www.digitalglobalsystems.com

Enhancing Crowdsourced Measurements for Wireless Network Optimization

October 2023

Dr. Armando Montalvo

CTO, Digital Global Systems

Table of Contents

ABSTRACT

In this paper, we discuss the challenges inherent when using crowdsourcing to estimate the performance of wireless networks. In general, these methods give a measurement of the user experience on the wireless network that only in the best of circumstances can be related to the performance of the network, but in most cases, it requires substantial postprocessing of the crowdsourced data to relate to information about the performance of the network that could be used to optimize the network to support emerging services. We introduce concepts typically used in control systems to explain how to use persistent RF environmental measurements to overcome these challenges. We also give an example of how this approach can be used to optimize the placement of access points in a network to improve its performance.

INTRODUCTION

Wireless networks underpin many essential services in today's interconnected age, making their optimal performance paramount. Crowdsourcing has emerged as a prevalent method to capture user experiences using the network and gauge network performance relative to the user. In general, relying solely on this approach offers a limited and often skewed perspective of the wireless network performance. Crowdsourced data, harnessed from various devices across varied locations, offers insights rooted in real-world scenarios. However, it is inherently influenced by myriad variables such as device types, usage patterns, and other external factors [1]–[4]. This variability often necessitates substantial post-processing to yield actionable insights on the wireless network's performance. Accurate network performance measurement is not just determining speed or bandwidth. It encompasses understanding latency, jitter, packet loss, coverage, and other vital metrics that can significantly affect user experience [11]. An accurately measured network can be efficiently optimized, ensuring seamless user experiences, especially for time-sensitive applications like video conferencing, online gaming, and telemedicine [12].

Furthermore, accurate measurements guide infrastructure investments, ensuring they are directed where they will yield the most significant improvements. Crowdsourcing, in this context, has emerged as a popular method to gather data about network performance as a function of the user experience [13]. Leveraging the power of the masses, it offers an opportunity to gather real-world user data from a plethora of devices across diverse locations [13]. This vast data set can offer insights that are difficult to obtain through traditional measurement methods. However, as with any tool, it comes with its unique challenges, which this paper aims to address and provide solutions for.

BACKGROUND

At its core, crowdsourcing involves soliciting contributions from a large group of people, especially an online community, rather than traditional employees or suppliers [14]. In wireless network performance measurement, crowdsourcing primarily means utilizing the vast number of wireless device users to gather data about their real-time network experiences [13], [14]. A comprehensive and geographically diverse dataset can be built by having numerous users run speed tests, report coverage gaps, or simply share passive data while they use their devices [5]. This information from real-world scenarios provides invaluable insight into the daily user experience across different device types, locations, and times of day.

The history of wireless network assessment has been primarily dominated by expert-driven approaches, with technicians conducting site surveys, specialized equipment gauging signal strength, and proprietary software monitoring network health [15]. However, as the number of wireless devices exploded in the late 2000s and early 2010s, a new paradigm was needed. This surge coincided with the rise of social media and the gig economy, both of which normalized the act of sharing personal experiences and data online [16]. It did not take long for innovators to realize that the millions of smartphones, tablets, and other connected devices could act as individual data collection points, each offering a snapshot of the network's performance from a user's perspective. Several platforms and applications, like Speedtest.net and OpenSignal, capitalized on this concept, allowing users to actively test their connection speeds and, in the process, contribute to a global dataset [17]. Over time, these crowdsourced measurements became pivotal, not just for consumers looking to compare service providers but for the providers themselves, eager to identify network usage shortcomings by the users.

While crowdsourced data is expansive and rooted in real-world user experiences, it has flaws. First and foremost, it is inherently uncontrolled [4]. The data comes from a myriad of device types, each with its hardware and software configurations, which can introduce variability in measurements. External factors, such as a device's battery level, running applications, or even the user's location (indoors versus outdoors), can skew the results [18]. Another challenge is the inconsistency in data density. Urban areas, dense with users, might produce a deluge of data, while rural areas could have sparse data points [19]. This discrepancy can lead to biased optimizations favoring regions with higher data contributions. Lastly, the voluntary nature of most crowdsourced platforms means there is a self-selection bias. Users experiencing issues might be more inclined to report or run tests, leading to overrepresenting problematic areas while well-functioning zones remain underrepresented [19]. In the face of these challenges, it becomes clear that while crowdsourced data is a potent tool to identify the user experience, it cannot be taken at face value to determine the performance of the wireless network without sophisticated post-processing and validation mechanisms that can distill this raw information into actionable insights into the actual performance of the wireless network.

