MIT Communications Futures Program

Mobile Broadband:

Toward a Sustainable Ecosystem

Mobile Broadband Working Group
William Lehr, Chair
Marie-Jose Montpetit, Editor
Computer Science and Artificial Intelligence Lab (CSAIL)

DISCLAIMER

This White Paper is the product of multiple contributors. Any views or opinions presented herein are solely those of a subset of the contributors and do not necessarily represent the views of their employers or consensus views of all contributors. See Appendix 1 for list of contributors.

MITSloan MANAGEMENT

CSAIL

MIT Media Lab

Massachusetts Institute of Technology

Executive Summary

The growth of mobile broadband presents the entire Internet ecosystem with important challenges and opportunities. It holds the promise of always-on/everywhere access to services, applications, and content. The realization of this vision has the potential to redefine the value proposition for broadband services, but not without challenging legacy business and policy models.

Meeting the needs of exponential traffic growth requires significant investments in new capacity all along the value-chain. More capable devices, more interactive applications, richer multimedia, and faster networks are all needed to enable new and better services. Users across the economy have grown accustomed to continuous improvements in the range and performance of broadband Internet enabled services.

While traffic has been growing exponentially, revenues have not. To insure efficient investment continues and resources are not wasted, we need to better understand the relationship between traffic and usage costs, and how best to manage and recover those costs. While this challenge existed in the earlier world of fixed broadband services, the growth of mobile broadband adds complexity. It is clear that responding to these challenges will require a multiplicity of solutions engaging participants all along the value chain, from application/content providers to fixed and mobile network operators, and from end-users to policymakers.

This paper explains how the growth of mobile broadband adds complexity to the broadband Internet investment challenge, explores some of the ways in which this challenge may be addressed, and identifies areas where further research is needed.

1.1. Scope and Goals

Enabling broadband mobility necessarily implies greatly expanding traditional cellular networking as well as Wi-Fi and other forms of wireless networking. It will require investment and innovation in all parts of the broadband Internet value chain, including:

- Access and transit networks for both fixed and mobile provider networks to support seamless connectivity and provide an exponential response to the exponential rise in mobile broadband traffic;

- Content and applications that can take advantage of the expanded devices, networks and access capabilities while striving for more efficient energy and bandwidth consumption;
- Consumer and edge-network devices with on-board computing, storage, and multiple radio access with expanding capabilities becoming readily available;
- Customer, operators, and content creators developing new business relationships that foster innovation especially in the services and applications areas;
- Policy frameworks to promote competition while protecting universal access to essential infrastructure, and ensuring security and privacy.
- The growth of cloud infrastructure and services, which will affect the cost of implementing and operating networks, and will create opportunities for new business models and services. Clouds offer a means to share costly infrastructure and reduce energy consumption, and provide more robust support for thin-client devices.

The purpose of this white paper is to define what these innovations and investments entail for the broadband industry as a whole.

Specifically the goals of paper are to:

- Articulate our vision of the mobile broadband future as an essential element of economic growth, a generator of infrastructure challenges and a reflection of societal changes;
- Identify key trends driving this future, including new costing/pricing model and low cost access networks;
- Identify the important economic and technological challenges;
- Frame the discussion around the mobile broadband growth phenomenon and its challenges.

Some of the identified challenges relate to pricing and traffic management models. Others relate to changes in industry structure, business relationships, communications regulation, and others to the technological requirements for ubiquitous content distribution and consumption. Finally the relationship between traditional operators and their customers will be seen to be changing when Wi-Fi and cellular both complement and compete with each other.

1.2. Intended Audience

The intended audience for this paper includes but is not limited to:

- Business decision makers in the Internet supply chain, including content/application providers, equipment and software vendors, ISPs and other network service providers;
- User/consumer advocates and enterprise customers;
- Policymakers including regulatory authorities;
- Researchers and analysts from both academia and industry concerned with technical, economic, and legal/policy issues;
- Application developers

The issues we attempt to address in this paper exist at the intersection of technology, economics, and policy. To make this accessible to readers with diverse backgrounds, we have included some tutorial discussion.

Contents

2. Introduction

This white paper is concerned with the role of mobile broadband in the overall Internet ecosystem and how it is helping to shape and is shaped in turn by the continuously evolving technical, market and regulatory environment. It is generally recognized that broadband is essential basic infrastructure for our society and economy, and the rise of mobile broadband is changing how this service is provisioned, used and managed.[1]

Internet traffic, and more recently mobile Internet traffic, continues to grow exponentially. That growth creates a need for an exponential expansion in capacity, which in turn fuels additional growth in traffic as new capacity enables new applications, such as richer multimedia. As the market of users and uses grows larger the Internet becomes ever more attractive as a platform for social and economic engagement for businesses across the economy. Dealing with exponential Internet traffic growth requires significant new investment in both fixed and wired access networks.[2] Additionally, because most mobile access is wireless, the growth of mobile broadband helps drive demand for expanded access to scarce radio frequency spectrum resources. Since broadband is viewed as essential basic infrastructure and because it is an input to so much activity across the economy, there is a desire and need to ensure that access remains affordable, both to meet universal service goals and to avoid choking off growth of nascent applications (e.g., location-based data services). For mobile networking, strategies to cope with growing demand include expanded deployment of Wi-Fi and smaller-coverage-area cellular base stations. These trends drive the convergence of wired and wireless networking.

Investment in mobile wireless infrastructure enables the broadband Internet to support everywhere, always available, interactive, rich data services. This includes legacy services such as mobile telephony and text messaging, but increasingly also newer services such as mobile streaming video, social networking (Facebook, Twitter), and still newer emerging services such as machine-to-machine ("m2m") or "Internet of Things" automation.[3] Increased penetration of more capable devices such as smart phones and tablets with higher resolution screens, faster processors, and more on-board memory enable new applications and services that are more resource intensive (power, bandwidth, computing cycles) and that make use of higher-resolution, interactive multimedia to provide a richer and more compelling user experience. Video replaces text, richer and more interactive video replaces lower resolution video; video teleconferencing replaces voice teleconferencing; and picture messaging replaces text messaging. Thus, the investment in devices, applications, and network infrastructure are co-dependent and fuel a virtuous cycle of demand and traffic growth. As bottlenecks in edge devices, networks, applications or content are uncorked, new capabilities are enabled that fuel new demand. The increased traffic resulting from this demand creates new potential capacity bottlenecks. Only if investment all along the value chain keeps pace with the traffic growth can this virtuous cycle of increasing economic activity and value creation continue. A failure to invest adequately at any point poses a threat to the whole system, but incentives to invest depend on the investors having a reasonable expectation of recovering their costs.[4] As the markets and technologies evolve, so also must the business models and policy frameworks. This evolution will require adjustments and responses all along the Internet value chain, from mobile to fixed broadband providers, from equipment vendors to application developers, and from cloud service providers to end-users.

Broadband Internet traffic is growing exponentially at 50% CAGR (Compound Annual Growth Rate). At the same time, traditional sources of revenue for fixed and mobile broadband access providers are growing more slowly, or actually decreasing. For example, as broadband access (either fixed or mobile) approaches 100% of the population, revenue from new subscribers stagnates, and for the industry as a whole, disappears.[5] Legacy services such as cable television, fixed and mobile telephony, and text messaging are increasingly confronting lower priced or higher quality alternatives that are siphoning off demand and margin contribution. Historically, these services have been priced significantly above incremental cost in order to sustain a significant contribution to the recovery of fixed and shared costs. To the extent these revenue sources are threatened, IAPs and ISPs are looking to new sources of revenue to fuel the virtuous cycle. At the same time, everyone is looking to new technologies, architectures, and business models to reduce costs.

