Is There WiFi Yet? How Aggressive Probe Requests Deteriorate Energy and Throughput

Xueheng Hu, Lixing Song, Dirk Van Bruggen, Aaron Striegel
University of Notre Dame
Notre Dame, Indiana 46556
xhu2@nd.edu, lsong2@nd.edu, dvanbrug@alumni.nd.edu, striegel@nd.edu

ABSTRACT

WiFi offloading has emerged as a key component of cellular operator strategy to meet the rich data needs of modern mobile devices. Hence, mobile devices tend to aggressively seek out WiFi in order to provide improved user Quality of Experience (QoE) and cellular capacity relief. For home and work environments, aggressive WiFi scans can significantly improve the speed at which mobile nodes join the WiFi network. Unfortunately, the same aggressive behavior that excels in the home environment incurs considerable side effects in crowded wireless environments. In this paper, we analyze empirical data collected from large (stadium) and medium (classroom) venues, and show through controlled experiments (laboratory) how aggressive WiFi scans can have significant implications for energy and throughput for mobile nodes. We close with several thoughts on the disjoint incentives for properly balancing WiFi discovery speed and crowded network interactions.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Miscellaneous

Keywords

WiFi; Probe Request; Energy; Performance

1. INTRODUCTION

The past several years have seen a veritable explosion of data consumption on mobile devices. Smartphones, tablets, and more recently the Internet of Things (IoT) have created a nearly insatiable demand for ubiquitous wireless connectivity. While the peak speeds for cellular (LTE) have risen impressively, dense and indoor environments remain challenging scenarios. Although LTE-Advanced (LTE-A) will offer relief with the introduction of small cell support, questions remain with regards to small cell economic viability and management complexity [1].

IMC'15, October 28-30, 2015, Tokyo, Japan.
© 2015 ACM. ISBN 978-1-4503-3848-6/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2815675.2815709.

For dense or crowded environments, WiFi offloading has emerged as a cornerstone of wireless network operator strategy. Despite the unlicensed nature of WiFi and potential issues with unpredictable Quality of Experience (QoE), the peak speeds of WiFi and more importantly the offloading of demand from the cellular network remain irresistible. Hence, nearly all mobile devices aggressively push the user onto WiFi networks. Whether the mobile device is configured to prompt the user anytime WiFi is available or certain services are restricted to WiFi only, the desire to offload is quite clear. Further efforts by standards bodies on protocols such as ANDSF (Access Network Discovery and Service Function) [2], Hotspot 2.0 [3], ANQP (Access Network Query Protocol), and 802.11ax (successor to 802.11hew) only reinforce that notion.

However, unlike cellular service, WiFi is neither pervasive nor contiguous. Although ANDSF can effectively steer the user to WiFi and Hotspot 2.0 can streamline the user joining WiFi, the mobile node must still find the channel in the WiFi spectrum where the intended WiFi service is located. The root of this discovery process can be found in the 802.11 Probe Request (PR) whereby a mobile node will actively scan across the WiFi space (2.4, 5 GHz) for viable 802.11 access points (APs). Access points, if inclined, can respond with Beacon Responses. The entire process allows the mobile device to quickly locate and join the WiFi network rather than passively waiting to discover an AP. Through the aggressive employment of active scans, mobile nodes can be rapidly directed to WiFi, satisfying both user QoE and decreasing cellular network load. For the ideal case of the home and work environment where the density of mobile devices is relatively manageable and the SSIDs are well known, such a setup tends to work fairly well.

Unfortunately, the tuning that is wonderful for the home and workplace tends to fare quite badly in the crowded or ultra-dense environment, namely venues such as sports arenas [4], large conferences [5, 6], or large classrooms where there is a significant density of individuals. Moreover, the ultra-dense environment is where WiFi offloading is needed most. In our paper, we show that not only do most mobile nodes excessively waste energy trying to find WiFi, we also show that aggressive scans have significant secondary effects on the legitimate users of any established WiFi networks. In short, the purpose of our paper is to argue that aggressive Probe Requests in the ultra-dense cases (hundreds or thousands of nodes) are the wireless equivalent of *'Are we there yet?'*, just as annoying, wasteful, and infuriating but with significant implications for overall network health and performance. While most practitioners in the wireless field would surmise such excessive Probe Requests to be a problem [7, 8], the degree to which such Probe Requests clutter the network in today's wireless devices is simply stunning. To that end, the contributions of this paper are three-fold:

