# Chapter 99
# Resilience of 5G Mobile Communication Systems to Massive Disruptions

Tibor Cinkler, Akos Ladanyi, Jacek Rak, Christian Esposito, Gianluca Rizzo

## 99.1 Introduction (Challenges to Resilience of 5G Systems)

It has been around 40 years since the introduction of the first generation (1G) mobile communication systems followed by deployment almost every next ten years of the second (2G), of the third (3G) and of the current fourth-generation architecture. 4G networks providing service to many bandwidth-demanding applications, including gaming or voice/video transmission or [4], have become an indispensable part of our everyday life. We are now entering the era of 5G systems expected to offer the users the communication possibilities significantly exceeding those known for LTE systems concerning, e.g., the effective data rate, the latency, the energy consumption and the availability.

It is widely expected that the effective data rates seen by individual users in a 5G network should be up to 10 Gbps, which is a hundredfold increase compared to 4G communications to mitigate the explosive data traffic growth [6]. The end-to-end latency should, in turn, go down to about 1-5 ms (which denotes an improvement close to one order of magnitude compared to 4G systems) to be able to respond to

Tibor Cinkler
Budapest University of Technology and Economics, HU; Gdansk University of Technology, PL, e-mail: cinkler@tmit.bme.hu

Akos Ladanyi
Budapest University of Technology and Economics, HU, e-mail: ladanyi@tmit.bme.hu

Jacek Rak
Gdansk University of Technology, PL, e-mail: jrak@pg.edu.pl

Christian Esposito
University of Napoli Federico II, IT, e-mail: mailto:christian.esposito@unina.it

Gianluca Rizzo
University of Applied Sciences of Western Switzerland – HES SO Valais, CH,
e-mail: gianluca.rizzo@hevs.ch

stringent requirements of critical services such as Industry 4.0, V2V/V2I communications and edge-cloud solutions.

Following the green technology paradigm, the overall energy consumption of 5G systems is anticipated to be reduced by 90% compared to the reference 4G systems. Another expectation refers to the improvement in terms of the lifetime of device batteries [4] which may be challenging in the context of high data rates and the global connectivity concept driven by the deployment of new communication schemes, including the IoT, D2D / M2M, as well as the almost full coverage assumed anytime and anywhere [16].

Following e.g., [4, 6, 18, 29, 7, 2, 9], the increase of performance expected for 5G systems may be reached by:

- utilization of massive MIMO and NR for enhancements in spectral efficiency (with the reference value of spectral efficiency in the range of 2-3 bits/s/Hz per macrocell),
- migration to the millimetre wave (mmWave) spectrum in the 3-300 GHz range,
- traffic offloading denoting performing many operations closer to the user (edge computing).

As the size of cells is (apart from bandwidth and spectral efficiency) one of the dominant factors concerning the effective capacity, improvements of data rates in 5G systems are expected to be achieved by introducing the small cells characterized by lower transmitter power (as add-ons to macrocells) to reduce the interference [18] and shifting to the 3–300 GHz frequency range [4]. However, 57–64 GHz and 164–200 GHz bands are excluded due to oxygen absorption and water vapor absorption problems, respectively, as shown in Table 99.1.

Table 99.1: Availability of millimetre-wave spectrum for 5G communications (bands potentially available for 5G communications are marked by *****)

| Frequency Band $[GHz]$ | Description of the use and of the limitations of the considered Band |
|---|---|
| up to 3 | *(Used by current wireless communications)* |
| 3 – 57 | ***** **Potential band for 5G communications** ***** |
| 57 – 64 | — *((Band excluded due to Oxygen Absorption))* — |
| 64 – 164 | ***** **Potential band for 5G communications** ***** |
| 164 – 200 | — *((Band excluded due to Water Vapour Absorption) )* — |
| 200 – 300 | ***** **Potential band for 5G communications** ***** |

However, as discussed in [34], despite possible high data rates, millimetre-wave frequencies are characterized by high atmospheric attenuation. In particular, communications at frequencies over 3 GHz are susceptible to weather-induced disruptions such as rainfalls, fog, or icing, which, depending on the intensity of weather factors and link frequency, are likely to substantially reduce the effective capacity of multiple wireless links during the time of a given weather factor influence [34, 28, 21]. Therefore, serving the 5G mobile traffic under weather disruptions seems questionable without any weather-oriented solution.

It is common that weather events often impact a specific geographic area, and thus result in *region disruptions* affecting multiple 5G links in a correlated manner located in a given region [34]. Although duration of weather disruptions is typically shorter than of failures following from natural disasters (e.g., hurricanes, or earthquakes leading to permanent failures of the networking equipment [30, 12]), even these relatively short periods of time of weather disruptions, if occurring frequently, can have a remarkable negative total impact on the availability of communication services [34].

The chapter is structured as follows. In Section 99.2, we explore the issue of assessing the resilience of a specific 5G network deployment. In Section 99.3 we present the Frequency Fallback technique for enhancing availability and survivability of 5G services, and in Section 99.4 we discuss the Segment Interleaving approach for enhancing the resiliency of the communications between base stations and the rest of the network. Section 99.5 introduces those approaches to resiliency based on the redundancy in infrastructure across different operators serving a same set of users. Section 99.6 describes how energy-aware techniques for dynamic management of base station configuration and of load distribution can play a key role in maximizing network survivability in case of disasters. Finally, Section 99.7 concludes the chapter.

## 99.2 Dependability Assessment of 5G Networks

The 5G vision encompasses a series of novel technologies and solutions based on the network softwarization aiming at providing ultra-fast, efficient, secure and reliable communications means, by dealing with the increasing data sharing requirements imposed by the novel ICT concepts of the Internet of Things, Smart Cities and/or Smart Factories. The Software Defined Networking (SDN), Network Function Virtualization (NFV), Network Slicing (NC) and so on represent those technologies realizing such a network softwarization by letting the network be programmable, software running on commodity hardware, or defining virtual networks sharing the same physical infrastructure. Within the context of these technologies, a series of novel mechanisms have been designed so as to theoretically cope with failures and/or attacks by supporting a slightly degraded QoS than the nominal conditions.