Obviously, no one in the value chain wants to bear more than a fair share of the investment/cost recovery burden, or to cede more than a fair share of the profits to be captured from value creation. Indeed, competition induces everyone to strive to capture as large a share as possible. To have an informed dialog about where investment is needed and how to apportion the costs and benefits of such new investment, we need to understand the overall ecosystem, the forces shaping it, and the options for responding to those forces. This includes having a more granular understanding of traffic growth trends (what applications and uses are adding to capacity bottleneck pressures? How is usage distributed across users? Who is in the best position to address the bottleneck pressures?). Not all traffic is equally valuable or costly to address. Thus, we also need to understand better what the costs are of different traffic scenarios, how those costs impact value chain participants, and what the relative trade-offs may be from managing those costs with different technical or business solutions. Finally, we need to explore business models, including pricing options, and policy responses that might help facilitate and sustain both the necessary investment and efficient competition.

In considering potential futures, especially ones premised on the continuation of exponential growth, we must remember that such growth is endogenously determined. If prices/costs are too high, if too little investment occurs at any point along the value chain from services to devices to networks, if spectrum resources are too scarce, or if inappropriate regulatory policies constrain market growth, the forecasted demand and traffic will fail to materialize.

The framing of this future will result in competitive tussles at all levels within the value chain: among ISPs and among application/content providers for retail customers; between ISPs and application/content providers for value shares; and across technologies and business model solutions to the investment challenges. As an example, and as a recurring theme in this paper, we consider the way in which wireless access is influenced by the competition and potential convergence of the Wi-Fi and cellular (increasingly LTE) wireless ecosystems. Mobile broadband growth has been enabled by the expansion of both cellular 3G/4G (LTE) and Wi-Fi infrastructures, but each has been associated with very different business models and cost structures. The cellular operators have relied on expensive licensed spectrum resources and wide-area coverage networks that are necessary to support high-speed seamless mobility (telephony in cars). On the other hand, Wi-Fi has grown mostly as a way to expand local access to fixed broadband access services within buildings and homes, and more recently in public hot spots. Wi-Fi operates in less expensive unlicensed spectrum.[6] As 3G/4G operators migrate to smaller sized cells (in part, to more efficiently utilize scarce spectrum resources) and Wi-Fi base stations proliferate outside of homes, both become more dependent on fixed broadband infrastructure to backhaul base station traffic.

Wi-Fi and 3G/4G technologies are simultaneously complements and substitutes. For operators the rise of Wi-Fi provides an opportunity to off-load mobile traffic from more expensive (in terms of spectrum resources and to the operator) cellular networks to customer or third-party provided Wi-Fi base stations. While this may reduce the total costs of supporting mobile traffic, it complicates and potentially threatens the cellular providers' relationship with its mobile subscribers. Elsewhere in the value chain similar tussles over business models and technical designs have surfaced. For example, the question of whether it is better to provision more functionality within the network ("clouds") or at the edges fuels the debate over the role for in-network services. While in-network services represent new sources of revenue for the operators, they threaten the end-to-end, or "dumb pipe" model favored by the Over the Top (OTT) applications and services. For example, legacy cable television revenue is threatened by the growth of OTT video. ISPs are responding with their own OTT offerings, and considering new architectures as resources shift from legacy video distribution to Internet-based platforms.

3. Key Drivers Shaping the Mobile Broadband Future

In this section we discuss three key economic factors that shape the future: *growth* of and changes in the nature of the demand; *cost* of the infrastructure leading to user pricing models; and sources of *revenue*. These trends are altering business relationships and helping restructure the Internet value chain.

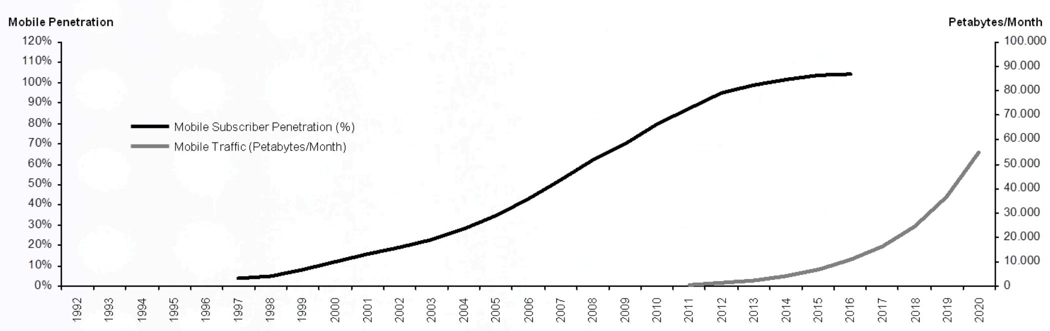
We start by providing a short overview of the growing mobile broadband landscape. We explore user behaviors and emerging applications in order to paint a complete picture of the current state of mobile broadband as well as a number of avenues for the future.

3.1. Growth in Demand

Internet traffic of all kinds and mobile broadband traffic in particular is growing exponentially, and in turn this growth both demands and drives investment. But linking growth and investment into a foregone conclusion is ill advised. Growth depends on pricing models and frictions in the business models: too high prices will throttle demand and lead to under investment in complementary goods; too low prices will encourage excess demand (wasting resources) and deter investment because of the threat to adequate cost recovery. In the remainder of this section, we discuss how patterns of demand have been changing, and in following subsections address how this impacts costs and revenue prospects.

The global mobile communications industry has seen rapid growth in penetration with the addition of significant numbers of new subscribers adopting services each year (see Figure 1). With less expensive devices and mobile tariffs, and more applications and content to choose from, and with a larger number of subscribers to share with (positive network externalities), penetration has expanded to beyond 100% in many markets.[7] In recent years, however, subscriber growth has slowed down in more mature mobile markets.[8] In spite of the slow down in subscriber growth, traffic continues to grow exponentially as per-subscriber usage for all usage tiers continues to grow. More capable devices, faster networks, and more resource intensive applications and content, as well as expanded consumer digital (mobile) literacy help drive per subscriber and aggregate traffic growth.

Figure 1 Mobile penetration and the growth of traffic[9]



Skipping PCs and going to Mobiles

The rise in mobile traffic is fueled by consumer behavior. 75% of teenagers now use smartphones for all their communication and computing needs, and yesterday's teenagers are today's young professionals. Retiring baby boomers are increasingly consuming their digital services on tablets and other mobile devices. And for those who have shied away from connectivity altogether (mobile or fixed) either because they are intimidated or simply see no value in it, it's easier and/or more compelling to leapfrog directly to mobile devices. There is plenty of anecdotal evidence that e-readers and tablets serve as a "gateway" technology for the elderly, and smartphones seem like a more natural evolution from portable or feature phones.

Already, in many developing countries wireless has already surpassed wired traffic because of technology leapfrogging. Consumers are skipping the wired phase of desktops and laptops and going directly to smartphones.[10] As can be seen in Figure 2, growth was exponential between 2008 and 2012, at which point mobile traffic represented 12% of *all* Internet traffic. (This highlights the fact that mobile traffic is growing more rapidly than other Internet traffic, but remains a small portion of total traffic. The growth of mobile broadband does not mean the end of fixed broadband or wired infrastructures, but it does change the relationship between the two both in terms of how networks are designed/managed and how consumers may use and perceive the services.)

