EcoCell: Energy Conservation through Traffic Shaping in Cellular Radio Access Networks

Paper #275, 12+2 pages

ABSTRACT

Cellular networks contribute significantly to global energy demands and carbon emissions due to the millions of base stations deployed worldwide. We characterize the energy consumption of production base stations by performing fine-grained power and network telemetry measurements using off-the-shelf base stations. Our measurements reveal unique insights about how variations in temporal-usage patterns affect base station energy consumption. Based on these insights, we design EcoCell, a software-only solution that introduces energy-efficient traffic patterns in network flows. EcoCell can be implemented either as a traffic scheduler in the radio access network or as an independent middlebox. We evaluate EcoCell with five popular networked applications on a production basestation. We demonstrate savings up to 32% in dynamic energy consumption of a base station, without drops in application-level quality of experience.

1 INTRODUCTION

Cellular network infrastructure is a significant contributor to global carbon emissions and energy consumption. One base station is estimated to produce 30 tons of carbon emissions in one year of operation [60]. Given 6 million base stations worldwide [66], the combined carbon emission is almost comparable to that of the aviation industry [3, 27]. The cost of powering such a large network has also increasingly become a concern for cellular network providers [28] – conservatively assuming 250 W power consumption [42] per base station and 0.1 USD/kWh power costs puts the annual global power expenditure estimate at 1.3 billion USD.

In spite of high energy and carbon costs of operating cellular network infrastructure, research in optimizing a base station's energy consumption is largely limited to theoretical models [14, 27, 30], simulations [13, 33, 52], and energy-efficient component design [26, 47, 59]. There has also been extensive research in optimizing the energy consumption of user devices like smartphones [31, 35, 55, 65] to improve their battery life. In contrast to a client, a base station transmits more energy, serves multiple clients, and is a more complex system. In general, a system-level understanding of base station energy consumption and its implications have been under-explored in past work.

To fill this gap, we conduct extensive measurements on an off-the-shelf cellular base station (Sec. 4) deployed in the recently opened CBRS (Citizens Band Radio Service) band.

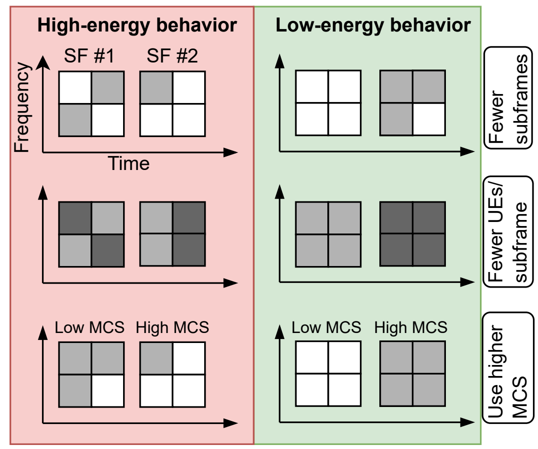


Figure 1: Energy-saving Opportunities

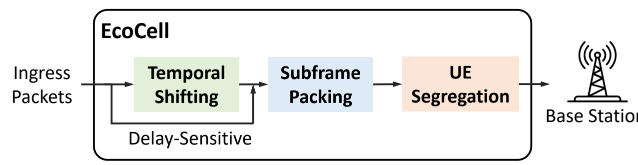


Figure 2: EcoCell Overview

Our measurements include millisecond-level power measurements in conjunction with fine-grained cellular network telemetry including Layer 1 and Layer 2 information (coding schemes, spectrum assignments, and retransmissions). In addition, we conduct long-term telemetry-only measurements for commercial traffic in a campus cellular network and in a large US city downtown. We make three observations from our measurements, as shown in Figure 1.

- *Packing data into fewer subframes saves energy.* Cellular protocols like 4G and 5G send data in one millisecond subframes. Each subframe is further divided into multiple physical resource blocks (PRBs), which are smaller time-frequency chunks. We observe that it is energy-efficient to use more PRBs in a single subframe compared to using few PRBs spreading across multiple subframes, in part because transmitting fewer subframes incurs lower control overhead. Prior work, e.g. [7], has made this observation, but we are the first to empirically analyze and exploit it for energy conservation at the base station.
- *Serving fewer UEs per subframe saves energy.* When multiple user equipments (UEs) are active in a base station, the base station's scheduler may allocate multiple UEs

in one subframe. We empirically measure that, for the same amount of traffic, serving more UEs per subframe consumes more energy compared to serving one UE per subframe; we hypothesize that this is due to higher control signaling overhead when serving more UEs in a single subframe (need per-UE control information in each subframe) and more complex processing.

- *Higher-order modulation reduces energy consumption.* Expectedly, higher-order modulation and coding schemes, which pack more bits per wireless symbol, are more energy-efficient because they can transmit more data per symbol.

Based on these observations, we design EcoCell (Fig. 2), a new traffic-management solution that reduces the energy consumption at the base station, without requiring any protocol-level or hardware changes to either the base station or the client. EcoCell can be implemented either in the RAN scheduler or as a middlebox that sits right before the base station and shapes traffic going into the base station. EcoCell's goal is to reduce energy consumption with minimal effects on the performance of applications or underlying protocols, such as the throughput, delay, and QoE (Quality of Experience). EcoCell works for both delay-sensitive traffic (such as video streaming, teleconferencing) and delay-tolerant traffic (such as large file transfers). EcoCell can be deployed *today* with existing base station deployments because it can be implemented without any software or hardware modifications to the client or the base station. Specifically, EcoCell leverages the three observations above as follows:

(i) Subframe packing : Our first goal is to reduce the number of transmitted subframes by packing more data into a subframe. Base station downlink capacity is under-utilized in practice. For example, even a video streaming flow consumes a few Mbps of bandwidth, compared to a downlink capacity of up to 100 Mbps. In practice, this results in a large fraction of subframes with only a small number of PRBs being utilized. We sniffed subframe utilization in a commercial campus network and found that the median utilization of active subframes was just 16%.

To enhance subframe-level utilization, EcoCell introduces artificial burstiness in traffic across the time span of a few tens of milliseconds. Specifically, EcoCell intercepts and buffers all the incoming packets at the base station. It periodically releases all the buffered packets to the base station. By making the traffic bursty, we can *pack* more data in each subframe. Such subframe packing leads to fewer subframes being sent from the base station to the client and results in energy savings.

(ii) UE segregation: Each subframe can contain traffic from multiple user equipments (UEs) leading to energy-inefficient behavior as discussed before. In our measurements on a production network, we observe over 20% of the subframes being shared by multiple UEs at peak traffic hours. To maximize energy efficiency, EcoCell attempts to reduce the fraction of shared subframes. Specifically, we implement an egress scheduler in EcoCell that maintains per UE queues. EcoCell drains these queues in order such that all data packets for one UE are sent together to the base station, followed by all data packets for another UE. Our scheduler design forces the base station to segregate UEs in separate subframes when possible and improves energy efficiency.

(iii) Temporal shifting : Since higher-order modulations are more energy efficient, it is natural to shift traffic to higher-order modulations. However, higher-order modulations are only feasible at better channel quality, and channel quality variation happens at the order of seconds instead of milliseconds. Therefore, we design a new temporal traffic shifting method solely for delay-tolerant traffic (such as large file transfer, software updates, etc.). We envision that this technique can be used by base station operators to enable a low power slice (e.g., by using explicit priority assignment in cellular networks [8]) for high-volume delay-tolerant traffic.

EcoCell's goal is to opportunistically shift traffic to times with improved channel conditions. However, the base station does not know if channel quality will improve in the future or get worse. Therefore, it is non-trivial to decide how to shape the traffic without any future information. We take a heuristic approach to solve this problem. EcoCell adjusts the traffic egress rate linearly according to the recent downlink signal strength while meeting the slice requirements.

We have designed and implemented EcoCell as a middlebox that intercepts and shapes traffic going into a off-the-shelf CBRS base station. We evaluate EcoCell in five major traffic classes: web-browsing, videoconferencing (e.g., Google Meet), large file downloads, live streaming (e.g., sporting events), and video streaming (e.g., Netflix). Our evaluation demonstrates that EcoCell can reduce dynamic energy consumption at the base station by up to 32% while maintaining application QoE's (and higher if a degradation in QoE is acceptable). *We plan to release all the datasets and source code in this project.*

Our work makes the following contributions:

- We create the first fine-grained real-world power measurement datasets with cellular network telemetry and application layer information.
- We observe specific traffic patterns that lead to power savings in cellular base stations.
- We propose EcoCell, a new algorithm that enhances subframe packing, UE segregation, and channel quality selection to reduce power consumption at the base station.
- We implement and evaluate EcoCell across various real-world applications and reduce dynamic energy usage by up to 32%.