As 5G will be an enabler of many novel applications and a booster for some of the existing ones, where some of them are particularly mission/business critical (such as those related to healthcare, infrastructure protection or manufacturing), it is of pivotal importance to have empirical evidences of the provided dependability degree. In particular, providing the adequate levels of disaster-resilience is particularly crucial for future network paradigms, such as 5G [11], and having solutions for testing the resilience of a 5G deployment is needed. Specifically, such a test could certify the provided resiliency degree and verify if the demands are met and if some further enhancement mechanisms should be introduced.

Unfortunately, the current literature is scarce on resilience assessment for 5G. However, given the strong relationship of disaster-resilience with dependability [20], in the rest of this section methods for dependability assessment in the context of 5G will be presented, so that in future research activities they may be exploited and evolved for campaigns of disaster-resilience assessment [5].

Dependability is defined as the ability of a system to provide a certain service or functionality whose behavior can defensibly be trusted within a time-period [38]. It is a complex concept made of multiple attributes, i.e., availability, reliability, safety, maintainability and security (limited mainly to the integrity and availability aspects, neglecting the confidentiality). Dependability assessment consists in a test of the provided dependability by means of qualitative or quantitative measures. In fact, on the one hand availability and reliability are typically quantified by means of objective measures by using model-based methods or fault-injection. On the other hand, the other attributes are verified in a more subjective manner by analyzing the internal design of a system.

Within the context of the 5G technologies mentioned above, it is important to determine the dependability in a rigorous manner, and to determine their behavior in the case of possible faults and/or attacks. To this aim, some recent related works have been focused on the applications of quantitative methods of dependability assessment to SDN, NFV, NC and other 5G enabling technologies [14, 35, 27, 23, 25, 26, 10, 24, 19, 36].

In SDN, there is a clear distinction between the data plane, where routing and other networking mechanisms are performed, and a control plane, where critical control functions are run so as to decide about the network behavior and are offloaded to a software entity known as the SDN controller, rather than the traditional hardware components [39]. The dependability of the SDN is mainly bounded to the one exhibited by the control plane and the SDN controller itself [14]. To this aim, the centralized architecture of a single SDN controller has been progressively left behind for a more effective approach where the controllers are federated or assuming a distributed approach [40]. However, despite the benefits brought by the softwarization, SDN controllers may be affected by bugs or software related outages, which represent a severe limitation for the application of SDN to industrial contexts. The authors of [35] presented the application of software reliability growth modeling (SRGM) to determine the reliability of software components being used to device SDN controllers. SRGM consists in detecting and removing software-related faults during the early usage of the software, where bug manifestation and corrections are treated as a stochastic process parametrized on past experiences, and used to quantify reliability metrics so as to determine when a given softwarised SDN controller is matured to be deployed and put in use. In [35] the framework of SRGM is applied to some open source implementations of SDN controllers by using real data on software-related failures within SDN controllers.

A stochastic reward net has been applied in [27] for the assessment of the fault tolerance capabilities of SDN in the case of hardware and software faults, based on the steady-state availability under default parameters. This is an example of the application of the theory of the Petri nets and other mathematical modeling languages on the prediction of faults and/or the assessment of the dependability in case of various kinds of faults [23]. The basic idea is to model all the components of a given architecture in terms of a directed bipartite graph representing transitions and pre- post- conditions on the transition occurrence, where token moved within the net alongside the transitions and used for the quantification of certain metrics of interest. In [27], the overall hardware and software components are described, and the authors found that the network topology is a key aspect to consider when high dependability must be provided. Moreover, certain mechanisms, such as VM live migration, and a proper infrastructure configuration allow achieving high fault-tolerance features.

In [25], another example of the modelling of SDN is proposed by applying the Markov models to describing the whole SDN backbone, and found that a key aspect to tolerate hardware and software faults is to have a high replication of the processors in the cluster composing the SDN controller, and the Operation and Maintenance (O&M) failures (i.e., those related to the interface of the SDN switched receiving commands from the controllers) are the ones with the highest impact on the backbone availability. While, in the extension of this work presented in [26], the failure correlation is also studied, revealing the current SDN architecture being particularly vulnerable than traditional networks.

Aside the assessment of SDN controllers, also the NFV orchestrator found quite some attention, and its vulnerability to overload, software-related bugs and design faults have been recent objective of an intense investigation [13]. The work in [10] exploited a different dependability assessment method called software fault injection [24], where the deliberate injection of software faults are performed on the software component under test within the context of different NFV implementations. Specifically, for each physical and virtual entity in an NFV architecture, the root causes casing unavailability (i.e., entity being unresponsive), corruption (i.e., data managed by the entity being incorrect) and delay (i.e., entity being slowed down) are determined, and for each of them a fault model is provided by identifying how to emulate it by means of Software-Implemented Fault Injection (SWIFI) techniques [3]. It consists in emulating hardware and software faults by injecting their expected effects on the target software component. This work demonstrated that hypervisors are more dependable than container-based virtualization thanks to more sophisticated fault-management mechanisms. In fact, Docker leverages on the fault detection mechanisms of the Linux OS, which are less mature, making the NFV implementers exploiting it to add additional monitoring capabilities than the one provided by the OS. In addition, the proposed framework is a powerful tool in hand of the NFV providers to carefully configure the overall infrastructures to provide the suitable dependability degree meeting the customer needs in terms of availability and reliability.

Recently, also network slicing [41] has been studied with the intent of testing the isolation among the various slices, by determining if the fault propagation among the slices is avoided or not. Some works have been investigated the problem, but such an aspect is not mature and theoretical studies are currently being conducted by several research groups [19, 36], since it is still an open issue how orchestrating a network slice across such different proprietary virtual platforms by considering dependability needs of the various slices. Therefore, on this topic the research is still related to the ways of achieving such a feature, rather than verifying one of its open-source implementations.