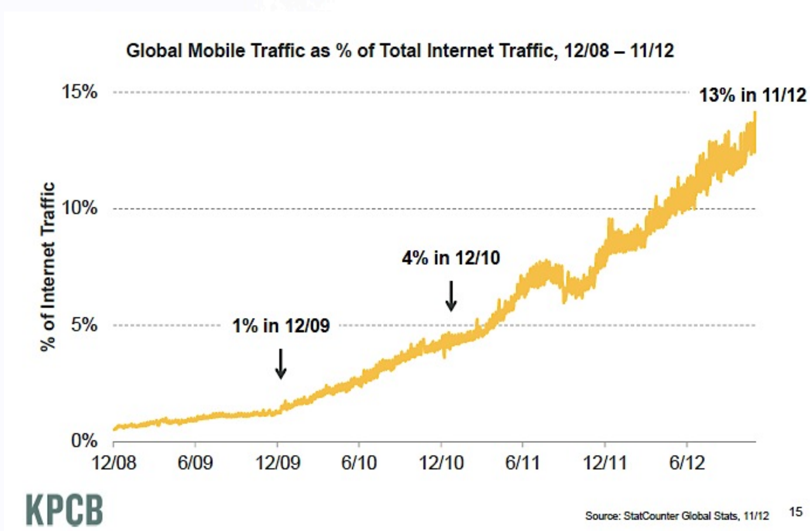


Figure 2 Mobile vs. Fixed Access (2008-2012)[11]

Figure 3 illustrates how changing usage behavior is shifting the mix of mobile traffic and the role that data services (non-voice telephony) are playing in driving growth. While voice telephony traffic remains approximately level,[12] more resource intensive services such as video and interactive applications (e.g., gaming, social networking, mobile file sharing) on more capable devices (tablets and mobile PCs, and smartphones) are accounting for significant traffic growth.

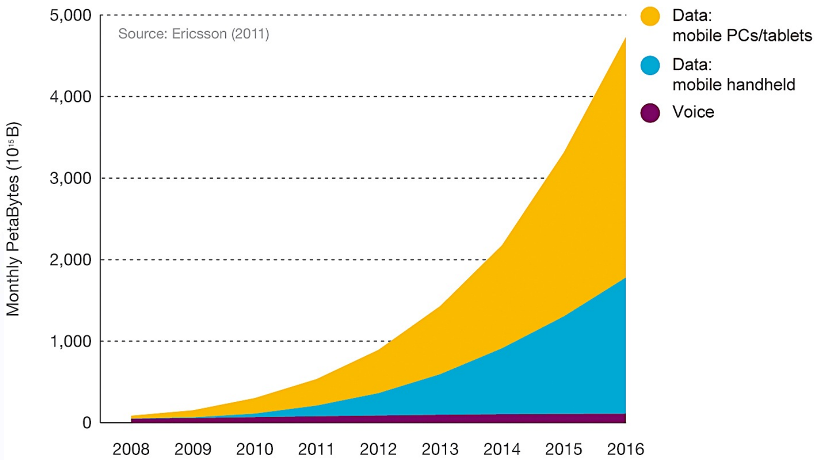


Figure 3 Voice vs. Mobile Data Traffic (2008-2016)[13]

In Figure 4, traffic volumes for 2011 are shown per device type. It is interesting to consider that whatever the platform, the traffic patterns are fairly similar across all devices, and dominated by video streaming and web access.

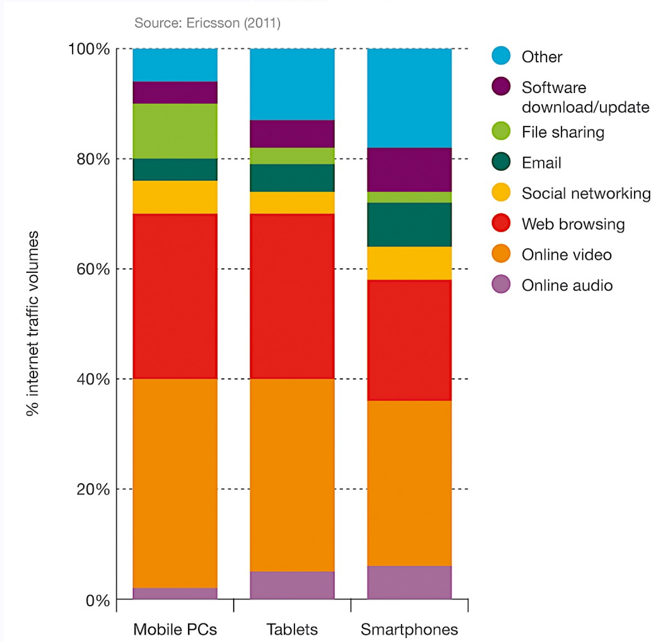


Figure 4 Traffic per Device Type 2011[7]

Figure 4 also highlights how social networking and Internet video are becoming increasingly important on mobile platforms.

Figure 5 provides forecasts for traffic growth per device type for the period 2012-2017 from the CISCO Virtual Network Index (VNI).[14] The latest report shows that mobile data traffic grew 81% in 2013, and the current forecast is for mobile data to grow at a compound annual growth rate (CAGR) of 61% during 2014-2018. (This represents a reduction from the 66% CAGR previously predicted for 2012-2017).

> **The Growth of M2M Traffic**
>
> *Machine-to-Machine (M2M) mobile applications are multiplying, including for mapping, lifestyle, payment, health monitoring, smart watches, and home security. Also, the rise of vehicular networks and smart cars is expected to continue this trend. By 2020, each of us may be carrying hundreds of sensors that will both directly and indirectly contribute to generating traffic. New technologies such as Google Glass and other wearable devices could be significant in the rise of M2M traffic.*

The most recent VNI report confirms some previously noted trends.

- By 2018, global mobile data traffic will reach 15.9 exabytes per month or 190 exabytes annually.
- Smartphones will represent 66% of total mobile data traffic in 2018, compared to 62% in 2013.
- 4G connections will be 15% of the total mobile connections in 2018 and will account for 51% of mobile data traffic.
- Globally, the average mobile network connection speed increased 2.6-fold in 2013 (1.4 Mbps) and will nearly double by 2018, reaching 2.5 Mbps.
- 52% of global mobile data traffic will be offloaded in 2018, up from 45% in 2013; this confirms the complementarity of fixed and mobile infrastructure.
- By 2018, 69% of the world's mobile data traffic will be video, up from 53% in 2013.
- The Middle East and Africa will have the strongest mobile data traffic growth of any region over the forecast period, with a 70% CAGR, followed by Central and Eastern Europe at 68% and Asia Pacific at 67%.

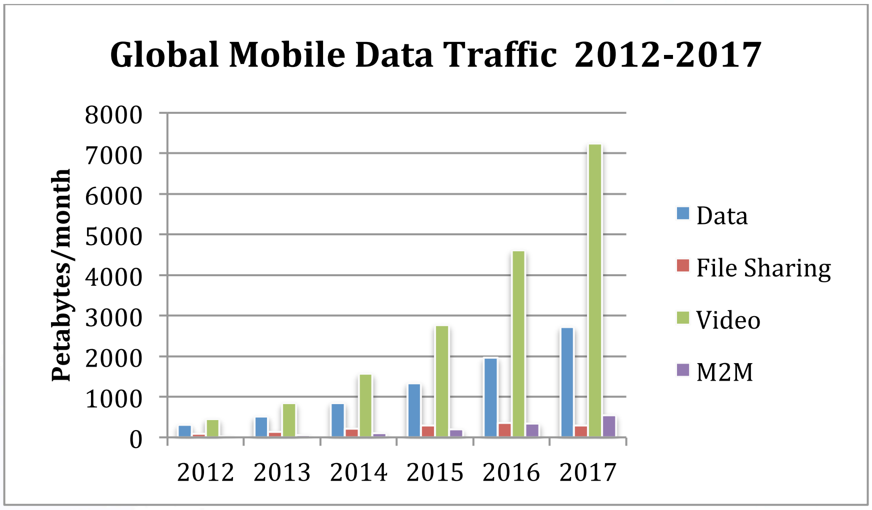


Figure 5 Growth of Mobile Traffic 2012-2017 (Cisco VNI)

As noted earlier, these forecasts presume that capacity will continue to expand and prices will not rise so as to choke off the demand growth. Meeting the growth in mobile traffic will require significant investments in expanding capacity across the Internet ecosystem. A significant share of this investment will include expanding the capacity of wireless access networks, which means deploying a larger number of Wi-Fi and cellular base stations. We can expect cellular operators to increase their deployment of pico cells and perhaps femtocells (smaller coverage area cell sites), and third parties including end-users to expand Wi-Fi-based hotspot access options.

