CrowdREM: Harnessing the Power of the Mobile Crowd for Flexible Wireless Network Monitoring

Andreas Achtzehn, Janne Riihijärvi, Irving Antonio Barria Castillo, Marina Petrova, Petri Mähönen
Institute for Networked Systems, RWTH Aachen University
Kackertstrasse 9, D-52072 Aachen, Germany
{aac, jar, iba, mpe, pma}@inets.rwth-aachen.de

ABSTRACT
High-speed mobile broadband connections have opened exciting new opportunities to collect sensor data from thousands or even millions of distributed mobile devices for the purpose of crowdsourced decision making. In this paper, we propose CrowdREM (crowdsourced radio environment mapping), a framework with the specific aim of monitoring and modelling wireless cellular networks. CrowdREM enables operator-independent and highly efficient collection of network performance data along all layers of the communications protocol stack. Such extensive information on network load, spectrum usage, or local coverage can help operators to optimize their networks and service quality and enable improved consumer decision making. In this paper, we introduce the CrowdREM mobile architecture and show first results from a prototype implementation on open-source mobile phones. We demonstrate the versatility of using commodity devices for network and spectrum monitoring, and present the challenges originating from the use of uncalibrated and low-precision measurement equipment. We have acquired an extensive data set from using our prototype implementation in a 21-day measurement campaign covering more than 1000 hours of measurement data. From this we present and discuss the potential derivation of tangible and relevant network performance and signal quality indicators, which could, e.g., be conducted by independent parties.

Categories and Subject Descriptors
C.2.3 [Computer-Communication Networks]: Network monitoring; C.4 [Performance of Systems]: Measurement techniques

General Terms
Measurement; Performance; Reliability

Keywords
Crowdsourcing; Mobile; Drive Testing; Cellular Networks

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

HotMobile ’15, February 12–13, 2015, Santa Fe, New Mexico, USA.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3391-7/15/02...$15.00.
http://dx.doi.org/10.1145/2699343.2699348.

1. INTRODUCTION

The densification and complexity explosion of wireless cellular networks is putting high burdens on network operators. Monitoring their networks for coverage holes or low performance becomes increasingly difficult, if not impossible. Traditional approaches of carrying out extensive drive testing campaigns with expensive personnel and equipment are reaching their limits, and it has become obvious that the “intelligence of the crowd” will need to be employed in order to manage networking complexities. LTE release 10 [8] has therefore foreseen a possibility to collect network performance statistics from user equipment (UE). A significant limitation of a proposed minimization of drive testing (MDT) is its focus on a single network and few user link performance aspects. For establish a truly comprehensive view of wireless network operations, extended means of collecting performance statistics from UEs will need to be employed.

Several projects already gather performance data on cellular networks for this purpose, see [7]. NetRadar [2] collects traces from mobile devices to acquire localized link throughput statistics; OpenCellID [5] collects cell tower information and correlates them with user locations to establish network density and coverage maps; and finally, OpenSignal [12] measures the signal strength of cell towers to estimate achievable throughputs. However, none of the current approaches yet takes a holistic view along all operators and protocol levels to enable advanced modeling, e.g. through spatial statistics and machine learning based filtering methods, to build highly relevant coverage and performance estimates.

In this paper, we present CrowdREM, a framework for collecting spectrum usage and protocol data with sensor fusion capabilities. CrowdREM’s uniqueness comes from its capability to jointly collect spectrum and performance data from several local network operators; both, at the level of pure signal level measurements as well as through monitoring of relevant control traffic. Such extensive data, if combined and cross-correlated with local information [3] and other higher layer performance metrics, can yield an enhanced understanding of the current state of the radio environment [14, 6]. This will help operators to conduct terminal-centric network tuning, and allows users to adapt their download schedules to reduce radio activity overheads [4]. Further, CrowdREM may be maintained, e.g., by a third party to offer real-time, independent information for consumers and local regulators.

To demonstrate its feasibility, we have implemented a proof-of-concept CrowdREM on an open-source mobile phone platform. We present the challenges of distributed spectrum sensing with inexpensive hardware such as the signal degra-
The power of a policy-based data collection lies in the flexibility of defining rules for the collection process. It constantly monitors relevant system sensors such as the brightness sensor or the gyroscope to determine where the UE is currently located, and whether deviations from the regular collection schedule are necessary.

Our architectures separates pure RF spectrum sampling from the collection of control traffic data. RF sampling comprises sweeps over the various frequency bands to determine power levels according to a scheduler-defined frequency plan and bandwidth selection. The scheduler needs to take particular care that this will not disrupt ongoing transmissions. In contrast, control traffic collection may be carried asynchronously and even while user interactions take place. While normally UEs discard control traffic not intended for their own operations, for CrowdREM the baseband system needs to pass through all messages. Further, the CrowdREM scheduler needs to enforce a channel sequence to be monitored and a dwell time for the collection of control messages.