2 SCOPE AND SIGNIFICANCE

Power consumption decomposition of base station. We focus on optimizing the base station's (BS) *dynamic* power consumption, which pertains to the component reliant on data usage patterns. BSes also incur a constant static power draw even during idle periods, determined solely by the vendor's hardware choices, and falls beyond the scope of EcoCell's software optimizations. Each of the BS's power-hungry devices has dynamic and static components of power draw. These include the radio unit (the RU, which contains the power amplifiers and wireless transceiver chains), the base band unit (the BBU, which contains the digital signal processing circuitry), and other supporting systems like cooling. Our approach centers on optimizing BS power holistically, rather than dissecting it among these individual components.

Base station	Dynamic power	Static power
Macro 5G SA (32TR)	413 Watts	287 Watts
Macro 4G (32TR)	390 Watts	103 Watts
Our micro 4G, 4TR	10.7 Watts	31.3 Watts

Table 1: Breakdown of base station radio unit power consumption into dynamic and static components

Table 1 shows that dynamic power can account for the majority of BS power. We reproduce numbers for macro 4G and 5G cells from a recent country-wide measurement of commercial BSes [43]. As one example, a macro 4G BS with 32 antennas consumes 103 W when idle, and an additional 390 W when transmitting at maximum power on all frequency resources. In our testbed's small base station (details in Sec. 4), the dynamic power is a smaller fraction of total power since (1) it is a micro BS with low transmission power to serve its shorter range, and (2) it has only four antennas. In this paper, we report primarily relative reductions in *dynamic* power draw and energy consumption. The actual Watts saved in practice will depend on the base station's size and architecture.

Large power denominator: Even small savings in dynamic power consumption (say 10%) can save 30–40 Watts (based on Table 1) of power. Extending such power savings to six million base stations worldwide [66] can significantly reduce the energy costs (by USD 160–210 million per year assuming 0.1 USD per kWh energy costs).

3 PRIMER ON CELLULAR RANS

We next provide backround on cellular RANs relevant for the base station's energy consumption.

Subframes: Cellular networks adopt Orthogonal Frequency Division Multiplexing (OFDM) for data transmission wherein

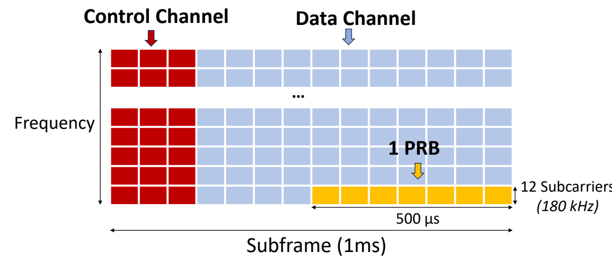


Figure 3: LTE subframe structure

a single data stream is split among several closely spaced yet orthogonal subcarriers (15kHz). In time, data transmissions in LTE and 5G are split into 1-millisecond subframes. A subframe consists of 14 OFDM symbols shown as each column in Fig 3, where the first one to three OFDM symbols are reserved for the control channel, and the remaining symbols are assigned for the data channel. The control channel transmits downlink control information (DCI) to tell UEs the necessary information for decoding, and this includes resource allocation, coding scheme, retransmission information, etc. The data channel carries the user payload data and is divided into small time-frequency blocks called physical resource blocks (PRB), spanning 12 subcarriers (180kHz) and 500 μs. The smallest time-frequency resource that can be scheduled to a device is one PRB pair that's in the same frequency of one subframe excluding the control channel[46]. Multiple UEs can be scheduled in the same subframe and PRBs can be allocated per base station policy. The control channel will contain the DCI for each UE separately.

DCI in control channel: The base station sends a DCI message to a UE on the control channel if it transmits data to that UE on the data channel within the same subframe. Each DCI encapsulates four critical parameters essential for our analysis: the allocated *physical resource block* (PRB) by the base station to each UE; the *transport block size* (TBS) indicating the number of bits transmitted to the UE within each subframe; the *modulation and coding scheme* (MCS) representing how data bits are coded and modulated based on radio link quality and the *radio network temporary identifier* (RNTI) uniquely identifying the UE. In scenarios involving spatial multiplexing (MIMO), where up to two data streams per millisecond share the same set of PRBs, each stream possesses a distinct set of (TBS, MCS) DCI content.

Downlink vs uplink traffic: Our work currently focuses on downlink traffic for two reasons: downlink traffic accounts for the large majority of cellular network traffic [32, 53]; and base stations consume much more power in transmission compared to reception [15].

4 MEASUREMENT METHODOLOGY

We conduct measurements using the following setup (Fig. 5):

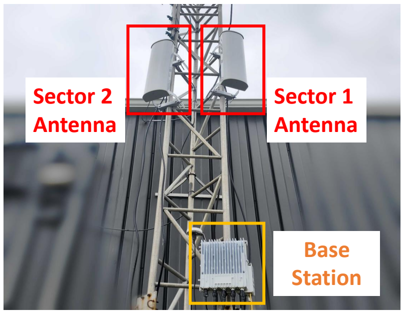


Figure 4: CBRS Base Station with 2 Sectors

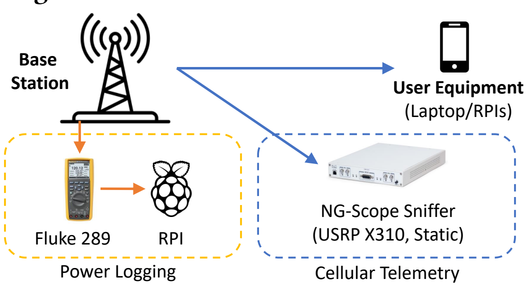
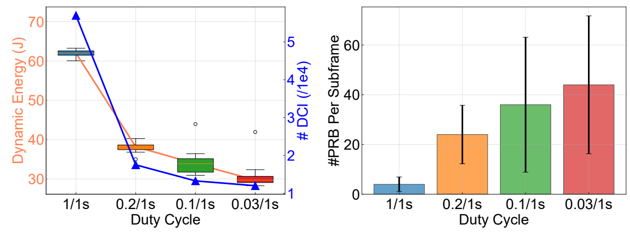


Figure 5: Measurement Setup

Base station power measurement: We deploy an off-the-shelf LTE CBRS base station (Fig. 4). The base station is a Celona LTE AP-11 [12], with two sets of KP Performance [5] 90-degree Sector antennas, each offering a gain of 16.7dBi. The maximum transmit power is 1000 mWatt/30 dBm. The two sectors work at two different frequencies of 3.59 GHz and 3.69 GHz. We have only black-box access to this base station, i.e., we cannot modify its hardware or software. We measure its power consumption by installing two Fluke multimeters [34] (Fluke 289) into the circuit for the DC current and voltage respectively; we use a Raspberry Pi to log the reading from multimeters through an IR189USB infrared adapter[22]. We measure power every 30 ms, with a multimeter-reported measurement accuracy of 0.05% [1].

Cellular network telemetry measurement: We run NG-Scope [62] on two USRP X310 [61] devices concurrently for monitoring both the 3.59 and 3.69 GHz bands. The USRPs decode the base station's DCI messages (Sec. 3) in every subframe. The USRPs act as only passive sniffers and do not transmit. During our experiments, we fix the location of USRPs to be 10 meters away from the base station. We choose NG-Scope among others [11, 19, 39] for its state-of-the-art decoding accuracy and its support for TDD cell and carrier aggregation. Tools such as MobileInsight [45, 56] aim at decoding for single user rather than the whole cell.

UE setup and signal strength measurement: We put SIM cards of our private network into MultiTech dongles [51] for cellular signal reception. We further get the signal strength measurement (RSSI/RSRQ/RSRP) by querying the dongle through the provided APIs every 1 second. For single-UE experiments, we connect the dongle to a Lenovo Laptop; for multi-UE experiments, we add another four Raspberry Pi 4 with Ubuntu22.04 and mount the same model of dongles.



(a) Energy with duty cycles **(b) PRB with duty cycles**

Figure 6: (a) Increased burstiness saves energy. (b) This is caused by higher subframe utilization and lower total control overhead.

Production base station sniffing: In addition, we collect: 1) an 8-day trace of network telemetry in a production FDD base station (2.175GHz with 20MHz bandwidth) at a university campus. 2) a 2-day trace in a production FDD base station (2.12GHz with 20MHz bandwidth) in the downtown area of one of the top 5 largest US cities. Since we do not have power measurements for the production base station, we use these measurements solely to motivate our observations about traffic patterns.