Challenges in Crowdsourced Wireless Network Measurements

Every individual's experience on a wireless network is unique, influenced by numerous factors ranging from physical location to the specific task they perform on their device. For instance, a user streaming a high-definition video in a high-rise apartment might have a different experience than another sending emails from a suburban home, even if they are on the same network. When collated, these diverse scenarios contribute to varied data points, introducing significant noise into the dataset [14].

Impact of Device Types, Usage Patterns, and Other Variables

The hardware and software landscape of wireless devices is vast. Measurements taken from a high-end smartphone might differ significantly from those from a budget model, given the disparities in their antennas, processing capabilities, and other hardware components [1]. Similarly, software configurations, such as operating system versions and background applications, can influence the performance metrics [2]. Usage patterns further compound these challenges. A device primarily used for web browsing will have different network demands than one used predominantly for online gaming or video conferencing [3], [4]. Other variables, like battery levels, can also play a role—some devices may throttle their network connectivity to conserve power at low battery levels, impacting the performance results [3].

Inconsistencies in Measurements

The resulting crowdsourced data can be inconsistent with such a diverse set of inputs. Two users in the same location but with different devices or usage patterns might report drastically different network experiences [20]. These inconsistencies can lead network providers down the wrong optimization path. For example, suppose most users in an area use older devices that report poor performance. In that case, the network might be erroneously flagged for optimization, even when the actual network infrastructure is robust and efficient. Furthermore, as previously noted, data density disparities between urban and rural areas can lead to over-optimization in densely populated regions, leaving less populated areas underserved [19].

Need for Postprocessing to Make Sense of the Raw Data

Given the challenges described, it is evident that raw crowdsourced data, in its unfiltered form, can be more misleading than informative. Substantial postprocessing is essential to extract value from this wealth of information [21]. This includes normalization techniques to account for device variations, algorithms to mitigate the effects of external variables, and spatial analyses to ensure geographically balanced insights [21]. After this rigorous refinement, crowdsourced data can be deemed reliable for guiding network optimization strategies.

In general, given the challenges mentioned above, crowdsourcing measurements using smart phones or other UE devices does give you a measurement of the quality of the user experience, but it does not give clear measurements of the quality of the network, nor does it give you information on where and how to improve the performance of the network.

Control Theory Concepts

To better illustrate this, let us consider the wireless communication network as the system under test, where some users (the UE equipment) inject signals and the collection server in the crowdsourcing approach is observing the output of the system under test. Here, the wireless communication network is a system with many internal nodes, some representing wireless links (between the UE and base stations and between base stations) and others representing internal wired links (like fiber paths between RANs and cores in the network, for example). In this model the performance of the network is directly related to the states of all the internal nodes in the network, thus the condition of the network and its capabilities to handle the input traffic injected by the users in the crowdsourcing approach. These states are constantly changing over time and are affected by the RF environment, for the wireless links, and the network health for the wired links and the initial state of the system. From this point of view, the wireless communication network acts as systems to be controlled for a particular performance given a set of inputs. This is the typical problem in control theory [15].

Borrowing several concepts from control theory, we know that the systems can be controlled if and only if the inputs can reach every internal state of the network. Otherwise, there are internal states that can cause problems (like instability, waste of resources, etc.) that the input will never the able to execute. Thus, we can never correct their impact in the network. Most of the crowdsourcing methods know nothing about the topology of the network (in relation to the wireless or wired links (i.e., the internal states of the system). They tend to make measurements on observations of the output of the network (the system under test). We also know that in such circumstances if the internal states of the network are not observed from the output measurements, then the system is not observable, and any measurements will not reflect the true performance of the system. For a system to be well behaved the system should be controllable and observable (a minimal system in control systems terminology). Therefore, wireless communication networks should be built to be both observable and controllable. For the most part, the typical assumptions of current wireless communication systems are that all the RF links are understood, and all wired links are known such that the internal states of the network can be reached by a set of inputs and be observed at the set of outputs, thus resulting in a minimal system. Unfortunately, this is not the case in real life, where the RF links between the UEs and the bases stations in the network are not well known and are changing with time. This is even more pronounced in cases where the RF environment is congested with multiple signals not expected by the system operators. If these conditions are observed by the network, then adjustments can be made to make the system minimal again. This highlights the need for persistent RF environmental measurements in any wireless communication system.

