UC BERKELEY

CENTER FOR LONG-TERM CYBERSECURITY

CLTC WHITE PAPER SERIES

Security Implications of 5G Networks

JON METZLER

CLTC WHITE PAPER SERIES

Security Implications of 5G Networks

JON METZLER
Lecturer, Haas School of Business

SEPTEMBER 2020

UC BERKELEY

CENTER FOR LONG-TERM CYBERSECURITY

Contents

Executive Summary

Network operators are rolling out fifth-generation (5G) cellular service worldwide. As of March 2020, 5G service was available in 24 countries.[1] GSMA, the global mobile operator association, predicts 1.8 billion 5G connections by 2025. The rollout of 5G has received much attention by consumers, industry, media, and policymakers.

5G has also become a geopolitical topic. The United States, for example, has advocated that global network operators exclude Chinese suppliers from their networks, based on perceived security concerns. The long-lasting nature of network investments means that supplier selection decisions will have implications for decades. Supplier decisions are not made lightly, nor are supplier changes.

This paper summarizes research and interviews conducted over a two-year period with support from the University of California, Berkeley Center for Long-Term Cybersecurity (CLTC). Our intent is to (1) help network operators — and their customers and partners — prepare for new risk vectors opened by 5G service, whether in terms of service models or network deployment models; (2) highlight security benefits of deploying both 5G RAN and core; and (3) help policymakers understand the economic and operational implications of 5G network deployment, including the switching costs of replacing suppliers and the site access needed to deploy robust, pervasive 5G networks.

This paper highlights the following key points:

- Networks persist. Network technologies and suppliers are used for decades once deployed. The switching costs that result from changing suppliers extend beyond capital investment. They also include re-training and changing operational practices. "Rip and replace" costs include these training and migration costs, and the transition itself may open security risks.
- 5G service will support a more diverse set of applications than traditional mobile service offered to consumers. This will add new value to 5G service as compared to prior generations. It will also raise the consequences of service outages. In this paper, this is referred to as "value at stake."

1 https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/03/GSMA_MobileEconomy2020_Global.pdf

- More diverse applications may mean more heterogenous suppliers, including device and service partners outside of the traditional set of operator suppliers. While mitigating supplier dependencies (single points of failure), working with unfamiliar suppliers may open new risk vectors. This will require operators or their partners to be able to test and verify new device partners quickly to validate their security practices.
- The three types of spectrum band (high-band or mm-wave; mid-band; and low-band) allocated to 5G have different implications for network topology. Mid-band and high-band service will necessitate significant densification of operator networks. This densification may open greater operational and physical access risks than do traditional cellular networks. Further, the increase of cell sites required with network densification will require robust network monitoring capability, and the ability to update and patch software on small cells and customer premise equipment.
- 5G networks have at least three security benefits relative to prior generations: improved authentication; distributed core; and network slicing, i.e., dividing a single network into different "slices" while using the same wireless spectrum and physical network infrastructure. Realizing these benefits requires deploying both 5G RAN and 5G core. These benefits are compelling reasons for customers to investigate 5G-only service.

Based on the above, this paper recommends that:

- Operators, their partners, and their customers investigate the viability of 5G-only service;
- Operators and their partners develop the ability to rapidly deploy software updates, including security patches, to small cells, customer premise equipment, and other connected devices;
- Operators and their partners develop the ability to rapidly test and verify devices from new partners from outside of the traditional telecom ecosystem;
- Policymakers act to facilitate rapid deployment of 5G networks, including implementing policies to facilitate cell site acquisition;
- Policymakers recognize the role of global standards bodies and rapid standards development, as well as the economic value of globally harmonized standards.

This paper utilizes a mix of sources, many of which are listed in the footnotes. Other resources include conference proceedings and conference sessions, such as panels at RSA and BlackHat and ETSI Security Week; interviews with network operators, equipment suppliers, and analysts and researchers, mainly done on background; a private, anonymized survey of network operators; and operator and network equipment provider public statements, such as in other media interviews and product and marketing documentation.

Glossary

Core: The portion of a mobile network operator's network that handles authentication, switching, interface with other networks, etc.

Customer premise equipment (CPE): Refers to equipment installed at a customer site, such as a Wi-Fi router or small cell.

Latency: The round-trip time between when the sender makes a request, such as tapping a button on an application, and gets a response.

Massive machine-type communications (MMTC): Use of 5G for highly dense user deployments (target = 1M devices per sq km).

Network slicing: Segmenting operator networks into different "slices" to support different applications, while using the same wireless spectrum and physical network infrastructure. For example, an operator could compartmentalize consumer wireless traffic and industrial wireless traffic on different slices.

Radio access network (RAN): The portion of a mobile network operator's network, including base stations, that provides the wireless interface with customer devices and manages related radio resources.

Ultra-reliable low latency communication (uRLLC): Use of 5G for low-latency sensor networks.

Introduction

The launch of the fifth-generation (5G) technology standard for cellular networks has received significant attention, not only from the traditional telecommunications industry, but also from governments, consumers, and other stakeholders. Equipment providers and mobile network operators have marketed 5G's capabilities to a much greater extent than during the migration from 3G to 4G, often highlighting the increased speed of data transmission and reduced latency. Through its marketing, the industry has done much to raise expectations about 5G among both individual and business consumers.

As defined by the 3rd Generation Partnership Project (3GPP),[2] the body that develops and codifies each wireless generation's capabilities, 5G will support greater service model flexibility than prior generations, which were focused largely on consumer applications such as traditional mobile cellular service. This flexibility will enable network operators worldwide to address consumers and industry with different service models. For example, depending on its spectrum holdings and whether it fully upgrades to 5G, an operator can deliver faster mobile broadband as an extension of current 4G service, or provide service with 5G-and-above capabilities, such as network slicing.[3]

This flexibility has led to the proposal of a broad number of applications: AT&T alone claims that its 5G services will be able to support augmented and virtual reality applications, autonomous vehicles, retail, health care, finance, and manufacturing.[4] Consulting firm McKinsey projects that 5G connectivity supporting applications in manufacturing, retail, mobility (i.e., current cellular), and healthcare alone could raise global GDP by $1.2 trillion to $2 trillion by 2030.[5] This illustrates the large expectations placed on 5G as an enabler of the digital transformation of industry.[6]

2 https://www.3gpp.org/

3 Network slicing is defined with 3GPP Release 16, approved by 3GPP on July 3, 2020. Release 15 largely defined mobile broadband capabilities; Release 16, also referred to as 5G Phase 2, defines a set of additional capabilities that can be delivered if both 5G RAN and core are deployed.

4 https://www.business.att.com/content/dam/attbusiness/reports/5g-for-business-whitepaper.pdf

5 https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/connected-world-an-evolution-in-connectivity-beyond-the-5g-revolution

6 The author moderated a panel in April 2020 featuring new entrant, Rakuten Mobile, which cited digital transformation as one of the key benefits of its 4G/5G network.

This greater diversity of use cases — in particular, the addition of mission-critical industrial, business, and public-sector uses on top of traditional mobile voice and data service — suggests there will be increased value at stake relative to traditional cellular services. As network operators generally are integrators of devices and software from other partners, increased application diversity also implies greater heterogeneity of partners, such as device manufacturers. The diversification of service models and end users also potentially means greater consequences for service interruptions. We have witnessed in recent years how cyberattacks can cripple municipal systems in a major metropolis like Baltimore.[7] Smart cities are an oft-touted use case for 5G. If cities are to look at 5G as potentially augmenting city infrastructure, then risks inherent to this new standard should be identified and mitigated. Conversely, if there are advantages — including security advantages — in rapid migration to 5G, those advantages should also be understood and harnessed.

The research detailed in this report stems from the hypothesis that the combination of these two elements—*greater value at stake* and *greater heterogeneity of devices and service models*—will create new risk vectors for mobile operators that were not seen with previous generations. Based upon research and industry interviews, this paper argues that (1) 5G will indeed lead to greater value at stake due its service model flexibility, in particular its support of industry; and that (2) greater heterogeneity in supported devices, including from device partners that are not traditionally suppliers to telecom operators, will require operators or their partners to be able to test and verify new device partners quickly to validate their security practices.

There is a counter-argument to the diversity-as-risk hypothesis: that greater ecosystem diversity reduces dependencies on any one supplier, and reduces the risk of single points of failure. It is likely that at least the network equipment market will remain consolidated as it is today, with diversity of end devices and services much more likely. To the extent security practices are clear and established, the argument for greater diversity as benefit bears noting. Harnessing that diversity, however, means developing robust supplier verification practices, and it is in the transition phase — i.e., on-ramp of new suppliers — that risks may be incurred, such as the risk of misconfiguration as staff get familiar with new equipment.