Applicability to 5G. Although we designed EcoCell from our observations on an LTE base station, 5G's subframe structure largely retains the characteristics we exploit for our design. We conduct limited experiments on an off-the-shelf commercial 5G standalone base station in Sec. 7.9.

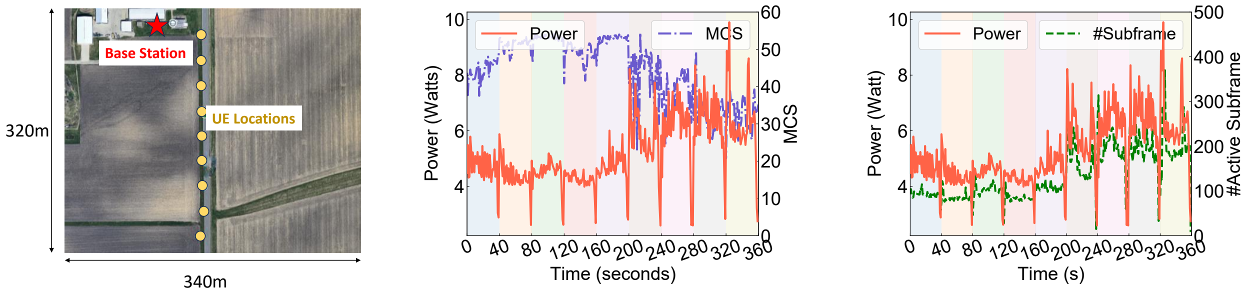
5 ENERGY SAVING OPPORTUNITIES

Based on cellular telemetry data and power measurements, we identify three key opportunities for energy savings.

5.1 Effect of Subframe Utilization

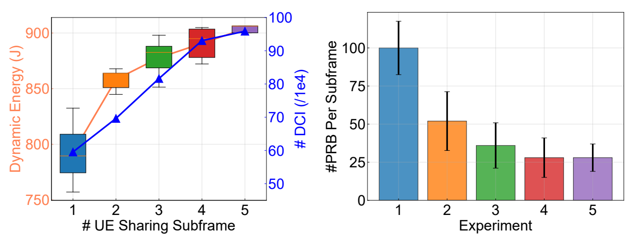
Observation: Our first observation is that reducing the number of subframes, by packing more data in a single subframe, reduces the energy consumption of a base station. This is because transmitting every sub-frame induces additional control (DCI) transmissions. Given static channel conditions, we can pack more data in a subframe by utilizing more PRBs in a single subframe.

Experiment: To validate this observation, we setup a microbenchmark. We use a server to transmit UDP data (200 Mb total) to a UE through the base station using different transmission patterns. We start by transmitting continuously at 2Mbps from the server. Then, we transmit at 10 Mbps for 0.2 seconds, followed by no transmissions for 0.8 seconds (i.e., 20% duty cycling). We repeat these experiments for 20 Mbps



(a) Experiment Map **(b) Power VS MCS** **(c) Power VS #Active Subframe**

Figure 7: Effect of MCS on energy consumption: (a) Base station and UE locations (b) Power shows inverse relationship with MCS (c) Power shows linear relationship with number of subframes transmitted



(a) **(b)**

Figure 8: (a) As more UEs share the downlink capacity, energy consumption increases.(b) PRBs allocated to each UE go down because subframes are shared across UEs.

(0.1 seconds on period every 1 second) and 66 Mbps (0.03 seconds on period every 1 second). These transmission patterns achieve the same effective rate, with different levels of burstiness. We plot the PRB utilization per subframe for each of these flows in Fig. 6b. As expected, with increasing burstiness, the PRB utilization goes up. Since our base station uses 20 MHz bandwidth, there are a total of 100 PRBs available for each subframe. We see a clear benefit of increased PRB utilization (i.e., more utilization per subframe, but fewer total subframes) in terms of energy savings in Fig. 6a. Across 10 runs of these experiments, the average energy consumption drops by 53.2% for the same amount of data.

Reason: This experiment demonstrates that transmitting more PRBs per subframe (and hence fewer subframes) leads to energy savings because of much less control signaling overhead. This is validated by the corresponding reduction in the number of DCIs (in blue in Fig. 6a) by 57%. The energy consumption drops linearly with the drop in downlink control signaling.

5.2 Effect of Subframe Sharing across UEs

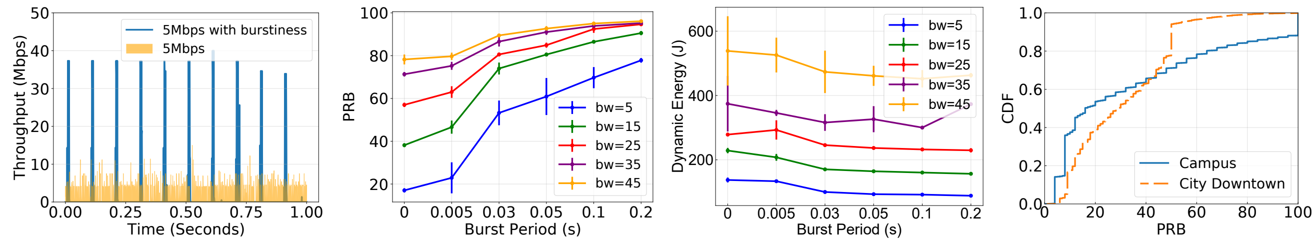
Observation: We observe that the distribution of PRBs across UEs significantly influences energy usage. In a cellular network, the base station's scheduling policy dictates the allocation of PRBs to multiple UEs in each subframe. For instance, consider two UEs requesting 100 PRBs each. One allocation strategy, **(i)**, evenly divides the PRBs across subframes, providing each UE with 50 PRBs in two different subframes. Alternatively, **(ii)**, one subframe could be entirely allocated to the first UE, and the next subframe entirely to the second UE. Our findings suggest that the latter approach leads to significantly reduced energy consumption.

Experiment: To verify this, we saturate the cellular link with UDP traffic at 100Mbps (higher than base station capacity). In each experiment, we divide this rate equally into different numbers of UEs to compare the energy consumption. Saturating the link with multiple UEs causes the base station to share a subframe's resource blocks with multiple UEs. We plot the energy consumption (in orange) in Fig. 8a; it shows that when number of UEs increases from 1 to 5, the energy consumption increases by 14%. Fig. 8b shows the number of PRBs assigned to individual UE per subframe. When the number of UEs increases, fewer PRBs are assigned to each UE in each subframe, demonstrating sharing of the subframe by the multiple UEs. For instance, with a single UE, the median PRB per UE is 100, whereas with four UEs, it drops to 28 due to sharing. These two results combined show the implication that: *more UEs shared per subframe leads to more energy consumption.*

Reason: Including multiple UEs in a single subframe requires control information for each UE in that subframe. Therefore, more UEs sharing a subframe leads to more control signaling overhead, as shown in Fig. 8a blue. The number of DCIs increases by 59% when the number of UEs increases to 5. Considering the allocation strategy examples mentioned above, there will be 4 DCIs in total (2 in each subframe for 2 UEs) in the strategy **(i)**, while only 2 in strategy **(ii)**, each in a subframe serving a single UE.

5.3 Effect of MCS

Observation: The modulation and coding scheme, or MCS, governs how much user data can be encoded into each resource block. The base station computes per-UE MCS values



(a) SF packing, with t_{burst} = 0.1s **(b) SF packing increases PRB** **(c) SF packing saves energy** **(d) Commercial sites' PRB**

Figure 9: EcoCell's subframe (SF) packing: (a) Bursty pattern with period of 0.1s is shown clearly, captured by NG-Scope. (b) Our subframe packing technique increases the number of PRBs per subframe. (c) Energy is saved at varieties of bandwidth due to PRB usage increasing, energy is averaged over 10 repeated experiments with 100 seconds of Iperf downloading. (d) Commercial base station measurement shows severe PRB under-utilization

based on the UE's downlink signal quality [2]. The base station energy consumption increases with decreasing MCS values.

Experiment: We vary the MCS value by placing one UE at nine different locations shown on the yellow dots in Fig. 7a, each separated by around 40 meters; The red star shows the BS's location. At each location, we ensure that the base station is not serving any other users, and then run an iperf TCP flow with a fixed throughput of 25 Mbps to download data through the base station for 40 seconds. We plot the aggregated MCS values by adding MCS values from both spatial streams and the result is shown in Fig. 7b. Different background colors from left to right represent UE moving away from the base station. As the UE moves away from the BS, its MCS decreases, and the BS's power draw increases.

