University of Pennsylvania Law School

ILE

INSTITUTE FOR LAW AND ECONOMICS

A Joint Research Center of the Law School, the Wharton School,
and the Department of Economics in the School of Arts and Sciences
at the University of Pennsylvania

RESEARCH PAPER NO. 19-17

5G and Net Neutrality

Christopher S. Yoo
UNIVERSITY OF PENNSYLVANIA

Jesse Lambert
UNIVERSITY OF PENNSYLVANIA

5G and Net Neutrality

Christopher S. Yoo[†] and Jesse Lambert[‡]

Abstract

Industry observers have raised the possibility that European network neutrality regulations may obstruct the deployment of 5G. To assess those claims, this Chapter describes the key technologies likely to be incorporated into 5G, including millimeter wave band radios, massive multiple input/multiple output (MIMO), ultra-densification, multiple radio access technologies (multi-RAT), and support for device-to-device (D2D) and machine-to-machine (M2M) connectivity. It then reviews the business models likely to be associated with 5G, including network management through biasing and blanking, an emphasis on business-to-business (B2B) communications, and network function virtualization/network slicing. It then lays out the network neutrality regulations created by the EU in 2015 as well as the nonbinding interpretive guidelines issued by the Body of Body of European Regulators for Electronic Communication (BEREC) in 2016 and assesses how they will be applied to 5G. Network neutrality's impact on 5G will likely to be determined by the way that the exceptions for reasonable traffic management and specialised services are interpreted. A broad interpretation should accommodate network slicing and other new business models needed to support the deployment of 5G, while a narrow interpretation could restrict innovation and investment.

[†] John H. Chestnut Professor of Law, Communication, and Computer & Information Science and Founding Director of the Center for Technology, Innovation and Competition, University of Pennsylvania. Professor Yoo is a member of the Federal Communications Commission's Broadband Deployment Advisory Committee (BDAC). Any views asserted in this paper are those of the authors and should not be attributed to the BDAC.

[‡] Member of the J.D. Class of 2019, University of Pennsylvania Law School. An early draft of this paper was presented at the 50th Freiburg Seminar on Network Economics on "The Future of the Internet – Innovation, Integration and Sustainability" held on July 10-11, 2017, at the University of Freiburg. The authors thank the organizers for making this even possible and the conference participants for their comments. Any errors remain the responsibility of the authors.

1 Introduction

One of the most striking changes in the Internet ecosystem in recent years is the growing importance of mobile broadband. Driven in large part by the advent of the smart phone in 2007, mobile has emerged as the primary source of Internet connectivity in most countries around the world. By the end of 2010, mobile broadband connections surpassed fixed-line broadband connections in the U.S. (Federal Communications Commission 2012). Some estimates calculate that mobile IP traffic has grown by more than 100 times over the past ten years and is forecast to continue to grow at an annual rate of more than 45% per year through 2021 (Cisco, 2007, 2017).

Current discussions about mobile broadband are dominated by the impending deployment of the next generation of wireless technology, commonly known as 5G. The Korean government showcased 5G at the Winter Olympics in Pyeongchang in February 2018. Later that same month, 5G was the focus of seemingly every talk and every booth at the leading trade show for the wireless industry, known as the Mobile World Congress. In the U.S., Verizon and AT&T are racing to be the first to deploy a commercial 5G networks, with both planning to do so before the end of 2018.

Despite all of this hype, there is little understanding of what sets 5G apart as a technological matter and what the business case for its deployment will be. In addition, discussions of how the regulatory system might impact 5G and vice versa are only just beginning. For example, GSMA (2018) recently released a report on megacities underscoring the need to streamline the process for deploying 5G, a topic that is also the focus of the U.S. Federal Communications Commission's Broadband Deployment Advisory Committee (BDAC).

This Chapter focuses on another concern raised by the "5G Manifesto" released by a group of European telecommunications providers and equipment manufacturers in July 2016. This document warned that that strict implementation of network neutrality might hinder networks' ability to experiment with the kind of flexible, elastic configuration of resources and other innovative specialised services, which in turn could dampen investment in 5G (Patterson et al., 2016). A number of other industry observers have made similar questions about the compatibility of net neutrality and 5G (Lawson, 2015; Rysavy, 2017; Manion, 2017; Morris, 2018).

This Chapter seeks to examine the unresolved questions about how 5G will deploy and how network neutrality regulation will affect it. It begins by describing the technical underpinnings of 5G and identifying what sets it apart from previous generations of wireless technology, paying particular attention to new business model known as *network slicing*. It continues by looking at the potential impact of EU network neutrality regulations on 5G.

2 Defining 5G

Because the standards for 5G have not yet been finalized, a precise definition of 5G is not yet available. That said, a broad consensus exists as to 5G's overall outlines and the technologies that are likely to be incorporated into the final standard. (For overviews appearing in the engineering literature, see, for example, Boccardi et al., 2014; Osseiran et al., 2014; Andrews et al., 2014).

The operational goals for 5G are fairly well established: support for multiple gigabit speeds and millisecond latencies, the higher degree of reliability needed to support increasingly critical infrastructure, a greater diversity of devices, and reductions in energy consumption and cost. But a new generation of wireless connectivity requires more than just the ability to meet enhanced performance measures. It requires a sufficiently large technological step forward to be regarded as a paradigm shift. In addition to technological changes, the deployment of 5G appears to depend on the adoption of new business models. Finally, the potential use cases motivating 5G suggest that deployment is more likely to be driven by the Internet of Things (IoT) than the ability to expand network coverage or replace previous architectures.

2.1 New technologies

A consensus has emerged about the technologies most likely to prove essential to the deployment of 5G: use of millimeter wave bands of the spectrum, reliance on massive multiple input/multiple output (MIMO), and extreme densification through small cells. In addition, the 5G network needs to support multiple radio access technologies (multi-RAT) as well as direct device-to-device (D2D) and machine-to-machine (M2M) connectivity.