Mobile Video Consumption

TV/video usage on tablets and smartphones continues to grow, where third-party applications dominate and the concept of a program as an "app" is growing. The use of Wi-Fi for video is now entrenched and is disrupting legacy models for home video distribution. New services such as Aereo that make over-the-air programming available on tablets and smartphones are further roiling markets for video entertainment services. Legacy television providers are responding with their own offers of on-line mobile programming access with services such as Comcast's XFinity or HBOGo.

3.2. Costs

The growth in traffic described in the previous section has important implications for the cost drivers for operators. On the one hand, significant continuing investment in capacity expansion is needed, but the relationship between traffic-related costs and aggregate investment expenditures is unclear. Parts of this investment benefit from Moore's Law-like technical progress that continues to drive down the costs of processors, storage, and other network components. Other capacity-related cost components appear to be on slower trajectories such as the labor-related costs of site preparation and installing new smaller-cell site infrastructure, or regulatory compliance costs. Spectrum scarcity may be adding to spectrum resource costs, and growing concerns over security, reliability, or privacy may be adding new cost components to the total costs of supporting mobile services. While most analysts seem to agree that unit-capacity costs are declining, the cost-declines appear to be significantly slower than Moore's Law and the exponential growth in capacity requirements suggest that aggregate costs are increasing. However, precise estimates or consensus on mobile costs are elusive and better insight into mobile costs is needed.

Broadband provider costs are comprised of multiple components, including but not limited to:

- Capital equipment and construction costs;
- Infrastructure operational costs (e.g., power, cooling, real estate leases);
- Operational staffing;
- Wholesale transfer costs;
- Spectrum licenses costs.

While innovation, scale economies, and competition have generally led to a reduction in the average cost per gigabyte (GB) of capacity, the benefits of such improvements have been realized inconsistently across service providers, cost components, and market segments. Some actual estimates of the cost of usage would be helpful in understanding the tensions in the ecosystem, but numerical estimates are hard to obtain, both because operators view their cost structure as proprietary, and because legitimate accounting choices can lump different costs into the “usage-related” basket. Consideration of the list above will suggest the difficulty of deriving a precise cost of usage.

A 2010 consultancy report estimated that the costs of usage (and the necessary investment) for fixed termination were approximately 0.05E/GB, and for mobile termination, were approximately 3.0 E/GB.[15] Other reports give a lower cost per GB for mobile, but there seems to be at least a 10-fold cost difference between the price of fixed broadband (perhaps $0.10 per GB in the U.S.) and the price of cellular (perhaps $1 per GB). This estimate, even if imprecise, helps to explain the advantage of deployment of Wi-Fi as both a complement and a replacement for cellular broadband access. It also suggests that unless the cellular industry can greatly drive down the cost of usage, cellular service will not be a close substitute for fixed service, except for users who make no use of streaming content or other high-traffic volume services. Another way to compare the cost structure of the industries is the claim that for fixed terminations, Internet connectivity is only 20% of CAPEX in 2008/09 data, whereas it was 49% for mobile in the same period.

Changing market dynamics and industry restructuring impact both the incidence and recovery of costs. For example, handset subsidies have historically played a significant role in cellular business models, but may be less important in the future.[16] The rise of Mobile Virtual Network Operators (MVNOs) alters sales and marketing costs, and changing conditions in Internet connectivity/industry structure are altering network termination/interconnection/roaming costs.

Part of the challenge in estimating broadband costs, and especially for mobile, is due to the fact that infrastructure costs are significantly determined by peak capacity needs, and thus are largely fixed. Mobile base stations have to be installed to ensure coverage, even when utilization is low or only temporarily high ("peak hour"). Moreover, while it is reasonable to assign the fixed cost of a fixed broadband subscriber line to the household it serves, per-subscriber allocation of mobile base station costs is less straightforward because subscribers are moving and thereby consuming traffic dynamically at different points. Exceptions are public Wi-Fi hotspots: they tend to have a cost profile similar to mobile broadband access, however, at the time of writing, public Wi-Fi is far from being able to support the near ubiquitous geographical coverage offered by traditional cellular, although this is rapidly changing.

Furthermore, the costs of mobile usage will depend on the mix of traffic. For example, whether the traffic is symmetric with respect to upstream and downstream, whether it is real-time or may be delayed, whether it is cost-effectively cacheable,[17] and/or whether the traffic requires special handling (e.g., to meet special security, privacy, or regulatory requirements).[18] In the past, the growth in subscriptions (access lines) and the expansion of coverage were closely correlated with the growth in total investment and operating costs for broadband providers. As subscription penetration and coverage saturate, traffic-related drivers for investment and operating costs become more important. As users adopt multiple and more capable devices and shift to richer media services it becomes increasingly important to track trends in GB usage and the mix of traffic[19] in order to forecast investment and operating cost requirements.

The changing nature of broadband traffic and the corresponding infrastructure requirements, along with the nature of the costs, together pose a difficult challenge for cost-recovery and optimal pricing. For the investment to be economically viable, prices should exceed incremental costs, and on average, revenues need to exceed total costs. Because of the significant share of total costs that are fixed and or shared, this means that some services need to be priced significantly above incremental costs in order for revenues to be sufficient to recover total costs. Pure marginal cost pricing would fail to recover average costs. Some argue that since peak traffic is what drives capacity investments, peak traffic ought to bear the principal burden for recovering capacity costs such that off-peak traffic would be free, or substantially less expensive. While there is merit to the argument, it begs the question of how to define "peak" and distinguish which costs are truly peak-related.

In the race to expand wireless access capacity, it is unclear who should make the investments in the larger number of lower-powered small cells that are needed; what is the best technology for investment (Wi-Fi v. LTE, fiber etc.); and how to control/manage investment. Traditionally, most of the investment in Wi-Fi base stations has been by end-users (mass market consumers and businesses), whereas most of the investment in cellular base stations has been by cellular operators. Wi-Fi has operated in unlicensed spectrum, whereas cellular has operated in more expensive, licensed spectrum. Wi-Fi is connected via the end-users' fixed broadband service, whereas the cellular base stations are connected via wired facilities managed directly by the cellular operator. Going forward, fixed and mobile operators may play a bigger role in deploying Wi-Fi base stations, while end-users may be induced to set up cellular femtocells to extend in-building/in-home reach of cellular networks. The move to smaller cell architectures raises the potential for shifting an increasing share of mobile broadband infrastructure costs, and potentially, control from operators to end-users, but whether this will happen is uncertain. It may depend on pricing, market and regulatory policy trends.

Finally, the overall system costs and the costs of specific factors depend, in part, on system-wide behaviors. For example, the cost of spectrum access depends in large measure on regulatory policy, which impacts the supply and hence relative scarcity of spectrum. However, it also depends on the wireless technology and business models employed to make use of the spectrum. The choice and pace with which cost-saving innovations are adopted depends on the expectations of others' behavior because scale, scope, and learning economies are so important.