Reason: As the MCS decreases, the base station needs to use more frequency and time resources (i.e., subframes and PRBs) to exchange the same amount of user data, increasing the BS energy consumption. We plot the number of non-empty subframes to verify this relationship in Fig. 7c.

6 ECOCELL DESIGN

We design EcoCell with three goals: (a) improving BS energy efficiency, (b) compatibility with off-the-shelf BSes, and (c) minimizing the impact on application QoE. Based on the observations above, EcoCell introduces energy-efficient patterns in the BS traffic through three modules: (1) subframe packing, (2) UE segregation, and (3) temporal shifting.

Our current design of EcoCell is implemented as a software middlebox that intercepts and shapes traffic going into the base station. This design does not require any software, hardware, or protocol changes at the UE or the BS. In emerging cloud-based or programmable RANs, EcoCell can directly be integrated into the BS itself, above the RAN's Packet Data Convergence Protocol (PDCP) layer.

6.1 Module 1: Subframe Packing

EcoCell's subframe packing module aims to aggregate more data into each subframe by introducing artificial burstiness in the downlink traffic to the base station.

We implement this by queuing ingress traffic at a base station for a short period of time (50 ms by default), and releasing it in periodic bursts to the base station. This bursty traffic pattern naturally results in some subframes being left empty while allowing more data to be packed into a smaller number of subframes. Note that this improvement is achieved without any modifications to the base station itself. We implement this burstiness with a Linux TUN device to intercept traffic. EcoCell captures and buffers all packets passing through the TUN device, subsequently releasing all buffered traffic at periodic intervals of duration t_{burst}.

Commercial site statistics: To validate that subframes are under-utilized in commercial BSes, we sniff traffic from a campus base station and a base station in a large city, using the setup described in Sec. 4). Figure 9d shows that over half of the subframes use fewer than 20% of the PRBs on campus and fewer than 29% of the PRBs in a major U.S. city downtown, showing high subframe underutilization and the potential to reduce energy consumption. Our measurements match the well-known observation that cellular networks are generally under-utilized [23].

Microbenchmark: We conduct the following microbenchmark to show how EcoCell can improve subframe packing. We attach one UE to our BS, and start a 5 Mbps UDP-based iperf download. We plot the on-air throughput observed using NG-Scope in Fig. 9a. With EcoCell configured with $t_{burst} = 0.1s$, the traffic pattern has large bursts every 100 ms, as expected.

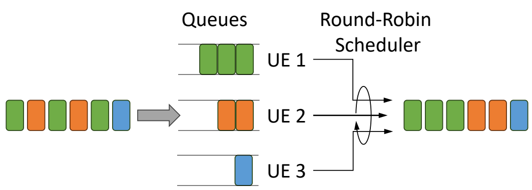


Figure 10: UE traffic segregation via per-UE buffer.

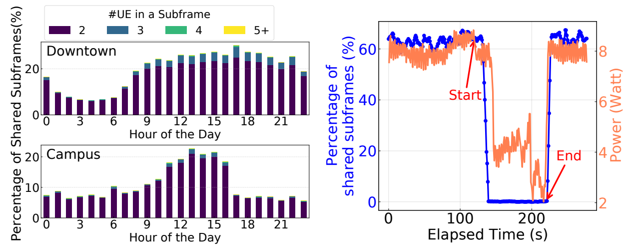
Figure 9b shows that increasing t_{burst} improves subframe packing, at the cost of higher added latency (assuming packets arrive uniformly, the added latency is $t_{burst}/2$). For example, for a 5 Mbps iperf download without EcoCell, non-empty subframes use fewer than 20% of their RBs on average. In contrast, bursting at 50 ms intervals increases the average RB utilization of non-empty subframes to over 60%. Figure 9c shows the corresponding reduction in BS's dynamic energy used for the download: 137J without EcoCell, and 93J with 50 ms burstiness. We note that we did not observe significant subframe packing benefits in iperf TCP flows which were naturally bursty. However, our experiments (Sec. 7) show that subframe packing saves power in other important TCP workloads such as video streaming over HTTP.

How should a cellular operator choose an appropriate burst interval t_{burst}? We believe that it depends on the energy-latency tradeoff that the operator is willing to make, which may vary across different cell sites and time of day. Operators already make such energy-performance tradeoffs with energy efficiency optimizations like turning off cell sites or carriers at certain hours of the day (e.g., post-midnight). We explore this tradeoff in detail in Sec. 7 for various applications. Based on our evaluation results, operators can design custom, application-adaptive algorithms in RAN Intelligent Controller(RIC) to directly provide millisecond-level bursty pattern adjustment. For example, URLLC traffic (such as drone, remote surgery), which requires ultra-low latency (<10 ms), can be assigned a small bursty period or be allowed to pass directly, while eMBB services could use a larger bursty period simultaneously. Empirically, we observe that t_{burst} = 50 ms works well for all studied applications with significant energy gains and negligible impact on application performance.

6.2 Module 2: UE Segregation

Our second goal is to reduce the number of UEs scheduled concurrently in a subframe, which we find empercially to be a more energy-effient traffic pattern.

EcoCell reorders the downlink traffic to the base station such that packets from the same UE arrive together. When packets arrive at EcoCell's TUN device, we place them into separate per-UE queues (Fig. 10). We, then, build a round-robin scheduler that dequeues packets from each queue one



(a) Subframe sharing ratio over hours of the day

(b) UE segregation reduces power consumption

Figure 11: UE segregation measurements and microbenchmark

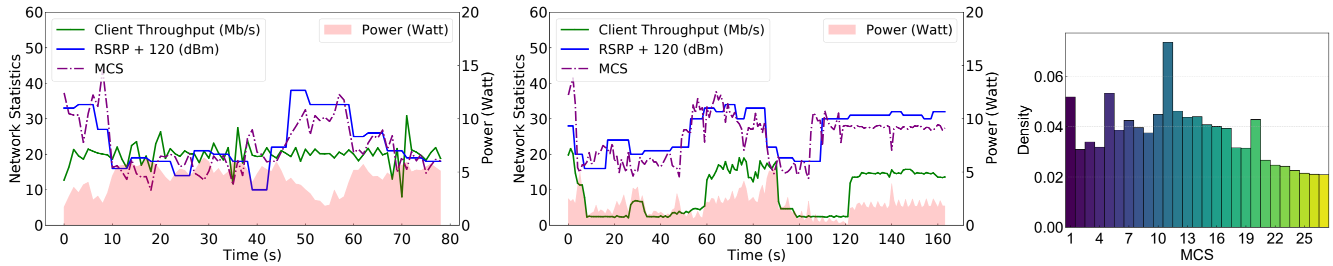
by one, i.e., it dequeues all packets from one UE followed by all packets from the second UE. To ensure that one UE does not starve the others, we implement a maximum quota for each UE within a t_{burst} period. The quota corresponds to the fair share for a UE within a bursty period and depends on the number of UEs and the length of the bursty period.

Commercial site statistics: How often are subframes shared across UEs and how many UEs share a subframe in practice? We analyze this question in our week-long trace from two commercial base stations as described before (Sec. 4). We plot the percentage of subframes shared by more than one UE in Fig. 11a. Over 20% of the subframes are shared by two or more UEs during the day because of high traffic demand. While there is less sharing at night, the ratio remains above 5% on campus and 10% in the city. These patterns show a significant opportunity for using UE segregation to reduce subframe sharing and save energy.

Microbenchmark: To show the effect of UE segregation, we conduct a microbenchmark experiment shown in Fig. 11b. We start by saturating the cellular link with 5 UEs each receiving UDP traffic at 20 Mbps. We find that over 60% of the subframes are shared by multiple UEs (blue line). We enable our UE segregation module 120 seconds into the experiment, which reduces the fraction of shared subframes to near-zero. Correspondingly, the BS's dynamic power draw decreases from roughly 8 W to 4 W.

6.3 Module 3: Temporal Shifting

Energy consumption tends to rise during poor channel conditions, characterized by lower MCS, and conversely decreases during favorable conditions. A fundamental question arises: Can we defer data transmission to a UE until channel conditions improve, thereby reducing energy consumption per bit? Given that channel conditions may persist unchanged for several seconds, especially for stationary users, it would be unwise to delay latency-sensitive data, such as video conferencing, by seconds. However, for delay-tolerant traffic like file transfers and software updates, we propose a traffic



(a) Dynamic Energy Consumed: 362 J **(b) Dynamic Energy with EcoCell : 318 J** **(c) Commercial Site MCS**

Figure 12: EcoCell's temporal shifting: (a)(b)12% of the energy is saved in a real-world file downloading example. (c) MCS distribution spreads out in commercial base station measurement, showing potential for leveraging EcoCell.

temporal shifting strategy. This approach involves dynamically adjusting throughput based on observed improvements in channel conditions, as indicated by the Reference Signal Received Power (RSRP).