The high level of reliability, performance, and other non-functional features promised by the technologies under the umbrella of 5G vision are not only related to the softwarization of the network, but also to a rethinking of the RF communication means. In fact, the adoption of mmWave spectrum with phased array antenna technologies has on the one side the benefit of an higher bandwidth, but has also certain weaknesses towards interference and corruptions, which will severely reduce the overall reliability. mmWave signal are easily attenuated by the rain and other atmospheric phenomena or RF pollution within the urban context, so that the signal level to the receiver can easily drop below a given signal to noise (SNR) to make the carried information corrupted or not understandable. This is exacerbated in particular conditions caused by natural disasters, like severe flashing or rains, as described in section 99.3. As an example, 3GPP Release 15 introduced some reliability improvements over 4G LTE aiming at significantly reducing packet data corruption and errors [42]. A proper assessment is needed to understand if the used wireless communication technology is able to offer the desired level of dependability, despite adversarial normal conditions or sudden exceptional situations (unintentional faults or intentional attacks). The starting point is the modelling of ultra-reliable low-latency communications (URLLC) as a repairable system based on a continuous time Markov process with a finite discrete state space, as done in [44], or by using similar reliability modelling formalisms and tools. Consequently, a reliability analysis should be conducted by computing the failure rate and failure mode of the elements of the URLLC solution so as to calculate the overall reliability prediction. Last, an on field statistical campaign on the real implementation of the solution under investigation is the last step to verify and validate the reliability study, by collecting statistics on the Bit Error Rate (BER) or Packet Error Rate (PER) during the use of the solutions by assuming a benchmarking application [45]. Such an assessment has the twofold aim of demonstrating the reachable dependability levels (and verify if the communication solution cope with an application requirement in terms of performance, reliability or availability), but also to find possible week points and propose dependability reinforcements to deal with them. The application of this methodology at 5G is still at its early stage, and few attempts and research results are being proposed recently, such as in [46, 47].

Last, 5G will witness to a radical change also at the hardware perspective, where telcom operators will shift from proprietary devices to cloud-hosted telecommunication services, and this will have a tremendous impact on the provided dependability. A recent research topic consists in the dependability assessment of this Telco clouds, and some initial results on such a result are available in [43]. In this work, a SDN-NFV architecture is assumed for Telco clouds and modelled by using Colored Petri Nets, the procedures that generate the most traffic load on the network has been modelled by using the queuing theory, and the respect of the three availability levels for telecommunication systems specified by ETSI have been studied. More research is needed by moving from a simulated/emulated enviornment to a real deployment, to include more failures types and to enlarge the investigation to also security and the protection against various kinds of possible attacks.

## 99.3 Frequency Fallback under Atmospheric Disruptions

The dependence of the attenuation due to water vapour in the air on the frequency of radio signals has been defined by ITU-R in Recommendation P.676-9 [17]. In particular, attenuation due to water vapours is two times higher than the respective attenuation for the case of dry air at frequencies just above 10 GHz, at pressure of 1013 hPa, temperature of 15 degrees Celsius and water vapour density of 7.5 g/m$^3$. For higher frequencies, the attenuation is even higher (for instance, at frequencies slightly above 20 GHz this attenuation is nearly twenty-fold) [17]. It is also worth mentioning that other forms of water, e.g., fog, mist, rain, rain drizzle, shower, sprinkle, snow, or thundershower may also have an adverse effect on attenuation.

According to the 3GPP Release 15 [1] defining the 5G New Radio (NR) mobile communications standard, new spectrum bands are reserved for NR ranging from 2.5 GHz to 40 GHz. In particular, the 5G NR will use multiple bands ranging from unlicensed ISM bands via licence-assisted use, through all the frequencies used by 4G LTE and range up to 40 GHz. As explained in the introduction part of this section, frequency regions with the excessive molecule (e.g., $O_2$) and particle (e.g., water vapour) absorption are excluded from 5G considerations.

To reduce the negative impact of water vapor on 5G communications, the fallback from 5G to LTE, LTE LAA (Licensed-Assisted Access), Wi-Fi or any other radio technology can be performed.

In particular, LTE typically uses 44 channels for both up- and downlink communications in the 452.5 MHz–3800 MHz range. LTE LAA (Licensed-Assisted Access), in turn, currently uses a frequency band 455–555 MHz within the ISM unlicensed spectrum of the 5 GHz band. Twelve ISM bands that do not require any licence are used for Industrial Scientific and Medicine applications. These are in the range of 6.765 MHz–246 GHz. However, in practice, only small bands from this range are used (only the frequencies up to 5.875 GHz are defined for terrestrial mobile and data communications; while frequencies above that are defined for satellite communications). Wi-Fi, in turn, typically uses distinct radio frequency bands at 900 MHz 2.4 GHz, 3.65 GHz, 4.9 GHz, 5 GHz, 5.9 GHz and 60 GHz [15]. Table 99.2 presents information on attenuation for exemplary frequencies discussed in this section.

Table 99.2: Attenuation due to water vapour at specific frequencies [17].

| Frequency $[GHz]$ | Attenuation $[dB/km]$ |
|---|---|
| 1 | 0.055 |
| 2.4 | 0.07 |
| 3.8 | 0.08 |
| 5 | 0.09 |
| 10 | 0.14 |
| 40 | 1.3 |
| 60 | 14 |
| 246 | 4 |

For scenarios of QoS flow quality deteriorated by atmospheric impairment, we propose to use the Frequency Fallback scheme in which communication is handed over to a channel of lower frequency, for which the atmospheric impairments affect the performance of a network to a lesser extent.

Possible frequency fallback schemes include:

– *intra-gNB intra-RAT Frequency Fallback* is the FFB (Frequency FallBack) within a single NR gNB (**N**ew **R**adio next **g**eneration **N**ode **B**) using the same RAT. This is the simplest scheme. It assumes that the RAT is the same, the gNB is the same, only the frequency band changes. In 5G, the CA (Carrier Aggregation) will be used. Carriers in distinct frequency bands are aggregated allowing smooth transition anytime. As an additional example, LAA (Licensed Assisted Access) must be mentioned. LAA is already part of the LTE Advanced Pro. Currently, it uses the 5 GHz unlicensed band and it is expected to be used in 5G as well where further frequency bands will be added,
– *inter-gNB intra-RAT fallback* is when using dual connectivity that allows the UE to be connected to two different gNBs via two distinct radio links at the same time using even quite distinct frequency bands (referred to as spectrum aggregation) to proceed with a handover of a given UE from one gNB to another. A similar scheme can be applied when macro cells and small cells of the same technology are used in the same network,
– *intra-gNB inter-RAT fallback* is when changing the technology (RAT) from, e.g., the 5G to the 4G or to the 3G or even to a non-3GPP technology such as Wi-Fi or any other while remaining connected to the same geographic base transceiver station,
– *inter-gNB inter-RAT fallback* is changing both the base transceiver station at different geographic location and the radio access technology. Examples include, e.g., selection of a 4G LTE eNB or a Wi-Fi HotSpot at different location instead of a 5G NR gNB, or utilization of Fixed-Mobile Convergence (FMC) for

a smooth handover from macro-cells to home-connected small cells or Wi-Fi access.