- *Ultra-dense probe request dynamics:* We capture and analyze the dynamics of PRs via packet sniffers at four home football games, measuring PR prevalence both at the exterior (gates) and interior (bowl) in various manners across those four games. We demonstrate that most mobile devices continue to unashamedly probe despite never finding WiFi. We show that the PR rate in the stadium approaches nearly an order of magnitude more prevalence than even reasonably dense classroom settings (2,500 per minute in 2.4 GHz, 900 per minute in 5 GHz vs. 350 / 100 in the classroom). We further characterize the scanning patterns of various devices observed in such ultra-dense environments.
- *Energy impact of probe requests:* While the bulk of our paper focuses on Probe Request dynamics, we characterize the energy cost of active WiFi scanning, exploring the energy cost of a complete WiFi scan (Probe Request across all channels with appropriate Preferred Network Lists (PNLs)). We show that aggressive scanning can burn up to 44% more energy with little to no adaptation in response to the success or failure of WiFi scans to find available WiFi.
- *Throughput impact of probe requests:* Finally, we isolate the negative effects of aggressive WiFi scanning across the 2.4 and the 5 GHz bands on network throughput. We demonstrate that even only a relatively few UEs (user equipments) under default settings can significantly reduce network throughput.

2. RELATED WORK

WiFi has received incredible attention from the research community. For the purposes of this paper though, we are chiefly concerned with works focusing on increased discovery speed [9] and most notably, improved efficiency or accuracy for WiFi scanning [5–8, 10–12]. Our work is unique in that we highlight the prevalence of Probe Requests in the ultra-dense venue (affirmed by [4]) as well as exploring the patterns of the various devices observed in said venues.

The ability to efficiently and quickly scan is a fundamental requirement for fast, seamless handoffs in WiFi. Teng. et. al in [9] proposed D-Scan, specifically targeted at improving scan efficiency in dense environments. Monitoring also plays a key role in distinguishing performance issues with Yeo in [10] and more contemporary work by Rayanchu et. al in [11] trying to pin down interference issues related to WiFi. In their work, Gupta and Mohapatra [12] focused specifically on the power consumption of WiFi on phones while the work by Raghavendra et. al in [5] and Gupta et. al in [6] looked at larger scale venues (i.e., conferences) and overall performance. The issue of ultra-dense venues and WiFi performance was recently discussed in a Cisco slide deck for the 802.11ax working group meeting in Athens in late 2014 [4].

Although the notion of needless Probe Requests is not a new or necessarily a misunderstood topic, the challenge of how to reduce such spurious PRs has seen some research attention. Wu et. al proposed the concept of Footprint [8] while Ananthanarayanan and Stoica proposed Blue-Fi [7] for the express purpose of using cellular (Footprint) or Bluetooth (Blue-Fi) to efficiently guide WiFi scans. Unfortunately, while admirable with regards to the focus of the respective solutions, deployment and in particular the construction of security relationships make such solutions exceptionally difficult to realize in practice. Industry standardization efforts such as Hotspot 2.0, ANDSF, and efforts by the 802.11ax (previously 802.11hew) working groups are making some progress but are still must also be deployed.

3. ULTRA-DENSE DATASET

In this section, we summarize the data collected from four football games and two large class periods at the University of Notre Dame. We begin with a general description of how the data was gathered and continue with in-depth analyses of the data.

3.1 Data Summary

The football data was gathered near gate entrances (during normal queuing prior to ticketed entry) and in the bowl of the stadium. Multiple Linux laptops (Ubuntu 14.04) were used with extended wireless adapters (TP-Link TL-WN722N, Airpcap NX-900) placed into monitor mode and running *tcpdump*. Notably, the stadium does not have publicly accessible WiFi. Several small WiFi APs line the edge of the stadium as used for the ticketing system but that WiFi is not publicly accessible. The stadium itself seats roughly 80,000 with five separate entrance gates (A-E). Crowds of up to 150,000 gather around the stadium and tailgate, resulting in an exceptionally overwhelmed cellular network where cellular network coverage is provided via the campus DAS (Distributed Antenna System) for each of the three major campus carriers (AT&T, Verizon, Sprint). For the Michigan, Stanford, and North Carolina games, data gathering commenced roughly one hour before the start of the game to coincide when most ticket holders began to arrive at the stadium. For the Northwestern game, data was collected inside the stadium bowl in the student section during game time. Two of the games (Stanford, Northwestern) had notable weather impacts with the Stanford game having the most significant impact due to consistent rain and abnormally cold temperatures. Furthermore, fan interest had waned over the course of the season following a mid-season defeat to Florida State (mid-October) and various subsequent losses.