Implementation: The challenge with an RSRP-based temporal shifting algorithm lies in the possibility that future RSRP values may deteriorate, resulting in unnecessary traffic delays and no energy savings. To address this challenge, we adopt a heuristic-based approach. We make decisions regarding traffic delays and their durations based on observed average RSRP values. Specifically, we adjust the throughput of each flow every three seconds in proportion to the observed RSRP. This adjustment is linear, ensuring that higher RSRP values correspond to increased throughput and vice versa.[1] We place two additional constraints on the throughput. First, throughput adjustments should not violate any deadlines, such as flow completion times in long-running flows (when they are available). Second, we scale the overall throughput deviation to be between the minimum and maximum throughput supported by the base station. In practice, without access to the RAN, we implement such traffic shaping through EcoCell's TUN interface.

Commercial site statistics: We present the MCS distribution of devices observed in our commercial traffic trace in Fig. 12c. The distribution showcases a broad range of MCS values, with a slight concentration observed between 10 to 16. This MCS landscape underscores the significant potential for generalizing our energy savings solution via temporal shifting in commercial cellular networks.

Microbenchmark: To demonstrate the impact of temporal shifting, we conduct a real-world experiment and present the results in Fig. 12. In this experiment, a user device attempts to download a 200MB file from the server, with the server download speed set at 20Mbps, We set the desired job finishing time to be 600s.

[1]In practice, the MCS (and hence the datarate) is mapped by CQI (Channel Quality Indicator), we use RSRP instead due to the data availability and its high correlation with CQI.

Fig. 12a illustrates the download process without EcoCell, where the client throughput remains constant despite fluctuations in RSRP values. For instance, between the 40 to 60-second interval, the RSRP increases from -110dBm to -90dBm, prompting a corresponding rise in transmission MCS (aggregated by 2 spatial streams) from 20 to 31. However, due to the constant throughput, energy consumption decreases by 62%. Conversely, during the 20 to 40-second interval, energy consumption is higher due to downloading at a constant throughput during a period of low RSRP.

In contrast, Fig. 12b demonstrates our dynamic throughput control mechanism. As RSRP changes, EcoCell adjusts the throughput accordingly. For instance, during the 10 to 50-second interval, with RSRP at -100dBm, EcoCell throttles the traffic to a lower throughput of 2.6Mbps, awaiting improved RSRP conditions. Overall, our method reduces energy consumption during file downloading from 362J to 318J, a reduction of 12%, while ensuring the file is downloaded within the specified time frame.

How do we identify delay-tolerant traffic? Such traffic classification has been explored in past work extensively [57, 64]. In general, such traffic classification is also increasingly available to operators through slicing indicators [8, 36]. Therefore, for our work, we assume that delay-tolerant traffic characteristics are available to us either through classification algorithms or through dedicated delay-tolerant slices.

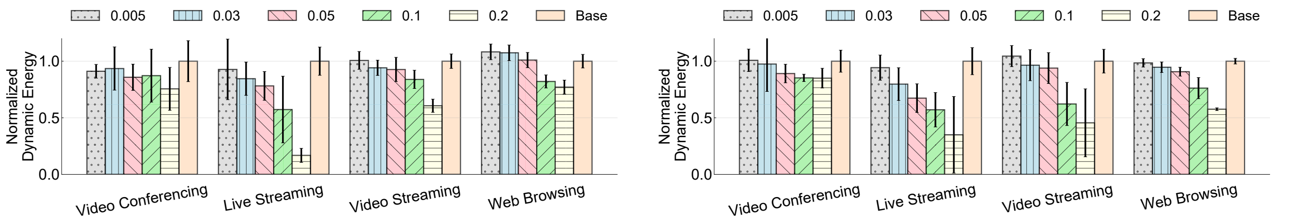
7 ECOCELL EVALUATION

We evaluate EcoCell's energy saving ability and its effect on the network and user experience in the real world with 5 representative applications. We test both single and multiple client cases on all applications.

7.1 Target Applications

We choose the following applications in our experiment representing varieties of latency and throughput requirements.

Video conferencing: We use Google Meet as a representative video conferencing application [17]. Video conferencing



(a) Single UE **(b) 5 UEs**

Figure 13: EcoCell reduces dynamic energy consumption of multiple applications across different values of t_{burst}

is an extremely delay-sensitive application. We set up a meeting room using Chrome on the server with a virtual camera which is fed with a pre-recorded video. A client joins the meeting through the cellular network with the same virtual camera setting as well. We use the WebRTC extension in Chrome to obtain the QoE metrics including frame per second, video resolution, etc.

Live streaming: Live streaming applications such as live games, TV, news, and events have more delay sensitivity than YouTube/Netflix but less urgency than Google Meet. To emulate such workloads, we stream the "EnvivioDash3" video using DASH from a server to the UE. This video is 193 seconds long with 48 chunks and has been commonly used as a benchmark for past work [44, 50]. The server performs Adaptive Bitrate (ABR) adaptation using the robustMPC algorithm [63] to dynamically select the video bitrate among {300, 750, 1200, 1850, 2850, 4300} kbps. The client player is a Google Chromium browser and we set the DASH playback buffer capacity to be 2 seconds in order to enforce the real-time effect. We achieve this by setting the dash.js parameters in the html: *setlivedelay()* and *setStableBufferTime()*.

Video streaming: Video streaming applications such as Netflix and YouTube account for a large fraction of network traffic. We adopt the same setup as above but we set the DASH playback buffer capacity to be 60 seconds [50] since it has less urgency.

Web browsing: The client downloads websites from the HTTP server through the cellular network. We choose 10 popular websites and obtain not only the HTML but CSS and multimedia assets as well and put them onto the server. The average website size is 6.5MB with the highest of 17.9MB and the lowest of 2.4MB. The client mimics the user browsing these websites with intervals of 5 seconds.

File downloading: We transfer a 400MB dummy file from the server. The client uses the wget tool to download from it through the cellular network.

7.2 Summary of Results

We begin by summarizing our key results in Table. 2, for $t_{burst} = 50ms$ and $t_{burst} = 100ms$. As we demonstrate later,

	EcoCell Dynamic Energy Saving (%)			
t_{burst}	50ms		100ms	
	1 UE	5 UEs	1 UE	5 UEs
Video Conf	14.21	10.92	12.79	14.77
Live Streaming	21.85	32.65	42.73	42.93
Video Streaming	7.35	6.11	16.08	37.80
Web Browsing	0.00	9.30	17.91	23.79
File Downloading	25.90	X	25.90	X

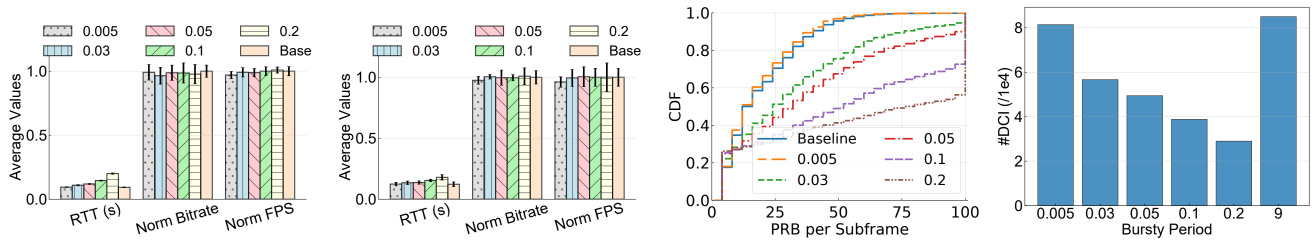
Table 2: EcoCell energy saving on 5 applications.

$t_{burst} = 50ms$ maintains QoE for all applications and $t_{burst} = 100ms$ leads to slight drop in QoE for some applications (e.g., Google Meet). We also plot detailed results across different configurations and applications in Fig. 13.

EcoCell saves energy across most configurations. The benefits for energy saving are the most pronounced for video streaming and live streaming applications, with power savings of 32% and 42% for with different burst settings. Web-browsing with a single UE is a low-bandwidth flow that doesn't improve from subframing packing at the 50 ms scale, but improves at higher granularity of packing. As noted in Sec. 2, even 10% energy savings in dynamic energy consumption of base stations significantly reduces energy costs. Note that, video constitutes about two-thirds of Internet traffic [58]. Therefore, the prominent benefits on various video applications are noteworthy. In addition, as expected, higher burst periods lead to more opportunities for subframe packing and UE segregation and lead to higher energy savings. Finally, the 5 UE case usually saves more energy than 1 UE because it can benefit from both UE segregation and subframe packing. We delve deeper into individual applications in subsequent sections. We choose up to 5 UEs in our experiments because we find modern base stations rarely share more than 5 UEs(<1%) in a subframe as indicated in Fig. 11a.