To conclude, FFB can be applied by decreasing the transmission frequency via changing to a channel of lower frequency or changing the technology, while keeping the technology and geographic location for the access unchanged.

Both the deteriorating weather conditions and all the kinds of disasters can cause the above fallbacks. If there is an increased water vapour concentration in the air (e.g., rain or fog), it typically affects a certain geographic area, and it commonly moves. Within the affected area the frequency and the RAT can be changed to minimise the impairments, however changing the gNB has sense at the edge of the area and close to the edge of the area. As this high vapour concentration area moves, the gNB connectivity changes dynamically.

In the case of disasters, a part of the network looses either power or connectivity (or both). If the area affected is not too large, 5G Coverage Extension can be used. Some examples of disasters where the affected area is such that coverage extension can help are as follows:

– close to the coast line in the case of a tsunami,
– close to a volcano in the case of eruption or along the lava flow,
– a strip of few tens or hundreds of meters in the case of a hurricane,
– a strip along a river in the case of a flood.

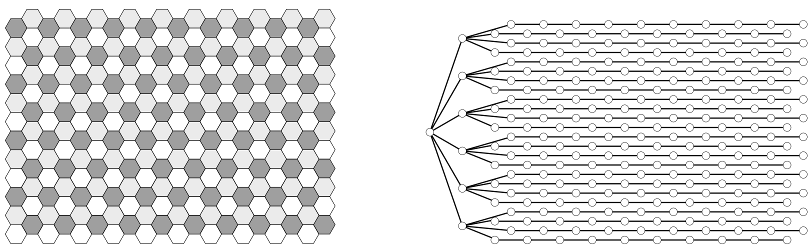
In all the above cases, these are not random single failures, but multiple failures, that are mutually correlated and geographically concentrated or aligned. 5G Coverage Extension can be supported by Dual Connectivity, Carrier Aggregation and Beam Forming via Massive MIMO. It is inevitable that in all cases network capacity will drop. Therefore, in some cases preemption may be used, where higher service classes (i.e., more critical services) can take network resources of lower classes of service.

## 99.4 Backhaul Segment Interleaving

High costs of deployment of a large number of service provider sites (such as 5G high-frequency small cells) visibly reduce the affordable level of redundancy in the network architecture. One of the possible enhancements of resilience in this context is a deployment of interleaving parts of the backhaul networks making it possible for service areas of cells affected by the failure to remain covered by other unaffected cells. Such a solution can naturally lead to reduced capacity offered in the affected areas, which is, however, obviously still better than no service at all.

In this section, we focus on improving the resilience of a 5G system for the case of failures occurring in the backhaul / metro-access part of the network. The Backhaul Segment Interleaving (BSI) technique is explained here for the example architecture shown in Figs. 99.1a and 99.1b consisting of a Passive Optical Network (PON) used as a backhaul for a set of 5G cells providing service in the residential area. In particular, Fig. 99.1a presents 24 rows (streets), with ten cells in each, resulting in 240 cells in a regular hexagonal grid.

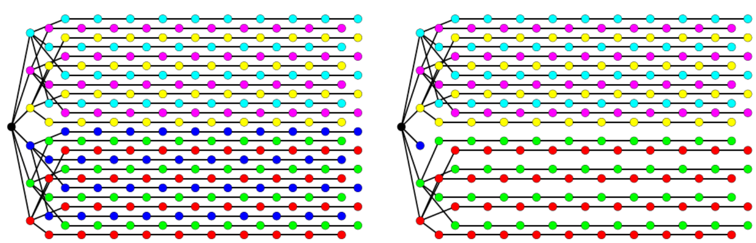
Figure 99.1b, in turn, shows the PON part of the network connecting the cells to the core network. The leftmost node in Fig. 99.1b represents a six-port optical line terminal OLT, while its six direct neighbors denote six respective splitters. The other nodes represent ONUs connecting cells to the PON. Each OLT port in Fig. 99.1b is thus to serve four streets (40 ONU/cells).



(a) The Cells to be interconnected (b) The one-side connected PON

Fig. 99.1: Simulation scenario for BSI: The cells and the PON that connects them to the core network.

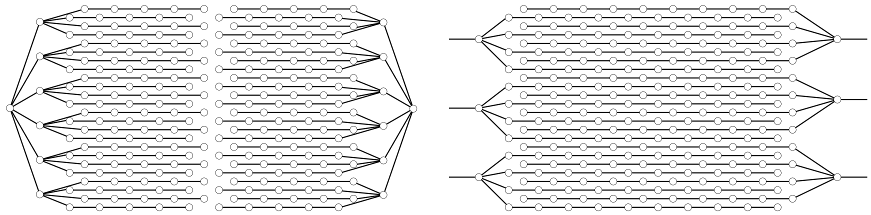
A disadvantage of this scheme is that a failure in the PON cuts off a whole tree rooted at an OLT port, which is equivalent to a massive failure of all cells connected to that tree (i.e., this is also a correlated multi-failure case). This interleaving approach can be very efficient in case of a disaster with small or narrow area affected. To enhance resilience in such a scenario, we propose to form a tree in a way that the cells of a single tree are physically separated, as shown in Fig. 99.2a.



(a) no failure (b) after failure

Fig. 99.2: Topology of a PON with BSI.

After a failure of an OLT (say OLT 3, counted from bottom up, marked in dark blue in Fig. 99.1b), availability of a network from Fig. 99.2a is less affected than of the network from Fig. 99.1b since the failed cells are spread out, and the streets with the failed cells are surrounded by other cells which belong to two different PON trees. Service availability is, in turn, significantly affected in Fig. 99.1b due to a large continuous area remaining uncovered after a failure of the OLT 3.