USING RF ENVIRONMENTAL AWARENESS

Conventionally, wireless communication systems were designed under the assumption of unique service licenses, where only the wireless network operator who owns the license is using their RF environment (only one set of signals exist and the wireless network operator has complete control of those signals in their licensed bands), thus no congestion exists that is not controllable and using conventional KPIs makes the licensed spectrum observable. But in the case of congested RF environments, where multiple signals are present in addition to the ones controlled by the wireless network operators, the current KPIs might not offer observability. This results in a system that is not minimal and is not well behaved. In such cases, the network observations collected using the typical KPIs cannot be properly related to the internal states of the network and therefore cannot be used to help mitigate and optimize the network. Again, this can be corrected if persistent RF environmental measurements are available.

In current crowdsourcing methods, even if all the challenges mentioned above were meticulously resolved, the measurements are from the perspective of the user injecting input signals to the entire wireless communication network (which acts as the system under test), and the output of the network as the observable measurements. Therefore, these measurements can only be related to the state of the wireless communication network (the system) if and only if the system is minimal. In many realistic circumstances this is not the case. Thus, such methods cannot be used by themselves as the basis to evaluate the performance of any wireless communication system. What they do provide is a measurement of quality of the end user experience, which is valuable information for network operators to indicate if there are problems in the network but does not indicate where in the network the problems are or how to mitigate them. If operators want to mitigate these types of problems and optimize their network, persistent RF environmental measurements are needed and should be used to evaluate and optimize the performance of the network.

Given the advantages of crowdsourcing mentioned above, it is prudent to used crowdsourced measurement data to determine if there are problems servicing the users of the network and then use a data collection campaign to determine the cause of the problem and how to mitigate it. In a conventional data collection campaign, a mobile vehicle with special equipment is used to measure all the internal wireless links of the network. The resulting information will provide a time-specific view of the system that might or might not reflect the cause of the problem. To avoid this problem, several time-specific views of the system are needed, resulting in a costly approach to solve the problem (negating some of the advantages of using crowdsourced measurements). If the cause of the problem is not attributed to input signals in the network that are directly or indirectly controlled by the network operator (e.g., outside interference), even this multiple collection approach is not sufficient unless the collection equipment can detect and analyze all signals present in the RF environment where the network (system) operates.

A more effective method is to persistently monitor the RF environment where the system operates, detecting and analyzing all signals present in the RF environment. This process of detecting and analyzing all signals present in the RF environment to understand their properties and interactions is defined as RF environmental awareness. DGS has developed tools to support persistent monitoring of the RF environment while obtaining RF environmental awareness. The tools consist of either independent or integrated sensors in the radio access network (RAN) that can be deployed with existing base stations in a network or around strategic areas in a network. They provide persistent RF environmental awareness measurements that can be used to locate causes of potential problem in the network and supply information about how to mitigates such problems. When this information is correlated with crowdsource measurements pointing to potentially low QoE in the network, causes and mitigation approaches can be quickly identified, helping the network operator optimize the performance of the network in a dynamic and ongoing basis. Furthermore, this could lead to an automated network optimization procedure based on services that the end users want the wireless network to support.

Let us consider a real-world example. We deployed three sensors in an area covering Tysons, Virginia (Figure 1), where persistent measurements were collected, and RF awareness information was obtained over a period of several months.

Figure 1

Three main wireless networks were detected, along with other signals in the frequency bands adjacent to the bands the three networks were utilizing (700 MHz, 1900 MHz, 2100 MHz, 2300 MHz and 2509 MHz). Using RF awareness information obtained from the DGS sensors, several determinations were made about the performance of these three wireless networks. For example, we could quickly compare the composite area of coverage from each wireless network relative to the others. Using the strongest signal strengths measured by the DGS sensors across all bands, a composite coverage was easily constructed as shown in Figure 2, where a composite coverage comparison between pair of wireless networks is presented.