For the classroom venue, data was gathered at various points in the largest classroom on campus (DeBartolo 101) across multiple class periods. DeBartolo 101 can hold up to 450 individuals though attendance at each of the observed periods ranged in multiple hundreds. Each classroom observation began prior to the start of the class and ran through the end of the class session which lasted roughly 75 minutes. The classroom venue offers an alternative view whereby most UEs should have already established WiFi connectivity with the well-established WiFi infrastructure. In contrast to the stadium venue where UEs were especially keen to search for WiFi due to the overloaded cellular network, the classroom venue should demonstrate considerably less scans once a mobile device is settled and joined to the network.

Table 1: Probe Request Data Summary

Venue		Michigan		Stanford		North Carolina		Northwestern		Class 1		Class 2	
Date		2014/09/06		2014/10/04		2014/10/11		2014/11/15		2015/04/20		2015/04/22	
Time Duration		27 min		42 min		60 min		35 min		72 min		76 min	
Weather		Warm		Rainy,Cold		Good		Cold		N/A		N/A	
Location		Stadium Gate		Stadium Gate		Stadium Gate		Stadium Bowl		DB 101		DB 101	
Band		2.4G	5G	2.4G	5G	2.4G	5G	2.4G	5G	2.4G	5G	2.4G	5G
# of PRs		75,791	7,977	86,195	4,335	82,174	21,244	87,322	32,897	28,010	8,676	27,867	6,839
# of Source MACs		4,785	1,458	6,813	805	7,144	3,191	12,802	3,853	1,087	522	1,051	443
PRs / Min	Mean	2,778	294	2,098	103	1,362	353	2,568	967	379	118	366	90
	Max	3,721	596	3,029	223	1,855	603	3,505	1,344	673	227	722	249
	Median	2,690	295	2,168	97	1,330	349	2,612	1,005	346	105	334	78
	Stdev	460	119	663	56	202	79	584	266	115	42	107	45

In each of the data gathering scenarios, multiple laptops were used as noted earlier with each laptop possessing multiple external wireless adapters. Individual laptops were configured to monitor multiple channels both within a band and between bands, i.e., either monitor multiple 2.4 GHz channels (Channel 1, Channel 11) or across multiple bands (2.4 GHz Channel 1, 5 GHz Channel 153). The intuition behind the setup was to capture both the scan width (how long does it take for a device to complete a scan across the 2.4 GHz band) as well as dual-band capabilities (2.4 vs. 5 GHz). Notably, 5 GHz adoption rose significantly even over the course of this study making our dataset likely to underestimate the net PRs / minute with the most recent slate of mobile devices. Data was processed through a combination of *tshark* and Python with PR information stored in a MySQL database. Following processing, data files are discarded and only anonymized header information is preserved for the purpose of analysis.

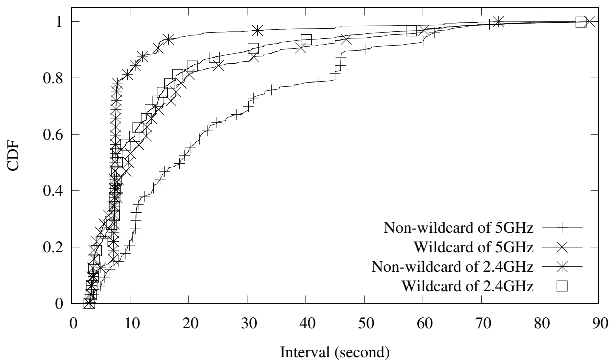
Table 1 shows the key characteristics observed across each of the respective venues. The data is broken for each band and represents the impact on only one particular channel (Channel 1 for 2.4 GHz, Channel 153 for 5 GHz). Taking the Michigan game as an example, for the 2.4 GHz band, the average density of PRs comes in at 2,778 per minute, just over 46 PRs per second. The 5 GHz channel sees remarkably fewer PRs (294 per minute) but it is also notable that many devices were still not fully 5 GHz capable at the time of our experiment as noted earlier. While we had expected to see a bump in 5 GHz PRs at the Stanford game due to the recent release of iPhone 6, the inclement weather had clear impacts in terms of attendance for the game (upper 30°F, rainy). Even with the reduced fan turnout, the number of PRs for the Stanford game still averaged 2098 per minute in Channel 1 (nearly 35 PRs per second).