(a) two-side connect (b) Interleaved two-side connect

Fig. 99.3: One-side and two-side PON connection of cells to the core network.

Figure 99.3 shows an approach to further enhance the availability of the network upon a multi-failure. While the cable cost is increased due to the increased cable length by a few percent only (up to 5-10 %) when instead of the one-side connectivity (Fig. 99.1b) two-side connectivity (Fig. 99.3a) is used, or rather the Interleaved two-side connectivity (Fig. 99.3b), the availability and quality (throughput) enhancement can be significant. The two-side scheme of Fig. 99.3b can be combined with any of interleaving schemes, e.g., with the one shown in Fig. 99.2a to further enhance the availability of the infrastructure.

BSI scheme is of interest in case of any failure that affects a smaller area or if it has a narrow shape, that can be any long. The system capacity upon a failure is typically reduced.

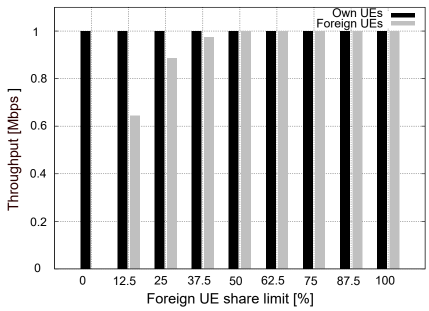
## 99.5 Multi-Operator Protection

The resilience of 5G systems can also be enhanced by using the resources of other network operators providing services in the same considered area. Indeed, if there is more than one network operator in a given area, then they can offer backup services to their foreign customers in failure scenarios affecting any one of these networks.

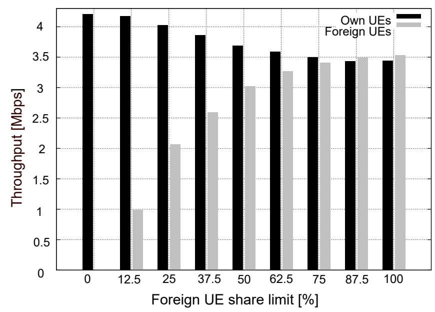
In this section, we present the evaluation of the characteristics of the considered Multi-Operator Protection (MOP) scheme assuming that there are two operators providing service in a given area. The first one is supposed to be operational, while the second one is supposed to have failed, implying that user equipment (UE) units of the second operator have to be attached to the first operators network in this post-failure scenario. These “roamed” UEs are here referred to as foreign UEs. The number of own UEs is the same in each network. In a post-failure situation, the resource usage of foreign UEs is limited. This is modelled by a resource restriction parameter ($p$) defined as a percentage of the Resource Blocks (RBs) that can be allocated to these foreign UEs by the operational operator.

It is thus clear to see that for UEs of the failed network, the only possibility is to connect via the network of the other operator and that the resource usage for these UEs is restricted. In particular, in simulations there were five cells, the number of UEs was 40. The throughput requirement of each UE was either $1 Mbps$ or

5*Mbps* (for the low and high utilization case, respectively). Results of evaluation are presented in Fig. 99.4.




(a) low utilization




(b) high utilization

Fig. 99.4: Throughput (a) for low ($1 Mbps/user$) and (b) for high ($5 Mbps/user$) network utilisation (load) as a function of allowed resource share of foreign users ranging from 0 to 100 % in case the foreign network fails.

Figure 99.4a shows that in the case of low network load (for UE requirements defined as 1*Mbps*), after a failure of the second network operator, when providing the backup service to foreign UEs, the throughput of the own UEs remains unaffected since the network's capacity is high enough to serve UEs of both networks. However, in the case of high network load (for UE requirements equal to 5*Mbps*) as shown in Fig. 99.4b, with the increase of resource share, the throughput of foreign UEs increases at the price of decreasing the throughput of own UEs as there is not enough capacity in the unaffected network to provide full service for both groups, i.e., for all the UEs of both operators.

Clearly, for single failures any single operator can react, and single failures will not be noticed by users at all. Weather-caused impairments, as well as disasters will likely impact all the networks in the same way, since often the mobile service providers place their equipment onto the same location or onto the very same antenna mast. These multiple failures are correlated. They are correlated even between multiple network operators. Therefore, the network of one operator may often not be able to overtake the traffic of another operator because they will be likely affected by the same group failure or a disaster. However, if there is a software upgrade, or an error of it, or some configuration error, or a DDoS[1] attack or a sabotage against a single operator, it will cause correlated multiple failures in the network of a single operator only. In this case, the networks of other operators can serve the UEs from the failed network.

[1] DDoS: Distributed Denial of Service, e.g., the one caused by Mirai botnet in Autumn 2016

## 99.6 Power-aware Load Balancing for Energy Efficiency and Survivability

Interruptions in power supply due to the effects of disasters of various origin (such as earthquakes, floods, fires) or to extensive blackouts due, e.g., to cascading failures in the power distribution network, are one of the main cause of unavailability of base stations in Radio Access Networks (RANs) [31].

At the same time, the reduction of the carbon footprint of RANs has recently received a lot of attention from the research community. Indeed, for Mobile Network Operators (MNOs), a large share of operational expenditures (OPEX) is represented by power supply costs [8]. In order to decrease the carbon footprint of the network and its operational costs due to power supply, while enhancing service availability by providing power during interruptions in the supply from the grid, in recent times operators have considered the possibility of equipping Base Stations (BSs) with batteries and dedicated renewable energy sources (RES), such as solar panels, and with a system for dynamic power management, capable of managing battery charge levels, in addition to the supply from the power grid. Besides reducing operational expenditures, these base station configurations are important for 5G network deployments in those countries where power supply from the grid is either not available at all, or only intermittently available [22]. As a side effect, these base station designs allow the network to be, at least in part and over limited time intervals, autonomous from the power grid, and hence potentially capable of providing some connectivity services even after a power grid failure or shutdown due to a disastrous event.

