Energy and Cost Efficient Ultra-High Capacity Wireless Access

Sibel Tombaz, Anders Västberg, Jens Zander
Wireless@KTH, Royal Institute of Technology, Stockholm, Sweden

Abstract

Mobile communication networks alone today consume 0.5% of the global energy supply. Meeting the rapidly increasing demand for more capacity in wireless broadband access will further increase the energy consumption. Operators are now facing both investing in denser and denser networks as well as increased energy cost. Traditional design paradigms, based on assumptions of spectrum shortage and high cost base station sites, have produced current cellular systems based on 3G and 4G (LTE) standards. The latter ones are characterized by very high spectrum efficiency, but low energy efficiency. Deployment has favoured strategies with few, high power bases stations with complex antenna systems. The key method for indoor coverage has so far been to literally "blast signals through walls" - a solution that is neither energy efficient, nor very sound from a radiation perspective. As environmental aspects maybe perceived as important from a societal perspective, the cost remains the short to medium term concern for operators of future mobile broadband systems. What becomes evident now is that the so far mostly neglected energy cost will be a major concern. Future system deployment has to balance infrastructure deployment, spectrum and energy cost components.

Ongoing incremental improvements in electronics and signal processing are bringing down the power consumption in base station. However, these improvements are not enough to match the orders-of-magnitude increase in energy consumption cause by demands for more capacity. It is clear that solutions to this problem have to be found at the architectural level, not just by increasing the efficiency of individual components. In this paper we propose a framework for a total cost analysis and survey some recent, more radical, "clean slate" approaches exploiting combinations of new spectrum opportunities, energy efficient PHY-layers and novel deployment and backhauling strategies that target minimizing overall system cost. The latter involve network deployment tightly tailored to traffic requirements, using low power micro base stations tailored specifically to decrease the power consumption compared to today’s high power macro base stations schemes.

To illustrate our findings, a power consumption model for mobile broadband access networks taking into account backhaul is presented and main trade-offs between infrastructure, energy and spectrum cost are analyzed. We will demonstrate optimal deployment strategies in some simple scenarios where a certain capacity has to be provided in a dense, interference limited scenario.


I. INTRODUCTION

Cellular networks of today provide good coverage and service in many countries, both in urban and rural areas. The key challenge for the industry is the rapid proliferation of smartphones, laptops and tablet-PCs with built in cellular access that is rapidly driving the demand for increased capacity. Forecasts range between a 100-fold to a 1000-fold increase in traffic volume before 2020. State-of-the art technology, e.g. LTE, can certainly meet this demand, but it is not obvious that it can be done in a cost effective (e.g. profitable) manner. More capacity in wireless broadband access will require the deployment of several orders of magnitude more base stations which, investment cost and acquisition of new spectrum aside, also will significantly increase the energy consumption. Traditional design paradigms, based on assumptions of spectrum shortage and high cost base station sites, have produced current cellular systems based on 3G and 4G (LTE) standards. These systems are designed for extreme spectrum efficiency. 3G (HSPA) has a spectral efficiency of approximately 10 times that of a 2G (GSM) system while LTE is 3-4 times more efficient than 3G systems. The drawback of these schemes is that they require very linear power amplifier that have to be backed-off to very low energy efficiency. Calculations have been made show that the energy consumption for an LTE network has to increase about 60 times compared to a 2G network to offer the same level of coverage [9]. The key energy problem has been heat dissipation and cooling, but as the fraction of energy cost of the total cost has been comparatively low, this factor has been neglected in traditional designs. An exception is cellular networks in rural areas outside the power grid, has even today up to 50% of the OPEX cost related to energy. The major part of the energy consumption of a cellular network is related to base stations and backhaul.
With the increasing rate of base station deployment, we suggest that energy soon will become more of a limiting factor for the operators also in urban areas. Over the last decade global warming has become important issue on political agenda and various “green” initiatives to reduce power consumption have attracted significant attention. Besides these long-term goals of the cellular industry to reduce carbon emission, there is a strong, shorter term, economical motivation for network operators to decrease the power consumption of the network. As energy becomes a more scarce resource and subject to increased taxes (e.g. carbon dioxide outlet rights), energy prices will increase.