The introduction of the iPhone 6, iOS 8, and newer versions of Android also added peculiarities to the experiment as the source MAC on WiFi scans could no longer safely be viewed as an indicator of the device density [13]. Android also introduced similar functionality. Hence, while the number of PRs is accurate, the number of unique source MACs from the Stanford game onwards is reasonably suspect.

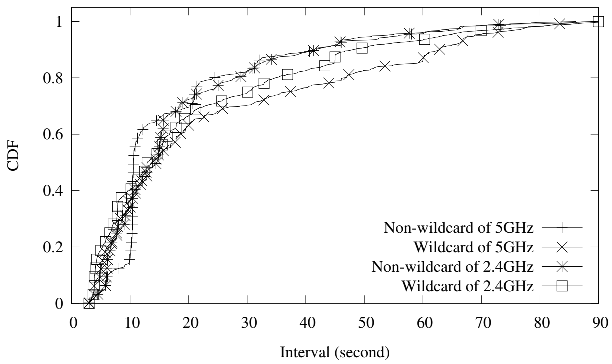
Regardless of the issues with device counts via unique source MACs, PRs can serve as an alternative indicator of device density as PRs are a function of existing WiFi connectivity, device type / configuration (ex. Android, iOS), and device usage (screen on, WiFi scanning screen). In contrast to the stadium, each of the class periods exhibits dramatically different prevalences with regards to PRs. As would be expected with an established WiFi infrastructure, most UEs are likely to have been associated and hence do not nearly as aggressively probe the WiFi spectrum. Most importantly as noted earlier in the contributions, the prevalence of PRs in the reasonably populated large classroom (several hundred individuals) is dwarfed by nearly an order of magnitude versus the stadium environment. Critically, while a legitimate WiFi deployment in the stadium would be likely to reduce the prevalence of PRs, there still would likely be a non-trivial number of unaffiliated UEs who will continue to aggressively probe in such venues.

Circling back around, from an overhead perspective for nearly all cases aside from initial entry to the venue, each Probe Request can be viewed as wasteful as either no public WiFi exists (stadium) or the infrastructure is unlikely to change (classroom). The waste manifests itself from a foundational perspective by virtue of time consumed on the primary channel as well as secondary consumption on interfering channels. Hence, a simplistic way to evaluate the waste of PR is to explore the time consumption for each request. From a distributional analysis, the most common rate setting (92%) for PRs in the 2.4 GHz spectrum was 1.0 Mb/s with speeds observed for PRs up to 11.0 Mb/s. PRs in the 5 GHz spectrum were universally set to 6.0 Mb/s. If we assume a rough PR size of 100 bytes, a perfect PR (ignoring CIFS, DIFS, 802.11 headers, DCF effects) would be 800 microseconds. The reality though is that the PR consumes an impact of much more than the 800 microseconds of air time due to, for example, the exponential back-off introduced by DCF.

First, frequent PRs are highly likely to impact the DCF of any mobile nodes affiliated with WiFi. While the stadium does not offer WiFi, we could view the mobile nodes as captured as being indicative of nodes without ANDSF policies / uncooperative mobile nodes, the extent of which would require careful observation that is beyond the scope of this paper. Second, for the 2.4 GHz channels, the lack of channel orthogonality means that as a mobile node iterates through an active scan, it may cause issues as it traverses nearby channels (ex. Channel 2, Channel 3, Channel 4, Channel 5 on Channel 1). If we are lucky, the overlapping channels have a minimal impact. Otherwise, the air time consumed on the overlapping channels may create significant bit errors for the primary users. Third, while PRs are relatively short, the low data rates of the PRs means that the actively scanning nodes tend to clutter / slow down the higher speed / affiliated nodes (ex. 1 Mb/s vs. 54 Mb/s). This disparity of speed is only amplified in the 5 GHz bands with higher