3.3. Pricing Models and Revenue Sources

The growth in traffic helps motivate the need to invest in capacity expansion, which impacts investment and operating costs. These, in turn, impose constraints on pricing, which in turn affects demand growth. Thus, there is a close but complex relationship between traffic/demand growth, investment costs and network provisioning, and pricing and business models. In the following sub-sections, we review trends and forces shaping pricing and revenue models.

3.3.1. Pricing

In the U.S., most wireline broadband access service is priced as a flat monthly charge for cumulative use of the network resources by speed tier. That is, users pay for a flat monthly fee for a nominal peak data rate and with an aggregate traffic volume that is less than some volume cap (expressed in GB per month). Furthermore, in the U.S., the volume caps for fixed broadband subscribes have been sufficiently generous as to constrain only a small subset of heavy users. The volume caps employed for fixed broadband in Canada, Europe, and other markets have been tighter and more likely to impact behavior and usage charges.

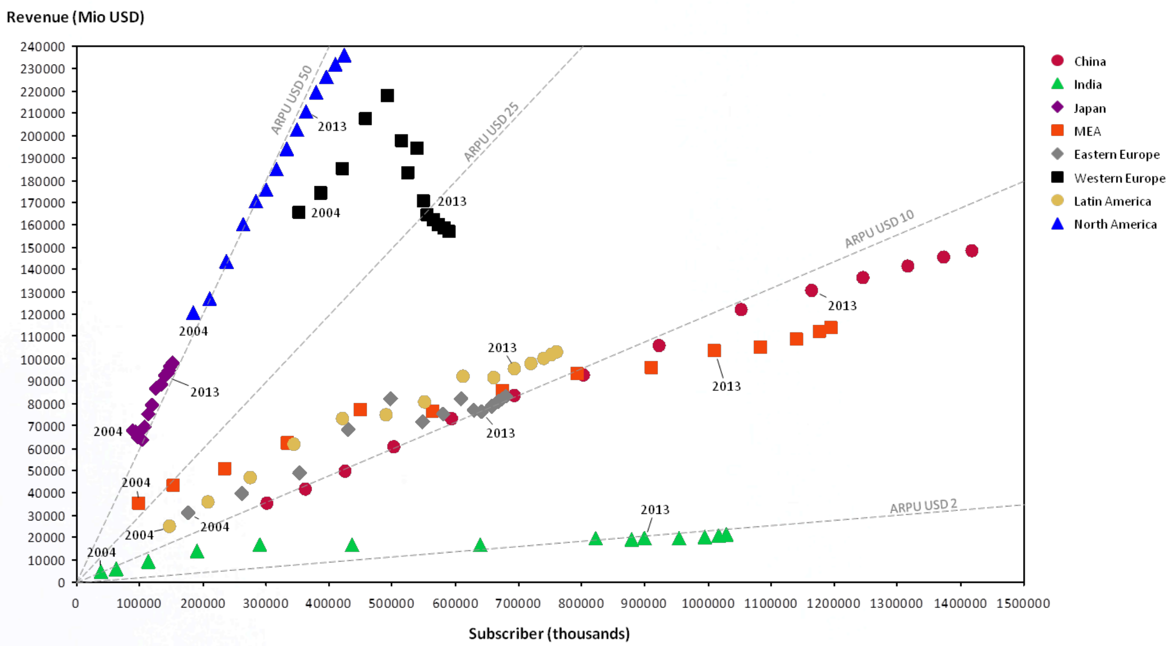
Volume-based usage rate elements have been more prevalent for mobile broadband service pricing. Cellular operators have been shifting to usage-tier pricing where service is priced on the basis of a flat connection charge plus a usage-related charge that is keyed to the size of the usage-bucket in terms of GB. Services are less likely to be differentiated in terms of nominal data speeds as with fixed broadband. Instead, data rates are more likely to vary as a consequence of the user's device (3G vs. 4G handset) or local network conditions (coverage, carrier, congestion). The move to GB tiered pricing for cellular would seem to align revenues with usage costs, but at the same time, it may inhibit the growth of the cellular market, and lead users to take explicit steps to avoid using the cellular system. The trend of using Wi-Fi as a means to offload traffic from the cellular system can be seen, in part, as a consequence of the higher usage-related pricing for cellular service (as well as other factors, such as performance). Whether cellular operators view Wi-Fi as a competitor or complement may be a matter of perspective. On the one hand, the ability to offload cellular traffic to Wi-Fi may facilitate faster growth in mobile broadband subscriptions and aggregate usage that might otherwise be suppressed by cellular network capacity constraints and higher-usage-related end-user pricing. On the other hand, the shift from cellular to Wi-Fi might lessen the cellular operator's ability to manage its subscribers and its cellular network. Figuring out precisely what share of Wi-Fi traffic is traffic that would otherwise have ended up on the cellular network or is just fixed broadband traffic that relies on Wi-Fi for local mobility (e.g., within the home or business) is not easy and is subject to significant disagreements.[20]

The connection between usage pricing—whether based on traffic volumes (over the month), use of specific data services (e.g., SMS versus other data), or time-of-day (weekend or off-peak discounts)—is only tenuously connected to peak-provisioning and cost causation. However, as already noted, since the allocation of costs is ambiguous, there is no general agreement on how usage-costs should be recovered via usage-sensitive pricing.

Historically, the question of optimal usage-based pricing for data was less of an issue because the principal focus was on the pricing for basic telephony and text messaging services. Competition for telephony services from texting and Internet calling and other OTT services threatens legacy broadband provider revenue models. At the same time, a shrinking share of incremental investment is required to meet the on-going needs of legacy voice and text messaging traffic. Service providers have responded to these challenges in multiple ways. For example, a growing share of users purchases communication services as a bundle that includes telephone, text messaging, and broadband. Fixed bundled offerings now can include unlimited local and long distance telephony, basic television programming, and broadband at a given speed tier. Mobile broadband bundles are typically sold as a fixed bundle that includes a telephone, text messaging, and aggregate-usage of data traffic, with special fees for content. And mobile-fixed bundles are also becoming prevalent with services. This pricing strategy enables operators to implement multi-part tariffs, which may be effectively used to implement price discrimination strategies, so long as competitive market conditions and regulatory policies permit such pricing. Whether such pricing is welfare enhancing depends on the circumstances, but it is inappropriate to presume that price discrimination is always contrary to consumer interests. Indeed, price discrimination may be necessary in order to efficiently recover total costs. However, identifying beneficial from harmful price discrimination practices remains a contentious and difficult issue to address.

3.3.2. Global Price Levels

Price levels (average revenue per user or ARPU) differ between countries (see Figure 6). North America and Japan seem to have been able to sustain relatively high ARPU levels, whereas ARPU levels in Western Europe are lower and have declined over time. Some have attributed the European price trends to the increase in competition from MVNOs. Coupled with slowing revenue growth (due in part to saturation of subscriptions), operators in Western Europe might be facing the prospect of a decline in aggregate revenues as well. In contrast, while ARPU levels are low in India and China, penetration rates are also low, and so subscription growth provides ample room for significant revenue growth.[21]



Source: Global Mobile Yankee Group Sep 2013, dynamic currency exchange

Figure 6 Global Revenue per Subscribers in 2013. Source: Yankee Group's Global Mobile Forecast, September 2013 (dynamic currency exchange).[22]

Higher operator revenue leaves more money available for investing in broadband infrastructure, but is only part of the picture (Figure 7). Other factors impacting investment incentives include:

- Level of competition, which, for example, impacts uncertainty about capacity utilization (market shares) and non-network expenses (sales & marketing);
- Substitutes for Internet access;
- Available broadband backhaul (sunk cost of optical fiber roll out for example);
- Country GDP (per capita income differentials).