Industry is actively improving equipment based on current standards to reduce the energy consumption of access networks. However, we believe that these improvements, although necessary, are not enough to match the orders-of-magnitude increase in energy consumption cause by demands for more capacity. It is clear that solutions to this problem have to be found at the architectural level, not just by increasing the efficiency of individual components. Instead we should take a “clean slate” approach, asking ourselves how should one design the architecture of a broadband access network that meets user requirements at lowest total cost? How are various cost factors and constraints such as energy, infrastructure investments and spectrum affecting the design?

There are several ongoing projects involving academia and industry partners, that study the energy efficiency of cellular networks. For example EARTH [1], an FP7 European Commission research project, takes a holistic view on the problem, with a goal to reduce the energy consumption of mobile broadband networks by 50%. An overview of the different proposed methods to increase the energy efficiency on all levels of the network including components, links and network levels are given in [2]. Other work focuses on the architectural level. Richter et al. study the energy efficiency in heterogeneous networks for uniformly distributed traffic in [3]. Reductions in energy consumption are achieved by turning transceivers on and off, matching the diurnal variations in traffic demand. Examples of such studies are [4] and [5], where it is claimed that 50% energy saving is possible in low traffic periods without compromising the network performance. Johansson et. al conducted techno-economic studies, focusing on the infrastructure cost of wireless broadband networks [6], [7]. Here different deployment scenarios for heterogeneous networks for inhomogeneous traffic demands were studied. The energy cost, however, is only marginally considered. Chen et al. have in a recent paper [8] combined the problem of energy efficiency and deployment cost. Although the backhaul cost is included in this work when the deployment efficiency is considered, the energy efficiency analysis is restricted to the base stations, not the whole system.

In this article, we propose a high-level framework for comparing various architectures for future wireless broadband access network in terms total cost. We propose a simple power model for broadband wireless network, including the power consumption of backhaul. A detailed cost analysis is completed and impact of each cost factors on the design of future very high capacity wireless access network is studied. The goal of this paper is to overview main trade-offs between energy-infrastructure cost-spectrum for cellular mobile radio networks.

II. MODELLING POWER CONSUMPTION

Where and how much power is consumed in mobile radio networks? We believe that a precise characterization of the power consumption of the network is crucial for any energy efficiency analysis to arrive at valid conclusions. This section thus first reviews some power consumption. We then propose a new power consumption model that is not limited to the base stations but also includes backhaul and network components of the access network. The power consumption will be a function of the total traffic and the total number of base stations.

In the literature, the average power consumption of a base station is modelled as a linear function of average radiated power and is given by [1]

\[ P = aP_{tx} + b. \]

Here \( P \) and \( P_{tx} \) denote the average total power per base station and the power fed to the antenna, respectively. The coefficient \( a \) accounts for the part of the power consumption that is proportional to the transmitted power (e.g. RF amplifier power including feeder losses) while \( b \) denotes the power that is consumed independently of the average transmit power, (e.g. signal processing, site cooling, backhaul etc). The total power consumption of the network is calculated based on the deployment strategy, i.e. type and the number/density of base stations in the network. The backhaul part has been ignored in most previous work. The deployment strategy will indeed affect the implementation of the backhaul and consequently its power consumption. Therefore, we propose a more detailed power consumption model that considers also the backhaul and the network power consumption:
\[ P = N_{BS}(aP_{tx} + b_{radio}) + N_{BS}(b_{backhaul} + y\overline{R}) + d = \\
= N_{BS}(aP_{tx} + b_{radio} + b_{backhaul} + y\overline{R}) + d \] (1)

Here, \( b_{backhaul} \) represents the power consumed by the backhaul transceiver and the uplink interface, \( \overline{R} \) is the total traffic (data rate) per base station. \( N_{BS}\overline{R} \) is the total network traffic and \( y \) denotes the traffic dependent power consumption of the backhaul and network switches. The part of the switch power that is independent of the traffic and hence the number of base stations, is denoted by \( d \). (This latter part is typically assumed to be 80% of the backhaul switch power). It should be noted that in this model there are two power components that are independent of the average transmit power, \( b_{backhaul} \) (backhaul) and \( b_{radio} \). The latter one corresponds to \( b \) in the simplified model above. The numerical values of all these coefficients strongly depend on the actual equipment used and the state-of-the-art in technology. In the following, we will use estimates frequently used in the literature and sensitivity analysis will be performed to analyse the effect of different parameters to the total power consumption.