(a) Michigan



(b) Northwestern

Figure 1: CDF of Scan Interval

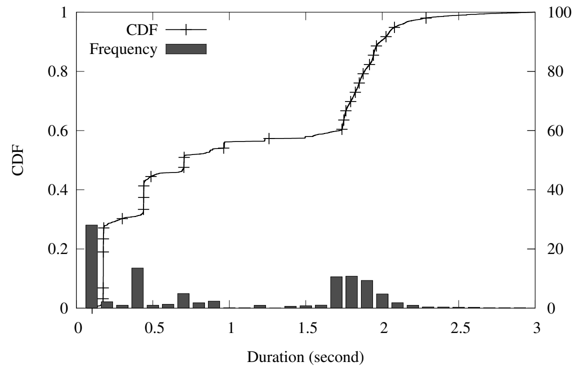


Figure 2: Distribution of Scan Duration

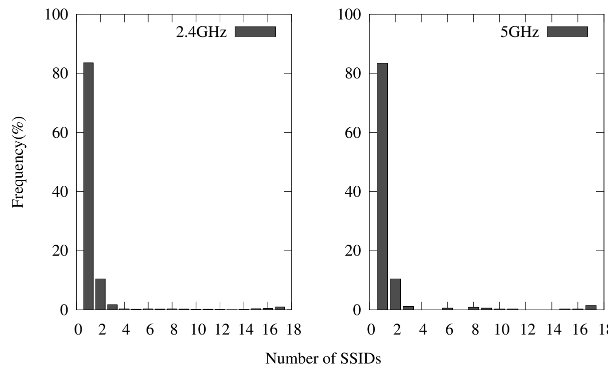


Figure 3: Number of SSIDs per Active Scan

potential peak speeds afforded by 802.11n (270 Mb/s) and 802.11ac (1.3 Gb/s). Later in this paper, we explore the impacts on 802.11n via small scale experiments. Fourth, as indicated by the observations from the classroom, mobile nodes may continue to scan even once affiliated with an AP if AP performance is insufficient or simply if the mobile nodes hopes for observing 'better' WiFi.

3.2 Scanning Behavior Analysis

To explore the data further, we dive into the data as observed from the games of Northwestern and Michigan. The reason for exploring the Northwestern game is that observations from the bowl allowed us to experience a much more representative sample of the ultra-dense environment at scale. In contrast to the gate-based observations for the prior three games with somewhat transient behavior (queuing at the gates), the Northwestern game presented observations from a fixed vantage point. We also include the Michigan game (before iOS 8) for the purpose of investigating how iOS 8 MAC randomization could potentially impact our measurements on WiFi scanning behaviors, particularly, the scan interval [13].

Figure 1 plots CDFs of the inter-scan interval that represents the average wait time between successive scans initiated by mobile nodes across both the 2.4 and 5 GHz bands. Each CDF is also broken out by the requests for empty SSIDs (interval between unknown SSIDs) and known SSIDs via the PNL. A low-pass filter is applied with a floor of three seconds as observed by the data distributions which means the interval is only counted if the node has at least three seconds of idle time between requests from the same MAC with the same SSID. SSIDs from the PNL may only be counted once for a given MAC (ex. BestBuy matching with BestBuy counts and then precludes any subsequent matches in the same scan for that UE). The interesting result is that at the 2.4 GHz spectrum many nodes (60% for Michigan and 40% for Northwestern) scanned quite frequently at intervals around or smaller than 10 seconds. The frequent scan may in part be driven by the cellular network on campus being overwhelmed on game day (nearly 150k individuals can be on campus) as well as individuals turning on their phone trying to find WiFi while waiting in line to enter the stadium.

Next, Figure 2 measures the average time duration for a particular active scan. The duration is recorded by measuring the occurrence of the first PR for a UE in Channel 1 followed by the appearance of the last PR for that same UE in Channel 11. Time synchronization is provided by running each of the monitors on the same laptop in monitor mode. The data for Figure 2 was focused only on the 2.4 GHz band. Both the CDF and the Frequency (PDF) are plotted in the figure. Interestingly, there are two clusters that can largely be attributed to differences between the respective mobile operating systems. On the left side, Android devices tend to frequently try to cap the maximum scan width ranging typically less than 800 milliseconds. Alternatively, iOS devices tend to fan out over a wider period of time by scanning for up to 2 seconds at a time. Critically, Android devices tend to squeeze as many PRs as possible into a shorter pe-

Table 2: WiFi Scan Interval

	In WiFi Settings Screen	
Device Type	**Yes**	**No**
iPod touch	3 s, 8 s, 10 s	15 s → 480 s
Dell Venue 7	10 s	43 s
Nexus 10	10 s	10 s
Galaxy S4	10 s	60 s

riod of time while iOS devices tend to spread out the PRs over time.