The use of renewable sources for powering base stations has also been proposed in 5G small cell deployments. Indeed, one of the key features of the 5G paradigm in urban scenarios is the extremely dense (up to hundreds of units per km$^2$) small cell deployment required to provide high capacity and low latency services in the most crowded parts of city centers. In addition to high OPEX, such high BS density leads to high deployment costs, mainly due to BS power wiring (assuming wireless backhaul). The use of RES and batteries in at least part of the small cell base stations in these scenarios may enable BS deployments in locations where wired power supply is unfeasible, but it may also mitigate the impact of such deployments on the overall power consumption of the mobile network, by complementing grid power supply in periods of high energy costs and low power availability [33].

The network scenario which emerges from such new base station architecture, and which is likely to become increasingly common in future 5G networks is composed by a variety of base stations (implementing macrocells and small cells) with a heterogeneity of power requirements induced by the differences in hardware architecture and in the amount of users served, and with a different combination of power sources available induced by local availability, power cost considerations, and by constraints due to installation and power wiring.

As already mentioned, a key enabler of such scenarios is a base station power management system capable of efficiently managing battery charge in order to minimize the power expenditure while minimizing (in those base stations which are disconnected from the grid or for which power supply from the grid is not available due to power outages) the likelihood of base station switch-off due to a battery depletion [37].

However, in scenarios with uneven spatial distribution of users and of unavailable BSs, *system level* techniques are required in order to allow the surviving BSs to take over the traffic load of those BSs which are not operating. To this end, the surviving BSs may tune the number of users associated to each BS, possibly acting on BS transmit power for scenarios in which users associate to the BS with the strongest received signal, or acting on user association criteria, with the goal of guaranteeing a minimum QoS to all users, while maximizing the availability of surviving BSs. This may be done by taking into account in computing a feasible load distribution configuration, of the power budget available at each BS, as well as of the power model of each device (i.e. of the relationship between BS configuration, amount of traffic served, and consumed power), and of the specific device-level strategy for power management.

As an example, authors in [33] have considered the impact of a system level techniques on a setup for which BS locations, antenna orientations, and angular beamwidths, as well as traffic information in terms of number of active calls over time handled by each BS, are derived from the data of a large Italian operator. The considered geographical area corresponds to the center of the city of Pisa, Italy (Fig. 99.5). The area contains 16 BS sectors which we assume each associated with an independent LTE BS, and distributed over 6 locations. All BSs in a same location are served by the same power supply system, and hence share batteries and solar panels. We adopt the COST 231-Hata model, which is widely used to model urban environments, along with lognormal shadowing with a standard deviation of 4 dB, to model propagation conditions. We assume a frequency reuse factor of 3.

A spatially inhomogeneous Poisson process has been used to model the spatial load at any given time. All BSs have been assumed to have a maximum nominal power consumption of 500 W, except for the BS with omnidirectional antenna, consuming a maximum of 1500 W. Coherently with models proposed for recent LTE BS HW, the energy proportionality ratio ([33], Section II-A) has been assumed to be 60%. For the total amount of illumination in a day, the solar data for Pisa has been considered for to the month with less illumination (December).

The operational window has been assumed to start at 12PM and lasting 24h, subdivided into seven time slots. The choice of time slots number and duration has been done in order to obtain an acceptable accuracy of spatio-temporal patterns, while minimizing the number of network reconfigurations. In each time slot, the values of mean illumination and mean request arrival rate have been set equal to, respectively, the average and the maximum (over the 30-min average) over the time slot duration. For more details about the setup, please refer to [33].
The results of the technique proposed for tuning of the BS operating point are summarized in Fig. 99.6, where they are compared with a CAPEX optimized for a configuration with no tuning. The reduction in CAPEX of the power supply system of the network is of 69%, slightly larger than the reduction in the total energy con-

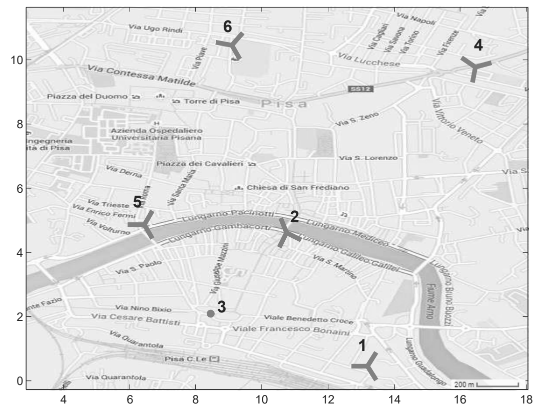



Fig. 99.5: Location, antenna orientation, and angular beamwidth for the considered BSs in downtown Pisa ([33]). The red point indicates an omnidirectional antenna.

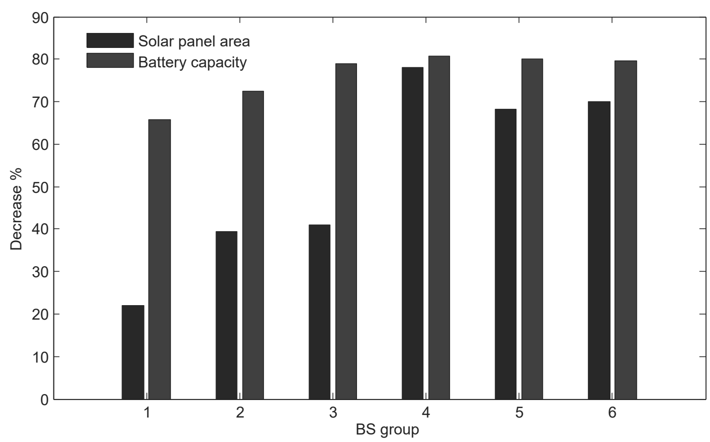



Fig. 99.6: Percentage of decrease in total panel area and battery capacity with respect to the CAPEX-optimal configuration without tuning of BS activity status and transmit power ([33]).

sumed during 24h (62%). The larger CAPEX saving is due to BSs sharing resources within each group, which compensates the (generally) larger battery costs due to differences in BS consumed energy between consecutive slots.