III. DEPLOYMENT FOR MINIMUM ENERGY

Providing wireless access services energy efficient, i.e. at minimum energy consumption, has been the focus of most work in the field. Two main directions can be seen in the literature: a) reducing the power consumption of the main consumer, e.g. by using more efficient equipment or by using more advanced software to adapt the power consumption to the traffic situation, and b) employing power saving network deployment strategies. In approach b) investigated here, power is saved by carefully adapting the deployment of small, low power base stations to the traffic requirements in various areas. In “hot-spot” areas with large traffic volumes, the density of base stations will be high, whereas in rural areas only few macro base stations are needed to provide coverage. This is a well known strategy to manage spectrum and reduce infrastructure cost in the access network, but which density of base stations should be used to minimize power consumption in each area type is not known. Even though the transmitter power can be reduced in small cells, this is counteracted by the increased number of base stations (cells) and it is not obvious how the total power consumption is minimized. In the following we will first find the base station density (cell size) that minimizes the power consumption for a given service area and quality of service (QoS) requirement. Further we study the trade-offs between power consumption, spectrum allocation, and cell size.

In the downlink direction of wireless communication, the received power at distance, \( d \), from the base station can be modelled by[10]

\[ P_{rx}(d) = \frac{cGP_{tx}}{d^\alpha}, \]

where \( P_{tx} \) is the transmitter power, \( c \) and \( \alpha \) are the path loss coefficient and exponent respectively and \( G \) is the antenna gain. Without considering the interference, signal-to-noise ratio (SNR) is simply written as

\[ \gamma = \frac{cGP_{tx}}{\sigma^2 d^\alpha}. \]

Here, \( \sigma^2 \) is the noise power on total bandwidth (\( \sigma^2 = N_0 W \)) and \( N_0 \) is the power spectral density of additive white Gaussian noise. Under a minimum achievable transmission rate target per base station, \( \overline{R} \), the required transmitter power can be calculated for a worst case scenario (at the cell edge) using Shannon’s formula;

\[ \overline{R} = W \log_2(1 + \gamma_0) \]

where \( W \) is the system bandwidth and \( \gamma_0 \) is the SNR of the cell-edge user. Inserting \( d = R_{cell} \) and solving for the transmit power yields:
We observe from Eq. (2) that if we consider the dependence on the path loss only, the required transmit power for a certain coverage increases with cell range, \( R_{cell} \). To compare the total transmit power in the network we need to multiply the power in (2) with the number of base stations

\[
N_{BS} = \frac{A}{\pi R_{cell}^2},
\]

where \( A \) is the size of the service area. Based on this relationship, the total required transmitter power in access network to satisfy the coverage requirement can be written as a function of number of base stations as

\[
P_{tx,tot} = \left[\frac{\pi}{2^\frac{\alpha}{2}} - 1\right] \frac{N_0 W}{cG} \frac{A}{\pi} N_{BS}^{\frac{\alpha}{2}}.
\]

From the above expression, the total required transmitter power of mobile radio network with a fixed service area is decreasing as we increase the number of base stations in the network for any \( \alpha > 2 \). We can conclude that if we consider only transmit power and power components directly proportional to the transmit power, the cell sizes should be as small as possible, i.e. a high density deployment strategy with many micro base stations is the most energy efficient.