We find that control traffic information is well suited to determine network load, because it alleviates the hidden node problem purely power sampling approaches as used e.g. in [4]. The extra power required for control traffic sampling itself is minimal, since the UE constantly needs to monitor broadcast transmissions anyway; any overheads thus originate from storing and processing steps. The collected measurements are augmented with regular system sensor information to determine, e.g. the relative orientation of the UE towards the serving cell tower. They are stored in a local database, which is synchronized asynchronously and depending on bandwidth availability with the cloud-based part of the CrowdREM architecture.

All heavy processing tasks of CrowdREM have been designed to operate in the cloud to provide better battery lifetime of the UEs, maximize the crowdsourcing benefits, and minimize maintenance overheads. A central CrowdREM controller manages the integrated UEs by distributing the high-level collection policy. The controller uses the available data in the CrowdREM data mining backend. For example, it would increase the sampling cycle for UEs located in regions of high network load volatility or limited sampling device density. The CrowdREM controller would then adapt the policy of a selection of UEs to create more fine-grained sampling results. As the CrowdREM controller has a global view of the radio environment, its decision making is drawing directly from the distributed nature of the crowdsourcing approach. Several data views can be defined to study the various performance aspects of the network. Thereby, different stakeholders can review the CrowdREM data according to their analysis needs.

2. CROWDREM ARCHITECTURE

In the following we provide a brief overview of the envisioned CrowdREM architecture. CrowdREM differs from approaches such as [9] as it is designed to operate on regular UEs, e.g. smartphones, tablets, or laptops with wireless connections. Fig. 1 shows that its design is divided between local data collection, cloud data processing, and storage.

At the core of the UE components, a CrowdREM scheduler manages the collection of radio environment and system data. Its main purpose is to balance the requirements of data storage and sampling costs in terms of energy consumption and service level, with the purpose of minimizing the impact of CrowdREM on the UEs regular operations and battery lifetime. As noted by Lane et al. [10], advanced prediction-based schedulers for smartphone crowdsourcing can reduce energy consumption by up to 90% depending on the scenario. For this purpose, the scheduler runs a dynamic collection policy, which can be adapted if, e.g., more extensive data collection becomes necessary. The CrowdREM scheduler furthermore monitors user behavior, in order to preempt collection tasks when user interactions take place.

Figure 1: The CrowdREM architecture.
the used APIs are OpenMoko-specific, their capabilities are implemented in every mobile phones. We presume that vendors may disclose their APIs given the benefits, based also on more recent discussions in the 3GPP MDT context.

After installing a modified Linux on the phone, the baseband processor firmware was replaced by OsmocomBB [13], a software suite which provides layer 1 to 3 functionalities with a fully open API. We have designed a wrapper to control OsmocomBB, collect data, and store it in the local data repository. We note that, while these steps are necessary for this particular platform, commercial smartphones may expose the same API without firmware modification. The required sniffing and power measurement components are necessarily implemented in any mobile phone, thus we consider exposing them e.g. to an Android/iOS app would be reasonable given the benefits for the various stakeholders.

We have conducted an extensive measurement campaign with the prototype in the mid-sized city of Aachen, Germany. To gather network statistics, we distributed identically configured OpenMoko phones to university students. The campaign lasted for 21 days, from which we acquired approximately 13,000 measurement session sets. We manually selected a simplified collection policy that allowed for a particularly broad data set to be generated at the cost of shorter battery lifetime. Each measurement session consisted of an initial phase in which the UE\(^1\) scanned across all physical E-GSM channels (ARFCNs) to yield power levels and network/cell identification information. In the city area, four different network operators are active, thus we selected for each of them the strongest ARFCN for collecting control traffic. The individual per ARFCN dwell time was set to 45 seconds and all decodable messages in the logical broadcast and common control channels (BCCH and CCCH) were stored along with additional system level sensor information.

3. CELLULAR NETWORK
RADIO ENVIRONMENT ANALYSIS

Using examples from our prototype implementation we will in the following show the challenges and opportunities in crowdsourced measurement setups. In this section, we focus on those physical layer measurements that are used to determine the radio environment of a UE.