2.1.1 Millimeter wave bands

New services need spectrum. However, the spectrum bands traditionally used for wireless communication have already been fully allocated to other services. Improvements in semiconductor cost and power consumption have made available higher frequencies, such as millimeter wave (mmWave) spectrum in the 28, 38, 60, and 73 GHz ranges, that were previously regarded as unusable due to poor propagation characteristics. This increase in the range of usable frequencies allows the allocation of spectrum to 5G without having to displace any preexisting uses (Boccardi et al., 2014; Osseiran et al., 2014; Andrews et al., 2014Agiwal et al., 2016).

2.1.2 Massive multiple input/multiple output (MIMO)

Wireless communications are plagued by a problem known as *multipath propagation*, which occurs when a signal from a source reaches the same point via multiple routes, such as when one signal arrives directly and one or more other signals arrive after reflecting off some nearby surface. A good example occurs when an echo in a room with bad acoustics causes listeners to hear the same message twice, with the signal travelling by the indirect path arriving slightly after the signal arriving by the direct path. Receivers that are unable to distinguish between these signals perceive the reflected signal as capacity-reducing interference or noise. The field of information theory, based on the seminal work of Shannon (1949), has long recognized that the usable capacity of a bandwidth-limited channel depends on its signal-to-interference-plus-noise ratio (SINR). Thus, multipath propagation has generally been regarded as reducing network performance (Yoo, 2016b).

A technique known as multiple input/multiple output (MIMO) can solve this problem. Deploying multiple antennas to both transmit and receive allows the systems to use the distance between the antennas to distinguish between signals that travel by direct paths from those that travel by various indirect paths. In addition, MIMO allows receivers to combine multiple signals that individually would be too weak to be a usable signal. Indeed, the latest generation WiFi standard (802.11n) and recent 3G and 4G technologies use MIMO employing two-to-four transmission and receiving antennas to improve throughput and deal with highly reflective environments. Massive MIMO uses a technique developed in 2007 to increase the number of antennas in use to more than one hundred (Boccardi et al., 2014; Osseiran et al., 2014; Andrews et al., 2014Agiwal et al., 2016).

2.1.3 Ultra-densification through small cells

Another central technical commitment underlying 5G is the use of smaller, more densely packed cells. Increasing network capacity by making cells smaller is an approach that is hardly unique to 5G, as prior generations of wireless technology have long allowed more users to rely on the same spectrum by splitting cells and having the resulting base stations operate at lower power. That said, 5G takes densification to a new level, sometimes estimated to be as high as one base station for every twelve homes. In the limit, 5G base stations could be as small as a home router, with each one serving a single user. The use of smaller cells operating at lower power allows the same spectrum to be reused by a larger number of users (Andrews et al., 2014).

2.1.4 Multiple radio access technologies (multi-RAT)

Because microcells cover relatively small areas, they are not viable in areas where customer densities are too low to generate sufficient revenue to support the necessary investment. As a result, instead of being a replacement for 3G and 4G technologies, 5G will likely serve as a complement to legacy network deployments, relying on 3G and 4G macrocells to provide baseline connectivity and overlaying 5G microcells in those areas that require additional capacity (Osseiran et al., 2014; Andrews et al., 2014).

To deliver on its promise of reliable and ubiquitous high-speed service, 5G needs to be able to integrate multiple radio access technologies (multi-RAT). The standard must support the ability to connect to legacy 3G and 4G technologies as well as other technologies such as Wi-Fi and 5G. The 5G network must be able to manage handoffs seamlessly as users migrate from one radio technology to another. Finally, as discussed below in section 2.2.1, the need to optimize capacity in areas served by both microcells and macrocells requires extensive network management (Osseiran et al., 2014; Andrews et al., 2014).

2.1.5 Support for device-to-device (D2D) and machine-to-machine (M2M) connectivity

Lastly, 5G deviates from prior approaches by enabling the related, but distinct, concepts of device-to-device (D2D) and machine-to-machine (M2M) communications. D2D is an approach to network management that deviates from the traditional centralized architecture by permitting nearby devices to communicate directly with each other without having to connect through a base station. Direct D2D communication can reduce the number of hops, transmission distances, power consumption, and path loss. As discussed below in Section 2.2.1, the elimination of a central infrastructure point to oversee individual connections poses complex management challenges (Boccardi et al., 2014; Osseiran et al., 2014; Andrews et al., 2014Agiwal et al., 2016).

M2M differs from D2D in that it is not concerned with the mechanics of how end-nodes establish connections with one another. Instead, it focuses on the type of entity being connected, the type of data being exchanged, and the level of human involvement needed to initiate it. Prominent examples include connected cars, road and traffic management, health care, and smart cities. M2M differs from legacy architectures in that it envisions much larger numbers of connected devices, the need for very high link reliability and low latency, and the tendency to generate short blocks of highly sporadic data. M2M need not be D2D and can follow the traditional model of connecting through base stations. Unlike D2D, M2M communications need not involve devices that are located close to one another (Boccardi et al., 2014; Osseiran et al., 2014; Andrews et al., 2014Agiwal et al., 2016).

2.1.6 Impact on spectral efficiency

These innovations both increase the amount of usable spectrum and the intensity with which the spectrum is used. Other technological improvements, such as slot aggregation, reductions in guard band size, elimination of the need of constant transmission of cell-specific reference signals (CRS), the use of code block groups to support more efficient transmission of large transport block sizes, and improved channel coding, promise to enhance efficiency still further (Khalid, 2018). Industry estimates that together these technological improvements can increase spectral efficiency by up to four times (Ghosh, 2017; Qualcomm, 2018).

2.2 New business models

In addition to employing new technologies, 5G systems are expected to incorporate new business models that have the potential to organize resources in a manner that is fundamentally different from the ways that prior networks were organized. It is also expected to place new functional demands on and require new functionality from the architecture.

2.2.1 Network management

5G deployments require more extensive and innovative forms of network management. As noted earlier, 5G applications are likely to place greater demands on the network, requiring the connection of many more devices as well as latency and reliability guarantees that are significantly more stringent than the current best-efforts architecture can support (Osseiran et al., 2014; Andrews et al., 2014).