More diverse applications illustrate a new potential form of diversity, which is data diversity — consumer data versus industry; mission-critical data versus fault-tolerant data. Transporting data

7 https://baltimore.cbslocal.com/2019/06/12/baltimore-ransomware-attack-inches-closer-to-normal/

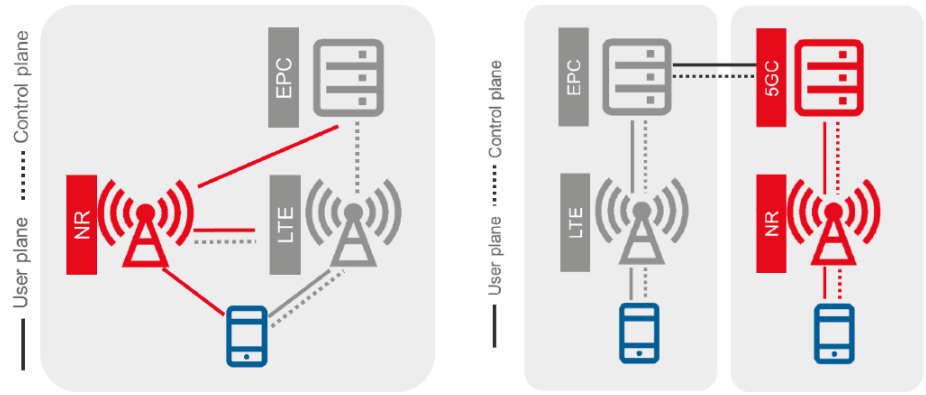

FIGURE 1: 5G DEPLOYMENT SCENARIOS

As depicted in this GSMA reference diagram, the mobile industry has defined two standards-based types of 5G deployment: Standalone 5G (SA), and Non-Standalone 5G (NSA). For Standalone 5G, 5G RAN connects to the 5G core. In the Non-Standalone scenario, 5G RAN and 4G RAN connect to a 4G core. (*NR = 5G New Radio. EPC = Enhanced Packet Core, or 4G core. LTE = Long Term Evolution, or 4G.*)

from different customer sets with different requirements may require policies around the treatment and compartmentalization of that data.

As another key finding, 5G presents certain security advantages over prior generations, particularly for network operators that upgrade both their radio access network (RAN) and core network to 5G.[8] Indeed, operators that rapidly migrate both their RAN and core to 5G may be able to gain competitive advantage relative to those who do not, because upgrading the core to 5G will enable key security-enabling features such as more secure authentication between devices and base stations.

In practice, however, many operators will upgrade to 5G RAN first, and operate it simultaneously with 4G RAN and 4G core. There are logical reasons for this, such as the desire to amortize 4G infrastructure as much as possible, and the need to support customers of multiple generations of network technology simultaneously. For example, all U.S. wireless carriers currently provide some level of 5G service, based on 5G RAN combined with 4G core.[9]

8 Verizon provides a explanation of RAN and core at https://www.verizon.com/about/our-company/5g/5g-radio-access-networks

9 In July 2020, Verizon Wireless announced successful testing of 5G service with standalone 5G core. https://www.fiercewireless.com/tech/verizon-readies-initial-shift-to-5g-standalone-core-after-successful-trial

This hybrid approach may negate many of the security benefits enabled by full upgrade of both RAN and core network to 5G.

A new entrant, starting from standalone 5G service, could begin from a position of advantage by providing more secure end-to-end 5G service. New entrants may incur significant startup costs, such as network equipment, cell site acquisition, and access to licensed spectrum. Additional costs include handset procurement and staff. Incumbent network operators without these startup costs could also stand up compartmentalized (area-specific or even building-specific) 5G networks to gain those same benefits. Localized or private 5G networks are discussed later in this paper.

Network Generations and Upgrades

In understanding the security implications of 5G network deployment and operation, it is important to first understand how networks are deployed; how long generations of network technology may be in operation; the implications of management of multiple generations of network technology; and how customers are migrated from one generation to the next. All of these factors can influence network operator decision-making.

First and foremost, network upgrades do not happen overnight. A mobile network operator does not flash-cut from one generation of network technology to another. Rather, a new generation will co-exist with prior generations for years, even decades. The existence of prior generations of network technology also can mean the foundation on which a new generation can be deployed is already defined. Network upgrades are akin to adding a new floor to a building while lower floors are still in use. Migrating tenants upstairs happens slowly and sometimes requires incentives.

This is in part due to the device upgrade cycle. Consumers in the United States, for example, were educated to expect "new every two," i.e., to upgrade their device every two years. More recently, the typical handset upgrade cycle has lengthened to roughly every three years. Some devices, such as automobiles or home alarm systems, may be in operation for much longer. For example, a typical car in the United States is on the road for 11 years. What if a car has embed-

ded cellular capability? Automakers and network operators do not want to strand incumbent customers. The migration of such equipment to new generations can create switching costs, such as the expense of sending a technician to switch out an alarm system. Embedded systems may not necessarily require the capabilities that come with newer generations, another reason operators may prefer to amortize their current systems as much as possible before upgrading. Here we note that T-Mobile USA plans to shut down its 2G network in December 2020, having maintained it in part to support customers of embedded systems.

In sum, customers migrate over to new generations gradually. GSMA, the global mobile industry association, projects that 20% of worldwide mobile handsets in use will be 5G-capable by 2025.[10] From a security perspective, this means that capabilities or fixes added to new generations of network technology have to co-exist with the limitations of prior generations. This can leave open loopholes that network operators would otherwise want to close. Bidding-down (downgrade) attacks are an example of this and are addressed later in this paper.

NETWORK OPERATIONAL LIFE

Substantive network generation upgrades happen about once per decade, with iterative improvements (e.g., 2.5G, 3.5G, 4.5G) occurring in between. In the US, AT&T launched its 4G network in 2011, but did not phase out its 2G network until 2017. Verizon launched 4G in 2010, and 5G in 2019, but it only plans to shut down 2G/3G in 2020. Simply put, networks persist. Equipment from new generations can coexist with that from prior generations for decades.

This overlap between generations creates advantages for incumbent firms. Network operators have incentive to buy multiple generations of equipment from the same supplier, not only for equipment interoperability reasons, but also for reasons of training and operational practice. Buying equipment from a new supplier and operating it in tandem with that from a different supplier adds operational complexity. Further, ripping out a prior generation's equipment from one supplier and simultaneously buying multiple generations of equipment (e.g. 4G and 5G) from a new supplier represents another set of costs.[11] It can also impact service quality. Thus it is not a decision that a network operator would normally make for economic reasons.

10 https://www.gsma.com/mobileeconomy/

11 The United Kingdom recently announced the decision to remove Huawei from operator networks. Britain's Culture Secretary commented that the removal of Huawei equipment would both delay 5G rollout and add to rollout cost. Sprint's prior removal of Huawei equipment in 2013 cost an estimated $1 billion. https://www.fiercewireless.com/wireless/report-sprint-could-pay-1b-to-rip-out-huawei-s-kit-from-clearwire-s-network

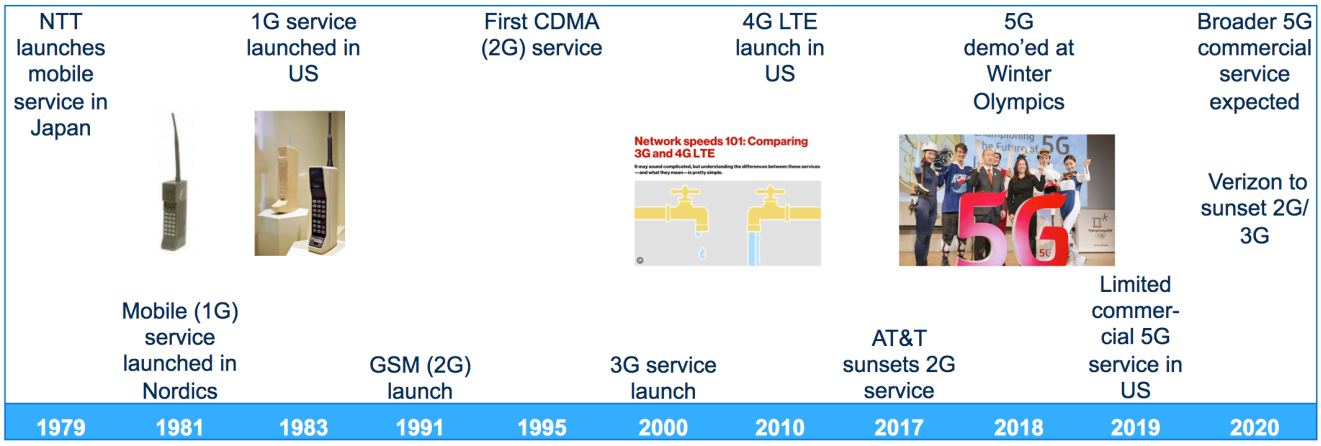

FIGURE 2

A timeline of network generation milestones. Diagram by the author.