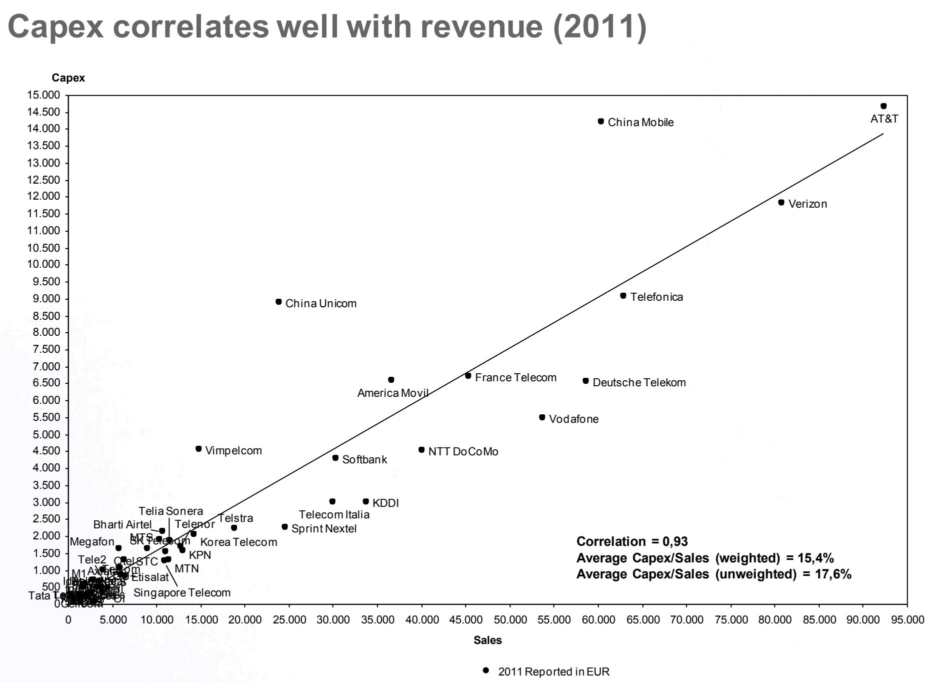


Figure 7 Capex Correlations with Revenue (2011)

3.4. Revenues

Historically, voice, SMS and television (basic and pay-per-view cable services) subscription and service pricing have significantly contributed to recovering the total costs of mobile and fixed broadband service providers. As long as these legacy services continued to pick up a major share of the total fixed and shared costs of broadband infrastructure, broadband usage-related prices could be kept relatively low. However, prices for these legacy services have been falling for decades and competition from alternatives have become increasingly important. Over-the-top (OTT) applications for voice telephony (e.g., Voice-over-IP, VoIP), SMS, or video/music programming are siphoning off legacy revenues.

Both fixed and cellular operators are impacted by the OTT services but the problem is more severe and complex in the mobile world. Firstly, since the resources for cellular are more expensive, any OTT service—especially resource intensive services like video—are more costly. Second, mobile applications seem to play a bigger role in driving operator costs (e.g., the use of "keep alive" messages or other always-on mechanisms) than they do in fixed broadband. Thirdly, and from another perspective, cellular operators are potentially less constrained by issues like net neutrality and potentially freer to impose constraints and negotiate payments with content and application providers. The shift to OTT and the shift to a more complex networking ecosystem with new device and application platforms (with iOS and Android ecosystems but also Facebook and Twitter) suggest a loss of control over how resources are used.

To address the on-going threat to legacy revenue streams, ISPs are responding in a variety of ways. One approach is to try and follow Google's success and try to capture a larger share of mobile advertising dollars. While this might provide a viable revenue stream in certain circumstances, there does not appear to be sufficient advertising revenue available to simultaneously offset the loss in legacy revenues and meet the need for expanded capacity investment.

Ultimately, there is a need to address the question of costs and the potential disconnect between total revenues and total costs. As long as total revenues exceed total costs,[23] including a fair risk-adjusted return on invested capital (and dealing with the disagreements about what that is), operators' concerns are understandable but less worrisome for the ecosystem as a whole. On the other hand, if costs do exceed revenues (a) in total and (b) in particular contexts (particular markets, or for particular customers or usage scenarios) there is a real incentive problem. If the problem is (a) then there is a fundamental problem of economic viability and in the long-term the market is unsustainable. If not true for (a) but true for (b) then there are implicit cross-subsidies and we may still have a problem, but a different sort of problem. For example, we may collectively agree that we want/need universal access to mobile broadband and that we are willing to subsidize such access if necessary for certain classes or customer or usage (e.g., rural/poor, emergency services and basic communications), while not willing to subsidize it for other classes of customers or usage (e.g., very heavy users or for entertainment).

4. Business, Market, and Policy Challenges

In earlier sections, we identified overall trends and the challenges those posed for the ecosystem as a whole. In the following sub-sections we provide a partial examination of several important questions that are adding to the complex landscape of issues in our evolving mobile broadband ecosystem.

4.1. Who Controls the User-Experience?

Much of the progress in next generation Internet architectures and evolving ecosystems of devices and applications tilts the landscape toward end-user customized services and toward an increased role for end-users in managing the end-to-end experience. More capable end-user devices (faster processors, more storage), growing options for multi-homing and home network management (more capable home routers and modems, better in-home wireless and home DASD), and more capable applications and services (Web2.0 and beyond, mobile applications, enhanced iOS and Android platforms) make it easier and create more opportunities for end-users to shift functionality and control to edge-devices and to play an active role in customizing their experience.

At the same time, network operators are adding functionality to their networks including storage and intelligence. These comprise the ingredients for expanded cloud services that may include Software-as-a-Service (e.g., hosted applications), Infrastructure-as-a-Service (e.g., data centers, enhanced VPNs), and other options. Together these options have the potential to enable network operators to provide greater customization and more advanced, higher quality services to application and content providers and end-users.

There is also the potential for expanded roles for new types of players. These may include new models for provisioning broadband access (e.g., community-based networking) or novel service models (e.g., leveraging the capabilities of social networking applications to create new value-added services).[24]

In this environment of expanded options and more complex service offerings, the potential for tussling over control of the end-to-end user experience is expanded. Advanced users might prefer to manage their network services themselves, while most of the mass market may choose a one-stop-shopping solution. When problems arise, it is not always clear where the source of the problem is and who is responsible. For example, when a user cannot view a YouTube video on their smartphone at home, it may be because they are using a misconfigured home Wi-Fi network, because of a problem with their broadband service, because of congestion on the cellular network, or because of some other reason.

At the same time as end-users may be poised to play a larger role in managing the end-to-end experience, end-user behavior which contributes to determining market demand is also changing. Some of the trends/issues associated with changing end-user behavior include:

4.2. Incentives to share network resources efficiently

To encourage optimal investment in networks, it is important that all stakeholders have appropriate incentives to use network resources efficiently.

For example, setting a marginal price for usage (traffic) equal to zero helps create incentives for users to use more mobile broadband services and for developers to invest. However, a zero (or too low) marginal price for usage may also encourage excessive, and ultimately, wasteful usage of network resources. Mobile network operators and application, content, and equipment vendors are exploring ways to provide mutually enforcing incentives to induce each other to optimize designs to lower total costs. For example, mobile applications that abuse keep-alive messaging may congest mobile network signaling and transport networks, resulting in a poorer end-to-end quality of experience. One response might be for network operators to simply expand capacity, but this begs the question of how the costs of this additional investment should be recovered. A better response might be for the applications to communicate with the network differently so as to economize on such keep-alive messaging. The ability to make such accommodations may require new protocols and pricing models.