If we want to fix the service requirement as the average total network throughput

\[
\bar{R}_t = N_{BS} \bar{R}.
\]

we can rewrite the above expression as

\[
P_{tx,tot} = \left[\frac{\pi}{2^\frac{\alpha}{2}} - 1\right] \frac{N_0 W}{cG} \frac{A}{\pi} N_{BS}^{\frac{\alpha}{2}}.
\]

Now, returning to the extended power consumption in Eq. (1) we will also consider power components that are not directly related to the transmit power (backhaul, signal-processing, cooling etc). We plug in expression (3) in (1) to get

\[
P_e = \left( N_{BS} \left[ \frac{\pi}{2^\frac{\alpha}{2}} - 1 \right] \left( \frac{A}{\pi N_{BS}^{2-\frac{\alpha}{2}}} \right) + b_{radio} + b_{backhaul} + y \frac{\bar{R}_t}{N_{BS}} \right) / A.
\]

This expression illustrates how the power consumption in the access network is related to the amount of spectrum allocated \( W \) and required number of base station (related to the infrastructure cost). We will also use the average total throughput per unit area

\[
\bar{R}_{area} = \bar{R}_t / A
\]

to measure the performance of the network. It should be noted that for simplicity, we assume a linear capacity growth as the number of base stations increases in our derivations.

We observe in Eq. (4) that for \( \alpha > 2 \),
Figure 1. Area power consumption vs. number of base stations for different system bandwidth.

\[ \overline{P}_c(N_{BS}) \rightarrow \infty \quad \text{as well as} \quad \lim_{N_{BS \to \infty}} \overline{P}_c(N_{BS}) \rightarrow \infty, \]

which means that there is always a non-null and finite \( N_{BS}^* \) that minimizes the area power consumption of the network. This minimum will depend on the required network area throughput, the service area, path loss coefficient and the power consumption coefficients of the network which, in turn, are dependent on used equipment. Fig.1 illustrates the relationship between the area power consumption and the number of base stations with its shallow minimum. The figure also show the significant effect of the amount of spectrum allocated on the area power consumption and the optimum number of base stations. A very small spectrum allocation requires the network to run at a very high spectrum efficiency

\[ S = \frac{\overline{R}_{tot}}{N_{BS} W}, \]

in order to meet the data rate (capacity) requirements. This is the cause of the rapid increase in power consumption in the left hand part of the diagram. Note that we see this increase even with constant RF amplifier efficiency \( a \). As further discussed below, in practice, this efficiency drops with higher values of \( S \), which would make the curves even more steep. It may no come as a surprise, that larger spectrum allocations lowers the total area power consumption and favours designs with fewer base stations. Further the area power consumption increases with the required data rate of cell edge user.

In Fig. 2, we illustrate the effect of the power components that are independent of the transmitter power e.g. \( b_{\text{radio}} \), \( b_{\text{backhaul}} \) etc., using \( b_{\text{radio}} \) as example. In the literature these components are usually referred to as the idle power of the network. For example, when all these idle power components are zero, when the power consumed by cooling, signal processing and backhaul is negligible, the area power consumption is monotonically decreasing with number of base stations. As soon as these components is non-negligible, the functional relationship becomes non-monotonic and we again have a non-null and finite \( N_{BS}^* \) that minimizes the area power consumption of the network.
Another dominant parameter in the power consumption model is \( a \) which accounts for the power consumption that scales with the transmitted power due to RF amplifier and feeder losses and its effect on area power consumption change is shown in Fig. 3. It can be observed that for \( a = 1 \), (100% efficient power amplifier and without any feeder losses) area power consumption is minimized by using long-range base stations. On the other hand, as the \( a \) increases (lower power amplifier efficiency + higher feeder losses), the power minimizing number of base stations shifts to the right, i.e. denser networks are needed.
IV. DEPLOYMENT FOR MINIMUM TOTAL COST