Consider the airline industry, another mature industry with a concentrated supplier base. Airlines create operational efficiencies by standardizing their fleets on equipment from one airplane supplier. The high switching costs (and long order lead times) involved explain why airlines have patiently waited for one major airplane supplier to recover from its struggles, rather than switch to competitors.[12]

Consequently, the major equipment suppliers for 4G network equipment, such as Ericsson, Huawei, and Nokia, are also the dominant 5G equipment suppliers.[13] The benefits of incumbency are high, and switching costs go well beyond equipment costs. Further, any discontinuities of service during a network equipment switchover could impact customers and adversely impact operator reputation and business results. Sprint's many years of challenges in integrating multiple different network standards and equipment suppliers are a cautionary tale.[14]

Another factor contributing to the incumbency advantages enjoyed by network equipment suppliers is that there are few new mobile operator entrants, due to saturated markets and

12 To continue the analogy, much as many airlines lease their planes, some network operators lease (outsource) both network operation and cell towers.

13 https://www.delloro.com/the-telecom-equipment-market-2019/

14 CDMA operator Sprint acquired Nextel, an operator using iDEN network technology, in 2005. It then acquired WiMAX operator Clearwire, then ultimately migrated all of these network technologies to LTE. For a time, Sprint had to support all of these technologies in parallel. After receiving poor ratings in Consumer Reports in 2013, Sprint asked customers to "pardon our dust," highlighting the challenges it had in integrating network technologies and suppliers. https://www.scientificamerican.com/article/sprint-dead-last-in-consumer-reports-phone-service-survey/

high startup costs.[15] Capital barriers to entry are high. Worldwide, with some significant exceptions, the trend is toward consolidation, not new operator entry. Some refer to the "rule of three," i.e., that the number of mobile network operators in a country eventually consolidates to three. There are exceptions: Reliance Jio, in India; Rakuten Mobile, in Japan; and Dish, in the U.S.,[16] are exceptions to the general trend toward consolidation.[17] However, as markets mature, the number of network operators tends towards consolidation, and long-term relationships develop between network operators and network equipment suppliers. This can make it challenging for new equipment suppliers to break into the market. When they do, it is often as a specialized subcontractor to a prime supplier.

Thus, when issues of security are raised, or when more secure capabilities are introduced, network operators are obliged to look at a live network of customers using multiple generations of technology, and consider benefits, overall switching costs, and the risk of disenfranchising current customers using earlier generations of technology. The potential downside consequences of adopting new technologies, such as the impact of any outages on consumers dependent on mobile service, must be taken into consideration. It should be noted that many mobile subscribers are mobile-only. The mobile phone is their primary phone, and it may be their primary means of internet access.[18] This inevitably leads operators to be cautious about switching suppliers. In solving one problem — mitigating a perceived national security threat — new problems and risks may be created.

Enhancements with 5G

Each new network generation introduces additional capabilities that go beyond faster speeds. Per-bit transport costs go down, and spectral efficiencies improve. To use a highway analogy, more cars can be fit into the same width highway, at lower cost per car.[19]

15 Starting a new national carrier in the US could easily cost more than $10 billion, including spectrum and network build-out costs. A national carrier may need 50,000-80,000 cell sites to reach national coverage.

16 Rakuten Mobile launched service in Japan in 2020. Dish has not launched service in the US.

17 Both Reliance Jio and Rakuten Mobile represent new looks at the network operator business model. Recently, Reliance Jio claimed to have developed its own 5G equipment. https://www.developingtelecoms.com/telecom-business/operator-news/9306-jio-reveals-self-developed-5g-network-equipment.html

18 https://www.pewresearch.org/internet/fact-sheet/mobile/

19 To extend the analogy, in the 5G case, lane width and throughput may vary based on the type of service being provided. Low data-rate sensor networks likely will not need much network bandwidth.

New generations of network technology are also often introduced in tandem with new spectrum bands. Enhancements associated with 5G are listed below.

Table 1: 5G enhancements.[20]

ATTRIBUTE	5G IMPLEMENTATION
Improved spectral efficiency[21]	30 bits/s/Hz[22] theoretical peak downlink efficiency (15 b/s/Hz used in this paper to reflect real-world conditions).
Cost per bit	Down 50-80% from LTE[23]
Lower latency	Target user-experienced latency of 4ms
New licensed spectrum bands	In the US: 600 MHz; 28 GHz; 39 GHz. Planned: 3.5 GHz, 24 GHz, 37 GHz, 39 GHz, 47 GHz.
More unlicensed spectrum	6 GHz; above 95 GHz for 5G
User densification	Target 1M devices supported per square-km for sensor communications.[24] Referred to as MMTC or Massive Machine-Type Communications.

Further, 3GPP's specifications also define the following performance requirements:

- Enhanced Mobile Broadband (eMBB): In essence, substantive enhancement of current mobile broadband (4G). Over 10 Gbps peak data rates on the uplink and downlink, and over 50 Mbps user-experienced data rates on the uplink and downlink.
- Ultra-reliable Low-Latency Communications (uRLLC): Latency below 1 millisecond.
- Time-sensitive Networking (TSN): Support for deterministic networks, Ethernet over 5G (5G LAN), quality of service control, and accurate device time synchronization.
- Hybrid Positioning (HyPos): As an alternative to satellite GPS for underground, urban canyon, and other areas where GPS reception can be poor.

20 Developed based on materials from the FCC, ITU, 5G Americas, and Rysavy Research, in addition to author research.

21 Network equipment provider Ericsson explains Shannon's Law, spectral efficiency, and techniques such as MIMO in this blog post. https://www.ericsson.com/en/blog/6/2020/economics-of-5g-deployments

22 b/s/Hz or (bits/sec)/Hz is a measure of spectral efficiency. A wireless system that can send 1000 bits of data per second over 1 kilohertz of wireless spectrum would have an efficiency of 1 b/s/hz.

23 Operator cost-per-bit can be influenced not just by capital equipment expense and spectral efficiency, but also by running costs such as power consumption.

24 URLLC: ultra reliable low latency communications.

Commentary on 5G enhancements

- When 4G was introduced in 2010 , a rule of thumb for 4G spectral efficiency was 4 b/s/Hz. Subsequent improvements were made, such as multiple-input, multiple-output (MIMO) antenna techniques, and increased cell site sectorization.[25] In this paper, 15 b/s/Hz is adopted for simple estimates of bandwidth available for 5G systems. Estimates are made to develop service model hypotheses for US wireless carriers, given their current 5G spectrum holdings.[26] Spectrum influences network topology and application; topology and application in turn influence potential risk vectors.
- User-experienced latency refers to lag experienced by the user between query and response. 5G targets a user-experienced latency of 4ms. This is a round-trip measurement, meaning 2ms in each direction. In practice, total system latency will likely be higher. Verizon achieved round-trip latency of 15ms in recently publicized testing results.[27] The requirement of reduced latency necessitates network architecture design that has potential security implications and benefits.

Further details on 5G target specifications are provided by the ITU.[28] Network slicing, which is a new capability if 5G core is also deployed, is addressed later in this paper.

Value at Stake

5G is being marketed not only to consumers but also to industry and the public sector. This speaks to 5G service having higher value at stake — providing greater customer and societal value, but also making service outages more consequential. Total mobile industry revenues reached $1.03 trillion in 2019.[29] GSMA estimates that mobile technologies and services had $4.1 trillion in economic value added worldwide, or close to 5% of global GDP, including $2.5 trillion in productivity benefits.

25 Traditionally, operators have divided cell sites into three sectors. More recently, some have divided cells into six narrower sectors to support more traffic off the same mast. Increased backhaul is required for this to be effective.