Another challenge is posed by the fact that 5G service relies on a combination of microcells operating at relatively low power and macrocells operating at relatively high power. The overlapping operation of low-power and high-power transmissions raises difficult management problems. As noted above, SINR determines the amount of usable bandwidth. The low-power microcells perceive the signals of high-power macrocells as interference or noise. Moreover, because microcells operate at lower power, networks based on them are more vulnerable to interference than previous architectures (Andrews et al., 2014).

Innovative solutions exist to solve these problems. One technique known as *biasing* requires a user to connect through a microcell even when connecting to the macrocell would provide greater bandwidth. The reduction in macrocell usage benefits all other microcells, but disadvantages the user subject to biasing (although the microcell can compensate by allowing that user to access more of the microcell's resources). Another technique known as *blanking* shuts off macrocells and shunts all of the traffic to the microcells even when utilizing the macrocells might be more efficient for individual users (Andrews et al., 2014).

Proper application of these techniques requires knowing the load each user is placing on the network as well as the load, SINR, and resources available at every base station. In addition, because interference is the product of pairwise interactions between users, the network must integrate these data into a massive combinatorial optimization problem. Moreover, the standard governing the network must have the authority to implement the optimal result in a D2D architecture in which each node can act independently outside the central control of the infrastructure (Andrews et al., 2014). The issues presented by these control problems are demonstrated eloquently by the fact that Google is employing the centralized management features of 4G LTE to support both its since-abandoned, balloon-based connectively project known as Project Loon and its wireless replacement for the now defunct Google Fiber instead of ad hoc, unmanaged routing technologies, such as Wi-Fi.

2.2.2 Significant role for business-to-business (B2B)

The primary focus of 3G and 4G wireless networks was to support broadband access by consumers. It is possible that 5G's greater spectral efficiency may provide sufficient cost reductions to support the deployment of 5G. If supply-side considerations are not sufficient, the value added needed to support return on investment must come from the demand side.

The data on consumer willingness to pay for 5G services are mixed. On the one hand, a survey by Ericsson ConsumerLabs (2018) found that, contrary to the belief that consumers are uninterested in 5G, 44% of users are willing to pay for 5G. Another survey jointly conducted by Qualcomm and Nokia (2017) found that 50% of respondents indicated they were likely to be early 5G adapters and that over 60% would be willing to pay an average of USD 50 for a 5G-enabled device.

Other surveys suggest the contrary. A study by Deloitte (2017a) of German consumers found that 61% did not regard 5G as important. It also found that willingness to pay for 5G was quite limited, with 58% unwilling to pay anything more for 5G and another 22% only willing to pay 5 € or less. A related Deloitte study (2017b) of U.S. consumers found that they were unwilling to pay much for IoT and speculated that "consumers are beginning to view IoT services more as another utility, one for which they want to pay little or nothing at all." This is consistent with prior findings that consumers have shown little willingness to pay for higher download speeds (Lee and Whitacre, 2017; Liu, Prince and Wallsten, 2018).

Uncertainly about the consumer demand for 5G has led many providers to focus their efforts on business-to-business (B2B) connectivity as well. A brief perusal of the business cases typically cited as the potential drivers for 5G deployments underscores the likely importance of B2B. Implementations such as the Internet of things (IoT), connected cars, smart cities, other forms of connected transportation, and health care all are likely to require coordination with other business entities rather than with consumers (Agiwal et al., 2016).

The reasons for the emphasis on B2B follow directly from 5G's emphasis on microcells. The small geographic coverage of each cell means that the service area served by each base station will encompass relatively few customers. The smaller number of consumers served by each cell means that the economic viability of microcells depends on tapping into as many other sources of revenue as possible. To the extent that latencies are important, sharing becomes critical. In a two-dimensional space, halving the distance to the nearest node requires quadrupling the number of nodes (Weinman, 2015). In a small cell world, it is unlikely that a single application would have sufficient volume to support the investment necessary to build out 5G.

2.2.3 Network function virtualization/network slicing

Each of the different Internet-connected industries that 5G needs to serve (sometimes called *verticals*) requires different clusters of services from the network. The 5G architecture could create a unique package of services for each vertical offered on an integrated basis by a single provider. Another approach would be for network providers to offer building blocks of network components that end users or third party integrators could assemble on a transactional basis to obtain the services that a particular vertical needs (Andrews et al., 2014).

This new business model is called *network function virtualization* or *network slicing*. These practices will enable the coexistence of independent virtual networks sitting side by side, much as cloud computing allows independent virtual machines to coexist on the same server. The comparison to cloud computing underscores the potential benefits of network slicing. In the traditional computing model, end users employ dedicated computing power and storage in the machines located on their desks and in their offices despite the fact that those resources were not in constant use and often lay fallow. The virtualization of computing and storage resources made possible by cloud computing allows multiple people to use the same servers. Instead of relying on dedicated machines, users instead obtain part-time access to virtual machines provided on demand on a set-up and take-down basis when they are needed (Weinman, 2015).

Network slicing applies the same rationale to networking. Rather than having network resources allocated to individual providers for extended lengths of time, network slicing allows network resources to be placed into new configurations on the fly in response to end users' immediate needs (Andrews et al., 2014). Network slicing thus results in multiple virtual networks sharing leased components in much the same way that cloud computing supports multiple virtual machines operating on the same server.

The resource sharing associated with network slicing yields a number of benefits, again illustrated by a comparison to cloud computing. Unless users' needs are perfectly correlated, computing power and storage can be shared in ways that allow the resources to be used more efficiently. Because users can expand their use of virtual machines to meet their demand means that they no longer need to maintain excess capacity to protect against potential surges in demand. These benefits can apply to networking resources to the same extent as to computer processing and storage. With the proper architecture, network slicing promises to allow users to access networking resources on a transactional basis to meet their immediate needs. This is done by decoupling the control and data planes and exposing the network capabilities to external applications through an application programming interface (API) (Andrews et al., 2014).

Network slicing is widely seen as critical to supporting the deployment of 5G. The small geographic areas served by microcells make it harder to maintain network components dedicated to any particular vertical. Creating configurable building blocks that can be assembled on a temporary basis maintains the flexibility of the network while allowing the network resources to be shared by all of the relevant verticals. So long as the usage of each different vertical is not perfectly correlated, network slicing should allow more intensive and efficient utilization of resources. Enabling multitenant usage of resources raises the possibility of parallel, independent virtual networks operating side by side (Andrews et al., 2014).