In addition, we have investigated the varying numbers of SSIDs contained within a PNL of individual PRs. WiFi scans are divided into sweeps representing a complete active scans with the number of unique SSIDs requested being mapped to each individual source MAC within that time frame. As demonstrated by Figure 3, most nodes (80%) do not make requests for SSIDs from their PNL but rather only make requests for the the empty (unknown) SSID. We comment a bit later on scan mechanisms and how the length of the PNL has only a minimal impact on the actual number of resulting scans (ex. some mobile nodes tend to scan for a timed duration rather than simple PNL coverage). Recent trends with regards to privacy indicate that the scan for only a single empty SSID is more likely to be the norm than scans with long and varied PNLs.

4. PERFORMANCE DETERIORATION

We now explore the energy and throughput costs of aggressive WiFi scanning in a controlled laboratory setting. Although it would be ideal to instrument the entirety of the stadium and to provide pervasive instrumentation, the laboratory experiments can shed some light on what might occur in the larger scale scenarios. To that end, we conducted a group of small-scale, controlled experiments with four types of handsets: the iPod touch (iOS 8.4), the Dell Venue 7 (Android 4.4), the Nexus 10 (Android 5.1), and the Samsung Galaxy S4 smartphone (Android 4.2). We are particularly interested in measuring the power cost of an active WiFi scan as well as the throughput impacts associated with WiFi scans. Moreover, the lab setting provides ideal cases for attempting to replicate the various intra-scan behavior observed in the stadium and the respective energy implications.

For our experiments, we investigated WiFi scanning behaviors for the aforementioned devices by configuring the devices with two different settings while at the same time running *tcpdump* in monitor mode for capturing PRs. The laptops utilized for packet capture were identical to the setups used for packet capture in the stadium environment. Scanning behaviors of the devices are summarized in Table 2. The distinguishing factor between the two columns refers to whether or not the WiFi settings screen was open (which implies a much more aggressive approach to scan). We observed for each active scan, all four types of devices typically sweep the PNL on one specific channel and then hop to the next channel.

As indicated by Table 2, for all four types of devices (WiFi enabled), if the listing of current WiFi is open, the intervals between two consecutive WiFi scans are roughly 10 seconds (3 s and 8 s were also observed for iOS device). If WiFi is still on but the user is not in the WiFi settings screen (and not connected to WiFi), the scan interval for different devices varies from 10 seconds to up to 8 minutes. Interestingly, iOS 8 tends to exponentially increase the scanning interval from 15 s to 480 s (i.e., 15 s, 30 s, 60 s, ... 480 s) when not in the WiFi settings panel. For all other experiments presented in this section, we forced the handset to stay in WiFi settings screen as this configuration allows us to better mimic denser environmental scenarios. We surmise though more difficult to measure, that the lack of cellular connectivity may also have caused more rapid scanning as the cellular network tended to be quite overwhelmed with the nearly 150k users in the local area (good perceived downlink signal but impossible to actually send data on the uplink).

Table 3: Energy Cost: Baseline vs. Scanning

	Power Consumption (uAh)	
Settings	**Average**	**Std Dev**
Baseline	12,333.42	7.18
Scanning	17,801.18	60.08

4.1 Energy Impact

For the purpose of evaluating energy cost per WiFi scan, we use the Monsoon power monitor and its PowerTool software [14]. We instrument the Galaxy S4 smartphone as it is the only device possessing a removable battery. Power for the phone is supplied by the Monsoon power monitor with the energy consumption recorded at a sampling rate of 5 KHz. The phone is evaluated in two different settings, *Baseline* where WiFi is off but the screen remains on and *Scanning* where the phone stays in the WiFi settings screen but remains unaffiliated with regards to WiFi. Power monitoring was run for an extended period of time (5 minutes). The average and standard deviation of power consumption with each setting are given in Table 3.