In the previous section we have demonstrated the characteristic of system where energy consumption is minimized, meeting some Quality-of-Service constraints. This is what wireless access systems would look like if energy was the only (or at least dominating) constraint. In practical system design and deployment, wireless operators have to deal several objectives simultaneously. From a strictly commercial perspective, things are quite simple – every operator will seek to maximize their profit and one important part that the operator can control is the total system cost. Let us make the reasonable assumption that the total cost for deploying and operating a wireless access network is dominated by cost for Spectrum (licenses), Energy and the Infrastructure cost. We can write the total cost

\[ C_{\text{tot}} = C_{\text{spectrum}} + C_{\text{infra}} + C_{\text{energy}} \]  

(5)

For sake of simplicity we consider both the annualized capital expenditure (CAPEX) as well as the operational expenditures (OPEX). Fig. 4 shows the relative impact of these cost factors on the design of the networks. Current Macro-cellular systems are represented with point A. Severe spectrum shortage has dominated the design. As spectrum licenses ran into the billions of dollars in early 2000 the main driver in current system design (e.g. LTE) has been spectral efficiency, i.e. squeezing as many bits/s out of a given small spectrum allocation. In the design phase energy was considered as basically a free resource (although high power base station equipment is more costly (which increases the CAPEX) and radiation hazard concerns has put limitations on the maximum power). With the “data explosion” stimulated by near-flat-rate data plans, large capacity improvements have been necessary. As we address this problem and install more and more (micro, pico, femto) base stations, the energy consumption (as shown in the previous section) actually goes down and the (fixed) infrastructure cost increases and will dominate the operator expenditure. Short range communication allows for dense frequency reuse, which relieves the pressure on spectrum and we move down to the left in the “constraint triangle”. Projections for the years to come, point at plenty of new spectrum becoming available, both licensed as well as secondary access spectrum (“white space”), as well as rapidly increasing energy cost. This will move us to point D, where energy and infrastructure costs dominate the design of “5G” systems. In this design paradigm where total cost is minimized, wireless infrastructures will have the following characteristics:

- Tailored deployment - matching deployed capacity carefully to the demand – the projected traffic: “Macro cells” for wide area low-capacity coverage and micro/pico/femto cells with high capacity. Do not waste infrastructure cost nor energy where not needed.
- Energy adaptive PHY layer: Modulation/coding schemes are more optimized for energy efficiency, not spectral efficiency, as spectrum is becoming more available – in particular at short ranges. Adaptive schemes would provide spectral efficiency in coverage cells and energy efficiency in smaller cells to minimize total cost and to adapt to varying spectrum availability.

In the following we will illustrate the above by investigating some designs that are the consequences of total cost minimization. Expanding Eq. (4) with the components of the previous section and adding the infrastructure and spectrum cost yields Eq. (6) below

\[
Cost = c_0 N_{BS} + c_1 \left\{ N_{BS} \left[ \frac{N_o W}{c G} \left( \frac{\pi R_{\text{area}}}{2 N_{BS}} - 1 \right) \left( \frac{A}{\pi N_{BS}} \right)^{\alpha/2} \right] + b_{\text{radio}} + b_{\text{backhaul}} + \frac{R_{\text{tot}}}{N_{BS}} \right\} + d + c_2 W \tag{6}
\]

Here \( c_0 [\text{€/BS}] \) is annual cost per base station which includes the annualized capital expenditures (CAPEX) and the annual expenditures (OPEX) excluding the energy cost. On the other hand, \( c_1 \) is the annual cost of energy (“electricity bill [€/energy unit]) and the \( c_2 \) is the annualized spectrum cost [€/MHz]. In the following we will study the impact of varying the three cost parameters \( c_0, c_1 \) and \( c_2 \) respectively in some interesting constellations.

First of all we investigate the trade-off between energy and infrastructure cost under the assumption that the cost for spectrum is fixed. In the previous section, we concluded that we need denser networks to minimize the power consumption. However, the required cost for additional base stations were not taken into account. Fig. 5 illustrates the various cost components for a given required total area throughput \( \overline{R}_{\text{area}} \). In this example the energy cost is much smaller than the infrastructure cost that increases linearly with the number of base station. The minimum total cost now occurs at a much lower number of base stations than in the energy-only minimization\(^1\).