26 CTIA, the US wireless carrier association, released a paper on 2019 on improvements in spectral efficiency. https://mma.prnewswire.com/media/944546/CTIA_Spectrum_Efficiency.pdf?p=pdf

27 https://www.rcrwireless.com/20190201/software-defined-networking-sdn/verizon-puts-low-latency-together-5g-mec-and-an-intelligent-edge

28 https://www.itu.int/pub/R-REP-M.2410-2017

29 https://www.gsma.com/mobileeconomy/

5G's value relative to prior generations (i.e., value over and above traditional mobile service) may be realized through the support of Internet of Things (IoT). Examples of industrial IoT markets include Industry 4.0, which refers to the digitization of industry, including manufacturing, smart cities, and smart energy. McKinsey predicts that roughly half of 5G IoT connections for business will be for Industry 4.0 applications.[30] Manufacturing added $16.8 trillion in economic value in 2019.

Smart cities as a concept refers to the digitization of city infrastructure and services. It also can refer to greater civic transparency, such as putting city data sets online. Thus, rather than unfolding as one market, smart city deployments have varied from city to city, with measures ranging from digitizing parking meters to installing video cameras to connecting city fleets. Project emphasis varies from transparency to efficiency to safety. Accordingly it is more of a *project* market than a *product* market. Cities put out system projects for bid. From a business model perspective, this means that network operators would likely partner with enterprise IT firms and integrators with experience in smart city deployments. From a value at stake perspective, it means that network operator systems would support key city infrastructures. Market size estimates for the smart city market range from $1 trillion to $2.5 trillion. This by itself represents a larger market than that of traditional mobile service, though not all those expenditures represent network operator transport fees.

When consumers lose mobile service due to network outages, there are substantial consequences, such as inability to call loved ones or public safety. This is the lifeline aspect of traditional mobile service. There are also inconveniences — the inability to check email or social media accounts, for example. The consequences for service outages in public-sector markets and industrial markets would also be substantive. Service outages impacting city infrastructure could impair citizens' ability to access information or public services. Service outages impacting high value machinery would harm industrial processes. Thus, the potential for making inroads into new end markets represents both new value added, and also new value at stake, for 5G service.

Network equipment providers and mobile operators both confirmed the hypothesis that new applications meant new value at stake and therefore new risk to prepare for. One network equipment provider noted that service-level agreements are common in the enterprise IT world (and between network operators and network equipment companies), and anticipated similar agreements would be required in these new end markets.

30 https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/the%205g%20era%20new%20horizons%20for%20advanced%20electronics%20and%20industrial%20companies/the-5g-era-new-horizons-for-advanced-electronics-and-industrial-companies.pdf

Spectrum, Network Topology, and Risk

Wireless spectrum provides a tradeoff. Lower spectrum (e.g. 600 MHz and 700 MHz, formerly used for UHF broadcast TV) propagates further; it also has less bandwidth and can require a longer antenna at the receiver.[31] Conversely, higher spectrum often has more bandwidth but does not propagate as far (e.g. 4G LTE at 700 MHz versus 4G at 2.5 GHz), necessitating more dense networks. In practice, wireless carriers often use a mix of spectrum, with lower frequencies often used for wide-area coverage (such as for voice service on highways), and higher frequencies used to provide capacity in local defined areas. Operators may refer to network overlays and underlays.

For 5G networks, spectrum being made available by regulators falls into three categories: low-band, mid-band, and high-band (mm-wave). There are significant differences between each of these in terms of propagation characteristics, bandwidth, and network density. These differences have service model implications, summarized in Table 2.

Table 2: 5G spectrum bands.

CATEGORY	BANDWIDTH	RANGE	SERVICE MODEL IMPLICATION	TOPOLOGICAL IMPLICATION
High-band (mm-wave)	400 MHz (Verizon @ 28 GHz)	10s of meters (Wi-Fi range or less); will not penetrate walls	Fixed broadband substitute; local / portable wireless in small defined areas	Customer premise equipment (CPE) on the customer site; densified small cells on street furniture (lamp posts, bus stops, etc.)
Mid-band	40 MHz (from FCC 3.5 GHz band plan)	100s of meters; will likely have indoor propagation issues	Increased local capacity; potentially, Wi-Fi substitution	Higher densification than today's cellular networks
Low-band	5 MHz (T-Mobile @ 600 MHz)	10s of kilometers; will propagate through walls	Faster cellular	No need to further densify; also useful for rural areas

31 Frequency-to-wavelength calculators are available on the internet. A half wavelength can be used to estimate antenna length requirements.

Commentary

- Low-band: From a rapid go-to-market perspective, low-band spectrum is advantageous. Operators can launch 5G using the same cell sites they use today, with no densification of towers required. From a service model perspective, low-band spectrum at current bandwidths (5 MHz @ 15 b/s/hz) enables an extension of current mobile broadband service. It also represents a cost-effective way to cover exurban or rural areas.
- Mid-band: The spectrum allocated for mid-band is higher than that used for current cellular (700 MHz, 850 MHz, 1700/2100 MHz, 1900 MHz, 2.5 GHz in the US). 2.5 GHz networks have had challenges propagating inside buildings. Thus, similar indoor propagation issues could be predicted for mid-band service. This will necessitate network densification relative to current networks, and may necessitate on-site infrastructure to enhance indoor reception.
- High-band (mm-wave): As shown with Verizon's city-level rollouts (e.g. in Sacramento), service using high-band necessitates significant network densification, and also significant fiber deployment to support network traffic. For Verizon's Sacramento home broadband service, customer premise equipment (a 5G-to-Wi-Fi router on the windowsill) was deployed. This is similar to traditional home fixed broadband in terms of service economics and customer experience. Sending a technician to a customer's home for equipment installation adds to the upfront costs of acquiring a new customer.

The three deployment models have a variety of security implications, particularly in the high-band and mid-band scenarios. From a topology perspective, low-band is essentially an extension of current cellular service. Potential security implications are listed below.

Table 3: 5G spectrum bands and deployment model security implications.

CATEGORY	DEPLOYMENT MODEL SECURITY IMPLICATIONS
High-band	- CPE potentially requires customer cooperation, and installation may require entering customer premises. - CPE requires updates and patches of CPE software. - Dense network topology requires site acquisition, deployment, maintenance, and updates.[32] - Network infrastructure will be physically much closer to street (e.g. on lamp posts and bus stops) and potentially within physical proximity of passersby. - Low-latency service will require distributed core or compute closer to the user.
Mid-band	- Potential propagation issues could necessitate user CPE, in turn requiring deployment, maintenance, and upgrades. - Dense network topology requires site acquisition, deployment, maintenance, and updates. - Network infrastructure will be physically much closer to street (e.g. on lamp posts and bus stops) and potentially within physical proximity of passersby. - Low-latency service will require distributed core or compute closer to the user.
Low-band	- Low-latency service will require distributed core or compute closer to the user.

32 Analyst EJL Wireless estimates that 4000 high-band cell sites would be required to cover a city of 500,000. EJL Wireless report for City of Sacramento, January 2019.

In practice, as they have with prior generations, operators will provide 5G service using a combination of spectrum bands. T-Mobile USA, for example, describes a "layer cake" combining low-band and higher bands of service.[33] Verizon is planning dynamic spectrum sharing across multiple 5G bands.[34] Thus, while the emphasis on one band over another may vary by operator, the implications described above will generally apply to operators using multiple spectrum bands.

Network operators have been criticized for being slow to distribute upgrades or patches to smartphone OS. Thus, if an operator plans to deploy CPE or densify its network, it will have to develop the capability to rapidly distribute security patches to such equipment.

Analysts have estimated that 4000 high-band cell sites will be needed to cover a city with a population of 500,000. Extrapolating to nationwide "NFL city" coverage[35] would mean adding over 100,000 high-band cell sites. This necessitates site acquisition, equipment installation, monthly rent, support and updates, and expansion of fiber networks. AT&T's former CEO has publicly commented that the company would add over 200,000 cell sites for 5G deployment.[36] Site acquisition and deployment alone are tremendous undertakings. Operation and updates (e.g. of firmware) will be additional challenges. Acquiring the real estate for hosting 5G cell sites, especially high-band service, represents one of the largest potential obstacles to broadly available 5G service. Significantly densifying networks, such as putting cell sites on light posts, would put cell site infrastructure in closer proximity to passersby than traditional hilltop or rooftop cell sites.