Different verticals require different types and levels of QoS. The configurability of these resources allows them to provide different levels of QoS as needed. In addition, the architecture must provide the primitives necessary to advertise, order, provision, meter, and bill resource usage. In addition, because slicing envisions short-term transactions that combine multiple components that are not part of a single, integrated system, the architecture must create interfaces to enable interconnection with other resources. Unless the information necessary to support such transactions are embedded in and conveyed by the network itself, such transactions will necessarily occur between firms in ways that reinforce the status quo, which would frustrate the vision of mix-and-match combinations of different elements among new providers on a dynamic basis (Yoo, 2015).

3. Assessing the potential tension between network neutrality and 5G

As noted in the 5G Manifesto, some commentators and industry observers have raised the concern that the EU approach to network neutrality may impede the business models needed to deploy 5G. EU network neutrality laws are embodied in Regulation (EU) 2015/2120 enacted on November 25, 2015, commonly known as the Telecoms Single Market (TSM) Regulation, which also included provisions addressing European roaming. The final language was the product of a protracted trilogue negotiation that resulted in a compromise that was less restrictive than the language initially adopted by the European Parliament. The Regulation calls upon the Body of European Regulators for Electronic Communication (BEREC) to issue nonbinding guidelines to help inform the National Regulatory Authorities' implementation of the regulation. BEREC issued these guidelines in August 2016.

The primary network neutrality obligations are laid out in Article 3 of the TSM Regulation. Article 3(1) recognizes end users'[1] "right to access and distribute information and content, use and provide applications and services, and use terminal equipment of their choice, irrespective of the end-user's or provider's location or the location, origin or destination of the information, content, application or service." Article 3(2) makes clear that these restrictions do not limit providers of internet access services and end-users from entering into agreements "on commercial and technical conditions and the characteristics of internet access services such as price, data volumes or speed" so long as these agreements do "not limit the exercise of the rights of end-users laid down in paragraph 1." The first subparagraph of Article 3(3) requires that providers of internet access services "treat all traffic equally . . . without discrimination, restriction or interference, and irrespective of the sender and receiver, the content accessed or distributed, the applications or services used or provided, or the terminal equipment used."

The TSM Regulation recognizes exceptions for *reasonable network management* and for what are commonly known as *specialised services*. Regarding the first exception, the second subparagraph of Article 3(3) specifies that the rules "shall not prevent providers of internet access services from implementing reasonable traffic management measures." It further specifies that to be reasonable, "such measures shall be transparent, non-discriminatory and

---

1 Unlike the repealed U.S. network neutrality regulations, which were limited in scope to consumer-facing Internet access services, consistent with the 2002 Framework Directive, the scope of EU network neutrality regulation appears to cover both individuals and businesses (BEREC, 2016, at 4)

proportionate" and "shall not be based on commercial considerations but on objectively different technical quality of service requirements of specific categories of traffic."

Regarding the second exception, Article 3(5) recognizes that providers of electronic communications[2] to the public "shall be free to offer services other than internet access services which are optimised for specific content, applications or series, or a combination thereof," which the BEREC Guidelines note is simply a new name for specialised services. Specifically, this category includes services that "are optimised for specific content, applications or services, or a combination thereof, where the optimisation is necessary in order to meet requirements of the content, applications or services for a specific level of quality." Specialised services "shall not be usable or offered as a replacement for internet access services, and shall not be to the detriment of the availability of general quality internet access services for end-users."

Whether network slicing and the other practices associated with 5G will fall more naturally within the exception for reasonable network management or specialized services is not clear. On the one hand, the recitals of the TSM Regulation note that "[t]he objective of reasonable traffic management is to contribute to an efficient use of network resources and to an optimisation of overall transmission quality" (Recital 9). Specialised services, in contrast, are "optimised for specific content, applications or services, or a combination thereof" (Article 3(5)). Given that the network slicing and other business practices are designed to support the particular applications and services needed by particular verticals, it seems unlikely to fit within the ambit of reasonable traffic management measures, which are focused on the performance of the

---

[2] Article 3(5) specifies that providers of electronic communications to the public include both providers of internet access services and providers of content, applications and services.

network as a whole. Instead, they seem more consistent with specialised services, which are focused on the performance of specific content, applications, or services.[3]

On the other hand, the distinction between the two categories may not be as sharp as the language would suggest. Traffic management measures that benefit the network as a whole do permit better operation of individual applications and services, whereas the benefits specialised services provide to specific applications and services also benefit the network as a whole. In many cases, the precise motivation may be difficult to discern.

On balance, practices such as network slicing appear to fall more naturally within the definition of specialised services than within the definition of reasonable network management. The idea of network slicing would appear to be more application oriented than network oriented. Moreover, footnote 26 to Paragraph 101 of the BEREC Guidelines acknowledges the link between specialised services and network slicing by noting, "Network-slicing in 5G networks may be used to deliver specialised services." But because of the ambiguity, this Chapter will analyze the application of both exceptions.

3.1 Reasonable traffic management

3.1.1 The EU Regulation

As noted earlier, subparagraph 2 of Article 3(3) of the TSM Regulation allows reasonable traffic management measures so long as they are "transparent, non-discriminatory and proportionate" and are not "based on commercial considerations but on objectively different technical quality of service requirements of specific categories of traffic." Subparagraph 3 of Article 3(3) specifies that traffic management measures "shall not block, slow down, alter, restrict, interfere with, degrade or discriminate between specific content, applications or

[3] Paragraph 75 of the BEREC Guidelines appears to offer a similar suggestion, but is somewhat opaque.

services." The Regulation does specify three exceptions to this rule: (1) compliance with national legislation or a legal order; (2) preservation of the integrity and security of the network, the services provided through it, and the devices connected to it; and (3) prevention of impending congestion and the mitigation of exceptional or temporary congestion, "provided that equivalent categories of traffic are treated equally."