7.3 EcoCell on Video Conferencing

Video conferencing features ultra-low latency and high bitrate UDP traffic. We test EcoCell with different bursty period settings and for each setting, we conduct a Google Meet session for 80s (repeated 10 times). Fig. 13a shows the normalized energy saving against our baseline on 1-on-1 meetings.



(a) 1-on-1 Video Conf QoE **(b) 5UE Video Conf QoE** **(c) EcoCell increases #PRB/UE** **(d) #DCI transmitted**

Figure 14: Video Conferencing (Google Meet) with EcoCell

For 1UE, EcoCell saves 9%, 7%, 15%, 13% and 25% of the energy with t_{bursty} values in [0.005, 0.03, 0.05, 0.1, 0.2]s.

As t_{bursty} increases, the energy savings increase because EcoCell's subframe packing technique can aggregate longer periods of data to better pack subframes. This is evident in Fig. 14c where the median number of PRBs per subframe increases from 12 to 60 when the bursty period increases to 0.1s. Better subframe packing leads to the reduction of 75% of the total transmitted control frames from the base station as shown in Fig. 14d because every subframe carries more data thus less DCI is needed.

We measure the influence of EcoCell on application experience in terms of latency, bitrate, and Frames Per Second(FPS) in Fig. 14a. EcoCell doesn't affect the bitrate and FPS but introduces small extra delays. The RTT increased from 92ms to 118ms when changing from baseline to EcoCell with 0.05s burst period. However, this delay is still within acceptable limits for video-conferencing applications [8, 10, 18, 49].

Similarly, for 5 UEs performing a group meeting on Google Meet, we show the energy saving achieved by EcoCell in Fig. 13b, 11% of the energy can be saved with 0.05s of the bursty period while only increasing the RTT from 123ms to 137ms and providing the same bitrate and FPS experience.

7.4 EcoCell on Live Streaming

In comparison to video conferencing, live streaming exhibits a less stringent latency requirement. Activities such as sports and live events can tolerate a few seconds of delay. In our approach, we employ DASH for live streaming, with the buffer size capped at a maximum of 2 seconds. This restriction implies that video chunks cannot be downloaded more than 2 seconds in advance.

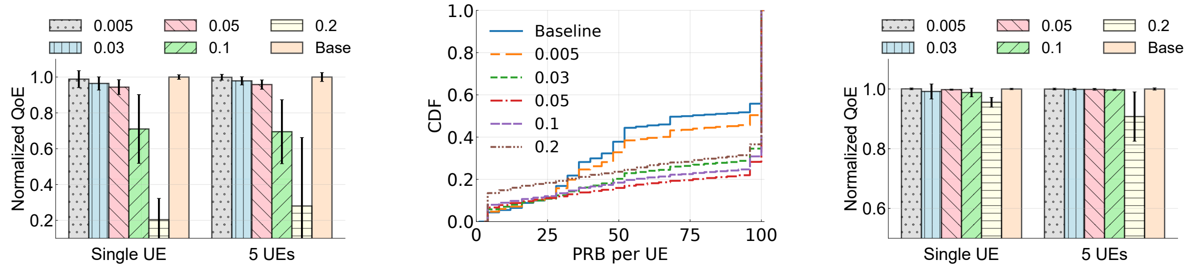
The energy consumption of averaged live streaming across 10 runs with different burst periods is shown in Fig. 13a. EcoCell leads to energy savings of 8%, 16%, 22%, 43% and 84% of energy consumption with 1 UE for the different burst periods. We plot the QoE experienced by users in Fig. 15a. QoE metric is adopted from [50, 63] and it's equal to the sum of the bitrate of each video chunk subtracted by smooth penalty (the total variation of chunk bitrate times 4.3) and total rebuffering time(video stalling time). It shows that EcoCell reduces the QoE on 1 UE live streaming by 2%, 4%, 6%, 29% and 80% with increasing bursty periods. One may wonder why the QoE drops for this delay-tolerant traffic even at 100 ms burst periods. This is because of two reasons: (a) the buffer is small (2 seconds), so larger delays lead to more rebuffering events and cause stalls, (b) we use RobustMPC [63] for bit rate adaptation. RobustMPC lowers its bandwidth estimate in response to longer delays and picks lower quality video to send to the client (leading to drops in QoE). However, at 50 ms delay, the QoE drop is minimal and still achieves 22% energy savings.

For live streaming to multiple simultaneous UEs, EcoCell saves even more energy due to subframe segregation. From Fig. 13, EcoCell saves 5%, 21%, 33%, 43%, and 66% energy across the different configurations which is more than single UE cases at the bursty period of 0.03 to 0.1s. Enabling UE segregation can significantly increase PRBs used for each UE (and reduce the number of shared subframes), as shown in Fig. 15b. Average PRB per UE increases by 25% from baseline to a bursty configuration of 0.05s. This configuration saves 33% energy at the cost of a mere 5% QoE drop.

7.5 EcoCell on Video Streaming

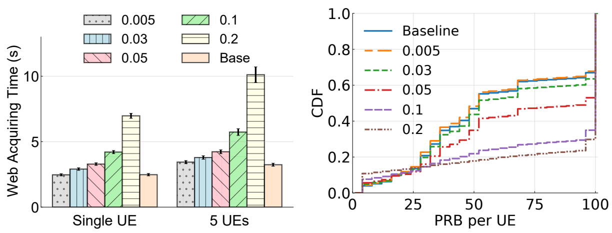
Next, we study the effect of EcoCell on DASH-based video streaming. For each burst configuration, we stream the same video 10 times. We show the average energy consumption per video normalized with the baseline in Fig. 13. The trends are similar to live streaming results discussed above. For both single UE and multiple-UE video streaming, EcoCell starts to save power with a burst period bigger than 0.03. With [0.03, 0.05, 0.1, 0.2]s bursty periods, we save energy by 6%, 8%, 17%, 40% with 1UE. For 5 UEs, we save 4%, 7%, 38%, and 55% of the energy consumption per video at [0.03, 0.05, 0.1, and 0.2]s burst periods. The increase in the 5UE case is due to EcoCell's UE segregation module.

However, compared to live streaming, the QoE drop is lower in the case of video streaming because of a larger buffer of 60 s. We find that EcoCell affects the video streaming QoE



(a) Live Streaming QoE **(b) 5 UE Live Streaming PRB** **(c) Video Streaming QoE**

Figure 15: Live Streaming and Video Streaming with EcoCell



(a) Web Fetching Time **(b) Web Browsing PRB Usage**

Figure 16: EcoCell on Web Browsing

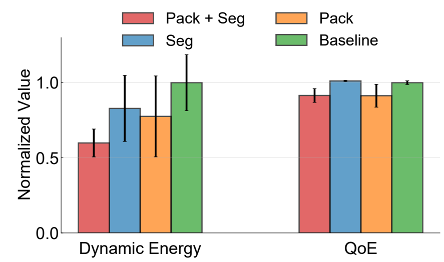


Figure 17: EcoCell Ablation Study – Contribution of different modules to EcoCell

very slightly, even with a bursty period of 0.2 seconds, QoE is only reduced by 4.5% (and has energy savings of nearly 40%). The QoE data is plotted in Fig. 15c.

7.6 EcoCell on Web Browsing

In our web browsing workload, each UE downloads multiple websites one by one at 5 second intervals. For the single UE case, EcoCell leads to minimal energy savings. This is because web-traffic is naturally short-lived and bursty. As shown in Fig. 13a, EcoCell saves no energy for burst periods of 0.05 s and lower, and saves just 17% energy at 100 ms. The added burstiness increases webpage load times (by 1.8s at 100 ms) as shown in ig. 16a.

Energy savings are higher for multiple UEs because of UE segregation (e.g., 24% on at 100 ms). This is demonstrated in Fig. 16b where the median PRB per UE increases from 48 to 100 when changing from baseline to bursty period of 0.1s.