Lowered latency is one of the most significant potential benefits of 5G networks. Achieving the 5G target of 4ms user-experienced latency has significant network design implications. Signals move at the speed of light.[37] With no delay (network, computational, overhead, etc.) whatsoever, assuming 2ms of "travel time" in each direction means the network core would have to be within 375 miles of the user. The core will have to be even closer when incorporating delay or more realistic deployment models, such as a user query going to the network core, then to an Amazon Web Services server, and back again.

From a security perspective, a distributed core could be beneficial. Outages could be more localized and contained. In a distributed denial of service (DDoS) attack, for example, outages

33 https://www.fiercewireless.com/5g/t-mobile-gets-5g-boost-from-2-5-ghz-nyc-layer-cake

34 https://www.rcrwireless.com/20200623/5g/verizon-completes-dss-tests-on-track-to-activate-this-year

35 In essence, cities large enough to support a major sports team.

36 https://www.vox.com/podcasts/2019/2/21/18233800/att-randall-stephenson-recode-media-peter-kafka-podcast-interview-5g-sports-nba-gambling-time-warner

37 As a rule of thumb, light travels at a speed of about one foot per nanosecond.

could be contained to a given metropolitan service area. For example, the current generation of smartphones launched in 2007, and Android in 2008. As Android devices grew more prevalent, in addition to a massive increase in mobile data, signaling traffic[38] also spiked. Network operators struggled to accommodate "signaling storms"[39] from smartphone applications. These were functionally equivalent to DDoS attacks (albeit over the control plane, rather than the data plane, as is the case for Internet DDoS attacks) and led to network outages, including broad regional outages.

5G is designed to support highly dense sensor applications if in URLLC mode. Such applications could create significant signaling traffic. Distributing core locally would help compartmentalize that traffic and keep any outages contained. From a service perspective, a distributed core could remove scalability bottlenecks, much as CDNs do on the internet by putting frequently-accessed content closer to the end user.[40] Interviews also indicated that at least one wireless carrier in the United States is embedding its point of presence (POP)[41] in the 5G base station to reduce latency. This puts interconnection between networks at a hyperlocal level, reducing latency.

Distributed core also can potentially open new attack surfaces. Previously, the network core was traditionally housed within secure network operator facilities. Distributed infrastructures, while lowering latency and potentially containing outages to a local level, increase the number of potential attack surfaces. Distributed core may be housed in leased buildings, or even in roadside cabinets, which could potentially open physical attack surfaces.

Network Slicing

Network slicing represents a major new capability of 5G networks. It is defined in the recently approved 3GPP Release 16, also known as 5G Phase 2. As the term implies, network slicing refers to segmenting operator networks into different "slices" to support different applications, while using the same wireless spectrum and physical network infrastructure. For example, consumer

38 Signaling traffic refers to traffic related to the control and management of the network, not the communications payload itself. Signals are sent over the control plane, not the user plane. A DDoS attack on the internet floods the user plane. Flooding the control plane for wireless networks has the same effect as denial of service.

39 https://arxiv.org/pdf/1411.1280.pdf

40 CDN: content delivery network. These are used for both fixed and mobile internet service. Frequently-accessed content is distributed to reside within local ISPs to minimize bottlenecks and improve customer experience.

41 POP: point of presence. Often used in reference to ISP and content delivery networks (CDNs). Can refer to a carrier hotel or colocation facility where networks interconnect.

smartphone traffic could be separated from industrial device traffic. Slices can have different quality-of-service levels; for example, slices can be designated for must-have, mission-critical applications versus consumer traffic. Networks can be optimized accordingly, rather than providing generic service. Realizing the benefits of network slicing necessitates upgrading both RAN and core to 5G, not just the RAN.

Table 4: 5G network slicing and security implications.

	DEFINITION	DEPLOYMENT CONSIDERATIONS	SECURITY IMPLICATION	OPERATIONAL IMPLICATION
Network slicing	Broadly, virtualization of one network into different slices, to enable partition of different types of traffic. For example, latency-sensitive traffic (e.g. telemedicine) can be separated from more latency-tolerant traffic (e.g. messaging).	Multiple virtualized networks on shared infrastructure. Network slices can be deployed on the same RAN and core. 5G core required. Different tiers of service based on traffic type and service level agreement. Partners and customers may vary by slice.	Separation and compartmentalization of traffic, processing, storage, management. Need to (or ability to) confine security issues within one slice. Potential need to handle different slices with different security levels with same device.	Ongoing zoning of traffic and data to ensure integrity of slices. Anomaly detection capability (should data be in a given slice). Policy consistency across slice roaming partners. Threat information sharing between slices (on the same network) and partners.

This list is only illustrative, not comprehensive. It does, however, identify a number of implications of network slicing for implementation and information sharing. What if, for example, a device attached to a mission-critical network slice roams overseas? Will roaming partners be able to provide the same level of service? Or, what if the device is an employee-owned smartphone switching between an enterprise slice (with rules around treatment of workplace data) and a consumer slice? Can suspicious devices or data be moved to a "quarantine" slice? Can the network proactively detect whether data belongs in a slice?

Roaming partnerships have existed for decades and are inherent to the appeal and utility of global mobile services. One can fly across the globe and enjoy wireless service without needing unique phones for each destination. Network operators or aggregators will have to add policy layers to their agreements consistent with network slice service-level agreements.
The potential benefits of network slicing are many. Slicing will help operators address the heterogeneity of applications and devices projected for 5G, and compartmentalize them at the appropriate service level. Any security issues can be restricted to one slice, rather than polluting the entire network.

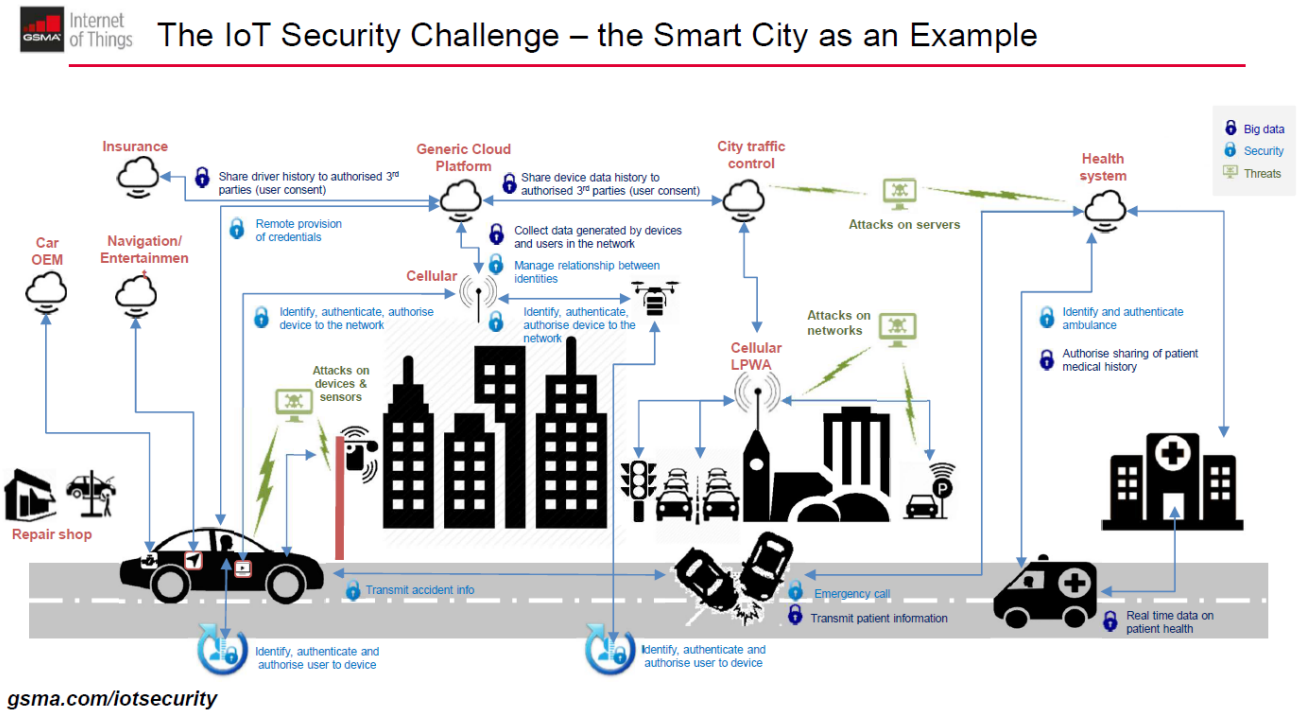

FIGURE 3: SMART CITY NETWORK AND DEVICE HETEROGENEITY[42]

As an example of device and data heterogeneity, this figure from GSMA illustrates the variety of devices a smart city deployment could entail. Note that not all connections shown are cellular. This further reinforces that 5G networks supporting city or industrial systems will have to be able to compartmentalize different categories of traffic, and potentially even contend with attacks injected from other forms of connectivity, such as Wi-Fi or low-power wireless systems.