In addition to repeating the operative regulatory language, the Recitals included in the TSM Regulation provide additional guidance as to what constitutes reasonable traffic management. Recital 9 provides, "The objective of reasonable traffic management is to contribute to an efficient use of network resources and to an optimisation of overall transmission quality responding to the objectively different technical quality of service requirements of specific categories of traffic, and thus of the content, applications and services transmitted." That said, "[t]he requirement for traffic management measures to be non-discriminatory does not preclude providers of internet access services from implementing, in order to optimise the overall transmission quality, traffic management measures which differentiate between objectively different categories of traffic." However, "[a]ny such differentiation should . . . be permitted only on the basis of objectively different technical quality of service requirements (for example, in terms of latency, jitter, packet loss, and bandwidth) of the specific categories of traffic, and not on the basis of commercial considerations." Also, "[s]uch differentiating measures should be proportionate in relation to the purpose of overall quality optimisation and should treat equivalent traffic equally. Such measures should not be maintained for longer than necessary."

Recital 11 makes clear that practices that go beyond the bounds of reasonable traffic management are prohibited unless they fall within the three exceptions specified above. In addition, "[t]hose exceptions should be subject to strict interpretation and to proportionality requirements." That said, prohibitions on altering content "do not ban non-discriminatory data compression techniques which reduce the size of a data file without any modification of the content," because "[s]uch compression enables a more efficient use of scarce resources and serves the end-users' interests by reducing data volumes, increasing speed and enhancing the experience of using the content, applications or services concerned."

Recital 12 makes clear that "[t]raffic management measures that go beyond such reasonable traffic management measures may only be applied as necessary and for as long as necessary to comply with the three justified exceptions laid down in this Regulation" specified above.

Recital 13 makes clear that compliance with law includes "requirements of the Charter of Fundamental Rights of the European Union . . . in relation to limitations on the exercise of fundamental rights and freedoms" and the procedural safeguards provided by the European Convention for the Protection of Human Rights and Fundamental Freedoms.

Recital 14 clarifies that protection of the integrity and security of the network includes "preventing cyber-attacks that occur through the spread of malicious software or identity theft of end-users that occurs as a result of spyware."

Recital 15 explains that congestion management is appropriate only "where such congestion occurs only temporarily or in exceptional circumstances" and includes both measures "necessary to prevent impending network congestion" as well as measures "to mitigate the effects of network congestion" after it has already occurred. The Recital defines temporary congestion as including "specific situations of short duration, where a sudden increase in the number of users in addition to the regular users, or a sudden increase in demand for specific content, applications or services, may overflow the transmission capacity of some elements of the network." Such congestion is likely to be more common "in mobile networks, which are subject to more variable conditions, such as physical obstructions, lower indoor coverage, or a variable number of active users with changing location." Even when such temporary congestion might be predictable, "it might not recur so often or for such extensive periods that a capacity expansion would be economically justified.

Recital 15 also characterizes exceptional congestion as "unpredictable and unavoidable situations of congestion" from causes such as "technical failure," "unexpected changes in routing of traffic," or "large increases in network traffic due to emergency or other situations beyond the control of providers of internet access services." "Such congestion problems are likely to be infrequent but may be severe, and are not necessarily of short duration." Measures to "prevent or mitigate the effects of temporary or exceptional network congestion should not give providers of internet access services the possibility to circumvent the general prohibition[s] contained in the TSM Regulation. Moreover, "[r]ecurrent and more long-lasting network congestion which is neither exceptional nor temporary should not benefit from that exception but should rather be tackled through expansion of network capacity."

3.1.2 The BEREC Guidelines

The BEREC Guidelines offer four pages and nineteen paragraphs of guidance on the implementation of the exception or reasonable traffic management. The provisions that are most important for 5G are those interpreting the regulatory requirement that reasonable traffic management measures "shall not be based on commercial considerations but on objectively different technical quality of service requirements of specific categories of traffic." Paragraph 62 reiterates the specific examples of what would constitute objectively different QoS requirements contained in the Regulation (latency, jitter, packet loss, and bandwidth). Paragraph 63 offers "real-time applications requiring a short time delay between sender and receiver" as an example of the type of application that is objectively sensitive to QoS requirements. Paragraph 66 offers that reasonable traffic management may be applied to "generic application types (such as file sharing VoIP or instant messaging)" only if it meets three requirements: (1) the generic application type "require[s] objectively different technical QoS," (2) "applications with equivalent QoS requirements are handled agnostically in the same traffic category," and (3) the justifications for the traffic management are specific to meeting those QoS requirements.

Most importantly for purposes of 5G, Paragraph 68 offers two examples of what the Guidelines believe constitutes a traffic management measure "based on commercial considerations" prohibited by the regulation. The first is "where an ISP charges for usage of different traffic categories." The second is "where the traffic management measure reflects the commercial interests of an ISP that offers certain applications or partners with a provider of certain applications."

Paragraph 75 distinguishes between what constitutes "categories of traffic" subject to reasonable traffic management on the one hand and specialised services on the other. Building on Recital 9 from the TSM Regulation, traffic management is intended to optimize overall transmission quality, presumably for the network as a whole, whereas specialised services are optimized "to meet requirements for a specific level of quality," presumably for specific content, applications, or services. Paragraph 21 specifies that traffic management practices do not require ex ante authorization.

3.1.3 Evaluation

Although the law is not completely clear, the best reading of the Regulation and the interpretive guidance put forth by BEREC suggests that the types of business models associated with 5G would not constitute reasonable traffic management. The recitals of the TSM Regulation note that "[t]he objective of reasonable traffic management is to contribute to an efficient use of network resources and to an optimisation of overall transmission quality" (Recital 9), as opposed to specialised services, which are "optimised for specific content, applications or services, or a combination thereof" (Article 3(5)). Given that network slicing and other business practices are designed to support the particular applications and services needed by particular verticals, it seems unlikely to fit within the ambit reasonable traffic management measures, because they are not primarily focused on the performance of the network as a whole. Instead, they seem more consistent with specialised services, which are focused on the performance of specific content, applications, or services.