Alternatively, user behaviors and application design might shift traffic away from traffic peaks (by time or location) to allow more graceful (and lower cost) network capacity management.[25]

Some Internet application providers are starting to use "push servers" to decouple applications from the low-level issues of continuous contact with the client. Other efforts, notably in distributed storage, seek to optimally locate caches closer to the edge to minimize network costs. Adaptive video transmission is essentially self-throttling in high traffic areas. Furthermore, with the trend away from unlimited data plans, there is a stronger incentive to design applications in a more conservative manner with respect to network resources.

There are also emerging models for sharing usage costs. For example, some third parties are experimenting with subsidizing usage directly in return for customer attention (e.g., purchases or watching advertisements may bring "free bytes" of wireless data).[26] Such subsidized usage models are not new, but using them to lower the costs (to the end-user) of mobile usage is.

The responsibility for encouraging more "network-resource-friendly" behavior is not limited to application developers. Some provisions in 3GPP standards like LTE-A[27] can be very inefficient for data and video applications, relying on acknowledgement and retransmissions to provide error-free packets.[28] This comes at the cost of added delay and poor use of bandwidth.

Finally, the mobile Internet ecosystem lacks consensus mechanisms for measuring and signaling congestion. Such mechanisms are needed to implement better real-time adaptive approaches for managing traffic.

> **Cross Subsidizing Example: The Montreal Winter Light Festival**
> *The City of Montreal wants to encourage its citizens to remain downtown after hours and to embrace the city's winter beauty. The Winter Light Festival organized multimedia installations in core downtown areas. Some of these required Internet access. To meet this need, the city provided free Wi-Fi access near those installations. Such subsidized access helped promote Festival attendance and also benefited local businesses, contributing to government economic development goals.*

4.3. Mobile and Fixed Broadband: Substitutes or Complements?

The question of whether mobile and fixed broadband are substitutes or complements matters for both competition and universal service policy. From the perspective of competition, the extent to which these are substitute technologies changes the way we may think about last-mile competition and fixed/mobile mergers or joint-marketing agreements. From the perspective of universal service, the question forces policymakers to decide whether it is important to ensure universal access to both fixed and mobile broadband, or only that everyone have at least one option. If the two services are complementary, then it suggests each should be defined/regulated distinctly, and that there may be significant benefits from bundled offerings. To date and in the mature markets, mobile broadband services have been more typically marketed as a complementary service, but the potential for this changing is increasing.

It is our contention that mobile and fixed broadband are both substitutes for and complements to each other. On the one hand, increased mobile usage may induce users to want to use more Internet services everywhere, stimulating demand for both mobile and fixed services that are each more useful when ubiquitously available. On the other hand, sufficiently high quality mobile broadband or cheap fixed broadband may induce a consumer to opt for a single service—mobile broadband or fixed broadband, but not both. In some developing markets, mobile broadband may leapfrog fixed broadband, while in many mature markets, the majority of customers may elect to subscribe to both fixed and mobile services. Thus, we can expect retail mobile and fixed services to be simultaneously complements and substitutes, but this will play out differently depending on the customer or market segment.

Each service needs to offer a unique benefit if fixed and mobile broadband are to continue to act as complementary services. Mobility is a key differentiator for mobile broadband, and the support for high-speed mobility is a key differentiator for cellular-based mobile broadband services in particular. However, it is usually easier for fixed broadband providers to scale data rates and provide lower pricing per GB. With the increased deployment of fiber optic infrastructure in last-mile networks, and potentially all the way to the home, fixed broadband providers are likely to continue to outpace the data rates supported by mobile providers. However, with the shift toward smaller cell architectures, the fixed and mobile services may converge since much of the backhaul will rely on fixed (wired) infrastructure. The off-loading of cellular traffic to Wi-Fi, which is typically connected to fixed broadband service, is an example of how these network markets are converging.

Combining LTE and Fixed Broadband

In rural or poor neighborhoods, provisioning of wireless broadband may be problematic. Solutions from OpenBTS (openBTS.org), Range Networks, and Vanu Inc. (CompactRAN) offer new ways to deploy "cellular" infrastructure at low cost by allowing low cost base stations to be easily connected to any available IP backhaul.

5. Research Challenges

In the following sub-sections, we highlight several areas where continued active research attention is needed. This research is inherently multidisciplinary. In addition to designing new technical solutions, we need new business model and regulatory approaches, and we need to better understand how these interact. We also need to better understand how user behavior is evolving, how best to commercialize new technologies/business models, and how to craft policies that simultaneously promote consumer welfare, efficient resource use, and competition.

5.1. Mobile "Big Data" Opportunities

Mobile services have the potential to generate large amounts of data of many kinds. That data can help us better manage mobile networks and services, as well as improve our understanding of user behaviors and customized services. But always/everywhere connectivity also poses a serious threat to personal privacy. Finally, managing, storing, and processing the distributed mobile data will pose significant challenges for Big Data analytics.

5.2. Metrics for Characterizing the User Experience

Whether or not a service will achieve success on the market depends ultimately on the users, their satisfaction and willingness to pay for the service. Subjective quality testing is difficult, and generalizing across market/user environments is challenging. Expectations and prior experience, as well as individual motivation and personal sensibility to certain artifacts resulting from quality degradation, may play a decisive role in subjective evaluation. Additionally, and as already noted, end-to-end services are likely to be composites of components that are managed/owned independently and figuring out how each contributes to the end-to-end experience is difficult. Thus, additional research on the user quality of experience is needed.

5.3. Impact of Small Cells

As was discussed previously in the document, the transition to smaller cell architectures for wireless access is likely to be an important feature in the future landscape of mobile broadband. This is likely to involve both Wi-Fi and LTE technologies, and a range of business models. It will be important to better understand how these models or technologies relate to each other for spectrum management policy, for network

Mobile Social Networking

Social Networking is moving to the mobile platforms. More than a quarter of Facebook traffic is now mobile. Facebook's mobile strategy is evolving. Instragram (purchased by Facebook) is used by more than 100 million users to share pictures directly from cell phones. It is now the fastest growing social application used by teenagers. From Twitter's own statistics, 60% of their 200 million active contributors tweet via a mobile device at least once a month. And check-ins from Foursquare's 40 million users have reached over 3.5 billion. Yelp mobile (and other recommendation engines like Google+ and SIRI) accounted for 40% of all searches.

interconnection, and for service pricing and control.

5.4. Future of Mobile Demand

Much of the attention paid to mobile broadband traffic has focused on the growth of mobile video entertainment services. While this remains an important and fertile area for continued research, the potential for other applications is huge. For example, M2M, vehicle-to-vehicle, and user-generated content are each likely to account for significant changes in mobile user behavior and future traffic growth, but our experience with such services is still in a sufficiently early stage that there is little consensus as to how the growth of such markets will impact the overall ecosystem.

5.5. New Models for Layered Regulation

Broadband wireless networks are in the process of becoming the “new PSTN.” Figuring out how to map (or whether to discard) legacy telecommunications regulation to the emerging mobile wireless world poses multiple questions for policymakers and industry stakeholders. More research is needed to establish the scope for what regulators might do but also to re-evaluate legacy regulations in an always connected layered world. Issues here include but are not limited to, regulation of facilities, regulation of pricing, and regulation of the control of resources.

5.6. Future Internet Architectures for Broadband Mobile

Some research is already addressing the development of an open information-centric network platform with flexible control over content storage. This research can be combined with, for example, in-network composition to support wireless multimedia experiences more efficiently (added compression, forward error correction etc.). Software-defined ICN networks can enable efficient content composition and enhance user’s experience when changing terminals and/or roaming.