In a private survey of network operators conducted for this research, the most frequently listed operator concerns were (1) the need to isolate different forms of traffic, and (2) the potential risks associated with attaching more heterogeneous devices to the wireless network. This is akin to risks seen in consumer IoT and industrial IoT, such as home cameras with poor security being hacked. A Nokia paper on 5G use cases confirms the importance of protecting networks from insertion of false information into a connected system.[43] This highlights the need for robust device approval processes and the ability to compartmentalize traffic from such devices. Education of partners and consumers also is imperative. What if a network customer attaches a poorly designed (i.e. not secure) device, simply because it is the cheapest

42 https://www.gsma.com/iot/iot-security/

43 https://www.ramonmillan.com/documentos/bibliografia/5GUseCases_Nokia.pdf

available? This raises the question of who is responsible for security in such a scenario, as the service provider with a customer relationship may be different from the network operator.

Network slicing and standards-setting

In March 2020, 3GPP delayed Release 16[44] (also referred to as 5G Phase 2) to June 2020 due to the impact of COVID-19, which limited the convening of meetings. Release 16 defines key capabilities like network slicing and ultra-reliable, low-latency communication. Release 16 was approved in July 2020.

Geopolitical tensions have also impacted the progress of standards development. In March 2020, President Trump signed the Secure 5G and Beyond Act of 2020. This makes multiple references to the role of standards-setting bodies.[45] Traditionally, standards bodies have focused on maximizing the reach and harmonization of standards; this enhances interoperability of wireless technology across the world, which in turn increases the utility of wireless service to end customers. Standards body participants are typically companies that are stakeholders in the wireless ecosystem. Prominent examples include Qualcomm, Nokia, and Huawei. Participants typically put forward technology to standards bodies like 3GPP so that they can be adopted worldwide. The standards development process requires harmonization of these proposals and eventual integration into one shared standard. The process is also a forum where companies that compete in the marketplace sit together to codify the technologies they will use in that competition.[46]

The question of whether governments should more closely observe standards body participation elicits a number of questions. Can a standards body (or participant, or observer) set limits on who gets access to a standard and any technologies included? Will standards body meetings start to look like trade negotiations, with trading blocs? The overall impact is to create ambiguity, which in turn can hamper deployment of 5G, and indeed the signing of Secure 5G and Beyond Act created questions about whether US companies could even participate in the appropriate standard-setting bodies. As of June 2020, the US Department of

44 https://www.3gpp.org/release-16

45 https://www.congress.gov/bill/116th-congress/senate-bill/893/text?overview=closed

46 Shapiro and Varian describe the dynamic of cooperation and competition in network industries as "coopetition." See Information Rules, Chapter 8.

Commerce has issued rules that will allow US companies to participate in the same standards body as blacklisted companies.[47]

Prior to 4G (LTE), multiple wireless standards such as CDMA and GSM coexisted, and what standard was in use varied around the world.[48] There were also national standards, such as TDS-CDMA, a 3G standard deployed in China, or PDC, a 2G standard primarily deployed in Japan. Global standards were only harmonized with 4G (LTE); a schism in standards could potentially lead to a return to multiple, coexisting standards. Such a split would cause a variety of diseconomies, such as adding additional components to devices so that they can roam, or consumers not being able to roam if their devices are not "global" devices. As noted above, networks, once deployed, persist. A schism in 5G could have decades-long impacts.

Other Security Improvements with 5G

Based on a review of 5G standards and capabilities, conference proceedings, and the work of security researchers, it is clear that significant effort has gone into mitigating security issues that were seen with prior generations. It is noteworthy that, even before the launch of 5G commercial service, outside security researchers had pointed out a number of security issues, thus giving industry an opportunity to rectify them in advance of ratification of standards and commercial deployments.[49] The increased ability of researchers to inexpensively emulate wireless networks is a factor behind this. Researchers have benefited from the availability of open-software implementations of the cellular protocol stack, for example, as well as software-defined radios. Security researcher white papers describe use of PCs, Universal Software Radio Peripherals, and the srsLTE open source project to emulate operator cellular networks.[50]

47 https://www.commerce.gov/news/press-releases/2020/06/commerce-clears-way-us-companies-more-fully-engage-tech-standards

48 As a rule of thumb, about 30% of the world used CDMA, and 70% used GSM.

49 https://softhandover.wordpress.com/2019/06/26/a-reflection-on-the-history-of-cellular-security-research-and-the-security-outlook-of-5g/

50 For an example of this test setup: Shaik, Park, Borgaonkar, and Seifert in *New vulnerabilities in 4G and 5G cellular access network protocols: exposing device capabilities*. Accessible at: https://dl.acm.org/doi/10.1145/3317549.3319728 Related presentation coverage: https://blog.3g4g.co.uk/2019/10/exploiting-possible-5g-vulnerabilities.html

Mobile device authentication

This paper has described two security benefits of 5G networks: network slicing, and the distributed core. An additional benefit is improved device authentication. This section describes security improvements in device authentication in 5G, and also shortcomings to those improvements that have been identified by security researchers.

When a cellular handset first connects to a mobile network, it goes through an authentication process to verify the connection. Up through 4G, this process is unencrypted, for a variety of historical reasons, until there is a cryptographic "handshake." A base station transmits information to a device (UE or user equipment) necessary to then register with the serving network through an authentication and key agreement (AKA) procedure. This information is sent without encryption.[51] While this process enables the base stations to authenticate devices, the devices are not capable of questioning the validity of the base station.[52] As a consequence, attacks using IMSI-catchers, i.e. fake base stations — some of which are quite inexpensive to make — have been broadly reported.[53] [54] Researchers have noted that the bulk of reported LTE security exploits are attributed to unencrypted pre-authentication traffic.[55]

All cellular handsets contain a universal subscriber identity module (USIM). Through 4G, this has been identified with the international mobile subscriber identity, or IMSI. Mobile handsets are associated with a home network operated by the network operator with which the subscriber contracted service. The actual serving network may differ. Thus, the authentication process allows the serving network to verify that a handset requesting service is associated with the network with which the serving network has a relationship.[56] This is how, for example, a US-based wireless subscriber can travel overseas and enjoy service without having to arrange for local service.

The home network provides the serving network with a one-time credential that is usable for authentication and generates a session key. The serving network then provides a temporary

51 https://www.usenix.org/system/files/sec19-yang-hojoon.pdf

52 Hussain, Echeverria, Singla, Chowdhury, Bertino, *Insecure Connection Bootstrapping in Cellular Networks: The Root of All Evil*, 2019. Available at: https://dl.acm.org/doi/10.1145/3317549.3323402

53 IMSI: International Mobile Subscriber Identity. An IMSI comprises a Mobile Country Code (MCC); Mobile Network Code (MNC); and a Mobile Subscription Number (MSIN).

54 IMSI-catcher: a device masquerading as a mobile network base station for the purpose of intercepting mobile phone traffic and identifying mobile handset location.

55 http://rogerpiquerasjover.net/5G_ShmooCon_FINAL.pdf

56 For further reading, the author suggests *Protecting IMSI and User Privacy in 5G Networks*, 2016, accessible at: https://dl.acm.org/doi/10.5555/3021385.3021415

identifier, or TMSI, to the authenticated handset. If no TMSI is available, the handset will fall back to providing its IMSI. This is what IMSI-catchers exploit. An active IMSI-catcher can send an identity request to all handsets connected to it, and thus capture those IMSIs. The use of IMSI catchers by law enforcement, the intelligence community, and foreign governments has been well reported.[57] Knowing a device's IMSI can enable monitoring of calls, SMS, and other traffic from the device, as well as tracking the device's location.