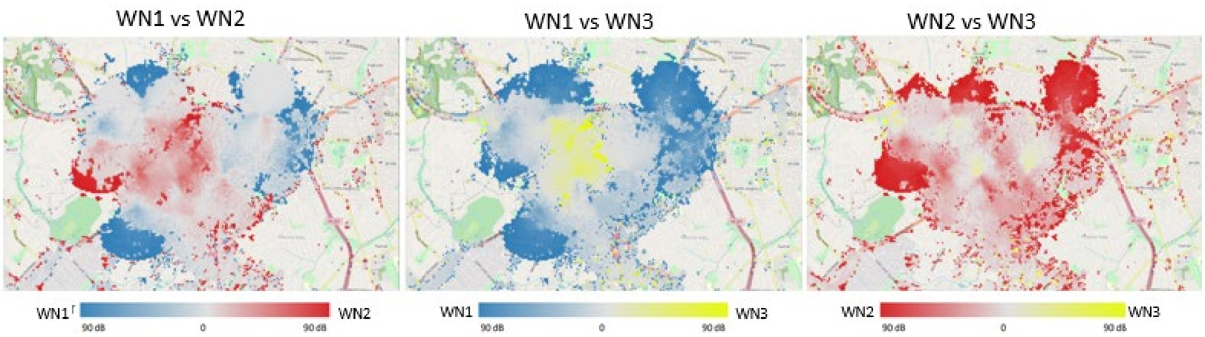

Figure 2

Other performance measurements were also determined for each wireless network detected in the area (e.g., quality of traffic served for each wireless network, total effective area of coverage for each wireless network, carrier to interference levels experience by each wireless network, among others).

From crowdsource measurements over WN1, poor connectivity was experienced during some of the measurement's intervals. Using the RF awareness information obtained from the DGS sensors and correlating this information with the events of poor connectivity, an area of poor connectivity coverage (denoted by low carrier-to-interference, C/I,) was identified (Figure 3). This figure also illustrates the typical baseline coverage of WN1, as well as the distribution and density of crowdsource data used and the average signal strength of the synchronization signals used by the WN1. Notice that the area of low C/I corresponds to where WN1 has a slightly lower signal strength, as well as where WN2 and WN3 have higher signal strengths. This conclusion, which could not be obtained without RF awareness information, can be used to guide WN1 to optimize its performance by either augmenting the power used in their synchronization broadcast channels or by considering deployments of new nodes (such as small cells) in strategic locations to improve the performance of the network.