A crucial aspect of such approaches to enhancing survivability of RANs is that it results into a network which is able by design of self organizing and of automatically adapting to available and forecasted power supply and traffic demand, both during normal operation, and after a disastrous event, and to automatically find its optimal mode of operation. Indeed, even during normal operation, future power distribution scenarios are bound to be characterized by significant fluctuations in power prices during the same day, due to the ever higher fluctuations in power production induced by the progressive integration of renewable power sources into the grid. In addition, the progressive realization of the 5G paradigm is expected to witness a rapid increase not only in average traffic demand, but also of the difference between peaks and lows in traffic demands. These aspects will force future 5G networks to operate in a way which is aware of such fluctuations and of their possible evolution in time, in order to optimize their operational expenditures.

## 99.7 Summary

In this section, some of the key techniques to enhance the resilience of 5G systems were discussed, with focus on the fronthaul and on the backhaul parts of the network. The FFB scheme was designed to enable a fallback to use frequency band or radio technology being less affected by temporal weather events. The BSI, in turn, incorporates various interleaving schemes in the access and in the metro-access parts of the network to increase the availability of services after failures occurring in the backhaul part of the system. The MOP scheme was proposed to address the resilience issues in a multi-operator scenario when one of the operators fails to deliver services, and the backup service can be provided to its users by another operator. Finally, system level techniques for dynamically managing the operating point of base stations enable the network to cope with outages by redistributing the load on surviving base stations in a QoS and resource aware manner, and taking into account the power budget of each base station, the forecast about energy availability (from battery and renewables), as well as predictions on the evolution of traffic demand.

All these approaches are expected to provide a remarkable enhancement of service availability in case of any correlated multiple failure scenario including bad weather conditions, disasters or attacks. These techniques can be used separately one-by-one, in any combination or simultaneously all. CAPEX for deployment of any of the methods is negligible, but a certain OPEX growth due to the added complexity in managing the network is inevitable. For the FBB, BSI and MOP approaches, the enhanced availability comes at a price of decreased throughput on the occurrence of a failure. However, only the rate of either less critical or of all services is reduced, without traffic losses.

**Acknowledgements** This chapter is based on work from COST Action CA15127 ("Resilient communication services protecting end-user applications from disaster-based failures – RECODIS") supported by COST (European Cooperation in Science and Technology).

## References

1. 3GPP Release 15 https://www.etsi.org/deliver/etsi_ts/123500_123599/123502/15.02.00_60/ts_123502v150200p.pdf"
2. N. Abbas, Y. Zhang, A. Taherkordi and T. Skeie, "Mobile Edge Computing: A Survey," in IEEE Internet of Things Journal, vol. 5, no. 1, pp. 450-465, Feb. 2018. doi:

10.1109/JIOT.2017.2750180
3. Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., Flinck, H.: Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions, IEEE Communications Surveys & Tutorials, 20(3), 2429–2453 (2018)
4. M. Agiwal, A. Roy and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," in IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1617-1655, thirdquarter 2016. doi: 10.1109/COMST.2016.2532458
5. Almeida, R., Madeira, H., "Evolving from Dependability to Resilience Benchmarks: Issues and Possibilities", in Proc. 7th Latin-American Symposium on Dependable Computing (LADC'16), 127–130 (2016)
6. J. G. Andrews et al., "What Will 5G Be?," in IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065-1082, June 2014. doi: 10.1109/JSAC.2014.2328098
7. F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta and P. Popovski, "Five disruptive technology directions for 5G," in IEEE Communications Magazine, vol. 52, no. 2, pp. 74-80, February 2014. doi: 10.1109/MCOM.2014.6736746
8. Budzisz, L., Ganji, F., Rizzo, G., Marsan, M.A., Meo, M., Zhang, Y., Koutitas, G., Tassiulas, L., Lambert, S., Lannoo, B., Pickavet, M., Conte, A., Haratcherev, I., Wolisz, A.: Dynamic Resource Provisioning for Energy Efficiency in Wireless Access Networks: A Survey and an Outlook, IEEE Communications Surveys & Tutorials, 16(4), 2259–2285 (2014)
9. Y. Cai, Z. Qin, F. Cui, G. Y. Li and J. A. McCann, "Modulation and Multiple Access for 5G Networks," in IEEE Communications Surveys & Tutorials, vol. 20, no. 1, pp. 629-646, Firstquarter 2018. doi: 10.1109/COMST.2017.2766698
10. Cotroneo, D., De Simone, L., Natella, R.: NFV-Bench: A Dependability Benchmark for Network Function Virtualization Systems, IEEE Transactions on Network and Service Management, 14(4), 934–948 (2017)
11. Esposito, C., Gouglidis, A., Hutchison, D., Gurtov, A., Helvik, B.E., Heegaard, P.E., Rizzo, G., Rak, J.: On the Disaster Resiliency within the Context of 5G networks: The RECODIS Experience", in Proc. 2018 European Conference on Networks and Communications (EuCNC'18), 1–4 (2018)
12. T. Gomes et al., "A survey of strategies for communication networks to protect against large-scale natural disasters," 2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM), Halmstad, 2016, pp. 11-22. doi: 10.1109/RNDM.2016.7608263
13. Gonzalez, A.J., Nencioni, G., Kamisinski, A., Helvik, B.E., Heegaard, P.E., Dependability of the NFV Orchestrator: State of the Art and Research Challenges,IEEE Communications Surveys & Tutorials, 20(4), 3307–3329 (2018)
14. Heegaard, P.E., Helvik, B.E., Mendiratta, V.B.: Achieving Dependability in Software-defined Networking – A Perspective, in Proc. 7th International Workshop on Reliable Networks Design and Modeling (RNDM'15), 63–70 (2015)
15. IEEE 802.11 Wireless Local Area Networks
16. GSMA Intelligence, "Understanding 5G: Perspectives on future technological advancements in mobile", white paper (2014)
17. ITU-R, "Recommendation ITU-R P.676-9, Attenuation by atmospheric gases", P Series, Radiowave propagation, 2012 Feb
18. V. Jungnickel et al., "The role of small cells, coordinated multipoint, and massive MIMO in 5G," in IEEE Communications Magazine, vol. 52, no. 5, pp. 44-51, May 2014. doi: 10.1109/MCOM.2014.6815892
19. Kasgari, A.T.Z., Saad, W.: Stochastic Optimization and Control Framework for 5G Network Slicing with Effective Isolation, in Proc. 52nd Annual Conference on Information Sciences and Systems (CISS'18), 1–6 (2018)
20. Laprie, J.-C.: From Dependability to Resilience, in Proc. 38th IEEE/IFIP Int. Conf. on Dependable Systems and Networks (DSN'08).
21. J. Luomala and I. Hakala, "Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks," 2015 Federated Conference on Computer Science and Information Systems (FedCSIS), Lodz, 2015, pp. 1247-1255, doi: 10.15439/2015F241