3.1 Determining UE Orientation

No UE antenna is perfectly omnidirectional, thus signal strength variations may be observed even if the position of the UE remains unchanged. This defines an initial requirement for a crowdsourced measurement approach, which is to implement means to determine the relative orientation of the UE to the monitored network infrastructure, e.g. by integrating gyroscope measurements. In order to illustrate the need for this additional information, we have carried out calibration measurement through running an open-source GSM-BS (OpenBTS) [1] with the popular USRP platform acting as a radio frontend. We mounted the UE to a tripod and took measurements while rotating the phone along its \(x\), \(y\), and \(z\) axis. Note that the UE antenna is located in the bottom part of the phone. Figure 2 shows that the measured signal strength varies by 4.4 dB when the UE is rolled (rotating the \(y\) axis), which can be considered minor and is

1 In the following we refer to the GSM mobile station (MS) as UE for the sake of conformity.

within the measurement precision requirements of a GSM phone. While this is an expected outcome due to the geometry of the device, larger deviations could be observed when the \(z\) axis of the UE was rotated. Here, an up to 10.7 dB difference in signal strength could be observed depending on whether the bottom of the UE was pointing towards the BS or away from it. Similar observations could be made when the phone was rotated while standing upright, i.e. the phone is considerably more sensitive from the back than from the front, which is a common antenna design for smartphones. This highlights the necessity to integrate this sensory data in the data collection process and compensate for resulting uncertainties in the analysis.

3.2 UE Localization

The second challenge in crowdsourced measurements arises from imprecisions in the localization process. Small deviations in the position of a UE can yield significantly different measurements, particularly when measurements are taken in complex indoor radio environments. This is especially challenging since traditional localization methods using GPS generally fail to work in such locations. A crowdsourced measurement setup such as CrowdREM must therefore collect auxiliary information from the UE sensors to augment GPS position estimates, e.g. through Wi-Fi fingerprinting.

In the following we illustrate how extensive these measurement deviations can become in a single UE scenario with GPS-only positioning. For this, we selected the data from a UE that was often located in the same building during our measurement campaign. In total, 14 distinct measurement periods (comprised of consecutive measurement sessions) could be identified where the phone reported to be placed in an area of approximately 40 m in or around the building. A sensible assumption is that all these measurements were taken inside the user’s apartment.

In Fig. 3 we show the mean measured signal strength on the 4 strongest ARFCNs for all measurement periods. Through readings from the control traffic analysis (see Section 4) we found that these ARFCN measurements always map to the same BSs. Up to 8 dB of relative deviation in

![Figure 2: Reported signal strength (in dBm) of the CrowdREM prototype with respect to its orientation towards the BS.](image-url)
the measurements between consecutive periods are reported, and only for period numbers 9 to 12 more stable results are found. However, as we can see from the relative values in the figure, there is an apparent correlation between signal strength values of the different ARFCNs. We can thus conclude that despite lower localization precision, it is possible to at least establish relative signal strength distributions for various operators and BSs.

An example of how CrowdREM’s cross-operator view of spectrum usage is beneficial is depicted in Fig. 4. Here, we have normalized the signal strength values of the 48 strongest ARFCNs to the value of the strongest ARFCN. We see in this figure the similarities between different measurement periods, e.g. 3, 5, 7, 11, and 13. We assume that the UE was located in the same place when these measurements were taken, because such result would generally only be observed with the same overall spatial geometry. By comparing these results to the absolute signal strength values of Fig. 3, we see that all values lie within the precision limits of the device. The CrowdREM architecture is thus capable of improving sensing consistency by taking the overall geometry of the radio environment into account.

4. CONTROL TRAFFIC ANALYSIS

While signal strength measurements yield an estimate of the physical characteristics of the radio environment, a detailed analysis of the control traffic enables us to derive relevant statistics on the performance and load of different network operators. To the best of our knowledge, CrowdREM is the first framework that explicitly uses control traffic data for such an extensive performance analysis. In the following we first discuss the issue of message losses and how it affects the monitoring accuracy, before providing example statistics on relevant network operations parameters.

4.1 Limitations of Single UE Traffic Measurements

Monitoring control traffic requires the UE to tune to the broadcasting channel(s) of the BS(s) and collect all decodable messages intended for all other UEs within the same location area or cell. Such collection does not require for the UE to be part of the operator network, because control traffic is generally sent through unencrypted broadcasting channels. However, since the UE itself may be located in a location with adverse signal conditions (low signal strength or high interference from other cells), not all messages can be necessarily decoded. Monitoring-derived statistics from single UEs may thus underestimate loads in cellular networks.

To illustrate the significance of this limitation we have conducted a small-scale measurement campaign with two CrowdREM devices. The first device was placed nearby a known GSM BS and set to permanently collect this BS’s control traffic. The other UE was moved to the fixed locations depicted in Fig. 5b which are within the decodability-determined cell range of approximately 1.3 km. At each location, control traffic was collected for 1 minute from which we derived the message loss rate by comparing the captured messages to those from the stationary device.