When a device connects to a network, it exposes a variety of attributes about itself, such as the type of device, manufacturer, and chipset (e.g. the baseband manufacturer and model). Researchers have identified three potential attacks based on these features:[58]

- Identification attack: discovery of software and hardware attributes of the targeted device
- Bidding-down attack: degrading the targeted device from a more advanced and secure standard (e.g. 4G LTE) to an older, less capable, less secure standard (e.g. 2G GSM)
- Battery drain attack: running down the battery of a targeted device

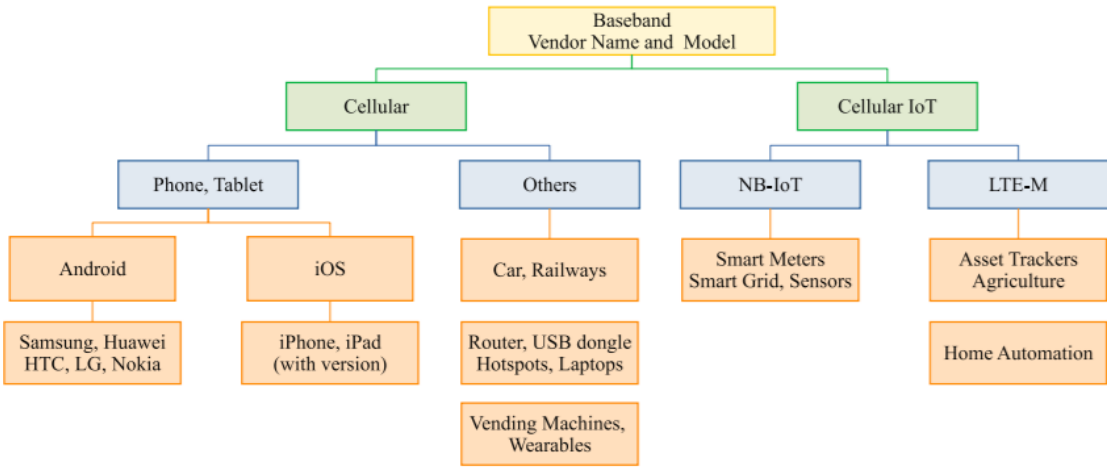

FIGURE 4: DEVICE-TYPE IDENTIFICATION LEVELS.[59]

Figure 4 (from Shaik, Park, Borgaonkar, and Seifert) shows the attributes an IMSI-catcher could capture about a device. All these factors can in turn be used to narrow the probability that a device is being used by a target of interest.

57 For specific examples of IMSI-catcher use by US law enforcement: https://www.aclu.org/issues/privacy-technology/surveillance-technologies/stingray-tracking-devices-whos-got-them

58 As described by Shaik, Park, Borgaonkar, and Seifert in *New vulnerabilities in 4G and 5G cellular access network protocols: exposing device capabilities*. Accessible at: https://dl.acm.org/doi/10.1145/3317549.3319728

59 Diagram from Shaik, Park, Borgaonkar, and Seifert in *New vulnerabilities in 4G and 5G cellular access network protocols: exposing device capabilities*. Accessible at: https://dl.acm.org/doi/10.1145/3317549.3319728

Based on the above, a unique or bespoke device (for example, running a military-grade OS build on an unusual chipset, in a certain geographic area) could be associated with a user of interest with some confidence, and targeted accordingly.

With 5G, the authentication process has been encrypted. Industry has implemented the Subscription Permanent Identifier (SUPI) in place of the IMSI from prior generations. Plain-text transmission of the SUPI is not allowed over-the-air interface, and, unlike with prior generations, plaintext transmission is not permitted even in the case of identification failure. A Subscription Concealed Identifier (SUCI) is used until the SUPI is verified. This is designed to prevent IMSI catchers. The SUPI is exposed in its entirety to the serving network.

Researchers have noted that, as of 3GPP Release 15, which defines much of mobile 5G, SUPI implementation is optional.[60] Further, these improvements do not prevent bidding-down attacks, which can be used to revert a 5G-capable handset to a prior generation.[61]

Researchers have also described activity-monitoring attacks that, over time, infer the Sequence Number (SQN) created after each authentication stored as part of the USIM. They have also identified possible exploits involving eavesdropping the SUCI and then fetching authentication tokens for that SUCI.[62] [63]

As a potential countermeasure, researchers have proposed that home network operators place less trust in their serving network partners. An example would be confirming with serving networks that a handset belongs to a legitimate subscriber of the home network, but not sharing the full identification number associated with the handset.[64] Currently, for reasons such as facilitating lawful intercept, the full SUPI is shared with the serving network. Another suggestion put forward would require base stations to transmit identifiers so that handsets can validate the base station, rather than trust the base station through a validation process. Such a measure could pose an interesting technical challenge, as devices would need to be capable of

60 http://rogerpiquerasjover.net/5G_ShmooCon_FINAL.pdf

61 Khan, Dowling, and Martin, *Identity Confidentiality in 5G Mobile Telephony Systems*, 2018. Accessible at: https://eprint.iacr.org/2018/876.pdf

62 Borgaonkar, Hirschi, Park and Shaik, *New Privacy Threat on 3G, 4G, and Upcoming 5G AKA Protocols*, 2019. Accessible at: https://eprint.iacr.org/2018/1175.pdf

63 Hussain, Echeverria, Karim, Chowdhury, Bertino, *5G Reasoner: A Property-Directed Security and Privacy Analysis Framework for 5G Cellular Network Protocol*, 2019. Available at: https://dl.acm.org/doi/10.1145/3319535.3354263

64 *Protecting IMSI and User Privacy in 5G Networks*, 2016. Accessible at: https://dl.acm.org/doi/10.5555/3021385.3021415

validating base stations before connecting to them. This could necessitate flashing devices with *a priori* information about cell sites deployed around the world.

The threat of bidding down a mobile handset to a less-secure previous generation of mobile technology points to the potential security benefits of a 5G-only network, one that does not need to support prior generations. For national carriers with massive customer install bases, this is a challenge. But for an area-defined network, such as on a campus, a 5G-only approach could provide more secure communications within the confines of that network. Licensed owners of 5G spectrum could provide a slice to a local network operator or partner, make spectrum available via a spectrum marketplace, or provide such private networks themselves.

Security researchers have noted that, while a 5G-only network would eliminate a bid-down attack, pre-authentication would still need to be encrypted to remove risk of attacks during the bootstrap process. An X.509 certificate has been suggested as a means of securing the authentication process.[65]

Implications and Opportunities

As described in the previous section, 5G provides multiple security improvements compared to prior generations. However, interviews conducted for this paper highlighted the potential for differing implementations of standards, and the importance of network operator awareness and practices.

Standards versus Implementation

Security researchers generally refer any exploits identified to both operators and standards bodies for remediation. This necessitates adoption by the relevant standards body (3GPP in this case) followed by implementation by network operators, which are collectively represented by GSMA.

65 Hussain, Echeverria, Singla, Chowdhury, Bertino, *Insecure Connection Bootstrapping in Cellular Networks: The Root of All Evil*, 2019. Available at: https://dl.acm.org/doi/10.1145/3317549.3323402

Standards describe requirements and recommendations for building a given system. Some attributes are required; others may be optional. Interviews for this paper highlighted the potential for variance in how network equipment providers implement standards. For their part, network equipment providers avoided any critique of fellow equipment providers and their development practices. However, the potential for supplier-to-supplier differences (for example, in software development practices) points to the need for operators (or their partners) to be able to test and validate network supplier equipment.

To the extent that equipment from one network equipment provider becomes regarded as more secure, the market could presumably then "solve" this issue by providing incentive for other network equipment providers to improve their security. Non-market forces, such as regulators, also are a factor, as has been shown in recent months with the decision by countries such as Japan and the United Kingdom to remove equipment from suppliers such as Huawei and ZTE from domestic operator networks.[66] [67]

Operator Awareness and Practices

Improved awareness: Interviews with network operators and their suppliers highlighted the need to improve awareness around security risks related to 5G. To one operator interviewed, network resilience referred to resilience to power outages and natural disasters (e.g. hurricanes such as Katrina and Sandy), rather than being resilient to the risks of cyberattack. Operators, particularly the former RBOCs (AT&T and Verizon), have well developed procedures to handle traditional outages, such as maintaining backup power sources at the cell site and prioritization of voice service and law enforcement. Another network operator noted that as a whole, the operational implications of diverse 5G service models were not getting sufficient attention, relative to the effort being put into marketing such applications.

Monitoring: One equipment provider commented that few network operators proactively monitor their networks for attacks. In practice, many rely on third-party tower operators (e.g. Crown Castle or American Tower in US) for cell site hosting, and network equipment providers (e.g. Ericsson) for network operation. Alerts about outages may be sent to tower operators or network equipment providers before being escalated to the network operator.