Even if 5G business practices are regarded as traffic management measures, subparagraph 2 of Article 3(3) of the TSM regulation draws a sharp distinction between traffic management measures that are "based on commercial considerations" and those that are based "on objectively different technical quality of service requirements," with the latter being permissible and the former being forbidden. Charging for access to network slices would seem to constitute differential treatment based on commercial considerations rather than differential treatment based on objectively different QoS requirements.

That said, the examples cited in the BEREC Guidelines of prohibited commercial considerations (an ISP "charg[ing] for usage of different traffic categories" or "offer[ing] certain applications or partners with a provider of certain applications") would appear to take selling access to network slices outside the ambit of reasonable traffic management measures (Paragraph 68). The Guidelines do permit differential treatment based on "generic application types (such as file sharing, VoIP or instant messaging)." Any such differential treatment must be based on objectively different technical QoS and must handle "applications with equivalent QoS requirements . . . agnostically in the same traffic category" (Paragraph 66). Presumably, the illustrations contained in Paragraph 68 of the BEREC Guidelines would prevent providers from charging for network slices even if they were offered as generic information types.

3.2 Specialised services

3.2.1 The EU Regulation

As noted earlier, the types of business models expected to underlie 5G appear to fit more comfortably within the definition of "services other than internet access services which are optimised for specific content, applications or services, or a combination thereof" (more commonly known as specialised services).

Recital 16 of the TSM Regulation recognized the existence of demand for "electronic communication services other than internet access services, for which specific levels of quality, that are not assured by internet access services, are necessary." Specific examples include "services responding to a public interest" and "some new machine-to-machine communications services." National regulatory authorities overseeing specialised services should make sure that they are not "simply granting general priority over comparable content, applications or services available via the internet access service and thereby circumventing the provisions regarding traffic management measures applicable to the internet access services."

As noted above, specialised services are governed by Article 3(5) of the TSM Regulation. Per subparagraph 2, providers may offer specialised services "only if the network capacity is sufficient to provide them in addition to any internet access services provided." In addition, "[s]uch services shall not be usable or offered as a replacement for internet access services, and shall not be to the detriment of the availability or general quality of internet access services for end-users." Recital 17 emphasizes the same points. When determining whether specialised services are having a detrimental effect on internet access services, "national regulatory authorities should assess the impact on the availability and general quality of internet access services by analysing, inter alia, quality of service parameters (such as latency, jitter, packet loss), the levels and effects of congestion in the network, actual versus advertised speeds, the performance of internet access services as compared with services other than internet access services, and quality as perceived by end-users."

The Recital notes that anticipating traffic volumes can be more difficult in mobile networks, because variations in the number of users can cause unforeseeable variations in the quality of internet access services. As a result, in mobile networks, the fact that specialised services may have an unavoidable, minimal, and brief negative impact on internet access services should not be considered a sufficient a detriment to bar the deployment of specialised services.

3.2.2 The BEREC Guidelines

The BEREC Guidelines offer six pages and twenty-nine paragraphs of guidance on specialised services. Focusing on the provisions most relevant to 5G, Paragraph 99 recognizes that "there can be demand for services that need to be carried at a specific level of quality that cannot be assured by the standard best effort delivery." Examples listed in Paragraph 113 include voice over LTE (VoLTE), linear broadcasting IPTV, real-time health services such as remote surgery, new machine-to-machine communications, and more generically "some services responding to a public interest." Paragraph 106 notes that certain applications have QoS requirements that are inherent in the nature of the application, while other applications may have QoS requirements that may vary. For example, real-time applications have an inherent requirement for low latency. The QoS requirements for video, in contrast, may vary depending on whether the resolution of the content is "standard definition with a low bitrate or ultra-high definition with high bitrate." Paragraph 114 observes that specialised services "might be especially important to corporate customers" and that such business services "have to be assessed on a case-by-case basis." As was the case for reasonable traffic management, Paragraph 21 specifies that specialised services do not require ex ante authorization.

The Guidelines provide language to help interpret the restrictions on specialised services contained in the Regulation. Paragraph 101 repeats the regulatory language that specialised services are "services other than [internet access services]," they are "optimised for specific content, application, or services, or a combination thereof," and "the optimisation is objectively necessary in order to meet requirements or a specific level of quality." Paragraph 102 repeats the Regulation's requirements that specialised services leave sufficient capacity for internet access services, are not usable or offered as a replacement for internet access services, and be to the detriment of the availability or quality of internet access services. Paragraph 123 echoed Recital 17's acknowledgement that QoS is harder to manage in mobile networks, where the number of uses and the amount of traffic are more difficult to anticipate, so that reductions in QoS that are "unavoidable, minimal and limited to a short duration" should not be considered detriments to internet access services sufficient to invalidate a specialised service.

Most importantly for 5G, Paragraph 110 indicates that specialised services "cannot be provided by simply granting general priority over comparable content." Instead, they can be offered, "for example, through a connection that is logically separated from the traffic of the IAS."

In addition, Paragraph 122 recognizes that end-user control can make a specialised service more justifiable. Specifically, enforcement officials should not consider competition for capacity between specialised services and internet access services to violate the regulations "[w]hen it is technically impossible to provide the specialised service in parallel to IAS without detriment to the end-user's IAS quality," when end users are informed about the specialised service's impact on internet access services and determine for themselves whether to use it, and when the use of the specialised service affects only the end users' own internet access services and not capacity that is shared by multiple end users.

3.2.3 Evaluation

The business models associated with 5G appear to fit more comfortably with the definition of specialised services. The BEREC Guidelines also acknowledge that specialised services will be characterized by M2M and B2B communications (Paragraphs 113, 114). Moreover, the examples of specialised services listed in the Guidelines (VoLTE, linear broadcasting IPTV, and real-time health services such as remote surgery) describe the types of verticals that 5G hopes to support (Paragraph 113). Furthermore, as noted above, the BEREC Guidelines specifically acknowledge that 5G may be delivered via network slicing (footnote 26 in Paragraph 101). This suggests that while network slicing does not itself represent a specialised service, it can provide the components from which a specialised service can be devised.

