic data and infrastructure**: it is important to recruit real users to use experimental infrastructure to provide realistic workloads and use cases. As concrete action items, the research community can build our data archive and create benchmark at least for use by the research community. Another focus can be to develop experimental infrastructure and open testbeds facilitating research in this area.

3.4 Data Collection and Privacy

Key open questions. How can we standardize the IRB process? Is there a way we can get to a middle ground where we can sign an NDA and get access to sensitive data? Maybe companies should be "forced" to give up their data if someone signs an NDA and go through the due process. Can we establish (at least within US) a standard privacy policy for collection of mobile data? How can we do risky things? We need better testbeds to do this. What can academia do to incentivize industry companies to share their data? What unique contributions academia can make in collecting and sharing data? How can industry and academia work together to support mobile measurement in privacy-protecting ways? How do we better work with the HTTPS protocol from a measurement perspective? Is there a differential privacy solution to sharing data?

Challenges for collecting data and maintaining privacy/security. Participants suggested that the key challenges are addressing the following topics:

- **Standardized mechanisms for allowing policy-based access to data.** Picking a data-sharing models can be challenging and models tend to be one of: anonymize everything and share with everyone, no sharing, or share it with anyone who signs NDAs. A key open challenge is that we are not lawyers and legal implications can vary by jurisdiction.

- **Community standards to ensure a level playing field for all researchers and keep us a step ahead of industry.** Results from existing industry/academia collaborations are also hard to verify and not conclusive, because there is no ground truth. For example, many studies leverage CDR data. But there are many black boxes inside (e.g., processing procedure/data coverage). In general, there can be many misuses of the datasets. There are many aspects that the community should be more scientific.

- **Incentives for industry to share data.** Government agencies can fund industry/academia collaborations that require open datasets. Another potential solution is to collect some "case study" or "benchmark" data from industry. We can collect the case study data set, not necessary conclusive, but can be used as benchmarking. It is similar to the Netflix challenge, but the incentive needs to be further explored. We also discussed how data sets with mobility, useful not only for understanding human mobility, but also used for mobile simulations and mobility management evaluations. In this context, industry has way more data, as well as data processing technology, but getting data from industry is very difficult.

- **A common ground or norm for data privacy/security accepted by the community.** A common thread among participants is that there is much confusion around IRBs, ethical standards and privacy — examples of which vary according to where you are and who you ask. One participant suggested that the community needs to develop a set of community 'norms' for ethical mobile measurement when human subjects are involved. The NSF and ACM could perhaps support the community in developing and disseminating these norms among research institutions and their IRBs, and brokering connections to relevant international organizations. Another common thread is the question of how much academia should be repeating work already done (often more thoroughly) in industry (but not shared openly).

Another challenge is improving IRB domain expertise so they are more effective/efficient in reviewing studies from CS. Last, a key challenge is that data privacy and access is a moving target.

4. RECOMMENDATIONS AND FUTURE DIRECTIONS

At the conclusion of the workshop, it was clear that mobile community measurement infrastructure is a broad area with many challenges to address across many disciplines and layers of the networking stack. We heard a wide range of talks about research efforts that were interesting individually, but there was a clear need for focus, communication, coordination, and collaboration across different research, industry, and governmental groups.

Through discussions and survey feedback, the following common themes emerged for recommendations to make progress toward sustainable, innovative mobile community measurement infrastructure:

- NSF funding of long-lived infrastructure to host collaborative testbeds for mobile measurement. In contrast with current efforts and solicitation, which focus on research products from measurements, there needs to be a program that funds long-term development and maintenance of measurement infrastructure as a service to the research community.

- Beyond support for a testbed, we recommend that researchers should work with vendors, service providers and government advocates to ensure measurements are integrated into wireless systems. There is a need for open and innovated testbeds, including low-layer tools for performing measurements, systems/architectures to simply measurement collection and techniques to use/combine the measurements effectively. Last, we need better instrumentation to understand the performance difference for new and current spectrum users.

- Incentives for researchers and industry to work together and share data. This can in part be solved by government policies or programs that provide "carrots" or "sticks" to encourage cooperation. Of course, it is also incumbent on researchers and industry to find mutually beneficial projects on which to collaborate.

- Instrumentation across all layers of the wireless stack, from spectrum to PHY to application layer, and we need to identify how to combine measurements across layers to address problems in today's wireless networks and to inform future network designs.

- Address data privacy and security issues in the mobile environment, particularly due to the increased risk for leaks of subjects' personally identifiable information (PII). We should develop community standards for gathering, securing, and sharing such data, and ensure that these policies are compliant with jurisdictional restrictions. Further, we should develop ways to "reward" those who comply with community standards for sharing data, to encourage the practice.

- We should improve IRBs, *e.g.,* creating a technology-focused IRB, and look into to how social scientists (*e.g.,* Census Bureau) deals with similar issues we are facing in the wireless measurement domain. Another participant suggested that we should establish best practices and policies for data sharing. We also should seek collaboration opportunities with industry and think deeply about what unique contributions academia can make. Last, NSF and societies (like ACM SIGMOBILE and SIGCOMM) should guide the community toward documenting and disseminating community norms, toward establishing data archive facilities, and toward mechanisms to reward data sharing and proper data citation practices.

- More research and policy effort need to be devoted to making testbeds and data set truly open and accessible to the community. It is important to work closely with industry to not duplicate their efforts and instead focus on challenges that lead to addressing more forward-looking aspects of the network design across different layers for next-generation mobile networks.

- Better communication between research groups, industry, policy makers, and pertinent government agencies. It was clear that many participants were unaware of salient testbeds, measurement approaches, and government initiatives presented at the meeting. We suggest future workshops that focus on bringing together more focused groups of participants who share stronger interests, and use these meetings to build a community around critical mobile measurement infrastructure.

Acknowledgments

This workshop was funded by the NSF under award number CNS-1455719. Aruna Balasubramanian, Mike Wittie, Chunyi Peng, and Aaron Schulman contributed valuable scribes of the proceedings. We thank the NSF and participants for their contributions to this workshop. In particular, we thank Thyagarajan Nandagopal from NSF for initiating the idea for the workshop and his valuable input.

5. REFERENCES

[1] NSF Workshop on Mobile Community Measurement Infrastructure (2014). `http://www.ccs.neu.edu/home/choffnes/nsf-meas-wkshp/index.html`.

[2] k. c. claffy, D. D. Clark, and M. Wittie. The 6th Workshop on Active Internet Measurements (AIMS-6) Report. *SIGCOMM Comput. Commun. Rev.*, 44(5):39–44, Oct. 2014.
[3] D. D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski. A knowledge plane for the Internet. In *SIGCOMM*, 2003.