7.7 EcoCell on File Downloading

For all our workloads so far, we solely relied on subframe packing and UE segregation modules in EcoCell. This is because these workloads cannot tolerate long term traffic shifting. On the other hand, large file downloads like operating system updates, background traffic, etc., can benefit from EcoCell's temporal shifting module. Therefore, for file downloads we evaluate the temporal shifting module. We hold the UE device and walk along 10 different trajectories while downloading a 200MB file from the server. We set the maximum download speed at the server side to 20 Mbps. Across these trajectories, EcoCell reduces the average dynamic energy consumption by 25.9%, while elongating the file download time from 78 seconds to 144 seconds.

EcoCell's energy savings come from the temporal shifting module. Since file download operates at large volumes and utilizes TCP which is naturally bursty, we do not see significant benefits from subframe packing since the subframe is already well-packed in the baseline case. Similarly, big window size and the burstiness of TCP makes the traffic designated to different UEs highly segregated. Therefore, the maximum benefits come from temporal shifting.

7.8 EcoCell Ablation Study

We explicitly measure the contribution of EcoCell's subframe packing and UE segregation modules independently in the context of live streaming with 5 UEs. We try different configurations with one or both of these modules turned on. For each configuration, we repeat the 193-second live streaming five times. Note that, the burst period is set to 0.05s throughout the experiments.

Fig. 17 shows the normalized energy consumption and streaming QoE under different configurations with 1 or 2 modules of EcoCell removed. With UE segregation alone, EcoCell saves 18% energy compared to the baseline. With subframe packing alone, the corresponding energy savings are 23%. Finally, EcoCell with both modules turned on saves

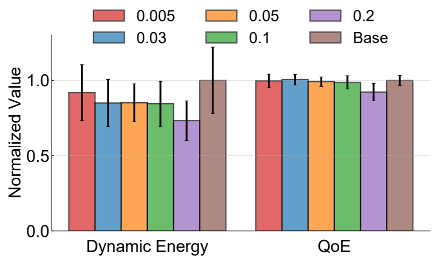


Figure 18: EcoCell saves energy on real-world 5G standalone base staion with different bursty periods.

41% energy compared to the baseline. This result validates that both modules contribute to energy savings.

7.9 EcoCell on Real-world 5G Base Station

We further show the generalization of EcoCell on a 5G Standalone (SA) commercial base station. The base station operates at 3.50 GHz (N78 frequency band) using TDD with a bandwidth of 100MHz. It has 2 omnidirectional antennas with a maximum gain of 6dBi, and its maximum power consumption is 35 Watt. The power measurement is conducted in the same way as Sec. 4. The UE is a Google Pixel 5 [25] and we connect it to a laptop with USB tethering. We have limited access to this testbed, so we conduct experiments in one setting to validate our findings. A detailed exploration is left to future work.

We conducted video streaming with different burst periods and repeated the experiment 10 times. The results are shown in Fig. 18. We observed that EcoCell starts to save 8.2% of energy even with a small bursty period of 5ms, which corresponds to 2.5ms of latency assuming uniform subframe arrival. With a bursty period of 30ms, energy savings increase to 15%. Additionally, we note a barely noticeable decrease in Quality of Experience (QoE) for video streaming, with less than a 1% reduction when the bursty period is under 0.1s. These results demonstrate that EcoCell's design generalizes well to 5G, effectively saving energy while meeting latency and quality requirements. This is feasible because 5G adopts frame structures similar to those of 4G, despite differences in waveform numerologies. For example, 5G data transmissions are also divided into subframes and include a control region that typically occupies 1-3 symbols. Thus, EcoCell can achieve energy savings in a manner similar to LTE.

8 RELATED WORK

Real-world base station energy saving: While numerous theoretical studies exist on base station energy analysis, practical implementations are scarce. [21, 43, 48] measure LTE base station power but lack cellular-level measurements. Our work differs from them by analyzing the power consumption jointly with PHY-level statistics such as encoding schemes and PRB usages. Another line of work by Ericsson [9, 16, 24, 29, 54] focuses on designing power-efficient base stations by either introducing sleep mode and dynamic sector selection or changing the base station PRB scheduling policy. We differ from them in 2 folds: 1) EcoCell does not require base station modifications, but utilizes traffic shaping for power savings rather than altering the base station's operation. 2) Our approach conserves power without compromising user experience.

Base station power modeling: Several lines of research have developed mathematical models to assess base station energy consumption, primarily through simulation [6, 13, 27, 30, 33, 47, 52]. These models often dissect base station components individually, including the power amplifier [14], baseband unit (BBU) [26, 59], cooling systems, *etc.* However, they lack real-world measurements to validate their models. In contrast, [4, 38] perform component-level measurements but do not propose power-saving strategies. Our work distinguishes itself by focusing on the energy incurred during signal transmission and offering practical solutions for energy conservation.

Client-side power saving: Previous studies have investigated power-saving strategies for cellular clients. For example, [31] develops an LTE client power model and introduces a power management scheme. Additionally, [35, 55, 65] examine the relationship between UE power consumption and device usage patterns, particularly focusing on LTE radio usage. Furthermore, research such as [20, 37, 40, 41] proposes novel scheduling algorithms or manipulates Radio Resource Control (RRC) connection states in simulations to decrease LTE UE power consumption. In contrast, EcoCell concentrates on energy-saving techniques from the perspective of the base station.

9 CONCLUSION

We design and implement EcoCell, an energy-aware traffic shaping system for cellular network power saving. Our work demonstrates energy savings in an off-the-shelf base station simply by traffic shaping. We evaluate EcoCell in many real applications and demonstrate that significant energy savings can be obtained with minimal impact on application throughput, delay, or QoE. In practice, this allows private network operators and base station operators to deploy EcoCell *today* and start saving energy. Our strategies can also be selectively applied to different slices such as delay-tolerant vs delay-sensitive to optimize energy use. In emerging software-based RANs, our schemes can be easily programmed into the scheduler and extract even finer control over base-station scheduling to increase energy efficiency.