6. Conclusion

As presented in this paper, the exponential growth of mobile broadband and the changing user behavior from fixed to mobile services presents important challenges and opportunities for the entire broadband Internet ecosystem. This includes fixed and mobile (cellular) network operators, as well as application and equipment vendors, Over-the-Top (OTT) service providers, end-users, and policymakers. Some of the salient themes include:

- Exponential traffic growth needs an exponential capacity response, but the growth and necessary response are contingent on future architectures (e.g., shift to smaller cell architectures), pricing/business models (e.g., more usage-based pricing), and regulatory policy;
- The relationship between revenues and costs is highly variable, but there appear to be legitimate concerns that infrastructure revenues may fail to keep pace with usage-related costs without wider implementation of usage-based pricing models, and cross-value-chain efforts to optimize behavior to use network resources efficiently;
- New models for cost-sharing across the value chain may be needed to sustain the requisite investment;
- No single response is sufficient to address growth, costs and revenues at the same time, since the mobile broadband environment is subject to complex forces, with many players entering the market. New pricing models, the redesign of applications, and the offloading of traffic to cells served by lower-cost fixed lines are all components of a solution.

7. Appendix: Working Group Participants

Mobile Broadband: Toward a Sustainable Ecosystem

A Communications Futures Program (CFP) White Paper
Prepared by the Mobile Broadband Working Group (MBWG)[29]

Bill Lehr (MIT) <wlehr@mit.edu> (Chair)
Marie-José Montpetit (MIT) <mariejo@mit.edu> (Editor)

Participants[30]
David Clark (MIT) <ddc@csail.mit.edu>
Raffa De Peppe (Telecom Italia) <raffaele.depeppe@telecomitalia.it>
Hannu Flinck (NSN) <hannu.flinck@nsn.com>
Ram Jagadeesan (Cisco) <rjagadee@cisco.com>
Natalie Klym (MIT) <nklym@csail.mit.edu>
Andy Lipman (MIT) <lip@media.mit.edu>
Karen Sollins (MIT) <sollins@csail.mit.edu>
Edwin Sutedjo (NSN) <edwin.sutedjo@nsn.com>
Tony Tauber (Comcast) <tony_tauber@cable.comcast.com>

[1] To clarify our terminology upfront, we distinguish "broadband" as the *access*, or last-mile, portion of the larger Internet, in contrast to *transit* or *backhaul*. And unless otherwise specified, we use "Internet service providers" (ISPs) to refer specifically to last-mile providers. (These are sometimes referred to in the industry as "Internet access providers" (IAPs).

[2] The capacity investment challenge posed by exponential growth in broadband traffic is not new. See CFP white papers "The Broadband Incentive Problem," (2005); "The Broadband Incentive Problem, Part II," (2006); and, "Vision of Personal Broadband," (2006).

[3] The Internet of Things refers to a vision of the emerging future in which embedded processors in almost anything and potentially everything allow closer integration of the real and virtual worlds. In addition to other things, this will enable Machine-to-Machine (M2M) direct communication and control, potentially without active human intervention. For visions of this future, see for example: Microsoft's, Cisco's, McKinsey's, or, the European Union's.

[4] Economic costs include an appropriate risk-adjusted return on invested capital. While cost recovery is not assured in competitive markets, even for efficient firms, the greater the risk of loss, the greater the potential return demanded by investors to induce them to bear that risk.

[5] As end-users or households churn among providers, new subscription revenue does not disappear; however this does not produce new revenue for the industry as a whole. Increasingly, the growth in new subscription revenue comes from existing subscribers adding additional devices, but this is not the same as adding a new subscriber. In the future, the potential growth of Internet of Things devices may rejuvenate growth in new subscriptions, where the subscribers are not human.

[6] While users do not pay to use unlicensed spectrum, they must bear the costs of potential interference. There are no free lunches.

[7] Penetration beyond 100% implies that a number of subscribers have multiple subscriptions (e.g., a business and a personal mobile phone, or a mobile phone and a tablet, etcetera). Although the number is smaller each year, there are still subscribers without mobile service services.

[8] There is still significant scope for subscriber growth and the shift to more advanced user devices such as smart phones in developing markets.

[9] Source: Mobile Penetration: 1997-2002 ITU, 2003-2016 Yankee Group; Traffic: Cisco VNI

[10] http://retailcommon.com/blog/thoughts-on-a-billion-smartphones/

[11] Mary Meekers, Internet Trends @ Stanford – Bases Kickoff – December 2013

[12] A voice telephone call requires far fewer bits to transmit relative to rich image data (pictures or streaming video).

[13] Ericsson forecasts:
http://www.ericsson.com/news/1561267?categoryFilter=reports_1270673222_c

[14]http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf. Currently, the VNI is being updated to include 2014-2018 forecasts (see, CISCO Virtual Networking Index (VNI): http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI-Forecast_QA.html).

[15] A. T. Kearney (2010), "A viable future model for the Internet," white paper, A.T. Kearney Management Consulting, Chicago, 2010, available at: http://www.atkearney.com/documents/10192/4b98dac5-0c99-4439-9292-72bfcd7a6dd1; see also,

[16] For example, see http://gigaom.com/2013/12/11/att-ceo-thinks-phone-subsidies-must-end/

[17] If the traffic is real-time (e.g., live sports) or the audience is highly fragmented (e.g., user-generated content) it may not be easily cached, which allows storage costs to be substituted for transport costs.

[18] For example, healthcare monitoring might require compliance with HIPPA rules that may make it more costly than other traffic (e.g., entertainment video).

[19] The mix of traffic is also important for cost estimation

[20] For a discussion of the issues associated with estimating mobile cellular off-load traffic, see Marcus, J. Scott and John Burns (2013) "Study on Impact of Traffic Off-Loading and Related Technological Trends on the Demand for Wireless Broadband Spectrum, Final Report," Study Prepared for the European Commission.

[21] At the same time, differences in per capita income (and what might be viewed as "affordable") make simple comparisons of ARPU levels problematic.

[22] Permission to use Yankee Group data has been given. They need to review prior to publication for final release

[23] Total revenue must be weakly greater than total costs for investment to be rational. But this is in a dynamic, expected value sense, an expectation of cash flows over the indefinite future. It may mean losing money in the near term to make more money in the long term and (excess) profits (or losses) may occur periodically; and in presence of uncertainty, forecasts may not pan out, so this is a broad average.

[24] Natalie Klym and Marie Jose Montpetit, "Innovation at the Edge: Social TV and Beyond," Communications Futures Program, 2008.

[25] Not all traffic is time-sensitive and shifting time-insensitive traffic off-peak may free up additional resources for time-sensitive traffic during the peak times, lowering total capacity requirements (and total costs), while protecting the quality of experience for both the time sensitive and insensitive traffic.

[26] For example, see aquto.com.

[27] For example, see TS 36.321 Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification, 3GPP, March 2008, available at: http://www.3gpp.org/DynaReport/36321.htm.

[28] S. Teerapittayanon, Fouli, K.,M. Médard, Montpetit, M.-J., Shi, X., Seskar, I., and Gosain, A., "Network Coding as a WiMAX Link Reliability Mechanism", *MACOM 2012*

[29] The opinions expressed in this paper reflect those of the working group participants, and do not represent official views or policies of CFP's sponsoring companies or universities. The Communications Futures Program (CFP) is a partnership between university and industry at the forefront of defining the roadmap for communications and its impact on adjacent industries. CFP's mission is to help our industry partners recognize the opportunities and threats from these changes by understanding the drivers and pace of change, building technologies that create discontinuous innovation and building the enablers for such innovation to be meaningful to our partners. For further information about CFP, see http://cfp.mit.edu.

[30] Participated in one or more WG meetings.