After recording the average power consumption for both the Baseline and the Scanning settings, the energy cost of an active WiFi scan can be approximated by calculating the delta between the consumption values of these two settings over the entirety of the monitoring period, yielding 5467.76 uAh for the 5-minute time window. Notably, this represents a 44.3% increase over the Baseline consumption despite the fact that the screen is on for both cases. We can further infer the power cost per scan since Table 2 has indicated WiFi scan is invoked every 10 seconds if WiFi settings screen is active. While somewhat crude in its approximation, each WiFi scan in an ideal scenario (no background traffic, no DCF issues) consumes roughly 182.26 uAh. An individual waiting in line for 10 minutes without WiFi while aggressively scanning could end up consuming nearly 10936 uAh extra energy, effectively 0.4% of a fully charged battery (2600 mAh).

4.2 Throughput Loss

While the energy losses may be tolerable (though still wasteful), aggressive WiFi scanning also has impacts on performance by introducing overhead to the wireless channels. For the purpose of measuring the potential network performance degradation, we designed a small-scale experiment using *iperf2* configured for UDP throughput. The components of this experiment are described follows:

iperf client and server: We instrumented a Dell EliteBook 8560 with a 802.11n dual-band network interface as the *iperf* client and a HP 3450 laptop with Ubuntu Linux installed as the *iperf* server.

Handsets: We used one iPod Touch, one Dell Venue 7 tablet, and thirteen Android smartphones as our handsets. All devices are dual-band capable. For each device, we de-associated the device from any known WiFi and created a PNL consisting of two hidden SSIDs. Screens were kept on and the devices were kept in the WiFi scan screen.

Wireless router: A Netgear AC1900 router (dual-band, 802.11ac capable) was used to set up the WLAN. The router provided up to 600 Mbps and 1300 Mbps WiFi down-link speeds on the 2.4 and 5 GHz bands respectively. The client was the only node associated with the AP and the server was directly connected to the router via GigE. A full 40 MHz of spectrum (full 802.11n speeds) was selected for the router with validation. Experiments were conducted in a basement with minimal interference from other devices. The distance between the client and the server was roughly three meters and all handsets were positioned between the client and the server.

The client was tuned in order to determine the maximum lossless send rate between the client and the server. Communications were unidirectional going only from the client to the server. On the 5GHz band, client performance topped out at roughly 180 Mb/s without any background traffic (see Figure 4). Each experiment setting was repeated multiple times with a typical test duration of five minutes. Once the baseline was established, the experiments were conducted by gradually increasing the number of UEs actively scanning. Notably, WiFi performance decreased by 16.7% once all fifteen devices were introduced, i.e., from the peak of 180.0 Mb/s down to 149.8 Mb/s. While degradation of performance is not entirely unexpected with WiFi, the fact that this performance decrease comes by virtue of 'useless' PRs is problematic. As noted earlier, stadium environments may have nodes both associated and un-associated with the WiFi infrastructure with the unaffiliated mobile nodes still chirping for WiFi. Furthermore, our limited lab experiments were actually quite benign entailing only roughly 360 PRs per min (observed via tcpdump). In the stadium case, such as the Northwestern game where 967 PRs were generated every minute, unaffiliated nodes could have significant performance issues even in the 'better' 5 GHz bands. As noted in Figure 4, the 2.4 GHz performance has a similar degradation pattern except with an obviously lower baseline and would potentially be much worse for the stadium case with dramatically higher numbers of PRs (360 PRs per minute versus 2568 PRs per minute).

5. DISCUSSIONS AND FUTURE WORK

The issue of how to solve the dilemma of aggressive WiFi speaks to the complexities of the wireless industry. On one hand, the solution would appear to be fairly trivial: slow down the WiFi scanning rate and scan only on a screen activation. The reality though is decidedly more complicated as applications largely do not wait to send data until there is WiFi available and the delay before locating WiFi at home could be considerable. Critically, the vast majority of a user experience tends to be dominated by the simple cases, ex. only a few devices and well-known SSIDs. In those cases, the first scan when in range tends to be successful and faster scanning means faster hopping onto WiFi.