One potential ambiguity lies in the restriction that specialised services are those "optimised for specific content, applications or services, or a combination thereof." The vision for 5G is that the network can be reconfigured for the different demands placed by different types of content, applications, and services. If so, the network at any particular moment may be configured to support a particular type of consumer demand. As a general matter, however, the overall network is designed to be flexible enough to support an arbitrary range of content, applications, and services rather than any one specific kind..

A further complicating factor is the catch-all phrase at the end of the list contained in Article 3(5) of the Regulation, which specifies that in addition to supporting "specific content, applications or services," specialised services may also support "a combination thereof." If interpreted broadly, this catchall could encompass services offered by a 5G provider designed to support any arbitrary combination of content, applications, and services. However, an interpretation this broad encompasses everything and thus would effectively render this language meaningless. A narrower interpretation limited to specified, pre-envisioned capabilities may not be broad enough to encompass the type of flexible building block architecture associated with 5G networks.

Equally problematic is the language in Paragraph 110 of the BEREC Guidelines, which prohibits "simply granting general priority over comparable content" and suggests that specialised services should be offered "through a connection that is logically separated from the traffic of the IAS." The emphasis on logical separation arguably suggests that physical separation through the use of dedicated bandwidth is not required, as had been proposed and rejected in earlier stages of the debate. Instead, it suggests that both the internet access service and the specialised service may share the same resource so long as they are segregated by using distinct code, name spaces, or other digital means.

Paragraphs 101 and 110 rest in uneasy tension, as prioritization through the use of differential types of service or through virtual channels or circuits represent traditional ways to achieve logical separation. The criticism of relying on simple grants of priority raises questions about traditional prioritization techniques such as Differentiated Services (DiffServ) implemented through the *type of service* field that has always been part of IPv4 and has been maintained in IPv6 through the inclusion of the *traffic class* field. Determining what constitutes logical separation that is not simply granting general priority remains unclear. Indeed, many of the uses of these fields, which have been part of the internet architecture since the very beginning, would seem to contradict the prohibition on simply granting priority to certain types of content. Questions remain about how these restrictions permit increasingly mainstream techniques that assert even more control over the physical assets, such as Multi-Protocol Label Switching (MPLS), which are enabled still further in IPv6 by the addition of the *flow label* field into the IPv6 header (Yoo, 2016a).

The requirement that specialised services not starve internet access services of resources, not serve as a replacement for internet access services, and not evade the net neutrality rules also raises potential problems. Those skeptical of the motives of network providers may raise difficult questions about what constitutes sufficient bandwidth for internet access services, how people are using different network services, and the nature of the network provider's motivation behind different practices. Ironically, these conflicts are likely to become more acute to the extent that specialised services turn out to be popular with consumers.

Finally, to the extent that network slicing is requested by end users, the language on end-user control contained in Paragraph 122 of the BEREC Guidelines offers some promise for enabling services supported by network slicing as a specialised service even when the specialised service does have some detrimental effects on internet access service. The logic underlying this provision is quite sensible: consumers are clearly in a position to ensure that the tradeoffs are made in a way that best suits their interests when they are the ones who dictate how those tradeoffs are made.

On closer inspection, the issue is not as simple as might be hoped. The language does not provide that end-user control automatically renders any completion between specialised services and internet access services unproblematic. Instead, end-user control provides a justification for the specialised service only "[w]hen it is technically impossible to provide the specialised service in parallel to IAS without detriment to the end-user's IAS quality." Read literally, this caveat has the potential to render end-user control irrelevant as a consideration. This is because, simply by deploying more bandwidth, it is almost always technically possible to provide specialised services in parallel with internet access services without degrading either. From this perspective, economic feasibility imposes the real limits, not technical impossibility. Thus, if read too broadly, the requirement of technical impossibility will never be met, and end-user control will never justify the deployment of specialised services despite the degradation

Even larger problems are raised by the requirement that the "use of the specialised service affect[] only the end users' own internet access services and not capacity that is shared by multiple end users." Networks have long moved past the day when each user connected to the Internet via dedicated twisted pair of coper wires that ran all the way to the telephone company's central office, the use of which was completely independent of and did not impose any detrimental impact on other people's usage. The increasing use of fiber nodes, in VDSL networks, hybrid fiber coaxial networks, and various fiber to the cabinet/node architectures means that bandwidth on the backside of the node is always shared. The advent of network slicing will accentuate the degree of resource sharing still further. Enabling users to access resources on a multitenant, transactional basis makes it almost inevitable that one person's usage will affect the use of others.

Lastly, it is far from clear whether end users will be able to exercise this discretion in an effective manner. Many end users may lack the expertise to make the necessary decisions. Moreover, end users care about services, not particular facilities, and the value of the services may depend on the behavior of other users both as competitors for limited resources and as potential partners whose demand can be aggregated to make resource usage more efficient. The issue is that end users can only see localized information and face collective action problems that place them in a poor position to make decisions based on what other network users are doing.

4. Conclusion

This Chapter explores the extent to which the European network neutrality regime is consistent with the innovative technologies and business models expected to be a part of 5G. The task is complicated by the fact that the technologies and business models that will comprise 5G have yet to be definitively determined. It is likely, however, that 5G will connect a significantly larger number of devices and that those devices will place demands on the network that are increasingly diverse. The idea that the increasing diversity of demand will require increasingly differential treatment rests in uneasy tension with the principle of equal treatment of traffic underlying network neutrality.

The regulatory provisions most likely to make such practices permissible are the exceptions for reasonable traffic management and for specialised services. Unfortunately, the regulatory language and the nonbinding interpretive guidance provided by BEREC do not completely resolve which approaches to 5G will be permissible, if any. Ultimate resolution of these issues will have to await the deployment of 5G, enforcement decisions and actions by the national regulatory authorities, and any subsequent judicial challenges to the regulatory decisions. The hope is that enforcement authorities and courts will enforce these provisions with enough flexibility to give innovation the room to experiment that it needs in order to thrive.

References

Agiwal, M., Roy, A., and Saxena, N. (2016). Next Generation 5G Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, 18(3), 1617-55.

Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C. K., & Zhang, J. C. (2014). What Will 5G Be? *IEEE Journal on Selected Areas in Communications*,

Boccardi, F., Heath, R. W., Jr., Lozano, A., Marzetta, T. L., & Popovski, P. (2014). Five Disruptive Technology Directions for 5G, *IEEE Communications*, 52(2), 74-80.