22. Marsan, M.A., Bucalo, G., Di Caro, A., Meo, M., Zhang, Y.: Towards Zero Grid Electricity Networking: Powering BSs with Renewable Energy Sources, in Proc. ICC'13 Workshops, 596–601 (2013)
23. Malhotra, M., Trivedi, K.S.: Dependability Modeling using Petri-nets, IEEE Transactions on Reliability, 44(3), 428–440 (1995)
24. Natella, R., Cotroneo, D., Madeira, H.S.: Assessing Dependability with Software Fault Injection: A Survey, ACM Computing Survey, 48(3) 1–55 (2016)
25. Nencioni, G., Helvik, B.E., Gonzalez, A.J., Heegaard, P.E., Kamisinski, A., Availability Modelling of Software-Defined Backbone Networks, in Proc. 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop (DSN-W'16), 105–112 (2016)
26. Nencioni, G., Helvik, B.E., Heegaard, P.E.: Including Failure Correlation in Availability Modeling of a Software-Defined Backbone Network, IEEE Transactions on Network and Service Management, 14(4), 1032–1045 (2017)
27. Nguyen, T.A., Han, K., Min, D., Choi, E. Thang, T.D., Choi, Y.-J.: A Stochastic Reward Net-based Assessment of Reliability, Availability and Operational Cost for a Software-Defined Network Infrastructure, The Journal of Supercomputing, 1–27 (2018)
28. J. Rak, Resilient Routing in Communication Networks, Springer, 2015.
29. T. S. Rappaport et al., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," in IEEE Access, vol. 1, pp. 335-349, 2013.
30. J. Rak et al., "RECODIS: Resilient Communication Services Protecting End-user Applications from Disaster-based Failures," 2016 18th International Conference on Transparent Optical Networks (ICTON), Trento, 2016, pp. 1-4. doi: 10.1109/ICTON.2016.7550596
31. Reuter, Ch.: Communication between Power Blackout and Mobile Network Overload, International Journal of Information Systems for Crisis Response and Management, 6(2), 38–53 (2014)
32. Rizzo, G., Marsan, M.A.,: QoS-Aware CAPEX Minimization in Urban Off-grid Radio Access Networks, in Proc. 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS'16), 179–188 (2016)
33. Rizzo, Gianluca and Marsan, Marco Ajmone: QoS-Aware CAPEX minimization in urban off-grid radio access networks, in IEEE 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 179–188 (2016)
34. M. Tornatore, J. Andre, P. Babarczi, T. Braun, E. Folstad, P. Heegaard, A. Hmaity, M. Furdek, L. Jorge, W. Kmiecik, C. Mas Machuca, L. Martins, C. Medeiros, F. Musumeci, A. Pasic, J. Rak, S. Simpson, R. Travanca, A. Voyiatzis, "A Survey on Network Resiliency Methodologies against Weather-based Disruptions", Proc. 8th International Workshop on Design of Resilient Communication Networks (RNDM 2016), pp. 1-12, IEEE Xplore
35. Vizarreta, P., Trivedi, K., Helvik, B., Heegaard, P., Blenk, A., Kellerer, W., Mas Machuca, C.: Assessing the Maturity of SDN Controllers With Software Reliability Growth Models,IEEE Transactions on Network and Service Management, 15(3), 1090–1104 (2018)
36. Wen, R., Tang, J., Quek, T.Q.S., Feng, G., Wang, G., Tan, W.: Robust Network Slicing in Software-Defined 5G Networks, in Proc. GLOBECOM'17 2017, 1–6 (2017)
37. Zhang, Y., Meo, M., Gerboni, R., Marsan, M.A.: Minimum Cost Solar Power Systems for LTE Macro Base Stations, Computer Networks, 112, 12–23 (2017)
38. Avizienis, A., Laprie, J. C., Randell, B., Landwehr, C.: Basic concepts and taxonomy of dependable and secure computing. IEEE transactions on dependable and secure computing, 1(1), 11-33 (2004)
39. Karakus, M., Durresi, A.: A survey: Control plane scalability issues and approaches in software-defined networking (SDN). Computer Networks, 112, 279-293 (2017)
40. Oktian, Y. E., Lee, S., Lee, H., Lam, J.: Distributed SDN controller system: A survey on design choice. computer networks, 121, 100-111 (2017)
41. Foukas, X., Patounas, G., Elmokashfi, A., Marina, M. K. Network slicing in 5G: Survey and challenges. IEEE Communications Magazine, 55(5), 94-100 (2017)

42. Sachs, J., Wikstrom, G., Dudda, T., Baldemair, R., Kittichokechai, K.: 5G radio network design for ultra-reliable low-latency communication. IEEE Network, 32(2), 24-31 (2018)
43. Abderrahim, W., Choukair, Z.: Dependability integration in cloud-hosted telecommunication services, IEEE Transactions on Dependable and Secure Computing, In Press doi: 10.1109/TDSC.2018.2875088 (2018)
44. Hler, T. , Scheuvens,L., Franchi,N., Simsek, M., Fettweis, G. P., Applying reliability theory for future wireless communication networks, Proceedings of the IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1-7 (2017)
45. Sattiraju, R., Schotten, H. D.: Reliability Modeling, Analysis and Prediction of Wireless Mobile Communications, Proceedings of the IEEE 79th Vehicular Technology Conference (VTC Spring), pp. 1-6 (2014)
46. Willmann, S., Krtzig, M., Rauchhaupt, L.: Generic dependability assessment approach in wireless automation, Proceedings of the 43rd Annual Conference of the IEEE Industrial Electronics Society (IECON), pp. 6039-6044 (2017)
47. Wang, C. X., Bian, J., Sun, J., Zhang, W., Zhang, M.: A survey of 5G channel measurements and models. IEEE Communications Surveys Tutorials, 20(4), 3142-3168 (2018)