The impact of increasing infrastructure cost is illustrated in Fig. 6. The infrastructure cost has a similar effect on the optimal number of base station as the idle power in Fig. 2 and the minimum is not that much affected. If we now also consider the spectrum cost, we first note in Fig. 5 that for a given system bandwidth \( W \), the spectrum cost is constant and only provides a level shift of the total cost. The architecture, i.e. the optimal number of base station minimizing the total cost, is not affected. To assess the amount of spectrum an operator should acquire, in Fig. 7 we show the various cost items as function of the system bandwidth \( W \). We note first that in an interference limited (high capacity) scenario, the infrastructure cost is now inversely proportional to the bandwidth \( W \). We simply need half as many base stations to provide the same capacity if we have twice the bandwidth available. From the previous chapter we have seen that the energy cost drops rapidly and for higher bandwidths, the system cost is completely dominated by the infrastructure cost. Again the cost has a distinct minimum that moves closer to the “energy asymptote” as the spectrum cost increase. For very high spectrum costs the infrastructure cost will not affect the optimal bandwidth which is the determined by the energy- and spectrum costs. Fig. 8 finally shows the dependence of the deployment densities on the energy cost for various spectrum availabilities. Here, severe impact of expected energy cost increase on network design can be seen more clearly. Ten-fold increase in electricity cost causes a 20-25% increase in optimum network deployment density.

\(^1\) In order to provide reasonably realistic numbers in the numerical evaluations, the following assumptions have been made regarding the costs based on [6], [8]. The annual cost for a macro base station which serves for a coverage with radius ranging from 500m to a few kilometers equals to 0.02M€/BS, and the annual electricity charge is 876€/kWh. For the spectrum cost, the 4G spectrum auctions in Sweden is taken into account where 210 M€ is paid for 15 years lease of 190 MHz in the 2.6 GHz band.
Figure 5. Cost vs. number of base station. ($\bar{R}_\text{area}=20\text{Mbps/km}^2$, $W=5\text{MHz}$, and $A=20\text{km}^2$)

Figure 6. The impact of increased infra cost. ($\bar{R}_\text{area}=20\text{Mbps/km}^2$, $W=5\text{MHz}$ and $A=20\text{km}^2$)
Figure 7. Total network cost vs. bandwidth. Here number of base stations is optimized to minimize the total cost $(\bar{R}_{area}=20\text{Mbps/km}^2$ and $A=20\text{km}^2$).

Figure 8. The impact of increased energy cost to optimum required number of base station for different system bandwidth. $(\bar{R}_{area}=20\text{Mbps/km}^2$ and $A=20\text{km}^2$)
IV. CONCLUSION AND FUTURE WORK

In this paper we have analysed some of the design constraints for future very high capacity wireless access systems and their impact on the system architecture. Whereas traditional mobile systems, primarily have been constrained by the available bandwidth, high capacity data systems are increasingly constrained by energy and infrastructure costs. This will be reflected in the design of future clean-slate approaches to “5G”-systems. Using simple models for these cost items, we can determine the main characteristics of such future infrastructures and how they depend on the various cost items. Exact results can only be determined when absolute cost figures are available. However, some fundamental conclusions can be drawn:

- Both the energy cost and the total access network cost are heavily dependent on the number of base stations (transmission range). When the energy cost is high, the total cost is minimized for dense base station deployments.
- For high density deployments, the idle power of the base stations and backhaul will become a significant factor.
- The energy cost is also strongly dependent of the amount of available spectrum. Significant cost savings, both in energy and in infrastructure cost, can be made in total cost if more spectrum can be made available.

Future work includes studying network deployment scenarios. This includes energy cost trade-off analysis of different heterogeneous networks, network optimization for non-uniform traffic scenarios and analysis of energy/cost reduction by using relays. Investigating how dynamic network control can minimize energy consumption has been done for average daily or weekly distributions. It could be interesting to study this for different specific traffic models (office, home users etc).

ACKNOWLEDGEMENT

The authors would like to acknowledge the contributions of Mats Nilsson and other colleagues in the eWIN project at KTH and ACREO.

REFERENCES