REFERENCES

[1] F. 289. https://www.fluke.com/en-us/product/electrical-testing/digital-multimeters/fluke-289.
[2] 3GPP36.213. 3gpp ts 36.213 version 14.2.0 release 14.
[3] U. E. P. Agency. Greenhouse gas emissions from a typical passenger vehicle. https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle#:~:text=2%20per%20mile.-,What%20is%20the%20average%20annual%20carbon%20dioxide%20(CO2)%20emissions,around%2011%2C500%20miles%20per%20year., 2022.
[4] N. Alliance. Green future networks: Network energy efficiency. *NGMN, Whitepaper*, 2021.
[5] K. P. Antennas. https://www.kpperformance.com/3ghz-45-degree-slant-90-degree-16-7-dbi-sector.
[6] O. Arnold, F. Richter, G. Fettweis, and O. Blume. Power consumption modeling of different base station types in heterogeneous cellular networks. In *2010 Future Network & Mobile Summit*, pages 1–8. IEEE, 2010.
[7] A. Balasingam, M. Bansal, R. Misra, K. Nagaraj, R. Tandra, S. Katti, and A. Schulman. Detecting if lte is the bottleneck with bursttracker. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2019.
[8] bectechnologies. Quality of service (qos) in lte, 2019.
[9] T. Berglund and H. Huynh. Energy efficiency of radio units and its impact on ran energy consumption. 2017.
[10] Broadband-forum. Can your network handle a hd zoom call?, 2023.
[11] N. Bui and J. Widmer. Owl: A reliable online watcher for lte control channel measurements. In *Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*, pages 25–30, 2016.
[12] Celona. https://www.celona.io/.
[13] K.-C. Chang, K.-C. Chu, H.-C. Wang, Y.-C. Lin, and J.-S. Pan. Energy saving technology of 5g base station based on internet of things collaborative control. *IEEE Access*, 8:32935–32946, 2020.
[14] A. Chatzipapas, S. Alouf, and V. Mancuso. On the minimization of power consumption in base stations using on/off power amplifiers. In *2011 IEEE Online Conference on Green Communications*, pages 18–23. IEEE, 2011.
[15] C. Desset, B. Debaillie, V. Giannini, A. Fehske, G. Auer, H. Holtkamp, W. Wajda, D. Sabella, F. Richter, M. J. Gonzalez, H. Klessig, I. Gódor, M. Olsson, M. A. Imran, A. Ambrosy, and O. Blume. Flexible power modeling of lte base stations. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2858–2862, 2012.
[16] T. Edler and S. Lundberg. Energy efficiency enhancements in radio access networks. *Ericsson review*, 81(1):42–51, 2004.
[17] emailtooltester. https://www.emailtooltester.com/en/blog/video-conferencing-market-share/.
[18] Evercast. What is low latency video conferencing and how to set it up, 2023.
[19] R. Falkenberg and C. Wietfeld. Falcon: An accurate real-time monitor for client-based mobile network data analytics. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7. IEEE, 2019.
[20] J. Fan and S. Alanara. Ue power saving with rrc semi-connected state in lte. In *2014 Wireless Telecommunications Symposium*, pages 1–5. IEEE, 2014.
[21] J.-L. Fan, Z. Li, X. Huang, K. Li, X. Zhang, X. Lu, J. Wu, K. Hubacek, and B. Shen. A net-zero emissions strategy for china's power sector using carbon-capture utilization and storage. *Nature Communications*, 14(1):5972, 2023.
[22] Fluke. Fluke ir189usb usb cable adapter.
[23] P. Frenger, C. Friberg, Y. Jading, M. Olsson, and O. Persson. Radio network energy performance: Shifting focus from power to precision. *Ericsson review*, 91(2):1–9, 2014.
[24] P. Frenger and R. Tano. More capacity and less power: How 5g nr can reduce network energy consumption. In *2019 IEEE 89th vehicular technology conference (VTC2019-Spring)*, pages 1–5. IEEE, 2019.
[25] GSM Arena. Google pixel 5, n.d.
[26] H. Guo, K. Wang, H. Ji, and V. C. Leung. Energy saving in c-ran based on bbu switching scheme. In *2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pages 44–49. IEEE, 2016.
[27] A. Gupta, I. Jain, and D. Bharadia. Multiple smaller base stations are greener than a single powerful one: Densification of wireless cellular networks. In *1st Workshop on Sustainable Computer Systems Design and Implementation (HotCarbon)*, 2022.
[28] L. Hardesty. 5G base stations use a lot more energy than 4G base stations: MTN . https://www.fiercewireless.com/tech/5g-base-stations-use-a-lot-more-energy-than-4g-base-stations-says-mtn, 2022.
[29] L. G. Hévizi and I. Gódor. Power savings in mobile networks by dynamic base station sectorization. In *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 2415–2417. IEEE, 2011.
[30] H. Holtkamp, G. Auer, V. Giannini, and H. Haas. A parameterized base station power model. *IEEE Communications Letters*, 17(11):2033–2035, 2013.
[31] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. A close examination of performance and power characteristics of 4g lte networks. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 225–238, 2012.
[32] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck. An in-depth study of lte: Effect of network protocol and application behavior on performance. *ACM SIGCOMM Computer Communication Review*, 43(4):363–374, 2013.
[33] I. Humar, J. Zhang, Z. Wu, and L. Xiang. Energy savings modeling and performance analysis in multi-power-state base station systems. In *2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*, pages 474–478. IEEE, 2010.
[34] F. Instruments. https://www.fluke.com/en-us.
[35] A. R. Jensen, M. Lauridsen, P. Mogensen, T. B. Sørensen, and P. Jensen. Lte ue power consumption model: For system level energy and performance optimization. In *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pages 1–5. IEEE, 2012.
[36] B. Khodapanah, A. Awada, I. Viering, D. Oehmann, M. Simsek, and G. P. Fettweis. Fulfillment of service level agreements via slice-aware radio resource management in 5g networks. In *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pages 1–6, 2018.
[37] S. Kim, K. Jung, J. Choi, and Y. Kwak. Improving lte/lte-a ue power efficiency with extended drx cycle. In *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, pages 1–5. IEEE, 2014.
[38] E. Kolta, T. Hatt, and S. Moore. Going green: benchmarking the energy efficiency of mobile. Technical report, GSMA Intelligence, Tech. Rep., 2021.(Cited on pages 4 and 79.).
[39] S. Kumar, E. Hamed, D. Katabi, and L. Erran Li. Lte radio analytics made easy and accessible. *ACM SIGCOMM Computer Communication Review*, 44(4):211–222, 2014.
[40] Y.-W. Kuo, L.-D. Chou, et al. Power saving scheduling scheme for internet of things over lte/lte-advanced networks. *Mobile Information Systems*, 2015, 2015.
[41] M. Lauridsen, A. R. Jensen, and P. Mogensen. Fast control channel decoding for lte ue power saving. In *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2012.

[42] T. Li, L. Yu, Y. Ma, T. Duan, W. Huang, Y. Zhou, D. Jin, Y. Li, and T. Jiang. Carbon emissions and sustainability of launching 5g mobile networks in china. *arXiv preprint arXiv:2306.08337*, 2023.
[43] T. Li, L. Yu, Y. Ma, T. Duan, W. Huang, Y. Zhou, D. Jin, Y. Li, and T. Jiang. Carbon emissions of 5g mobile networks in china. *Nature Sustainability*, pages 1–12, 2023.
[44] W. Li, J. Huang, S. Wang, C. Wu, S. Liu, and J. Wang. An apprenticeship learning approach for adaptive video streaming based on chunk quality and user preference. *IEEE Transactions on Multimedia*, 25:2488–2502, 2023.
[45] Y. Li, C. Peng, Z. Yuan, J. Li, H. Deng, and T. Wang. Mobileinsight: Extracting and analyzing cellular network information on smartphones. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, pages 202–215, 2016.
[46] O. Liberg, M. Sundberg, E. Wang, J. Bergman, and J. Sachs. *Cellular Internet of things: technologies, standards, and performance.* Academic Press, 2017.
[47] K. Liu, J. He, J. Ding, Y. Zhu, and Z. Liu. Base station power model and application for energy efficient lte. In *2013 15th IEEE International Conference on Communication Technology*, pages 86–92. IEEE, 2013.
[48] J. Lorincz, T. Garma, and G. Petrovic. Measurements and modelling of base station power consumption under real traffic loads. *Sensors*, 12(4):4281–4310, 2012.
[49] K. MacMillan, T. Mangla, J. Saxon, and N. Feamster. Measuring the performance and network utilization of popular video conferencing applications. In *Proceedings of the 21st ACM Internet Measurement Conference*, pages 229–244, 2021.
[50] H. Mao, R. Netravali, and M. Alizadeh. Neural adaptive video streaming with pensieve. In *Proceedings of the conference of the ACM special interest group on data communication*, pages 197–210, 2017.
[51] MultiTech. Multitech dongle mtcm2-l6g1-b03-kit.
[52] M. S. Mushtaq, S. Fowler, and A. Mellouk. Power saving model for mobile device and virtual base station in the 5g era. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2017.
[53] P. F. Pérez, C. Fiandrino, and J. Widmer. Characterizing and modeling mobile networks user traffic at millisecond level. In *Proceedings of the 17th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization*, pages 64–71, 2023.
[54] L. Prasad. An energy efficient radio base station. 2018.
[55] F. Qian, Z. Wang, A. Gerber, Z. Mao, S. Sen, and O. Spatscheck. Profiling resource usage for mobile applications: a cross-layer approach. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 321–334, 2011.
[56] Qualcomm-qxdm.
[57] S. Rezaei and X. Liu. Deep learning for encrypted traffic classification: An overview. *IEEE Communications Magazine*, 57(5):76–81, 2019.
[58] Sandvine Inc. Sandvine global internet phenomena report. https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2023/reports/Sandvine%20GIPR%202023.pdf?hsCtaTracking=3cbc04da-fd44-481b-ad03-811d23b7b2c5%7C131df09f-dbdd-41a0-9dce-aac8879403ff, 2023.
[59] T. Sigwele, Y. F. Hu, and M. Susanto. Energy-efficient 5g cloud ran with virtual bbu server consolidation and base station sleeping. *Computer Networks*, 177:107302, 2020.
[60] L. Williams, B. K. Sovacool, and T. J. Foxon. The energy use implications of 5g: Reviewing whole network operational energy, embodied energy, and indirect effects. *Renewable and Sustainable Energy Reviews*, 157:112033, 2022.
[61] X310n. https://www.ettus.com/all-products/x310-kit/.
[62] Y. Xie and K. Jamieson. Ng-scope: Fine-grained telemetry for nextg cellular networks. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 6(1):1–26, 2022.
[63] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli. A control-theoretic approach for dynamic adaptive video streaming over http. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, pages 325–338, 2015.
[64] C. Zhang, M. Fiore, C. Ziemlicki, and P. Patras. Microscope: mobile service traffic decomposition for network slicing as a service. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, MobiCom '20, New York, NY, USA, 2020. Association for Computing Machinery.
[65] L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang. Accurate online power estimation and automatic battery behavior based power model generation for smartphones. In *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*, pages 105–114, 2010.
[66] X. Zhang, P. Ciais, X. Jian, X. Liu, R. Wang, K. Chen, Y. Huang, T. Huang, H. Gao, Y. Zhao, et al. The carbon footprint response to projected base stations of china's 5g mobile network. *Science of The Total Environment*, 870:161906, 2023.