Our comparison of the fractional message loss and the received signal strength of the messages in Fig. 5a indicates that if the signal strength value drops below approximately -60 dBm, the mobile UE is no longer able to decode all messages. The message loss becomes more severe when the signal strength is further lowered, e.g. if the signal strength is at on average -90 dBm, only 50% of the messages are still retrieved. Near the overall decodability threshold, virtually all messages are lost. We can see from the plot that there is a region between approximately -60 and -75 dBm where there is a clear relationship between the received signal strength and fraction of lost messages. This relationship can be used to compensate for the naturally occurring censoring bias in the data, and improve the quality of our estimates of network load provided that sufficient calibration information is available. Gathering of such calibration information is enabled through a crowdsourced architecture such as CrowdREM, which allows to cross-correlate message rates at various monitoring UEs. For lower signal strengths no clear relationship to the fractional message loss is observed, i.e. there are further deviations in reception quality not sufficiently captured by the signal strength value.
4.2 Crowdsourced Network Load Estimation

A further benefit of using multiple coordinated UEs to collect control traffic is that message loss and limited UE availability do not affect the overall collection capability of the crowdsourced data mining architecture in CrowdREM. Since only a smaller number of devices are needed if good reception quality is experienced, one may extend the battery lifetime of the crowdsourcing UEs by scheduling less frequent data collection. In the following, we discuss relevant network load indicators that can be easily derived thereby. The results are illustrative for an advanced data mining that is enabled by using the CrowdREM architecture.

In our first analysis example we study the network load by means of quantifying the paging load observed for a single operator. In the GSM system architecture, paging messages are used to request channel reservations for UEs, e.g., to initiate a network-initiated voice call. In Fig. 6a we plot the rate of different paging messages for the location area of a single operator as measured by various UEs over 17 days in our campaign. We see that paging messages are dominated by type 1 requests, which are messages sent to up to 2 devices. Other paging messages for more than two devices (type 2 and type 3) do not contribute to the overall paging rate. Surprisingly, there is no apparent correlation between the time of day and the paging rate, which would have been expected from increased day-time user activity. Occasionally, higher paging rates can be observed during the night, which seems counter-intuitive. On deeper inspection we found that the operators intentionally insert blank paging messages, presumably to allow UEs to synchronise more frequently, which also explains the relative steadiness of ARFCN measurements (see Section 3.2) over time. Thus, we conclude that paging messages are a weak indicator of network load.

A better indicator for network activity is the number of individual UEs queried in paging messages. Fig. 6b exhibits the expected day-time dependency of network load. Furthermore, we can observe that during weekends the number of paged UEs drops, as seems intuitive. Through comparison of the different days we can make further derivations on the user behavior, e.g., increased activity is observed on Monday mornings, whereas the network load steadily drops already on Friday afternoons.

The time of day dependency can also be found in an analysis of the immediate assignment message rates of a single BS. These messages are sent to allocate channels to UEs within the cell, thus they are only observable by a smaller group of monitoring UEs in nearby locations. We take the example of a single cell that was monitored by a single CrowdREM UE. We note that such analysis is only possible if the UE is not moved during the day, and when beneficial signal quality levels are observed. For lower signal levels, means of message loss compensation as discussed in Section 4.1 need to be applied. A plot of the probability density function of the inter-arrival rate in Fig. 7 shows that during night times the inter-arrival time increases, i.e., fewer requests are made per unit time, while it noticeably decreases starting at around 5am in the morning. We can further observe lower rates during lunch time, when supposedly less calls are made. Such figures allow for the derivation of tangible load models for application in network traffic modelling as we, e.g., can already infer from the data set used to create Fig. 7 that heavy-tailed distributions are more appropriate to model the observed time variability.
5. CONCLUSIONS

In this paper, we have presented CrowdREM, a framework for the crowdsourced monitoring of cellular networks. CrowdREM enables the distributed collection of spectrum usage and network information from inexpensive mobile terminals to build a comprehensive and holistic view of the structure and use of wireless networks. Contrary to earlier works, CrowdREM is not limited to a single operator or networking aspect, but it provides means to independently derive performance statistics relevant to network operators, customers, and regulators.

Our practical analysis with a prototype implementation has showcased the benefits of using a crowdsourced data mining, where limitations in positioning accuracy, orientation, and message loss experienced for individual measurements can be compensated for through combined sampling from multiple sensor entities. We have furthermore used results from an extensive measurement campaign covering more than 1,000 hours of measurement data on live cellular networks to demonstrate the derivation of relevant network performance indicators such as the paging rate, number of active users, and the time dependency of network load. While our proof-of-concept implementation runs only on 2G networks, the principal idea of CrowdREM extends also to more current network generations.

In the future we plan to fully integrate and port CrowdREM to other platforms and 3G/4G networks to enable more extensive data collection, whereby we are also interested in other cross-layer metrics such as browser performance. Further, we are interested in studying social and economic incentives for participating in a crowdsourced data collection, e.g. through more precision localization from extended environment fingerprinting.

6. REFERENCES


Figure 6: Paging message rates and number of individual paged UEs per second over a period of 17 days for a single location area (LA). Rates depicted here are calculated individually for each 45-second measurement session.