66 https://www.washingtonpost.com/world/asia_pacific/japan-effectively-bans-chinas-huawei-zte-from-government-contracts-joining-us/2018/12/10/748fe98a-fc69-11e8-ba87-8c7facdf6739_story.html

67 https://www.bbc.com/news/technology-53403793

New skills: Network operators and network equipment providers have indicated that the virtualization and software-defined networking described in 5G requires a new set of skills, and that this shift potentially plays more to the strengths of traditional enterprise IT players (e.g Cisco, Dell/VMware, Oracle). Security working groups in telecom recognize this, and note the skills IT enterprise security has acquired in years of addressing constant attacks, and in automation of response.[68]

Misconfiguration: Another risk is that of misconfiguration. A famous historic example of this is AT&T's long-distance service outage in January 1990. A switching station in Manhattan crashed, and the ripple effects led to broad outages nationwide.[69] The crash was ultimately attributed to a software bug. A network equipment provider commented that telecom networks have generally become fault-tolerant — that if one base station goes down, networks can re-route, and that industry has also learned to address physical attacks such as cable cutting. This provider also commented that cell networks have been brought down by simple syntax errors during configuration, and that the greater risk today lies in misconfiguration or more subtle attacks. The densification of networks planned for 5G — with the potential addition of an order of magnitude more cell sites, with an increased number of small cells and/or customer CPE — could increase the potential for misconfiguration.

Rapid security updates: Network operators have been criticized during the smartphone era for delays in pushing out security updates to Android handsets. If 5G networks do indeed support more heterogeneous devices, the ability to rapidly push out security updates to all connected devices, not just commonly used devices, will be essential. If the service provider is different from the network operator, then which party has responsibility for security patches needs to be made clear.

68 https://docbox.etsi.org/Workshop/2018/201806_ETSISECURITYWEEK/5G/S02_SECURITY_5G_INTER-NWK_SIGNALLING_/SECURE_INTERWKG_NWK_5G_SERVARCH_NOKIA_Holtmanns.pdf

69 http://www.mit.edu/hacker/part1.html

5G Service Deployment and Recommendations

In addition to commonly marketed capabilities such as increased throughput and lower latency, 5G provides a number of potential security improvements over previous generations, such as network slicing, distribution of the core, and improvements in authentication. We note that distribution of core infrastructure has both potential benefits (reducing impact area for outages) but also may increase the number of eligible attack surfaces.

As this paper has noted, however, getting 5G service to market quickly has encountered a variety of challenges, from delays in the development of relevant standards, to the need for operators to support incumbent customers using prior generations of cellular technology, to basic topographical and operational challenges, such as cell site acquisition. Another dependency is the availability of 5G spectrum.

Historically, major network operators have launched new generations of cellular technology in major metropolitan centers and then moved to the suburbs and more sparsely populated areas. The challenges of site acquisition have led to efforts by some network operators to share cell sites.[70] Third-party tower operators in effect provide the same service. It is likely that the dependence on partners like tower operators will continue with 5G. The capabilities of 5G, especially those provided by mid-band or high-band spectrum, are still gated by basic access to real estate.

However, private 5G networks, such as the campus networks described earlier in this paper, may be less subject to the same constraints. Cell sites hosted by municipalities are subject to review and public comment.[71] This process can be lengthy, and led to the Federal Communications Commission promoting "shot clocks" to put a cap on how long the approval process could take after a network operator submits an application to install a cell site.[72] Private 5G networks hosted by enterprises or campuses would likely be more amenable to hosting of 5G infrastructure for their own purposes. This may also align 5G service with the delivery of enterprise IT service. Further, as operators look to build reference cases for the

70 The three incumbent carriers in Japan plan to share cell sites.

71 Traditionally concerns from municipalities have including environmental, aesthetic, and safety concerns. More recently, misinformation related to 5G networks has further complicated this review process.

72 https://www.commlawblog.com/2012/01/articles/cellular/fcc-shot-clock-presumptions-for-wireless-tower-permitting-upheld/ Also https://www.fcc.gov/document/fcc-facilitates-wireless-infrastructure-deployment-5g

value of 5G networks, private or locally defined 5G networks provide a way to develop use cases that other customers can review and learn from. 5G-only networks that also address pre-authentication security issues would enable operators to provide more secure service than has been possible with prior generations.

It is noteworthy that many announced private or local-area 5G networks involve solution partners with specific domain expertise, such as in industry verticals. In the enterprise IT world, this is common; IT equipment is often sold through value-added resellers or system integrators. Given the breadth of verticals potentially addressable by 5G service, such as manufacturing or health care or enterprise campuses, network operators will need different solution partners to address different verticals. It is the author's view that 5G service delivery and the historical enterprise IT market will start to overlap: operators will still serve consumers and business directly as they traditionally have with mobile broadband service, while also partnering with solution partners to address specific verticals. This will also be beneficial from a security perspective.

Starting hypotheses for this paper were:

- Use cases that include cities and industry will have higher value at stake than traditional consumer cellular, increasing the consequences of potential outages;
- Network densification required to provide low-latency service can create both physical access risk and necessitate rapid software update capability;
- More heterogeneous use cases can create new vectors of risk that have not been traditionally faced by network operators.

Research and interviews have confirmed these hypotheses, as well as means by which operators can adapt to these risks.

This paper recommends that:

- Operators, their partners, and their customers investigate the viability of 5G-only service;
- Operators and their partners develop the ability to rapidly deploy software updates, including security patches, to small cells, customer premise equipment, and other connected devices;
- Operators and their partners develop the ability to rapidly test and verify devices from new partners from outside of the traditional telecom ecosystem;

- Policymakers act to facilitate rapid deployment of 5G networks, including implementing policies to facilitate cell site acquisition;
- Policymakers recognize the role of global standards bodies; rapid standards development; and the economic value of globally harmonized standards.

For a variety of factors, such as spectrum holdings, there has been variance in how operators have gone to market with 5G service, especially when compared with prior generations. Holders of low-band spectrum, like T-Mobile, have acted to quickly maximize national coverage. Holders of high-band spectrum, such as Verizon Wireless, have rolled out on a city-by-city basis. This is likely because the wide-area coverage of T-Mobile's low-band spectrum means it can deploy 5G using current cell sites, whereas Verizon's high-band spectrum necessitates significant network densification (and commensurate deployment of fiber) and cell site acquisition.

While this variance is potentially frustrating for consumers and device makers, it is perhaps fortunate from a security perspective. Each new market allows operators to hone their craft and become more efficient with the next rollout. The market is still early in its development. It is the author's hope that the recommendations in this paper can be of value to operators as they build out their 5G services.

Acknowledgments

This paper was written with support from the UC Berkeley Center for Long-Term Cybersecurity. Heather Blanchard, Jay Goldberg, Rob Hull, Roger Piqueras Jover, Nitin Shah, and David Witkowski were generous with their time and feedback. Steven Weber, Ann Cleaveland, and Chuck Kapelke with CLTC made this a much better paper. Various network operators, security researchers, analysts, and equipment providers made themselves available on an on-background basis. This paper benefits from their time and insight.

About the Author

Jon Metzler is Lecturer at the Haas School of Business at the University of California, Berkeley, where he teaches on competitive strategy; competitive advantage in technology, telecom, and media markets; and international business. Jon teaches at the undergraduate and MBA levels. Jon is also associated faculty for the UC Berkeley Center for Japanese Studies. He has received research support from the UC Berkeley Center for Japanese Studies and the UC Berkeley Center for Long-term Cybersecurity. Jon is also faculty mentor at Berkeley Skydeck, an accelerator for Berkeley-affiliated startups. Jon has taught at UC Berkeley since 2014.

In 2008, Jon founded Blue Field Strategies, a consulting firm helping infrastructure clients such as network operators accelerate service innovation. Prior to founding Blue Field, Jon was director at Rosum Corporation (acquired by TruePosition), a pioneering location technology company augmenting the reach of GPS indoors. At Rosum, he was responsible for business development, government affairs, public relations, and standards. He secured development funding from DARPA and managed a multi-state E911 location trial. The 2006 Communications Act incorporates Jon's advocacy in E911 issues. He represented Rosum in the National Emergency Number Association (NENA) and Advanced Television Systems Committee, and successfully advocated for the adoption of what is now called ATSC Time, used in DTV broadcast networks today.

Contact: https://haas.berkeley.edu/faculty/metzler-jon/
@jonjmetz

CLTC

Center for Long-Term Cybersecurity

UC Berkeley

Center for Long-Term Cybersecurity
cltc.berkeley.edu
@CLTCBerkeley