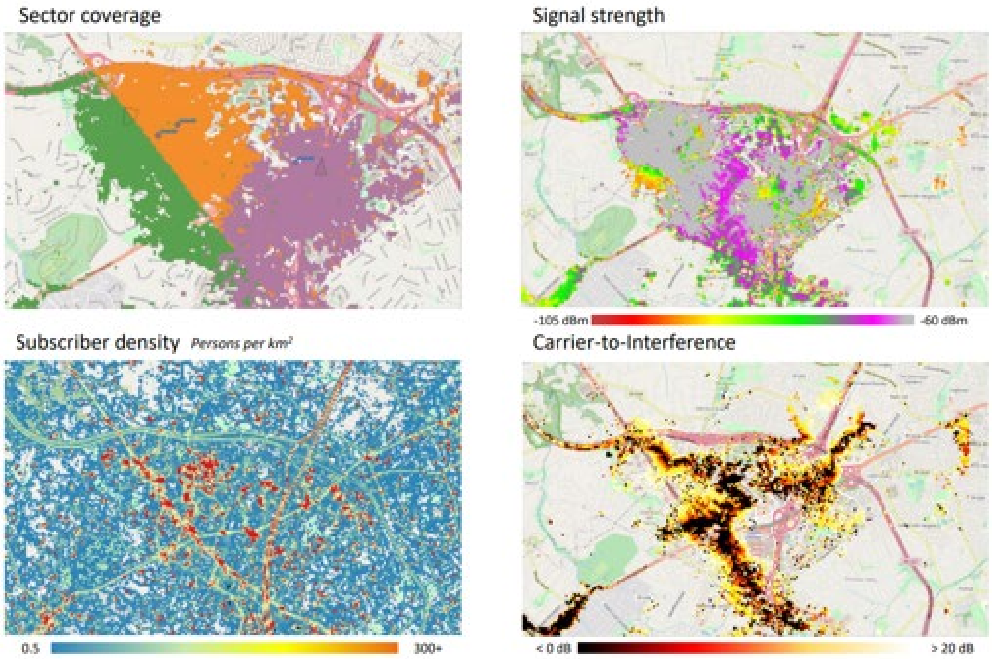

Figure 3

CONCLUSION

We discussed the advantages and challenges inherent when only using crowdsource measurements to estimate the performance of wireless networks. It was demonstrated that, in general, these methods give a measurement of the user experience on the wireless network. But only in the best of circumstances, where the only signals in the RF environment are controlled by the wireless network operator, does the deployed wireless network represent a minimal system (i.e., observable and controllable) where crowdsourced data collection is compensated, can crowdsource measurements accurately reflect the performance of the network. Unfortunately, this is not a realistic situation and therefore, in most cases: 1) substantial postprocessing of the crowdsourced data is required to compensate and calibrate for all the challenges mentioned earlier, and 2) data collection campaigns are required to obtain information about the internal states of the network (wireless and wired links of the network), which is done conventionally with data collection mobile vehicles. This is a very expensive proposition, negating many of the advantages of using crowdsource measurements. A more effective method was proposed in this paper where persistent RF environmental awareness data obtained using DGS tools is used to augment the crowdsource measurements. In this approach, the crowdsource data is primarily used to point to deficiencies in the user experience, which can then be correlated to RF awareness information to point out the potential causes for the low QoE and indicate potential mitigation approaches. One of the main advantages of this proposed approach is the ability to detect causes and point out potential mitigation approaches even if the system interference is from external signals not controlled by the wireless network operator. An example of this approach is discussed where crowdsource data exposed low quality of connectivity in a wireless network and the RF awareness data showed the causes in the network with low C/I due to both wireless network signal strength and interference from other signals in the same RF environment related to other wireless networks.

REFERENCES

[1] J. Howe. (2006, June) Crowdsourcing: A definition. [Online]. Available: http://crowdsourcing.com/cs/2006/06/crowdsourcing a.html

[2] A. Overeem, J. C. R. Robinson, H. Leijnse, G. J. Steeneveld, B. K. P. Horn, and R. Uijlenhoet, "Crowdsourcing urban air temperatures from smartphone battery temperatures," Geophysical Research Letters, vol. 40, no. 15, pp. 4081–4085, Aug 2013

[3] E. Gregori, L. Lenzini, V. Luconi, and A. Vecchio, "Sensing the internet through crowdsourcing," 2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), Mar 2013.

[4] A. Faggiani, E. Gregori, L. Lenzini, V. Luconi, and A. Vecchio, "Smartphone-based crowdsourcing for network monitoring: Opportunities, challenges, and a case study," IEEE Communications Magazine, vol. 52, no. 1, pp. 106–113, Jan 2014.

[5] S. Rosen, S.-j. Lee, J. Lee, P. Congdon, Z. Mao, and K. Burden, "MCnet: Crowdsourcing wireless performance measurements through the eyes of mobile devices," IEEE Communications Magazine, vol. 52, no. 10, pp. 86–91, Oct 2014.

[6] F. Kaup, F. Jomrich, and D. Hausheer, "Demonstration of network coverage - a mobile network performance measurement app," International Conference on Networked Systems (NetSys), 2015.

[7] W. A. Hapsari, A. Umesh, M. Iwamura, M. Tomala, B. Gyula, and B. Sebire, "Minimization of drive tests solution in 3GPP," IEEE Communications Magazine, vol. 50, no. 6, pp. 28–36, Jun 2012.

[8] D. R. Choffnes, F. E. Bustamante, and Z. Ge, "Crowdsourcing servicelevel network event monitoring," SIGCOMM Comput. Commun. Rev., vol. 40, no. 4, p. 387, Aug 2010.

[9] M. Stahr. (2015, Feb) Determine the environment of smartphones by an indoor/outdoor classifier. [Online]. Available: https://www.radioopt.com/ determine-the-environment-of-smartphones-by-an-indoor-outdoor-classifier/

[10] V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks, and R. Baraniuk, "Multifractal cross-traffic estimation," Proceedings of ITC Specialist Seminar on IP Traffic Measurement, 2000.

[11] M. Jain and C. Dovrolis, "Pathload: A measurement tool for end-to-end available bandwidth," Proceedings of Passive and Active Measurements (PAM) Workshop, pp. 14–25, 2002.

[12] L. Cottrell, "pathchirp: Efficient available bandwidth estimation for network paths," Apr 2003.

[13] J. Strauss, D. Katabi, and K. F., "A measurement study of available bandwidth estimation tools," Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement. ACM., pp. 39–44, 2003.

[14] N. Hu and P. Steenkiste, "Evaluation and characterization of available bandwidth probing techniques," IEEE Journal on Selected Areas in Communications, vol. 21, no. 6, p. 879894, Aug 2003

[15] "Using DGS RF Awareness tools to compare performance of wireless network providers, A case study in Tysons, Virginia, March 2021

www.digitalglobalsystems.com

Corporate and Sales Headquarters Digital Global Systems, Inc. 7950 Jones Branch Drive Studio 1A Tysons Corner, Virginia 22102 Phone: +1.240.477-7149
www.digitalglobalsystems.com

To purchase Digital Global Systems solutions or to learn more about Digital Global System's CLEARSITE™, please contact DGS Sales online or at the number above.