Body of European Regulators for Electronic Communications (2016). BEREC Guidelines on the Implementation by National Regulators of European Net Neutrality Rules, BoR (16) 127.

Cisco (2007). Cisco Visual Networking Index: Forecast and Methodology, 2006–2011.

———. (2017). Cisco Visual Networking Index: Forecast and Methodology, 2016–2021.

Deloitte. (2017a). Global Mobile Consumer Survey 2017 – Mobile Evolution: Ausgewählte Ergebnisse für den deutschen Mobilfunkmarkt. https://www2.deloitte.com/content/dam/Deloitte/de/Documents/technology-media-telecommunications/Global%20Mobile%20Consumer%20Survey%202017%20Study%20Deloitte1.pdf

———. (2017b). 2017 Global Mobile Consumer Survey: US edition – The dawn of the next era in mobile. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-2017-global-mobile-consumer-survey-executive-summary.pdf

Ericsson ConsumerLabs. (2018). Towards a Consumer 5G Future. https://www.ericsson.com/assets/local/networked-society/consumerlab/six-calls-to-action_report_screen_aw.pdf [Accessed 8 April 2018].

European Union (2015). Regulation (EU) 2015/2120 of the European Parliament and of the Council of November 25, 2015, laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union. *Official Journal of the European Union* (L 310), 1-18. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015R2120 [Accessed 8 April 2018].

Federal Communications Commission. (2012). Internet Access Service: Status as of June 30, 2011. https://apps.fcc.gov/edocs_public/attachmatch/DOC-314630A1.pdf [Accessed 8 April 2018].

Ghosh, A. (2017). 5G New Radio – Technology and Performance. 2nd Workshop of the NSF Millimeter-Wave Research Coordination Network, University of Wisconsin-Madison. mmwrcn.ece.wisc.edu/wp-uploads/2017/05/5G-NR-Ghosh-Nokia-Keynote.pdf.

GSMA (2018). Delivering the Digital Revolution: Will Mobile Infrastructure Keep Up with Rising Demand. https://www.gsma.com/publicpolicy/wp-content/uploads/2018/02/GSMA_DigitalTransformation_Delivering-the-Digital-Revolution.pdf [Accessed 8 April 2018].

Khalid, A. (2018). How 5G Wins the Spectral Efficiency Race. *Our Technology Planet*. https://ourtechplanet.com/how-5g-wins-the-spectral-efficiency-race/ [Accessed 8 April 2018].

Lawson, S. (2015). Suddenly, net neutrality doesn't look so great for 5G. *PC World*. https://www.pcworld.com/article/2893032/5g-net-neutrality-may-be-headed-for-a-showdown.html [Accessed 8 April 2018].

Lee, H. J., and Whitacre, B. (2017). Estimating willingness-to-pay for broadband attributes among low-income consumers: Results from two FCC lifeline pilot projects. *Telecommunications Policy*, 41(9), 769-80.

Liu, Y.-H., Prince, J., and Wallsten, S. (2018). Distinguishing Bandwidth and Latency in Households' Willingness-to-Pay for Broadband Internet Speed. *Information Economics and Policy*, forthcoming. https://www.sciencedirect.com/science/article/pii/S0167624517301609

Manion, P. (2017). Net Neutrality's fall is 5G's gain. *Electronic Products*. https://www.electronicproducts.com/Digital_ICs/Communications_Interface/Net_neutrality_s_fall_is_5G_s_gain.aspx [Accessed 8 April 2018].

Morris, I. (2018). Ericsson CEO: Net Neutrality Threatens 5G. *LightReading*. https://www.lightreading.com/net-neutrality/ericsson-ceo-net-neutrality-threatens-5g/d/d-id/740854 [Accessed 8 April 2018].

Osseiran, A., Boccard, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., Queseth, O., Schellmann, M., Schotten, H., Taoka, H., Tullberg, H., Uusitalo, M. A., Timus, B., & Fallgren, M. (2014). Scenarios for 5G Mobile and Wireless Communications: The Vision of the METIS Project. *IEEE Communications*, 52(5), 26-35.

Patterson, G., et al. (2016). 5G Manifesto for timely deployment of 5G in Europe. http://telecoms.com/wp-content/blogs.dir/1/files/2016/07/5GManifestofortimelydeploymentof5GinEurope.pdf [Accessed 8 April 2018].

Qualcomm. (2018). Predicting real-world performance of 5G NR mobile networks and devices. *OnQ Blog*. https://www.qualcomm.com/news/onq/2018/03/07/predicting-real-world-5g-performance

——— and Nokia. (2017). Making 5G a reality: Addressing the strong mobile broadband demand in 2019 and beyond. https://www.qualcomm.com/system/files/document/files/whitepaper_-_making_5g_a_reality_-_addressing_the_strong_mobile_broadban.pdf [Accessed 8 April 2018].

Rysavy, P. (2017). Analyst Angle: How "Title II" net neutrality undermines 5G. https://www.rcrwireless.com/20170503/opinion/net-neutrality-5G-tag9 [Accessed 8 April 2018].

Shannon, C. E. (1949). Communications in the Presence of Noise. *Proceedings of the I.R.E.*, 37(1), 10-21.

Weinman, J. (2015). Cloud Strategy and Economics. In C. S. Yoo & J.-F. Blanchette (Eds.). *Regulating the Cloud: Policy for Computing Infrastructure* (pp.21-60), Cambridge, Massachusetts: MIT.

Yoo, C. S. (2015). Cloud Computing, Contractibility, and Network Architecture. In C. S. Yoo & J.-F. Blanchette (Eds.). *Regulating the Cloud: Policy for Computing Infrastructure* (pp.115-33), Cambridge, Massachusetts: MIT.

———. (2016a). Modularity Theory and Internet Policy. *University of Illinois Law Review*, 2016(1), 1-62.

———. (2016b). Wireless Network Neutrality: Technological Challenges and Policy Implications. *Berkeley Technology Law Journal*, 31(2), 1410-59.