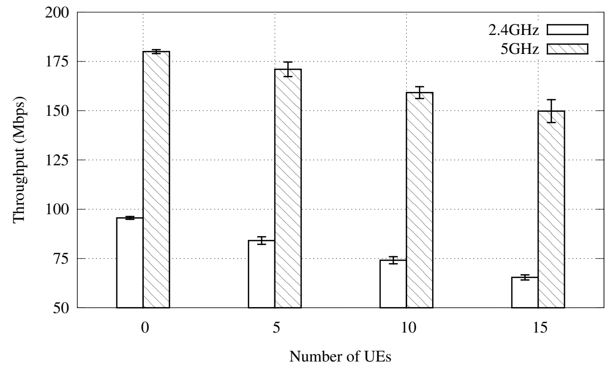


Figure 4: Throughput Reduction

Furthermore, the vendors at first glance most impacted by aggressive Probe Requests tend to be the WiFi infrastructure vendors who have little to no control over the mobile devices. After all, WiFi exists in the unlicensed bands which means that for all practical purposes, the equipment infrastructure vendors must simply endure. Recent discussions with the 802.11ax standard have noted that indeed, aggressive Probe Requests do create sizable issues in ultra-dense venues [4]. Handset and OS vendors are only marginally motivated by claims of reduced throughput as the vast majority of throughput scenarios are OK (the typical cases). However, we would argue that the energy cost of being aggressive is non-trivial and moreover, that energy cost burns worst when most users tend to be suffering energy issues (ex. the ultra-dense venue).

In conclusion, we believe that aggressive WiFi scanning has significant side effects on normal wireless network users, both with respect to energy and throughput. Moreover, we believe that our stadium analyses show that not only are aggressive Probe Requests wasteful in the ultra-dense case, the degree to which such wasteful Probe Requests occur far outstrips what is perceived in the literature. Moreover, despite research efforts to tamp down the impacts of Probe Requests, the perceived benefits with regards to WiFi joining speed seem to largely outweigh occasional problems in dense venues. Future work is needed to explore how one can bridge the conflicting goals of rapid WiFi detection with the cost of wasted WiFi scans in a deployable manner, likely with the assistance of handset or cellular carriers, a considerable challenge. For encouraging such efforts, the full anonymized stadium and classroom datasets are made publicly available.

6. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. IIS-0968529. Further support was provided through the University of Notre Dame and Sprint.

7. REFERENCES

[1] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What Will 5G Be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] "Cellular-Wi-Fi Integration," InterDigital, Tech. Rep., 2012.

[3] "The Future of Hotspots: Making Wi-Fi as Secure and Easy to Use as Cellular," Cisco, Tech. Rep., 2012.

[4] B. Hart, M. Swartz, J. Suhr, and M. Silverman, "Stadium Measurements," 802.11ax Working Group (Doc 1223), September 2014.

[5] R. Raghavendra, E. Belding, K. Papagiannaki, and K. Almeroth, "Unwanted Link Layer Traffic in Large IEEE 802.11 Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 9, pp. 1212–1225, September 2010.

[6] A. Gupta, J. Min, and I. Rhee, "WiFox: Scaling WiFi Performance for Large Audience Environments," in *CoNEXT 2012, ACM*, 2012, pp. 217–228.

[7] G. Ananthanarayanan and I. Stoica, "Blue-Fi: Enhancing Wi-Fi Performance Using Bluetooth Signals," in *MobiSys 2009, ACM*, 2009, pp. 249–262.

[8] H. Wu, K. Tan, J. Liu, and Y. Zhang, "Footprint: Cellular Assisted Wi-Fi AP Discovery on Mobile Phones for Energy Saving," in *WINTECH 2009, ACM*, 2009, pp. 67–76.

[9] J. Teng, C. Xu, W. Jia, and D. Xuan, "D-Scan: Enabling Fast and Smooth Handoffs in AP-Dense 802.11 Wireless Networks," in *INFOCOM 2009, IEEE*, April 2009, pp. 2616–2620.

[10] J. Yeo, M. Youssef, and A. Agrawala, "A Framework for Wireless LAN Monitoring and Its Applications," in *WiSe 2004, ACM*, 2004, pp. 70–79.

[11] S. Rayanchu, A. Patro, and S. Banerjee, "Catching Whales and Minnows Using WiFiNet: Deconstructing non-WiFi Interference Using WiFi Hardware," in *NSDI 2012, USENIX*, 2012, pp. 57–70.

[12] A. Gupta and P. Mohapatra, "Energy Consumption and Conservation in WiFi Based Phones: A Measurement-Based Study," in *SECON 2007, IEEE*, June 2007, pp. 122–131.

[13] B. Misra, "iOS8 MAC Randomization," accessed: 2015-05-05. [Online]. Available: http://blog.airtightnetworks.com/ios8-mac-randomization-analyzed/

[14] "Monsoon Power Monitor." [Online]. Available: https://www.msoon.com/LabEquipment/PowerMonitor/